

A study that puts human nature  
under the microscope p. 510

Sparse ice stresses  
polar bears pp. 514 & 568

HLA influences cancer  
therapy response pp. 516 & 582

# Science

\$15  
2 FEBRUARY 2018  
[sciencemag.org](http://sciencemag.org)

AAAS

## A COSMOLOGICAL CONUNDRUM

Motions of Centaurus A's satellite galaxies  
challenge structure formation pp. 520 & 534

# CONTENTS

2 FEBRUARY 2018 • VOLUME 359 • ISSUE 6375



## 516 & 582

Predicting immunotherapy response



Granting complete confidentiality, landmark study has followed New Zealanders from birth.

## NEWS

### IN BRIEF

**500** News at a glance

### IN DEPTH

#### **503 INDIA PLANS TO LAND NEAR MOON'S SOUTH POLE**

Chandrayaan-2 orbiter could slake scientists' thirst for detailed data on lunar water *By P. Bagla*

#### **504 JUDGE ORDERS UNMASKING OF ANONYMOUS PEER REVIEWERS**

Bitter legal battle between two physical fitness groups produces unusual court decision *By A. Han*

#### **505 NASA SEEKS TO REVIVE LOST PROBE THAT TRACED SOLAR STORMS**

Amateur found IMAGE while looking for spy satellite *By P. Voosen*

#### **506 FOREVER YOUNG? NAKED MOLE RATS MAY KNOW THE SECRET**

A new study claims the strange rodents flout the mathematics of mortality—but not everyone is convinced *By K. Kupferschmidt*

#### **507 ACCELERATOR BOOM HONES CHINA'S ENGINEERING EXPERTISE**

Bevy of particle accelerators could pave the way for future collider that would be world's largest *By D. Normile*

#### **508 DAMS NUDGE AMAZON'S ECOSYSTEMS OFF-KILTER**

Decline of giant catfish emblematic of habitat fragmentation *By B. Fraser*  
► SCIENCE ADVANCES RESEARCH ARTICLE BY E. P. ANDERSON ET AL. 10.1126/sciadv.aao1642

### FEATURE

#### **510 HUMAN NATURE, OBSERVED**

For decades, two psychologists have kept watch over 1000 New Zealanders, teasing out factors that shape a life's course *By D. Starr*

► PODCAST

## INSIGHTS

### PERSPECTIVES

#### **514 OUT OF BALANCE IN THE ARCTIC**

Polar bears have high energy requirements that rise further as a result of climate change *By J. P. Whiteman*  
► REPORT P. 568

#### **516 ENHANCING RESPONSES TO CANCER IMMUNOTHERAPY**

The capacity of antigen presentation influences responses to checkpoint immunotherapy *By P. Kvistborg and J. W. Yewdell*

► REPORT P. 582

#### **517 STEALTH REACTIONS DRIVING CARBON FIXATION**

New twists to bacterial metabolic pathways that contribute to the global carbon cycle *By S. W. Ragsdale*

► REPORTS PP. 559 & 563

#### **519 LIGHTING UP SUPERCONDUCTING STRIPES**

Terahertz radiation reveals pair density waves in underdoped cuprate superconductors *By E. Ergeçen and N. Gedik*

► REPORT P. 575

#### **520 GALAXY MOTIONS CAUSE TROUBLE FOR COSMOLOGY**

Satellite galaxy motions around a nearby galaxy seem to be at odds with dark matter models *By M. Boylan-Kolchin*

► RESEARCH ARTICLE P. 534

#### **521 TRACING SINGLE-CELL HISTORIES**

Origins of mutations in single cells during human brain development and aging are revealed *By J. H. Lee*

► RESEARCH ARTICLE P. 550; REPORT P. 555

### POLICY FORUM

#### **523 SCIENTIFICALLY ASSESS IMPACTS OF SUSTAINABLE INVESTMENTS**

Metrics can inform investors wary of "green washing" *By C. J. Vörösmarty et al.*

### BOOKS ET AL.

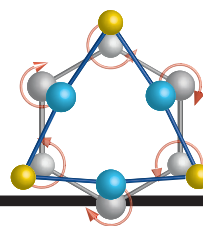
#### **526 PICTURING THE FUTURE**

A new film celebrates a mid-century illustrator who stoked imaginations with prescient comics *By D. Greenbaum*

#### **527 THE FETISHIZATION OF QUANTIFICATION**

A historian explores the dark side of metric-based performance evaluation *By T. M. Porter*

# CONTENTS



579

Phonons  
left and right

2 FEBRUARY 2018 • VOLUME 359 • ISSUE 6375

## LETTERS

### 528 MALARIA IN VENEZUELA REQUIRES RESPONSE

By *M. E. Grillet* et al.

### 528 INTEGRATED APPROACH TO MALARIA CONTROL

By *C. J. M. Koenraad* and *W. Takken*

### 529 RESPONSE

By *J. Hemingway*

## RESEARCH

### IN BRIEF

**530** From *Science* and other journals

### REVIEW

#### 533 EARTH SYSTEMS

Climate, ecosystems, and planetary futures: The challenge to predict life in Earth system models *G. B. Bonan* and *S. C. Doney*

REVIEW SUMMARY; FOR FULL TEXT:  
[dx.doi.org/10.1126/science.aam8328](https://doi.org/10.1126/science.aam8328)

### RESEARCH ARTICLES

#### 534 NEAR-FIELD COSMOLOGY

A whirling plane of satellite galaxies around Centaurus A challenges cold dark matter cosmology *O. Müller* et al.

► PERSPECTIVE P. 520; VIDEO

#### 537 STRUCTURAL BIOLOGY

Structure of a human catalytic step I spliceosome *X. Zhan* et al.

#### 545 BIOCHEMISTRY

Structure of the yeast oligosaccharyltransferase complex gives insight into eukaryotic N-glycosylation *R. Wild* et al.

#### 550 NEURODEVELOPMENT

Different mutational rates and mechanisms in human cells at pregastrulation and neurogenesis *T. Bae* et al.

► PERSPECTIVE P. 521; REPORT P. 555

## REPORTS

#### 555 NEURODEVELOPMENT

Aging and neurodegeneration are associated with increased mutations in single human neurons

*M. A. Lodato* et al.

► PERSPECTIVE P. 521; RESEARCH ARTICLE  
P. 550

#### CARBON FIXATION

**559** A primordial and reversible TCA cycle in a facultatively chemolithoautotrophic thermophile *T. Nunoura* et al.

**563** Reversibility of citrate synthase allows autotrophic growth of a thermophilic bacterium *A. Mall* et al.

► PERSPECTIVE P. 517

#### 568 ECOPHYSIOLOGY

High-energy, high-fat lifestyle challenges an Arctic apex predator, the polar bear *A. M. Pagano* et al.

► PERSPECTIVE P. 514; VIDEO

#### 572 SURFACE SCIENCE

Polarity compensation mechanisms on the perovskite surface  $\text{KTaO}_3(001)$  *M. Setvin* et al.

#### 575 SUPERCONDUCTIVITY

Probing optically silent superfluid stripes in cuprates *S. Rajasekaran* et al.

► PERSPECTIVE P. 519

#### 579 TOPOLOGICAL MATTER

Observation of chiral phonons *H. Zhu* et al.

#### 582 CANCER IMMUNOTHERAPY

Patient HLA class I genotype influences cancer response to checkpoint blockade immunotherapy *D. Chowell* et al.

► PERSPECTIVE P. 516

#### 587 MOLECULAR BIOLOGY

The piRNA targeting rules and the resistance to piRNA silencing in endogenous genes

*D. Zhang* et al.

514 & 568



#### 592 COLON CANCER

Patients with familial adenomatous polyposis harbor colonic biofilms containing tumorigenic bacteria *C. M. Dejea* et al.

## DEPARTMENTS

### 499 EDITORIAL

Our science, our society  
By *Susan Hockfield*

### 602 WORKING LIFE

In praise of slow  
By *Irene Nobel*

## ON THE COVER



Multiwavelength composite image of the elliptical galaxy Centaurus A, revealing a prominent dust lane and jets emanating from its central supermassive black hole. Dwarf satellite galaxies that orbit Centaurus A beyond the field of view are found to rotate in a coherent plane, perpendicular to the dust lane. That rotating plane of satellite galaxies challenges the standard cosmological paradigm. See pages 520 and 534. *Image: ESO/WFI (optical data); MPIfR/ESO/APEX/A. Weiss et al. (submillimeter data); NASA/CXC/CfA/R. Kraft et al. (x-ray data)*

Science Staff .....	498
New Products .....	598
Science Careers .....	599

SCIENCE (ISSN 0036-8075) is published weekly on Friday, except last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue, NW, Washington, DC 20005. Periodicals mail postage (publication No. 484460) paid at Washington, DC, and additional mailing offices. Copyright © 2018 by the American Association for the Advancement of Science. The title SCIENCE is a registered trademark of the AAAS. Domestic individual membership, including subscription (12 months): \$165 (\$74 allocated to subscription). Domestic institutional subscription (51 issues): \$1808; Foreign postage extra: Mexico, Caribbean (surface mail) \$55; other countries (air assist delivery): \$89. First class, airmail, student, and emeritus rates on request. Canadian rates with GST available upon request. GST #R125488122. Publications Mail Agreement Number 1069624. Printed in the U.S.A. Change of address: Allow 4 weeks, giving old and new addresses and 8-digit account number. Postmaster: Send change of address to AAAS, P.O. Box 96178, Washington, DC 20090-6178. Single-copy sales: \$15 each plus shipping and handling; bulk rate on request. Authorization to reproduce material for internal or personal use under circumstances not falling within the fair use provisions of the Copyright Act is granted by AAAS to libraries and others who use Copyright Clearance Center (CCC) Pay-Per-Use services provided that \$35.00 per article is paid directly to CCC, 222 Rosewood Drive, Danvers, MA 01923. The identification code for Science is 0036-8075. Science is indexed in the Reader's Guide to Periodical Literature and in several specialized indexes.

**Editor-in-Chief** Jeremy Berg

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

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

## Research and Insights

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

## News

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

**Executive Publisher** Rush D. Holt

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

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

**MARKETING MANAGER, PUBLISHING** Shawana Arnold **MARKETING ASSOCIATE** Steven Goodman **CREATIVE DIRECTOR** Scott Rodgerson **SENIOR ART ASSOCIATES** Paula Fry **ART ASSOCIATE** Kim Huynh

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

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

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

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

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

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

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

**AAAS BOARD OF DIRECTORS, CHAIR** Barbara A. Schaal **PRESIDENT** Susan Hockfield **PRESIDENT-ELECT** Margaret A. Hamburg **CHIEF EXECUTIVE OFFICER** Rush D. Holt **BOARD** Cynthia M. Beall, May R. Berenbaum, Carlos J. Bustamante, Kaye Husbands Fealing, Stephen P.A. Fodor, S. James Gates, Jr., Michael S. Gazzaniga, Laura H. Greene, Mercedes Pascual

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

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

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

**INFORMATION FOR AUTHORS** See [www.sciencemag.org/authors/science-information-authors](http://www.sciencemag.org/authors/science-information-authors)

## BOARD OF REVIEWING EDITORS (Statistics board members indicated with \$)

Adriano Aguzzi, *U. Hospital Zürich*  
Takuzo Aida, *U. of Tokyo*  
Leslie Aiello, *Wenner-Gren Foundation*  
Judith Allen, *U. of Manchester*  
Sebastian Amigorena, *Institut Curie*  
Meinrat O. Andrae, *Max Planck Inst. Mainz*  
Paola Ariotti, *Harvard U.*  
Johan Auwerx, *EPFL*  
David Awschalom, *U. of Chicago*  
Clare Baker, *U. of Cambridge*  
Nenad Ban, *ETH Zürich*  
Franz Bauer,  *Pontificia Universidad Católica de Chile*  
Ray H. Baughman, *U. of Texas at Dallas*  
Carlo Beenakker, *Leiden U.*  
Kamran Behnia, *ESPCI*  
Yasmine Belkaid, *NIAD, NIH*  
Philip Benfey, *Duke U.*  
May Berenbaum, *U. of Illinois at Urbana-Champaign*  
Gabriele Bergers, *VIB*  
Bradley Bernstein, *Massachusetts General Hospital*  
Peer Bork, *EMBL*  
Chris Bowler, *Ecole Normale Supérieure*  
Ian Boyd, *U. of St. Andrews*  
Emily Brodsky, *U. of California, Santa Cruz*  
Ron Brookmeyer, *U. of California, Los Angeles (\$)*  
Christian Büchel, *UKE Hamburg*  
Dennis Burton, *The Scripps Res. Inst.*  
Carter Tribley Butts, *U. of California, Irvine*  
Gyorgy Buzsaki, *New York U. School of Medicine*  
Blanche Capel, *Duke U.*  
Mats Carlsson, *U. of Oslo*  
Ib Chorkendorff, *Denmark TU*  
James J. Collins, *MIT*  
Robert Cook-Deegan, *Duke U.*  
Lisa Coussens, *Oregon Health & Science U.*  
Alan Cowman, *Walter & Eliza Hall Inst.*  
Roberta Croce, *VU Amsterdam*  
Janet Currie, *Princeton U.*  
Jeff L. Dangl, *U. of North Carolina*  
Tom Daniel, *U. of Washington*  
Chiara Daraio, *Caltech*  
Nicolas Dauphas, *U. of Chicago*  
Frans de Waal, *Emory U.*  
Stanislas Dehaene, *Collège de France*  
Robert Desimone, *MIT*  
Claude Desplan, *New York U.*  
Sandra Díaz, *Universidad Nacional de Córdoba*  
Dennis Discher, *U. of Pennsylvania*  
Gerald W. Dorn II, *Washington U. in St. Louis*  
Jennifer A. Doudna, *U. of California, Berkeley*  
Bruce Dunn, *U. of California, Los Angeles*  
William Dunphy, *Caltech*  
Christopher Dye, *WHO*  
Todd Ehlers, *U. of Tübingen*  
Jennifer Elisseeff, *Johns Hopkins U.*  
Tim Elston, *U. of North Carolina at Chapel Hill*  
Barry Everitt, *U. of Cambridge*  
Vanessa Ezenwa, *U. of Georgia*  
Ernst Fehr, *U. of Zürich*  
Anne C. Ferguson-Smith, *U. of Cambridge*  
Michael Feuer, *The George Washington U.*  
Toren Finkel, *NHLBI, NIH*  
Kate Fitzgerald, *U. of Massachusetts*  
Peter Fratzl, *Max Planck Inst. Potsdam*  
Elaine Fuchs, *Rockefeller U.*  
Eileen Furlong, *EMBL*  
Jay Gallagher, *U. of Wisconsin*  
Daniel Geschwind, *U. of California, Los Angeles*  
Karl-Heinz Glassmeier, *TU Braunschweig*  
Ramon Gonzalez, *Rice U.*  
Elizabeth Grove, *U. of Chicago*  
Nicolas Gruber, *ETH Zürich*  
Kip Guy, *U. of Kentucky College of Pharmacy*  
Taekjip Ha, *Johns Hopkins U.*  
Christian Haass, *Ludwig Maximilians U.*  
Sharon Hammes-Schiffer, *U. of Illinois at Urbana-Champaign*  
Wolf-Dietrich Hardt, *ETH Zürich*  
Michael Hasselmo, *Boston U.*  
Martin Heimann, *Max Planck Inst. Jena*  
Ykä Helariutta, *U. of Cambridge*  
Janet G. Hering, *Swiss Fed. Inst. of Aquatic Science & Technology*  
Kai-Uwe Hinrichs, *U. of Bremen*  
David Hodell, *U. of Cambridge*  
Lora Hooper, *UT Southwestern Medical Ctr. at Dallas*  
Raymond Huey, *U. of Washington*  
Fred Hughson, *Princeton U.*  
Randall Hulet, *Rice U.*  
Auke Ijspeert, *EPFL*  
Stephen Jackson, *USGS Southwest Climate Science Ctr.*  
Seema Jayachandran, *Northwestern U.*  
Kai Johnson, *EPFL*  
Peter Jonas, *Inst. of Science & Technology Austria*  
Matt Kaeblerlein, *U. of Washington*  
William Kaelin Jr., *Dana-Farber Cancer Inst.*  
Daniel Kammen, *U. of California, Berkeley*  
Abby Kavner, *U. of California, Los Angeles*  
Hitoshi Kawakatsu, *U. of Tokyo*  
Masashi Kawasaki, *U. of Tokyo*  
V. Narry Kim, *Seoul Nat. U.*  
Robert Kingston, *Harvard Medical School*

Etienne Koechlin, *Ecole Normale Supérieure*  
Alexander Koldokin, *Johns Hopkins U.*  
Thomas Langer, *U. of Cologne*  
Mitchell A. Lazar, *U. of Pennsylvania*  
David Lazer, *Harvard U.*  
Thomas Lecuit, *IBDM*  
Virginia Lee, *U. of Pennsylvania*  
Stanley Lemon, *U. of North Carolina at Chapel Hill*  
Ottoline Leyser, *U. of Cambridge*  
Wendell Lim, *U. of California, San Francisco*  
Marcia C. Linn, *U. of California, Berkeley*  
Jianguo Liu, *Michigan State U.*  
Luis Liz-Marzán, *CIC bioGUNE*  
Jonathan Losos, *Harvard U.*  
Ke Lu, *Chinese Acad. of Sciences*  
Christian Lüscher, *U. of Geneva*  
Laura Machesky, *Cancer Research UK Beatson Inst.*  
Anne Magurran, *U. of St. Andrews*  
Oscar Marín, *King's College London*  
Charles Marshall, *U. of California, Berkeley*  
Christopher Marx, *U. of Idaho*  
C. Robertson McClung, *Dartmouth College*  
Rodrigo Medellín, *U. of Mexico*  
Graham Medley, *London School of Hygiene & Tropical Med.*  
Jane Memmott, *U. of Bristol*  
Tom Misteli, *NCI, NIH*  
Yasushi Miyashita, *U. of Tokyo*  
Mary Ann Moran, *U. of Georgia*  
Richard Morris, *U. of Edinburgh*  
Alison Motsinger-Reif, *NC State U. (\$)*  
Daniel Neumark, *U. of California, Berkeley*  
Kitty Nijmeijer, *TU Eindhoven*  
Helga Nowotny, *Austrian Council*  
Rachel O'Reilly, *U. of Warwick*  
Joe Orenstein, *U. of California, Berkeley & Lawrence Berkeley Nat. Lab.*  
Pilar Ossorio, *U. of Wisconsin*  
Harry Orr, *U. of Minnesota*  
Andrew Oswald, *U. of Warwick*  
Isabella Pagano, *Istituto Nazionale di Astrofisica*  
Margaret Palmer, *U. of Maryland*  
Steve Palumbi, *Stanford U.*  
Jane Parker, *Max Planck Inst. Cologne*  
Giovanni Parmigiani, *Dana-Farber Cancer Inst. (\$)*  
John H. J. Petrini, *Memorial Sloan Kettering Cancer Center*  
Samuel Pfaff, *Salk Inst. for Biological Studies*  
Kathrin Plath, *U. of California, Los Angeles*  
Martin Plenio, *Ulm U.*  
Albert Polman, *FOM Institute for AMOLF*  
Elvira Poloczanska, *Alfred-Wegener-Inst.*  
Philippe Poulin, *CNRS*  
Jonathan Pritchard, *Stanford U.*  
David Randall, *Colorado State U.*  
Sarah Reisman, *Caltech*  
Félix A. Rey, *Institut Pasteur*  
Trevor Robbins, *U. of Cambridge*  
Amy Rosenzweig, *Northwestern U.*  
Mike Ryan, *U. of Texas at Austin*  
Mitinori Saitou, *Kyoto U.*  
Shimon Sakaguchi, *Osaka U.*  
Miquel Salmeron, *Lawrence Berkeley National Lab*  
Jürgen Sandkühler, *Medical U. of Vienna*  
Alexander Schier, *Harvard U.*  
Wolfram Schlenker, *Columbia U.*  
Susanna Scott, *U. of California, Santa Barbara*  
Vladimir Shalaev, *Purdue U.*  
Beth Shapiro, *U. of California, Santa Cruz*  
Jay Shendure, *U. of Washington*  
Robert Siliciano, *Johns Hopkins School of Medicine*  
Uri Simonsohn, *U. of Pennsylvania*  
Alison Smith, *John Innes Centre*  
Richard Smith, *U. of North Carolina at Chapel Hill (\$)*  
Mark Smyth, *QIMR Berghofer*  
Pam Solits, *U. of Florida*  
John Speakman, *U. of Aberdeen*  
Allan C. Spradling, *Carnegie Institution for Science*  
Eric Steig, *U. of Washington*  
Paula Stephan, *Georgia State U. & Nat. Bureau of Economic Res.*  
V. S. Subrahmanian, *U. of Maryland*  
Ira Tabas, *Columbia U.*  
Sarah Teichmann, *U. of Cambridge*  
Shubha Tole, *Tata Inst. of Fundamental Research*  
Wim van der Putten, *Netherlands Inst. of Ecology*  
Bert Vogelstein, *Johns Hopkins U.*  
David Wallach, *Weizmann Inst. of Science*  
Jane-Ling Wang, *U. of California, Davis (\$)*  
David Waxman, *Fudan U.*  
Jonathan Weissman, *U. of California, San Francisco*  
Chris Wickle, *U. of Missouri (\$)*  
Ian A. Wilson, *The Scripps Res. Inst. (\$)*  
Timothy D. Wilson, *U. of Virginia*  
Yu Xie, *Princeton U.*  
Jan Zaenen, *Leiden U.*  
Kenneth Zaret, *U. of Pennsylvania School of Medicine*  
Jonathan Zehr, *U. of California, Santa Cruz*  
Len Zon, *Boston Children's Hospital*  
Maria Zuber, *MIT*

# Our science, our society

**W**e live in a scientific golden age. Never has the pace of discovery been so rapid, the range of achievements so broad, and the changing nature of our understanding so revolutionary. Science today has extraordinary powers. It reveals fundamental phenomena of our universe, catalyzes new technologies, powers new businesses, fosters new industries, and improves lives. This month's annual meeting of the American Association for the Advancement of Science (AAAS, the publisher of *Science*) in Austin, Texas, celebrates this golden age. More than 10,000 scientists, students, teachers, business people, journalists, philanthropists, science enthusiasts, government officials, and others will come together to hear and discuss talks by experts on such topics as immunology, exoplanets, election polling, and much more. Today's advances and innovations presage a future that most of us have not yet imagined.

Lamentably, we also live in a new heyday of anti-science activism. Fake news and "alternative facts" abound. Climate-change deniers occupy political office and determine environmental policy. Fears of unsubstantiated dangers delay the deployment of genetically modified foods in starving nations. The risks of nuclear power are overstated rather than carefully weighed. The anti-vaccination movement endures, and there are claims that science is as culturally determined and subjective as any other endeavor. Public figures cynically dismiss scientific findings, fostering a popular distrust of expertise and experts. All this, too, presages a different future that most of us would not want to imagine.

In this environment, how can we ensure that science prevails and continues to flourish? What can be done to get the most from this scientific golden age? We can start by recognizing the critical role of institutions in nurturing the scientific enterprise. All too often science is viewed in terms of individual achievements: what someone did to win the Nobel Prize or a MacArthur "genius" award; what someone else did to achieve tenure or to launch a billion-

dollar business. This isn't surprising. The institutions that support scientific inquiry—universities, research centers, federal funding agencies, and private philanthropies—are designed to foster individual achievements, amplify individual abilities, and protect individual efforts. Stories of discovery and success tend to focus on individuals and individual accomplishments. But achieving success in science is a team sport, and a nation's institutions make it possible for its scientists to play. For example, in the United States, embedding much of the scientific enterprise into institutions of higher education has catalyzed productive collaborations between new and seasoned scholars, and between the discovery and the transmission of knowledge.

When the focus of science is placed on individual achievement, it can neglect the importance of the institutions that make the work of science possible. That leaves our institutions open to attack. And, indeed, both science and its institutions are under attack today, with rampant skepticism about the utility of the research enterprise and higher education. Also under attack are the core principles that unite scientists and science enthusiasts: that objective reality can be discovered; that anyone can compete in a game governed by ideas; that dis-

agreements are best resolved by assembling facts to test competing views; and that science and the application of scientific principles have the capacity to improve lives. What's more, science's universal truths call together people from any background, any nation, any phenotype or genotype. These principles have guided us for centuries along the road to discovery and understanding.

The very institutions that support individual inquiry also guard democratic principles and foster human advance. They convene people with shared purpose and amplify their impact. It is easy to assume that these institutions can stand on their own, but they cannot. None of science's successes is solely "mine" or "yours." They are all "ours," and it is our shared responsibility to actively defend the institutions that enable them.

—Susan Hockfield



*Susan Hockfield is president of AAAS and president emerita of the Massachusetts Institute of Technology, Cambridge, MA, USA. [hockfield@MIT.edu](mailto:hockfield@MIT.edu)*



***"...how can we ensure that science prevails and continues to flourish?"***

“I feel like a chump.”

**Jake McDonald**, an environmental health researcher in Albuquerque, New Mexico, to *The New York Times*, on learning that diesel-fume tests he conducted on monkeys used Volkswagens rigged to produce artificially low emissions.

## IN BRIEF

Edited by **Jeffrey Brainard**

### ECOSYSTEM CONSERVATION

## China to protect wetlands



China's decision to regulate construction may help stem the loss of habitat for migrating birds.

**C**hina plans to dramatically curb commercial development of coastal wetlands, habitat that is crucial for almost 500 species of migratory birds. “It is massive good news,” says Nicola Crockford of the Royal Society for the Protection of Birds, based in Sandy, U.K. Over the past 5 decades, building of sea walls and construction on reclaimed wetlands have destroyed more than half of the tidal mudflats in China (*Science*, 9 October 2015, p. 150). To help stop the losses, the State Oceanic Administration (SOA) announced on 17 January that authorities will halt unauthorized projects and tear down illegal structures. SOA will approve coastal wetland development only if it is important for a few specific purposes, such as national defense. Although scientists and conservationists would like additional protections, including national legislation with tough penalties, they note that the central government has been cracking down on offenders and designating new reserves.

## Funding boost for CRISPR

**BIOMEDICINE** | The U.S. National Institutes of Health (NIH) will put \$190 million over 6 years into developing gene-editing tools such as CRISPR to treat diseases caused by DNA mutations. Although companies are pouring money into CRISPR for new therapies, the NIH funding is meant to solve challenges still facing such editing. Among these are improving safety tests and successfully delivering CRISPR's components into cells. The NIH research will be limited to work on somatic cells, not sperm or eggs, which means any DNA edits could not be passed to subsequent generations. The money comes from NIH's Common Fund, which supports research that cuts across NIH's 27 institutes and centers.

## Safer cigarette claim rejected

**TOBACCO REGULATION** | Philip Morris International failed last week to convince a U.S. federal advisory panel that its new smokeless cigarette could reduce the risk of harm from smoking. The company's iQOS technology, already available in 30 countries, releases nicotine by heating tobacco without burning it. The firm had hoped that in addition to a still-pending approval for the product from the U.S. Food and Drug Administration (FDA), it might get the agency's blessing to market the innovation as safer than conventional cigarettes. But FDA's Tobacco Products Scientific Advisory Committee, which consists largely of public health experts from academia, has a history of challenging the tobacco industry. The committee agreed that the iQOS system would reduce users' exposure to harmful chemicals compared with conventional cigarettes. But it decided that Phillip Morris hadn't proved its claim that smokers can reduce their risk of tobacco-related diseases by making the switch. FDA can disregard the advisory panel's recommendations but rarely does.

## New leader for French science

**RESEARCH POLICY** | Computer scientist **Antoine Petit**, 57, was named president of the €3.2 billion CNRS, France's national

research agency, on 24 January. Petit, who previously was the president of the National Institute for Computer Science and Applied Mathematics, says he wants to make Europe's largest research organization more "agile and reactive" and offer its young staff researchers higher salaries. CNRS's interim president, Anne Peyroche, a cell and molecular biologist, left the agency prematurely earlier in January. The French research ministry declined to say why exactly, but her departure may be related to allegations of image manipulation in her papers. A spokesperson for the Alternative Energies and Atomic Energy Commission, Peyroche's main employer, says the agency is investigating those accusations, which were made on the website PubPeer.

## A photo album made with DNA

**COMPUTER SCIENCE** | Think of it as the world's smallest time capsule. Researchers at the University of Washington (UW) in Seattle announced last week that they have begun collecting 10,000 photos that they will digitize and encode as strands of synthetic DNA for future generations to access. The effort, known as the #MemoriesInDNA Project, is expected to be the largest ever to store digital data in DNA. Two years ago, the same group reported that it had recorded 200 megabytes of data, including 100 books and a Miles Davis jazz recording, using 2 billion nucleic acid bases—the adenines, thymines, cytosines, and guanines that make up DNA. The 10,000 photos will top that, depositing a gigabyte of data into tens of billions of bases. The point isn't just to increase the storage volume but also to help researchers look for new ways to search DNA-encoded information—such as finding all the images showing a red car—without first decoding each image and converting it into a digital format, says the group's leader, UW computer scientist Luis Ceze. To submit a photo of your own to be incorporated, visit the group's webpage at [memoriesindna.com](http://memoriesindna.com).

## FDA axes monkey study

**ANIMAL RESEARCH** | The U.S. Food and Drug Administration (FDA) in Silver Spring, Maryland, last week pulled the plug on a monkey study at its National Center for Toxicological Research because of animal welfare concerns and announced a broad review of the agency's animal research programs. The study, which observed the effects of nicotine withdrawal in young squirrel monkeys to better understand addiction in adolescents and

young adults, drew widespread attention after primatologist Jane Goodall wrote to FDA Commissioner Scott Gottlieb last September condemning it as unethical. FDA suspended the research pending an independent review. Research advocates, fearing the agency had unduly buckled to pressure from animal rights activists, demanded more justification for the decision. Last week's announcement explains that the suspension was prompted by the deaths of four monkeys; FDA said it found problems with the care and oversight provided by a "third-party animal welfare contractor." In a statement, Gottlieb announced plans for an independent, third-party investigation of

the agency's animal research programs, which include approximately 270 non-human primates. And he described a newly created oversight body within FDA, the Animal Welfare Council. Unlike existing, federally mandated committees that oversee individual studies, the council will track animal research across FDA and report to its Office of the Chief Scientist.

## Rector resigns in Ecuador

**RESEARCH UNIVERSITIES** | The divisive head of a university that founders hoped would become a research powerhouse in the Andes has resigned. Mathematical biologist **Carlos Castillo-Chávez** says he planned




James Reilly worked as an exploration geologist and flew on the International Space Station.

### SCIENCE AGENCIES

## Former astronaut picked to lead U.S. Geological Survey

**P**resident Donald Trump plans to nominate **James Reilly**, a former NASA astronaut and exploration geologist, to lead the U.S. Geological Survey, the White House announced last week. If confirmed, the 63-year-old would lead a \$1.1 billion science agency whose researchers monitor for earthquakes and volcanic eruptions, among other duties. Prior to NASA, Reilly worked for Enserch Exploration, an oil and gas company based in Dallas, Texas. Reilly, who holds a geoscience Ph.D. from the University of Texas in Dallas, retired from NASA in 2008 and currently serves as a technical adviser on space operations at the U.S. Air Force's National Security Space Institute in Colorado Springs, Colorado.



Sensors can switch  
off turbines before birds  
collide with them.

## RENEWABLE ENERGY

### AI saves birds from turbines

**A**t a Wyoming wind farm, an energy company will install a new kind of sensor that shuts down turbines when eagles are flying near to prevent collisions, the company announced last week. Wind farms each year kill thousands of birds of many species, including federally protected ones. The fatalities have aroused public opposition to the spread of this renewable energy source. Duke Energy, based in Charlotte, North Carolina, acted after pleading guilty in 2013 to criminal charges that turbines at two of its Wyoming wind farms killed more than 150 migratory birds of various species. Duke will deploy 24 bird detection devices at one of those farms, which has 110 turbines. The bird detection devices, made by Identiflight, based in Louisville, Colorado, use pole-mounted cameras and artificial intelligence to identify eagles and monitor their flight paths. When an eagle flies too close to a turbine, the networked devices will cause it to shut down within about 30 seconds.

to serve out his 2-year term as rector of Yachay Tech University in Ecuador's Urcuquí province but quit 11 months early to support his daughter through a medical crisis. The move comes 6 months after several scientists in leadership positions were fired abruptly, leading to accusations that Castillo-Chávez was undermining the university's mission (*Science*, 28 July 2017, p. 340). Castillo-Chávez calls the changes "necessary reforms" and says that Yachay Tech has "advanced a lot since I arrived." But Camilo Montes, a geologist who left last month, says, "The situation is just getting more and more unstable. ... The future of Yachay is done, with or without [Castillo-Chávez]."

### NIH requires 'life span' trials

**CLINICAL RESEARCH** | The U.S. National Institutes of Health (NIH) has expanded an existing rule that requires children's participation in clinical trials to embrace people across the life span, including the oldest Americans. NIH announced last week that, beginning in January 2019, applicants seeking funds for trials must present a plan for including subjects of all ages—or justify their exclusion, which must be for scientific or ethical reasons. The change will also compel researchers to report to NIH the age of every enrolled subject, numbers the agency will make public. The changes were mandated by a 2016 law, the 21st Century Cures Act. NIH has required participation of children in clinical studies since 1998, but advocates for pediatric research have been frustrated in their attempts to learn whether investigators were complying. Advocates for research on older Americans have also bemoaned data showing that they are too often underrepresented in clinical trials.

### NASA's GOLD hitchhikes a ride

**EARTH SCIENCE** | The first NASA science mission to be hosted on a commercial satellite was launched on 25 January on an Ariane 5 rocket. The Global-scale Observations of the Limb and Disk (GOLD) project is tagging along aboard a geostationary communications satellite named SES-14, rather than on a government satellite or the International Space Station. It's an economical approach NASA plans to use again, including for its Geostationary Carbon Cycle Observatory, which will be launched in several years on another commercial satellite host. The \$55 million GOLD project is designed to explore phenomena at the boundary between the upper reaches of Earth's atmosphere and space.

Downloaded from <http://science.sciencemag.org/> on February 1, 2018

PHOTO: IDENTIFLIGHT INTERNATIONAL

## IN DEPTH

## SPACE EXPLORATION

# India plans to land near moon's south pole

Chandrayaan-2 orbiter could slake scientists' thirst for detailed data on lunar water

By Pallava Bagla, in Bengaluru, India

Sometime this summer, a spacecraft orbiting over the moon's far side, out of contact with controllers on Earth, will release a lander. The craft will ease to a soft landing just after lunar sunrise on an ancient, table-flat plain about 600 kilometers from the south pole. There, it will unleash a rover into territory never before explored at the surface; all previous lunar craft have set down near the equator.

That's the ambitious vision for India's second voyage to the moon in a decade, due to launch in the coming weeks. If Chandrayaan-2 is successful, it will pave the way for even more ambitious Indian missions, such as landings on Mars and an asteroid, as well as a Venus probe, says Kailasavadivoo Sivan, chairman of the Indian Space Research Organisation (ISRO) here. Chandrayaan-2, he says, is meant to show that India has the technological prowess "to soft land on other heavenly bodies."

But lunar scientists have much at stake, too. "There has been a rebirth of lunar exploration across the globe, and India can't be left behind," says Mylswamy Annadurai, director of the ISRO Satellite Centre. Instruments aboard the lander and rover will collect data on the moon's thin envelope of plasma, as well as isotopes such as helium-3, a potential fuel for future fusion energy reactors. The orbiter itself will follow up on a stunning discovery by India's first lunar foray, the Chandrayaan-1 orbiter, which found water molecules on the moon in 2009. Before that, "It was kind of a kooky science to think that you'd find water" there, says James Greenwood, a cosmochemist at Wesleyan University in Middletown, Connecticut. "Now, we're arguing about how much water, and not whether it has water or not." Cameras and a spectrometer aboard the Chandrayaan-2 orbiter could help settle that question.

The \$150 million mission was originally meant to fly 3 years ago, but Russia failed to deliver a promised lander, prompting India to go it alone. Final preparations are underway on the Chandrayaan-2 spacecraft, which will launch from the Sriharikota spaceport on the Bay of Bengal aboard India's Geosynchronous Satellite Launch Vehicle.

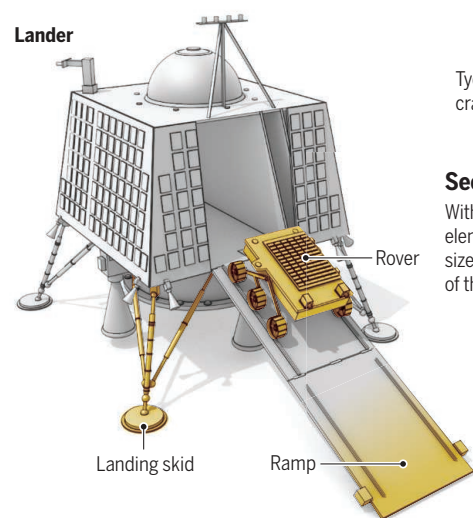
A landing so far from the lunar equator is especially tricky. "It is a difficult and complicated mission," says Wu Ji, director of the National Space Science Center in Beijing. Less sunlight reaches the poles, which means the lander and rover must be parsimonious with power. The plan is to set down in a high plain between two craters, Manzinus C and Simpelius N, at a latitude of about 70° south.

The lander will pack as much science as it can into its first lunar day—14 Earth days—as controllers may not be able to revive it after the long lunar night. The craft has a Langmuir probe to measure the moon's plasma—a wispy layer of charged ions that may explain why the lunar regolith, or dust, has a

tendency to float in the thin atmosphere. It also has a seismometer for recording moonquakes. Its seismic measurements would supplement those from the Apollo landings, because readings from high latitudes would be sensitive to signals passing through different parts of the moon. And if the seismometer is lucky enough to record a sizable quake during its operational lifetime, it might offer new evidence in a long-running debate over what the moon's core is composed of, and whether it's solid. "We just need more data to understand the lunar interior," says David Kring, a planetary geologist at the Lunar and Planetary Institute in Houston, Texas, who is not involved in the mission.

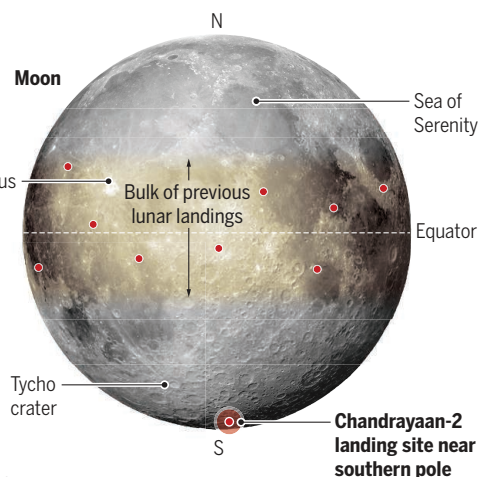
## Pole position

If all goes to plan, India's Chandrayaan-2 mission this summer will attempt a soft landing on an ancient high plain of the moon, some 600 kilometers from the south pole. It would be the first landing so far from the equator.



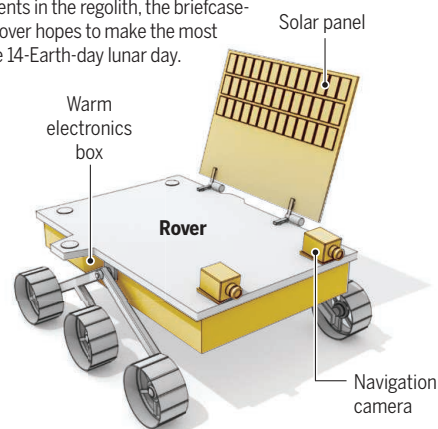
## Exploring lunar nova

The lander is equipped with a seismometer to listen for moonquakes and a Langmuir probe that will measure fluctuations in the wispy plasma enveloping the lunar surface.



## Seeking ground truth

With spectrometers for assaying elements in the regolith, the briefcase-size rover hopes to make the most of the 14-Earth-day lunar day.



The briefcase-size rover, weighing just 25 kilograms, will also carry two spectrometers for probing the lunar surface's elemental composition. The area is enticing, as it is thought to be made up of rocks more than 4 billion years old that solidified from the magma ocean that covered the newly formed moon. The data would be compared with those from Apollo-era missions that landed in other ancient highlands closer to the equator.

For some scientists, the most anticipated data will come from the orbiter's water mapper. Protons in the solar wind generate hydroxyl ions when they strike oxides in the regolith. The ions drift to the poles, where they are trapped in craters as water ice, which the orbiter will inventory. Shedding light on the moon's water circulation "is a worthwhile endeavor," says Carle Pieters, a lunar scientist at Brown University. Locating substantial water, adds Muthayya Vanitha, Chandrayaan-2's project director at ISRO, "could pave the way for the future habitation of the moon," as water is a limiting factor for operating a base.

Regardless of whether Chandrayaan-2 breaks new scientific ground, a successful soft landing near the south pole will be a technical accomplishment for India, as well as a proud moment for the country. It may even benefit other countries' moon programs. "One of NASA's main priorities is to go [to the south pole] on a sample return mission," Greenwood says, "so this could help us also later down the road as they give us more information as to what's there." ■

With reporting by Katie Langin.



To simulate the moon's anemic gravity during testing of its lunar rover, the Indian Space Research Organisation tethered a model to a helium balloon.



## SCIENTIFIC PUBLISHING

# Judge orders unmasking of anonymous peer reviewers

Bitter legal battle between two physical fitness groups produces unusual court decision

By **Andrew P. Han**, *Retraction Watch*

In what appears to be a first, a U.S. court is forcing a journal publisher to breach its confidentiality policy and identify an article's anonymous peer reviewers.

The novel order, issued last month by a state judge in California, has alarmed some publishers, who fear it could deter scientists from agreeing to review draft manuscripts. Legal experts say the case, involving two warring fitness enterprises, isn't likely to unleash widespread unmasking. But some scientists are watching closely.

The dispute revolves around a 2013 paper, since retracted, that appeared in *The Journal of Strength and Conditioning Research*. In the study, researchers at The Ohio State University in Columbus evaluated physical and physiological changes in several dozen volunteers who participated for 10 weeks in a training regimen developed by CrossFit Inc. of Washington, D.C. Among other results, they reported that 16% of participants dropped out because of injury.

In public and in court, CrossFit has alleged that the injury statistic is false. CrossFit also claims that the journal's publisher, the National Strength and Conditioning

Association (NSCA) of Colorado Springs, Colorado—which is a competitor in the fitness business—intentionally skewed the study to damage CrossFit. NSCA in turn has countersued, accusing CrossFit executives of defamation. Amid the legal crossfire, the journal first corrected the paper to reduce the number of injuries associated with CrossFit, then retracted it last year, citing changes to a study protocol that were not first approved by a university review board.

CrossFit suspects the paper's reviewers and editors worked to play up injuries associated with its regimen, and it has asked both federal and state judges to force the publisher to unmask the reviewers. In 2014, a federal judge refused that request. But last month, Judge Joel Wohlfeil of the San Diego Superior Court in California, who is overseeing NSCA's defamation suit against CrossFit, ordered the association to provide the names.

That order is an outlier, attorneys familiar with such disputes say. They note that over the past 3 decades at least three other parties have unsuccessfully asked courts to unmask reviewers, going up against publishers including the American Physical Society, Elsevier, and the Massachusetts

A paper reporting injuries associated with a popular fitness regime has sparked a yearslong court battle.

Medical Society, which publishes *The New England Journal of Medicine* (NEJM).

Academic publishers can't claim special protections for information provided in confidence, unlike journalists who can protect confidential sources under state shield laws. But judges have generally agreed with journals that unmasking reviewers would do more harm than good. In one notable 2007 case, pharma giant Pfizer subpoenaed NEJM for information about peer reviewers as part of a class action lawsuit related to the marketing of its painkillers Bextra and Celebrex. In an affidavit submitted on behalf of NEJM, Donald Kennedy, former president of Stanford University in Palo Alto, California, who was then the editor-in-chief of *Science*, stated that were Pfizer to win, "I have no doubt whatever that ... scientists would think twice about the next reviewing assignment, and that many would decide not to help." The judge quashed Pfizer's request.

Despite such high-profile decisions, journals are constantly defending reviewer confidentiality, says Paul Shaw, an attorney at Verrill Dana in Boston who has represented NEJM for 17 years. Over that time, Shaw estimates he has fielded 30 to 35 subpoenas asking for information about a peer-reviewed article—usually one related to drugs or medical devices at the center of class action lawsuits. "Invariably," he says, "one of the requests will be for the identity of and memoranda done by the peer reviewers." Usually, Shaw responds by sending a standard objection letter that puts the matter to rest.

So why did the CrossFit case take a different turn? One reason is that this time the publisher is the plaintiff, not the defendant, says Joshua Koltun, an attorney based in San Francisco, California, who has reviewed the case. Another is that the identity, conduct, and motivations of the reviewers could be key to CrossFit's defense. NSCA, he notes, "is saying: 'I'm going to sue [CrossFit] for saying I committed fraud, but I'm not going to let you see info that might prove [CrossFit's claim].' You can't have your cake and eat it, too."

Shaw, for one, believes the case is so unusual that it won't have broad ripple effects. "I don't think it has any precedential value whatsoever," he says. "For lack of a better legal descriptor, the facts seem very messy." ■

*Andrew P. Han is a reporter for Retraction Watch based in New York City. This story is the product of a collaboration between Science and Retraction Watch.*

## SPACE WEATHER

# NASA seeks to revive lost probe that traced solar storms

## Amateur found IMAGE while looking for spy satellite

By Paul Voosen

After years of silence, a NASA satellite is phoning home. Thought lost because of a power system failure, the Imager for Magnetopause-to-Aurora Global Exploration (IMAGE) was discovered last month, still broadcasting, by Scott Tilley, an amateur astronomer. NASA may find it possible to revive the mission, which once provided unparalleled views of solar storms crashing into Earth's magnetosphere. "I've discovered—or rather, 'recovered'—many satellites," Tilley says. "Nothing has resonated like this."

By day, Tilley designs marine power systems in Roberts Creek, Canada. But he spends his free time tracking the radio signals from spy satellites using antennas cribbed from his father, himself an amateur radio operator, and posting the results on his blog, Riddles in the Sky. On 20 January, Tilley was searching for evidence of Zuma, a classified U.S. satellite that's believed to have failed after launch last month. But he instead picked up a signal from a satellite identifying itself as "2000-017A," which he realized corresponded to IMAGE. Launched in 2000 and left for dead in 2005, the \$150 million mission was broadcasting. It just needed someone to listen.

After Tilley blogged about the detection, word spread to former members of IMAGE's science team, including Patricia Reiff, a space plasma physicist at Rice University in Houston, Texas. "The odds are extremely good that it's alive," she says. There also appear to be data in the radio signal beyond basic bookkeeping information, Reiff adds, a hint that some of the satellite's suite of six instruments are working.

Since Tilley's announcement, IMAGE mission scientists have been engaged in space archaeology, digging up command-and-control software backed up on obsolete tape cartridges. As of press time, NASA was continuing to try to communicate with the satellite using its deep space radio antennas. "The team is collectively holding their

breath waiting for some real information exchange between IMAGE and the ground," Reiff says.

Prior to its failure, IMAGE was already considered a successful mission. The half-ton satellite served as a sort of sentry, providing a global view of charged particles captured by the magnetosphere, the envelope around Earth dominated by its magnetic field. IMAGE's instruments detected the atoms kicked out when particles in solar storms crashed into a thin veil of hydrogen atoms cloaking the edge of Earth's atmosphere, boosting understanding and predictions of space weather. The observations have never been replaced, says Dan Baker, director of the Laboratory

for Atmospheric and Space Physics at the University of Colorado in Boulder. "We have sorely missed having the IMAGE capabilities."

IMAGE's signals first winked out in December 2005, well past its original 2-year mission. It had been working perfectly up to that point; NASA eventually attributed the loss to a misfire of the controller providing power to the satellite's radio. It remained possible, however, for IMAGE to reset itself. When Earth

blocked its solar panels from the sun for an extended time, its batteries could drain. Upon recharging after such eclipses, the spacecraft could reboot—but when it failed to do so after a 2007 eclipse, NASA declared the mission over. Several of these eclipses have occurred in the past few years, and indeed, another astronomer, Cees Bassa at the Netherlands Institute for Radio Astronomy in Dwingeloo, has now discovered in his records that IMAGE was broadcasting as far back as October 2016.

If IMAGE is revived, it will be well positioned in its orbit to monitor Earth's northern auroral zone, a region that space weather can hit hard. It's thrilling to think the spacecraft could be back, Reiff adds. Its rediscovery reminds her of the mission's motto: "The real voyage of discovery is not in seeking new landscapes, but in having new eyes." ■

***"The team is collectively holding their breath waiting for some real information exchange ..."***

Patricia Reiff,  
Rice University



## AGING RESEARCH

## Forever young? Naked mole rats may know the secret

A new study claims the strange rodents flout the mathematics of mortality—but not everyone is convinced

By **Kai Kupferschmidt**,  
in San Francisco, California

Over a lifetime of studying naked mole rats, Rochelle Buffenstein has often been surprised, even shocked, by the extraordinary features of these freakish-looking rodents. But when she recently tallied the life histories of thousands of animals born in her lab, she reached the most startling conclusion yet: There may be no such thing as a lifetime for naked mole rats. As the animals grow older (and they can grow very old indeed) their risk of dying does not seem to increase, making the concept of a maximum life span meaningless, she wrote in a paper published in *eLife* last week. “To me this is the most exciting data I’ve ever gotten,” says Buffenstein, who works at Calico, a Google biotech spinoff here that focuses on longevity. “It goes against everything we know in terms of mammalian biology.”

Although researchers who study aging

welcomed the trove of data in the paper, many cautioned against sweeping conclusions. “I think it’s too early to say naked mole rats are nonaging animals,” says João Pedro De Magalhães, a gerontologist at the University of Liverpool in the United Kingdom. Still, “It is clear that something very unusual is going on in these animals,” he says. “This study further substantiates that naked mole rats truly could unlock keys to healthy aging,” says Ewan St. John Smith, a pharmacologist who studies naked mole rats at the University of Cambridge in the United Kingdom.

Native to the Horn of Africa and Kenya, naked mole rats were already an outlier among mammalian species. Like some bees and almost all ants, the burrowing rodents are eusocial: Each colony has only a single breeding female. Their crumpled, sausage-shaped bodies are resistant to some types of pain, and they can survive 18 minutes without oxygen and rarely develop cancer.

Few have been studying their pecu-

For naked mole rats, death is “stochastic,” says Rochelle Buffenstein, who has studied the species since 1980.

liarities longer than Buffenstein. Born in Rhodesia—modern-day Zimbabwe—she started working with naked mole rats at the University of Cape Town in South Africa in 1980, where she studied with anatomist Jennifer Jarvis, who in 1981 first reported that the rodents are eusocial. On a field expedition to Kenya, the two caught many naked mole rats together, some of which became the seed of Buffenstein’s own colonies.

The animals have accompanied her to professorships in Johannesburg, South Africa; New York City; and San Antonio, Texas, where she studied their vitamin D metabolism, thermoregulation, and the energetic cost of pregnancy, among other things. “She really has been a catalyst to the field of naked mole rat biology,” says St. John Smith, not least by helping and encouraging colleagues. “For example, she was willing to provide me with an intact colony of animals,” he says.

Buffenstein now has some 3000 animals in 147 colonies, more than any other researcher; they’re housed in a basement an hour’s drive from her Calico office. The animals seem contented, noisily gnawing with their long teeth at their Plexiglas housing, made up of cylinders and boxes that each colony variously uses for storage, as a toilet, or to huddle together to sleep in a huge pink mass. Buffenstein herself is happy at Calico as well, she says. “Here you are liberated from writing grants [and] you are all fighting for the same goal, which is to find a way to slow aging.”

Researchers already knew naked mole rats can have very long lives. Mice in captivity live at most 4 years; based on their size, naked mole rats would not be expected to live past 6. But by the time Buffenstein moved to New York City in 2001, many of her animals were more than 15 years old. “I thought: ‘Wow, they are living a long time; we should start understanding why.’” For the new study, she pulled together the records of when each animal in her care was born and when it died, was killed for an experiment, or given away to other researchers.

The data suggest that it’s not just their longevity that is unusual. All mammalian species were thought to adhere to the Gompertz law, a mathematical equation that describes aging published by U.K. mathematician Benjamin Gompertz in 1825. He found that the risk of dying rises exponentially with age; in humans, for instance, it doubles roughly every 8 years after the age of 30. Scientists have subsequently shown that in most mammalian species, the law

kicks in when they reach an age two to five times their age of sexual maturity.

But not naked mole rats. After they reached sexual maturity at 6 months of age, each animal's daily chance of dying was a little more than one in 10,000, Buffenstein found, and it stayed roughly the same, and even went down a little, throughout their lives—even after they reached 25 times the age of sexual maturity. Dying, for naked mole rats, is “stochastic,” Buffenstein says. “It's like radioactive decay.”

Caleb Finch, a biogerontologist at the University of Southern California in Los Angeles, agrees that the animals' mortality is “remarkably low.” “At advanced ages, their mortality rate remains lower than any other mammal that has been documented,” he says. But more data on older naked mole rats are needed to be sure that their risk of dying really is flat, he says. Fewer than 50 in the study lived past 15 years of age, because many of the animals were killed or moved to other labs.

Buffenstein says her data are enough to show that the animals don't age. “If you look at any rodent aging study, 100 animals is all you need to see Gompertz aging. Here we have 3000 data points and we're not seeing it,” she says. “To me it says that these animals really have figured out a way to slow aging.”

Still, the oldest animal currently living in Buffenstein's lab is 35, and few are older than 30. Matthias Platzer, a biologist at the Leibniz Institute on Aging in Jena, Germany, says it's still not clear what happens when naked mole rats age beyond 30 years. “Maybe aging happens really fast then? Even Rochelle Buffenstein does not have the data on this,” he says.

The big question is how naked mole rats stay so young. Their low body temperature of about 32°C may help them avoid accumulating cellular and molecular damage, says Magalhães, and there is evidence that, compared with other mammals, they have better DNA damage repair and are more efficient at getting rid of misfolded proteins. Buffenstein says she hopes to identify a master switch controlling all these antiaging measures. “We're pretty certain that with so many things being different there has to be something upstream regulating the whole process.”

Although the naked mole rat is a special case, it could end up being the key to understanding aging in other mammals, including people, says Buffenstein, just as studying squid giant axons helped neuroscientists fathom how nerve cells work. “I would argue,” she says, “that most of our biggest discoveries in biology have been made using freak animals.” ■

## PHYSICS

# Accelerator boom hones China's engineering expertise

Bevy of particle accelerators could pave the way for future collider that would be world's largest

By **Dennis Normile**, in Dongguan, China

**A**iming a stream of protons at a target to produce neutrons for experiments is “so hard to do,” says Andrew Taylor, former director of the ISIS Neutron and Muon Source near Oxford, U.K. Most accelerators generate electron beams; accelerating and controlling protons, which are 2000 times as heavy, is far more complex and requires higher magnetic fields. The targets themselves pose challenges, too. In raising the curtain on the China Spallation Neutron Source (CSNS) here, China has joined just four other nations in having mastered the technology. The \$277 million facility, set to open to users this spring, is expected to yield big dividends in materials science, chemistry, and biology.

More world class machines are on the way. China is starting construction on four more major accelerator facilities this year (see table, below). The building boom is prompting a scramble to find enough engineers and technicians to finish the projects. But if they all come off as planned, the facilities collectively will put China “on par with or leading the world in the design, construction, and exploitation of accelerator-based research and applications,” says Wu-Tsung Weng, an

accelerator physicist at Brookhaven National Laboratory in Upton, New York.

That would position China to tackle the next global megaproject: a giant accelerator that would pick up where Europe's Large Hadron Collider (LHC) leaves off. With a \$6 billion price tag and a planned circumference of at least 50 kilometers, the Circular Electron Positron Collider (CEPC) would study the Higgs boson in detail. A possible future upgrade that would cost billions of dollars more, the Super Proton-Proton Collider, would search for physics beyond the standard model at several times the LHC's energy levels. The CEPC, which backers hope to see built by 2030, would “put China in a dominant position in fundamental physics in the world in the 21st century,” says Shing-Tung Yau, a mathematician at Harvard University. Many high-energy physicists in China and abroad think the CEPC is an obvious next step, though researchers in other disciplines worry it will monopolize available resources.

The current construction binge reflects a late start and pent-up demand. The Institute of High Energy Physics (IHEP) in Beijing inaugurated China's first particle accelerator, the Beijing Electron Positron Collider, in 1989; it was followed 2 years later by the Hefei Light Source, a facility designed to

## Accelerating Chinese science

China is ramping up its capacity in accelerator physics with several machines nearing completion or coming online in the next decade. Particle physicists hope the buildup will climax with the Circular Electron Positron Collider.

FACILITY (LOCATION)	RESEARCH TARGET	COST	STATUS/SCHEDULE
China Spallation Neutron Source (Dongguan)	Materials science, physics, chemistry, and life sciences	\$277 million	Opening to users in spring
Shanghai Soft X-ray Free-Electron Laser (FEL), Hard X-ray FEL (Shanghai)	Materials science and life sciences	\$110 million	Soft x-ray FEL opening to users in mid-2019 Hard x-ray FEL online in 2025
High-Energy Photon Source (Beijing)	Materials science, chemistry, and biomedicine	\$730 million	Mid-2025
China Initiative Accelerator Driven System (Huizhou)	Nuclear waste transmutation and future energy technologies	\$280 million	2024
High-Intensity Heavy ion Accelerator Facility (Huizhou)	Atomic and nuclear physics	\$240 million	2024
Circular Electron Positron Collider (site TBD)	Particle physics	\$6 billion	Under study

produce synchrotron radiation for studies of molecules and materials. The next major accelerator—another effort to catch up with other countries—was the Shanghai Synchrotron Radiation Facility, completed in 2009.

By then, China's physicists were thinking bigger. Next on the wish list was a powerful spallation neutron source. By the early 2000s, the United Kingdom and Switzerland had such machines in operation, and the United States, Japan, and a European consortium were either building or planning comparable facilities. It was clear that "neutron scattering would be very important for many fields," says Chen Hesheng, CSNS project manager. Neutrons interact weakly with other particles, making them better than synchrotron x-rays at probing cells or substances without damaging them. Just as the CSNS was getting underway, China's physicists got a windfall. In 2010, the National Development and Reform Commission (NDRC), China's primary funder of big science projects, agreed to build several more accelerators; four won funding under China's 13th 5-year plan, starting in 2016.

Technological development for the new facilities is ambitious, Taylor says. A hard x-ray free-electron laser, being built by the Shanghai Institute of Applied Physics (SINAP) and ShanghaiTech University, employs superconducting cavities to accelerate electrons that shed x-rays. That approach, used in only three other facilities in the world, yields higher energies, says Wang Dong, an accelerator physicist at SINAP. A heavy-ion accelerator now rising at a new complex in Huizhou, China, will have a higher beam intensity than comparable machines elsewhere, providing advantages in studying atomic and nuclear structures and the origin of heavy elements, explains Chen Xurong, a nuclear physicist heading the project for the Institute of Modern Physics in Lanzhou, China. The Huizhou complex will also host an accelerator-driven system that will experiment in nuclear waste transmutation and energy production.

Groundbreaking is slated this year for what would become one of the world's brightest x-ray sources, IHEP's High-Energy Photon Source in Beijing. It owes its intense brightness to a beam bending angle that cuts energy losses, says project manager Qin Qing. It, too, relies on superconducting cavities.

The CEPC likely would rely on similar cavities, says Lou Xinchou, an IHEP deputy director in charge of planning the megaccelerator. The High-Energy Photon Source, he notes, also requires know-how in cooling and control technologies that will pay off for the CEPC.

IHEP is now working on a conceptual design for the CEPC, due out in April. But some physicists wonder whether the scientific payoff will be worth the investment, even though China anticipates that international partners will bear a hefty share of the machine's cost. In 2016, NDRC rejected IHEP's request for \$100 million for R&D on the megacollider during China's current 5-year plan, forcing the institute to scrape together funds from other sources. "This is a great project for China and the world," Yau argues. He notes that 300 scientists from nine countries contributed to the preliminary design report.

In the meantime, the flurry of activity in planning and completing the approved accelerator projects has left IHEP staff "overloaded," says the CSNS's Chen Hesheng. His institute and others are recruiting from the Chinese diaspora to get the job done. Lou is one example. Mainland-born, he took a leave from a tenured position at the University of Texas in Austin to work on the CEPC. Adequate funding is another challenge, as construction spending squeezes budgets for operating the new facilities. The CSNS, for example, planned for five experimental instruments, but is starting with three because construction costs rose more than anticipated.

Still, as IHEP's Wang Sheng puts it: "It's a good time to be an accelerator physicist in China." ■

## ECOLOGY

# Dams nudge Amazon's ecosystems off-kilter

Decline of giant catfish emblematic of habitat fragmentation

By **Barbara Fraser**, in Lima

Once upon a time, thousands of dorados, a giant among catfish, would swim more than 3000 kilometers from the mouth of the Amazon River to spawn during the austral autumn in Bolivia's Mamoré River, in the foothills of the Andes. But the dorado, which can grow to more than 2 meters in length, is disappearing from those waters, and scientists blame two hydropower dams that Brazil erected a decade ago on the Madeira River.

"The dams are blocking the fish," says Michael Goulding, a Wildlife Conservation Society aquatic ecologist in Gainesville, Florida, who has been studying the dorado since the 1970s. They are "probably on their way to extinction" in Peru and Bolivia.

Most Amazon dams are in Brazil, where scientists have raised concerns about the displacement of local communities and emissions of greenhouse gases such as carbon dioxide and methane from large reservoirs. But as countries seek new energy sources to drive economic growth, a surge in dam construction on the eastern flank of the Andes could further threaten fish migration and sediment flows, Elizabeth Anderson, a conservation ecologist at Florida International University in Miami, and colleagues warn this week in *Science Advances*.

For the ecology of the western Amazon Basin, where the mountains meet the lowlands, the main consequence of proliferating dams is habitat fragmentation (see map, p. 509). Interference with spawning is one facet. Another is that dams hold back sediments and nutrients that nourish the Amazon Basin, Anderson says. Her team documented 142 hydropower dams that are operating or under construction on headwaters in the western Amazon Basin, and another 160 that are under consideration. If even a fraction of the planned projects



The new China Spallation Neutron Source is one of just a handful of comparable facilities worldwide.



Catches of dorado are in decline as the catfish is failing to navigate two dams on its epic spawning run.

is completed, the habitat disruption could have a cascade of ecosystem effects with devastating consequences, scientists say.

The disappearance of the dorado (*Brachyplatystoma rousseauxii*) from the Mamoré River suggests fragmentation is already taking a toll. And that's despite features of the dams that are meant to mitigate their impact. The Madeira dams, for example, are designed to allow fish to pass: The lower dam has a bypass channel and the upper dam has an enclosure in which fish are captured. They are then trucked upstream for release into the reservoir. Perhaps because of variations in currents or water chemistry, dorados are not using the channel, says Carolina Rodrigues da Costa Doria, an ichthyologist at the Federal University of Rondônia in Porto Velho, Brazil. The threat may not be limited to fish: Freshwater dolphins and river otters may also migrate along Amazonian rivers, and how dams affect their behavior is unknown, says Paul Van Damme, director of the Institute for Applied Research on Water Resources, a research center in Cochabamba, Bolivia.

During the rainy season—from November to May—water levels rise in the Amazon Basin, flooding large tracts of forest. Various species of fish swim into the forest, where they feed on fruits—and later disperse seeds. By blocking migration routes or changing water levels, dams change seed dispersal patterns, says Sandra Bibiana Correa,

a freshwater ecologist at Mississippi State University in Starkville. The fish-forest pas de deux “is a really delicate interaction,” she says. “It has been going on for tens of millions of years. We can disrupt that very easily.”

Some species are taking advantage of the disorder. Another kind of giant catfish known as the manitoa or piramutaba once rarely ventured upstream of rapids that predated the Madeira River dams. Unlike its cousin, this species (*B. vaillantii*) can make it through both dam bypasses and into the upper reservoir, and from there swims another

1000 kilometers or so upstream to Peru's Madre de Dios watershed. Whether the newcomer will fill the dorado's ecological role or prey on different fish and thus skew species assemblages is unclear, says Carlos Cañas, a river ecologist at the Wildlife Conservation Society here, who plans to monitor the migration of large catfish in the watershed.

Two other threats—climate change and the deforestation that accompanies road building during dam construction—could amplify the severity of ecological deterioration, Anderson says. At stake, she says, are the livelihoods of

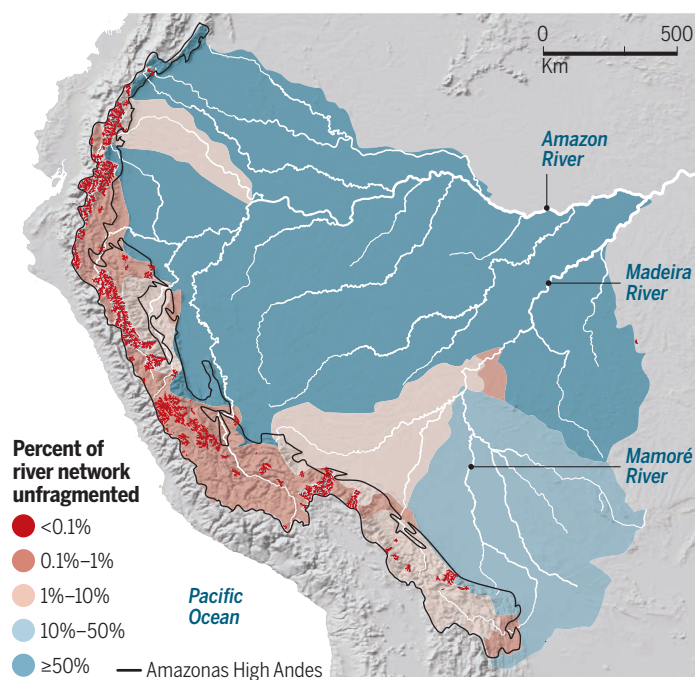
indigenous peoples who depend on fishing. Other projects, such as Peru's plans to dredge rivers to improve navigability, could exacerbate the ecological impact by changing flows, disturbing spawning sites, and disrupting the river-forest connection, says Fabrice Duponchelle of the French Institute of Research for Development in Marseille.

Amazon nations should work together to craft a basin-wide management plan for migratory fish, says Thomas Lovejoy, a tropical ecologist at George Mason University in Fairfax, Virginia. “You have to manage the Amazon as a system,” he says. It's not too late to preserve much of the ecology of the western Amazon Basin, where “we still have a lot of free-flowing rivers,” Anderson adds. “There is a huge opportunity to protect at least a subset of them by thinking at a regional scale.” ■

Barbara Fraser is a freelance journalist in Lima.

## Parting the waters

A new analysis forecasts severe habitat fragmentation in the western Amazon Basin if some of the 160 planned dams are built in the region, where 142 dams are already.



## FEATURES

# HUMAN NATURE, OBSERVED

For decades, two psychologists have kept watch over 1000 New Zealanders, teasing out factors that shape a life's course

By Douglas Starr

In 1987, Avshalom Caspi and Terrie Moffitt, two postdocs in psychology, had adjacent displays at the poster session of a conference in St. Louis, Missouri. Caspi, generally not a forward man, looked over at Moffitt's poster and was dazzled by her science. "You have the most beautiful data set," he said. Not one to be easily wooed, Moffitt went to the university library after the meeting and looked up Caspi's citations. Yep, he'd do. "It was very nerdy," Caspi recalls. "We fell in love over our data."

It's been a personal and scientific love affair ever since. For nearly 30 years, Moffitt and Caspi have been collaborating on one of the more comprehensive and probing investigations of human development ever conducted. Launched in 1972, the Dunedin Multidisciplinary Health and Development Study is as fundamental to human development as the Framingham Heart Study is to cardiovascular disease and the Nurses' Health Study is to women's health. From detailed observations of the life courses of about 1000 New Zealanders, Dunedin has spun out more than 1200 papers on questions from the risk factors for antisocial behavior and the biological outcomes of stress to the long-term effects of cannabis use. Moffitt, who joined the study in 1985, and Caspi, who followed, have led much of the work. They "have done

so much it's impossible to pigeonhole them," says Brent Roberts, a psychologist at the University of Illinois in Champaign who has collaborated with the now-married couple.

One early finding, on the transient nature of most juvenile criminality, was cited in the U.S. Supreme Court's 2005 decision to prohibit the execution of underage murderers. Moffitt and Caspi did pioneering research showing that self-control in early childhood predicts health and happiness in adults. They detailed how the genetic makeup of certain individuals can make them vulnerable to specific stresses, elucidating the complex interplay between genes and life experience. And last year, the Dunedin team published a study that drew on decades of data to show that, contrary to conventional belief, the vast majority of people experience mental health problems over the course of their lifetime.

"Their work has transcended psychology to influence thinking in psychiatry, genetics, criminology, epidemiology, sociology, and many other areas," the American Psychological Association said when it awarded Moffitt and Caspi its 2016 Award for Distinguished Scientific Contribution.

The field of psychology is rich with longitudinal studies, going back to 1946, when the United Kingdom's Medical Research Council began a survey of more than 5000 people from birth to old age. Other

researchers have followed identical and fraternal twins over time to tease out the influence of nature versus nurture. Such studies, although slow, make it possible to observe phenomena in real time instead of having to reconstruct them from subjects' memories or medical records, or by comparing disparate groups.

Dunedin is the Goldilocks of longitudinal studies. It's not the biggest or the longest; but its high retention rate—about 95% of the original cohort has stayed with the study since it launched—and the intimacy of the data-gathering process make the group one of the most closely examined populations on Earth. Every few years, the team conducts intensive cognitive, psychological, and health assessments. They interview every member of the research cohort as well as their teachers, families, and friends and review their financial and legal records, promising them complete confidentiality in return for the fullest possible picture of their lives.

As a result, Moffitt, Caspi, and their colleagues have been able to tease out previously unseen patterns in human development. "It's been a remarkable resource," Columbia University psychiatrist Ezra Susser says of the study. "There really is no equivalent," Roberts adds. "Time after time they've anticipated where we've needed to go and done the work earlier than the rest of us."



Downloaded from <http://sciencemag.org/> on February 3, 2018

PHOTO: GUY FREDERICK



In a New Zealand testing center, Terrie Moffitt (left) and Avshalom Caspi (right) probe their subjects' physical and mental condition.

pleted his Ph.D. in international studies. "The transition from a kibbutz to Berkeley was pretty amazing," Caspi recalls. "This was the '70s, and I remember dodging riots with my dad."

His experiences had made him curious about people and places, and he thought about becoming a journalist. But psychology won him over in college, and at Cornell University he did graduate work with sociologist Glen Elder, who wrote *Children of the Great Depression: Social Change in Life Experience*, a landmark study of how the economic environment influences human development. Caspi became fascinated by the interplay between people's dispositions and their social environment. He was presenting a study on how ill-tempered children become ill-tempered adults when a young woman with crimson hair set up her poster next to his.

A few years of long-distance courtship ensued, before the couple settled in at the University of Wisconsin in Madison. Colleagues warned against making their personal partnership a professional one, but they plunged ahead, and Caspi joined the Dunedin team.

His first collaboration with Moffitt involved a question that had long intrigued him: Do dramatic life events change people, or simply cement who they already are? The team focused on female puberty. Just over half of the Dunedin study subjects were girls, who by age 15 had been given physical and psychological exams since childhood. Previous studies showed that reaching puberty early is especially stressful, and Caspi and Moffitt found that the girls who had the most trouble adjusting to early adolescence were those who had shown behavioral problems in early childhood. That contributed to Caspi's "accultuation" hypothesis—that stressful transitions tend to accentuate who we basically already are. It was a radical notion in the early 1990s, when psychologists generally believed that situation, not temperament, was key to determining personality.

**SINCE THEN**, their work has often explored the darker side of human nature. Contrary to what the idyllic landscapes in the *Lord of the Rings* films might suggest, New Zealand is not heaven on Earth. Economic inequality is

**MOFFITT GREW UP** in central North Carolina, where her Scotch-Irish ancestors had settled in the 1700s. She helped with farm chores and rambled in the woods, collecting arrowheads and catching fireflies, often in the company of her beloved Grandma Macon. "Respect is earned, not demanded," Grandma Macon had taught her. "A woman's beauty is in her strength and in her laugh."

Red haired and blue eyed, Moffitt was the first person in her extended family to attend college. Having come from a "low-quality high school," she chose easy classes at the University of North Carolina in Chapel Hill, but when professors recognized her potential, she warmed to the challenge of advanced courses in psychology and behavioral pharmacology. She supported herself as a lab tech in nearby Durham's Research Triangle Park—a job she got because having grown up hunting rabbits and squirrels, she didn't feel queasy about sacrificing lab animals.

In graduate school at the University of Southern California in Los Angeles, working with the famed psychologist Sarnoff Mednick, she focused on the roots of criminal behavior. A few months before defending her Ph.D. dissertation, Moffitt and a few classmates decided to relax by taking sky-diving lessons. It didn't go well. Moffitt broke her leg in several places. Confined to a wheelchair, she spent several hours entertaining a

visitor from New Zealand who had come to speak about his study. Phil Silva had been tracking all 1037 children born in the Queen Mary Maternity Centre at Dunedin Hospital from April 1972 to March 1973, examining them every couple of years to assess their cognitive and psychological development.

Moffitt saw the population as a readymade laboratory to explore some of her ideas about early childhood and the roots of criminality. She joined the study when the children were 13 years old. Not long after earning her Ph.D., she made her first trip to New Zealand, gathering data about girls who got in trouble with the police. She presented her findings at that fateful St. Louis conference. Next to her was a guy with olive skin, soulful eyes, and brown hair in a ponytail.

Caspi, his hair now silver but still often in a ponytail, grew up on a tiny kibbutz in Israel's Negev Desert. The son of a Yemenite father and a Lithuanian mother, he was part of the "peopling" of the Negev envisioned by Israel's founding prime minister, David Ben-Gurion, who had retired to a cottage up the street. Every morning, Caspi and his five classmates would walk past Ben-Gurion's cottage on the way to their little school. "*Shalom*, David," they'd call out to him. "*Shalom yeladim* [children]," Ben-Gurion would reply. Caspi was 10 when his family moved to Berkeley, California, where his father com-

## A lifetime of discoveries

For decades, Terrie Moffitt, Avshalom Caspi, and other researchers have used a variety of tools to monitor a cohort of about 1000 New Zealanders. Multiple findings show that the course of their lives can largely be traced to their health, environment, and temperament as children. ● Key findings ● Study milestones

### 1 Temperament

Girls who have behavioral problems as children have more trouble adjusting to early puberty.

### 2 Personality continuities

Undercontrolled 3-year-olds tend to grow up impulsive and antisocial; inhibited 3-year-olds tend to become unassertive and depressed.

### 3 Violence and gender

Women are as likely as men to commit intimate partner violence.

### 4 Schizophrenia

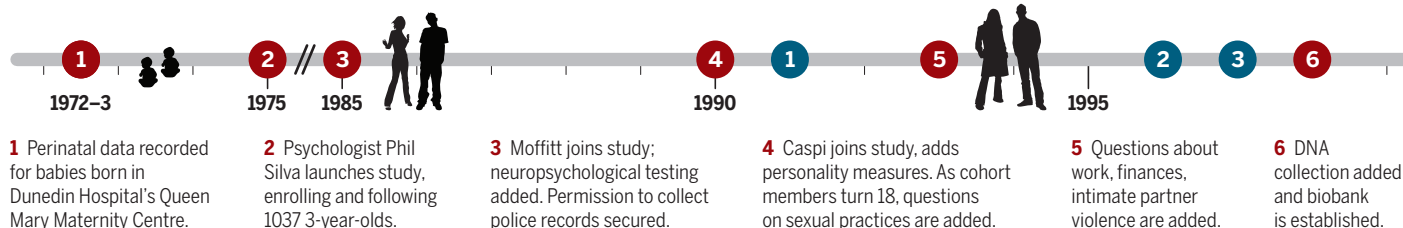
Hallucinations in childhood predict an increased risk of adult schizophrenia.

### 5 Persistent offenders

People who are violent and antisocial as children grow into violent adults more often than those whose antisocial behavior is limited to adolescence.

### 6 Genes and environment

Among maltreated children, those who produce low levels of the enzyme monoamine oxidase A are more likely to become violent adults.



comparable to that of the United States, and addiction, suicide, and assault rates are similar. By promising strict confidentiality, the research team gets shockingly frank confessions from their subjects, including about brutality and criminal behavior. (A startling example: Women admitted to physically abusing their spouses as often as men, although they generally inflicted less physical damage.) “We have a policy of never interfering and never ratting them out,” Moffitt says. “And the New Zealand police, who understand the value of our research, have never asked.”

This long, intimate surveillance enabled Moffitt to track a troubled subset of the Dunedin cohort. Delinquent behavior is known to peak between the late teens and mid-20s, mostly in men. By 15, about a third of the study's boys took part in some degree of delinquency, Moffitt found, while a subset offended more frequently. When she looked at data from early childhood, she found that the habitual offenders had been making trouble from age 3, and had arrest records starting before their teen years.

Over the years, Moffitt reported in a series of papers that these boys did poorly in neuropsychological tests (such as verbal skills and verbal memory), measured high for impulsivity, and were likely to engage in substance abuse as they grew older. In 2002, she reported that at age 26 this same group was committing most of the crimes in the community—a pattern that persisted well into their 30s. In short, whereas many boys exhibited “adolescent-limited” criminal behavior, about 5% were “life-course-persistent” offenders. The work had important implications for social work and law, and won Moffitt the 2007 Stockholm Prize in Criminology.

In 2000, after Silva retired, Richie Poulton, a psychologist at the University of Otago in Dunedin, became project director, with Moffitt continuing as associate director. Ev-

ery couple of years Moffitt and Caspi would visit Dunedin to spend a few months with Poulton collecting information, then return home to analyze data. (A few years earlier they had also initiated a study of 1100 U.K. families with twins that added prodigiously to their Dunedin data.)

The testing schedule in New Zealand, limited at first, expanded as they added new techniques and technologies, such as DNA analysis, retinal imaging (which can help gauge the brain's vascular health), and scans of brain activity. The study leaders trained staff to interview participants, their families, and associates, and they employed data managers to handle the masses of coded data. And they maintained a high retention rate by assiduously keeping track of the participants, a quarter of whom have emigrated from New Zealand, financing travel back to the study site, and visiting study members who were in prison or too sick to travel.



Data from the Dunedin subjects' childhood years make it possible to trace early influences on their later lives.

“They have gotten to know these people intimately, watched them age, get married, have children, and encounter the things that happen in middle life,” says James Tabery, a philosopher at the University of Utah in Salt Lake City who has written about the study. It can be harrowing, he says. “What do you do if one of them is on the verge of homelessness? Or if one of them is \$1000 short on rent? It's like David Attenborough watching a giraffe get eaten but it's a giraffe you've come to know for 40 years.” Researchers can unburden themselves in a safe room after listening to upsetting information. Outside of that room, anyone who violates confidentiality will be fired.

**THE STUDY'S FINDINGS** mostly point to patterns, not mechanisms. But in the mid-1990s, Moffitt and Caspi joined a group at King's College London looking for genetic roots of behavior. The result was what may be the highest profile and most controversial of their claims.

It's commonly known that children who are abused often become violent adults. In the Dunedin group, for example, about half the boys with abusive childhoods grew into men prone to committing crimes. Moffitt and Caspi thought that studying this group could shed light on the classic nature versus nurture question: Do experiences alone produce a life of crime, or are some individuals naturally prone to it? The idea had come to them on safari in Namibia, where they noted how even in that mosquito-infested region many people did not get malaria, undoubtedly because of a genetic resistance. Could genotype play a role in behavior as well, making certain children psychologically resistant to abuse?

Earlier research had linked aggression to low levels of an enzyme called monoamine oxidase A (MAO-A), which breaks down certain neurotransmitters in the brain. Laboratory mice bred without the enzyme showed

**7 Stress and illness**

People who were abused as children have higher markers of inflammation, indicating an elevated risk of cardiovascular disease.

**8 Self-control**

Childhood self-control predicts physical health, financial success, and a lack of criminal behavior among adults.

**9 Cannabis**

Long-term heavy cannabis use that begins in adolescence is linked to cognitive decline in adulthood.

**10 Early aging**

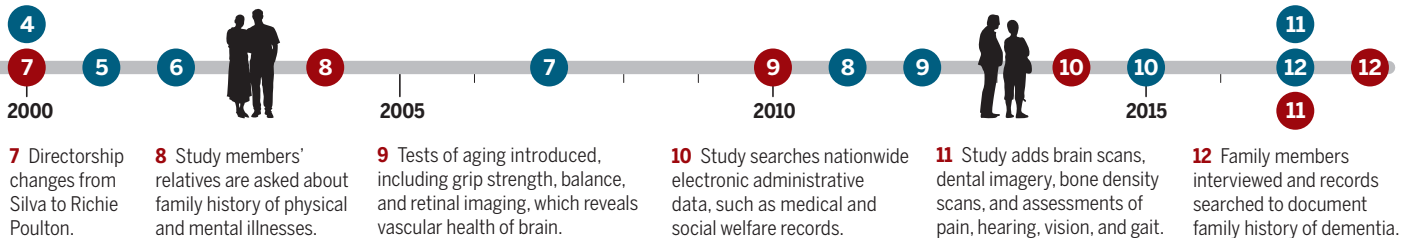
Multiple indicators reveal signs of rapid physical and cognitive aging in some adults in their 30s.

**11 Social ills**

Twenty-two percent of the cohort accounted for the bulk of welfare, medical, and insurance costs; fatherless child-rearing; and crime. Many had poor brain health at age 3.

**12 Mental health**

Dunedin and other studies show that most people have at least one episode of mental illness during their lifetime.



aggressive behavior, and in a famous human case, several male members of a Dutch family who were pathologically violent lacked the gene that encoded for the enzyme.

Moffitt and Caspi analyzed their study members' DNA for the gene. They found that among maltreated children, those who were genetically prone to make low levels of MAO-A were far more likely than their counterparts to become violent adults. It was the first substantial evidence that a specific gene could modulate the effect of a known cause of criminal behavior.

The finding, published in 2002 in *Science*, caused a media sensation, but some researchers questioned the utility of such studies. "Sure there's a genetic effect in people who have been exposed to violence," says Dean Hamer, scientist emeritus at the National Institutes of Health in Bethesda, Maryland, who has written widely on genes and behavior. "But the effects are so small and so variable" as to be dwarfed by the dangers of childhood abuse.

Since then, investigators moved away from individual genes. "We're long past the low-hanging fruit of Mendelian genetics," Tabery says. Geneticists now survey the genomes of thousands of people with a particular condition—asthma, for example—in search of gene/DNA variants whose small effects on risk combine. Often it turns out that hundreds or even thousands of genes contribute to the overall risk, not just one. Moffitt and Caspi have adopted that approach to search for genetic factors in conditions including tobacco addiction and obesity.

Recently, Caspi compared longitudinal research to farming—"sowing, nurturing, waiting, and harvesting." It's an apt comparison, given that his and Moffitt's grandparents were farmers. Year after year they have harvested data, and new revelations about human development have emerged.

In 1987, Silva was among the first to recognize that children can experience hallucinations, when he detected them while interviewing 11-year-olds. As the subjects aged and the study accumulated data, an ominous connection emerged. Follow-ups at age 26 revealed that the kids with hallucinations had developed full-on schizophrenia at more than 25 times the rate of the general population. (Schizophrenia is so rare that the sample numbers were small, however.)

The work, reported in 2013, added support to the emerging practice of "prodromal" psychiatry, in which symptoms are identified and treated in their earliest stages, years before they blossom into disorders (*Science*, 17 November 2017, p. 856). "I've had mothers write me letters all the time saying, 'My kid has some pretty crazy ideas, where can I go?'" says Moffitt, who describes the work as the most gratifying of her career. Now, she can direct them to clinics for treatment.

"It took a quarter of a century to see this unfold," Caspi says. "You just have to be patient. Longitudinal research is a big exercise in delayed gratification."

**WHEN MOFFITT AND CASPI** came back to the United States in 2007 to accept professorships at Duke University in Durham, it was a chance for Moffitt to return to her origins. They bought an old family property in central North Carolina, removed tons of old tires and debris, rebuilt structures, cleared trees, and built a new farmhouse.

Origins mean a lot in their work, which has shown that our nature as adults—our capacity for self-control, our propensity for violence—has deep roots in the children we once were. Recently, Moffitt and Caspi took a broad look at how their study members, now 45, are faring. It's a poignant time of life, when middle-aged adults face the real-

ity that many of their youthful dreams won't come true. The Dunedin data showed that just over a fifth of the population accounts for the bulk of the social costs: crime, welfare payments, hospitalizations, cigarette purchases, fatherless child-rearing, and other indicators of social dysfunction.

What's wrong with these people? Moffitt and Caspi went back and looked at their data from age 3. The target group seemed cursed from the beginning: They scored low on early language skills, fine and gross motor skills, neurological health, and self-control. Often they also grew up in poverty and suffered maltreatment. All through life their disadvantages haunted them. "They didn't get a fair start right out of the starting block," Moffitt says. "You can't expect people with this kind of childhood to do well."

They even seem to age faster than those who had a better start. In their cohort, Moffitt and Caspi have been finding signs of aging starting in the 30s. They're particularly struck by the effects of stress at an early age. Childhood abuse seems to erode telomeres—the caps at the end of chromosomes, associated with cell preservation—and that, in turn, may accelerate aging.

Moffitt and Caspi offer no grand unified theory of human development: Humans are too complicated, too irrational, to sum up in a principle. What their research gives them is not so much a conclusion about humans as a particular point of view.

"All people are not created equal," Moffitt says. "Some have real gifts and talents, and some have real problems right out of the starting block. Once we accept that, we can't dodge the responsibility for social action."

Watching people's lives unfold over decades, she adds, "obliges compassion." ■

*Douglas Starr is co-director of the program in science journalism at Boston University.*

# INSIGHTS

## PERSPECTIVES

### ORGANISMAL BIOLOGY

## *Out of balance in the Arctic*

Polar bears have high energy requirements that rise further as a result of climate change

By John P. Whiteman

As human activities lead to rising greenhouse gas concentrations in Earth's atmosphere, less incoming solar energy is released back into space, causing a net energy gain that increases global temperatures. The consequence of climate change for polar bears can likewise be understood in terms of an energy imbalance. Sea ice melting reduces the opportunities for polar bears to capture seals (see the photo), leaving them at risk of expending more energy in the pursuit of food than they can obtain. The magnitude of this imbalance is determined by

their rate of energy use. On page 568 of this issue, Pagano *et al.* (1) quantify the energy expense of wild polar bears and show that it is higher than previously estimated.

Animal energy use is typically described by two terms. The resting metabolic rate (RMR) includes the energy cost of basic organismal functions, such as blood circulation and breathing. The field metabolic rate (FMR) includes RMR as well as the energy cost of additional activities, such as movement and foraging. Because energy balance influences whether an animal survives and reproduces, RMR and FMR are critical variables in ecology and conservation. But they are notoriously difficult to quantify for

Polar bears look for food onshore in East Greenland, August 2016. The species needs sea ice as a platform to capture energy-rich prey, particularly seals. Polar bear population trends currently vary across the Arctic, but scientists expect broad declines if sea ice loss continues.

large, mobile animals, particularly in remote habitats.

Scientists have previously estimated polar bear energy requirements by calculating RMR with predictive body mass-based equations derived from other species and then multiplying RMR by a constant to yield FMR (2). In another study, a researcher extrapolated field observations of seal kill rates to estimate the energy intake and FMR of polar bears (3). These and other studies assumed that polar bears can spontaneously reduce their RMR during food deprivation, which would be beneficial while searching for sparsely distributed seals. A recent study,

Department of Biology, University of New Mexico, Albuquerque, NM 87131, USA. Email: jwhiteman@unm.edu

however, found that polar bear energy expense does not substantially decline during fasting (4), leaving the magnitudes of both their RMR and FMR uncertain.

In their study, Pagano *et al.* measured rates of oxygen consumption and carbon dioxide production in polar bears; these variables can be converted to energy use. First, the researchers assessed RMR by training a zoo polar bear to sit quietly in a chamber while its oxygen consumption was measured. Next, they quantified FMR in nine free-ranging bears on sea ice near the northern coast of Alaska, USA, by using doubly labeled water. To do so, the authors first captured the bears and injected them with water labeled with heavy stable isotopes ( $^2\text{H}_2^{18}\text{O}$ ); they also collected a baseline blood sample. About 10 days later, they recaptured the bears and took a second blood sample. In the interim, the injected  $^2\text{H}$  left the body in water molecules (for example, through urination), whereas the  $^{18}\text{O}$  left the body in both water and exhaled  $\text{CO}_2$  because oxygen atoms readily exchange between  $\text{H}_2\text{O}$  and  $\text{CO}_2$  in the blood. As a result, the difference in decline of  $^{18}\text{O}$  and  $^2\text{H}$  between sampling events represents the  $\text{CO}_2$  production rate, which in turn reflects FMR.

FMRs of the free-ranging bears were on average 1.6 times higher than previous estimates (see the figure) but were consistent with expectations for mammals that mainly consume other vertebrates; these mammals expend more energy than those consuming vegetation or mixed diets (5). Similarly, the zoo bear had a relatively high RMR, matching expectations for a carnivore. Although the mechanism for the high energy cost associated with meat-eating is unclear (6), these findings emphasize that polar bears have evolved as hunters, particularly of fat-rich marine mammal prey. This distinguishes them from other bear species, which include herbivores, insectivores, and omnivores. Polar bear activity patterns are also different from those of other bears: On average, the individuals sampled by Pagano *et al.* were active for 34% of their time, making them more similar to large terrestrial carnivores (39%) than to other bear species (50 to 60%) (7).

The high energy requirements of polar bears corroborate previous hypotheses that most terrestrial Arctic habitats, lacking prey as energy-rich as marine mam-

mals, cannot provide enough food for polar bears driven to shore by loss of sea ice (8). In addition, Pagano *et al.*'s results demonstrate that polar bears do not reduce their metabolic rate when food is scarce (4). Four of the nine bears lost  $\geq 10\%$  of their body mass during the days between blood sampling events; at least one of these bears lost substantial lean tissue (structural pro-

The data also help to resolve why the population-level response of polar bears to sea ice loss varies by region. For example, two Alaskan subpopulations of polar bears have recently experienced substantial ice loss, but only one of them is declining thus far (11). High and relatively inflexible RMRs and FMRs for polar bears suggest that population responses are mainly driven by local, ecological factors, such as primary productivity or overlap of seal distribution with remaining sea ice habitat. Ice loss, if unabated, will eventually cause the extinction of polar bears in the wild (12), but continued research is needed to understand the climate-related pressures that polar bears face. Such data-driven explanations of the biological consequences of climate change are critical for public understanding and action (13).

Important questions remain regarding the energy balance of polar bears. Pagano *et al.* captured free-ranging individuals during spring, when new seal pups and attentive adult seals are most vulnerable to predation. During winter, polar bear activity declines, and some individuals retreat into "shelter dens" for days to weeks at a time (excluding pregnant females, which hibernate). It is unknown whether there is a concurrent seasonal reduction in FMR. Energy expense in late winter is critical because body condition reaches an annual low in this period, making it the

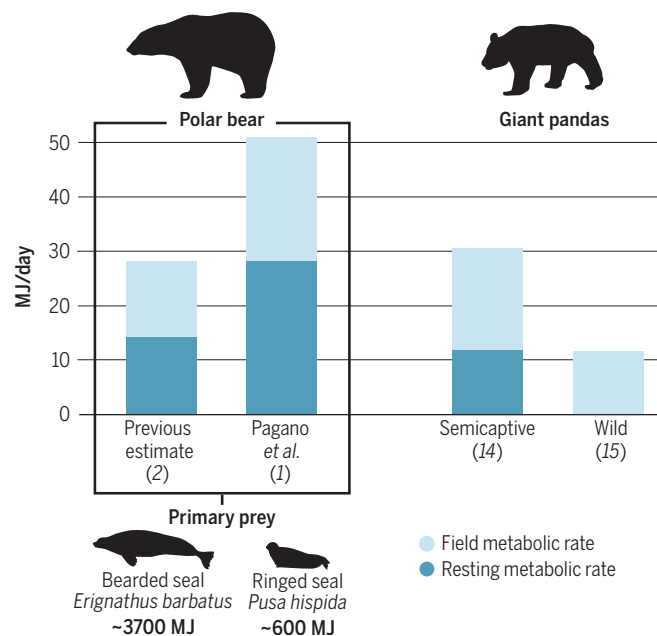
most likely time for climate change-related mortality. In addition, Pagano *et al.* sampled bears without dependent cubs, leaving the cost of rearing young (such as lactation) to be quantified. ■

## REFERENCES

1. A. M. Pagano *et al.*, *Science* **359**, 568 (2018).
2. I. Stirling, N. A. Øritsland, *Can. J. Fish. Aquat. Sci.* **52**, 2594 (1995).
3. M. C. S. Kingsley, *Ringed Seals in the North Atlantic* (North Atlantic Marine Mammal Commission, 1998), vol. 1, pp. 181–196.
4. J. P. Whiteman *et al.*, *Science* **349**, 295 (2015).
5. B. K. McNab, *Q. Rev. Biol.* **63**, 25 (1988).
6. A. Muñoz-García, J. B. Williams, *Physiol. Biochem. Zool.* **78**, 1039 (2005).
7. S. Paisley, D. L. Garshelis, *J. Zool.* **268**, 25 (2006).
8. K. D. Rode, C. T. Robbins, L. Nelson, S. C. Amstrup, *Front. Ecol. Environ.* **13**, 138 (2015).
9. P. K. Molnár, A. E. Derocher, G. W. Thiemann, M. A. Lewis, *Biol. Cons.* **143**, 1612 (2010).
10. G. M. Durner *et al.*, *Glob. Change Biol.* **23**, 3460 (2017).
11. K. D. Rode *et al.*, *Glob. Change Biol.* **20**, 76 (2014).
12. S. C. Amstrup *et al.*, *Nature* **468**, 955 (2010).
13. J. A. Harvey *et al.*, *BioScience* **10.1093/biosci/bix133** (2017).
14. Y. Fei *et al.*, *Sci. Rep.* **6**, 27248 (2016).
15. Y. Nie *et al.*, *Science* **349**, 171 (2015).

## Energy balance of adult female polar bears

A 175-kg polar bear needs far more energy than a herbivorous giant panda (*Ailuropoda melanoleuca*), scaled to the same body mass. Polar bears therefore require access to high-energy food sources, particularly pagophilic (ice-loving) seals.



teins such as muscle) as well as fat (the primary form of energy storage). These bears nevertheless had high FMRs.

Pagano *et al.*'s findings refine our understanding of how polar bears are affected by environmental change in the Arctic. Longer periods of ice melt are lengthening the time during which the bears must endure food deprivation in summer. Molnár *et al.* have previously used anatomical models of energy storage and body mass change over time to estimate that starvation mortality of adult male bears would increase from 6 to 48% if summer fasting were extended from 120 days to 180 days (9). Such models can now be modified to include empirical FMR estimates. Further, thinner sea ice drifts faster, and polar bears have had to increase their movement rates in order to remain in their preferred habitat, incurring higher energy costs (10). These costs can now be modeled more precisely because of the strong linear relationship between movement rate and FMR reported by Pagano *et al.*

## IMMUNOLOGY

# Enhancing responses to cancer immunotherapy

The capacity of antigen presentation influences responses to checkpoint immunotherapy

By Pia Kvistborg<sup>1</sup> and  
Jonathan W. Yewdell<sup>2</sup>

Immune checkpoint blocking therapies (ICBs) that target T cell inhibitory receptors (immune checkpoints) have been implemented in the clinic to treat a variety of malignancies. To exemplify the potential success of these therapies, the 3-year overall survival for advanced melanoma has increased from 12% before 2010, when standard of care was chemotherapy, to ~60% using ICBs (1). However, ICBs fail many patients (10 to 60% of treated patients respond, depending on cancer type), raising the obvious question of why. On page 582 of this issue, Chowell *et al.* (2) report that the success of ICBs is remarkably dependent on the ability to present diverse tumor antigens to T cells.

Human leukocyte antigen (HLA) class I molecules present short intracellular peptides (antigens) on the surface of all nucleated cells and play a key role in CD8<sup>+</sup> T cell immunosurveillance of viruses and other intracellular pathogens. On recognizing foreign peptides presented by HLA class I molecules, CD8<sup>+</sup> T cells are activated. Humans possess three *HLA* class I loci: *HLA-A*, *HLA-B*, and *HLA-C*. Each locus is among the most diverse in humans, each having thousands of different alleles encoding varying HLA class I molecules (allomorphs) (3). Within a population of common geographic origin, certain alleles are common, potentially decreasing the diversity of HLA allomorphs. Each allomorph presents a distinct, though often overlapping, repertoire of antigens, termed the immunopeptidome.

Immunosurveillance of cancer cells is restricted by central T cell tolerance, a process that removes self-reactive clones. Cancer cells, however, often express “neoantigens,” mainly resulting from peptide fragments of mutated proteins that arise from DNA damage,

which is commonly involved in carcinogenesis. Indeed, neoantigens are thought to be important for responses to ICBs (4). However, the impact of expressing different HLA allomorphs on response to ICBs has not been appreciated.

Studying more than 1500 cancer patients, Chowell *et al.* show a significant dependence of ICB responsiveness on *HLA-A*, *HLA-B*, and *HLA-C* heterozygosity. This trend is enhanced by (but not dependent on) the mutation load of the cells within the tumor, determined through exome sequencing of resected tumors. ICB efficacy was diminished by loss of *HLA* heterozygosity, presumably the result of immune selection for tumor cells with decreased expression of the cognate allomorph (see the figure). This builds on previous findings that ~40% of lung cancers display somatic loss of one or more *HLA* alleles (5), which results in decreased antigen presenta-

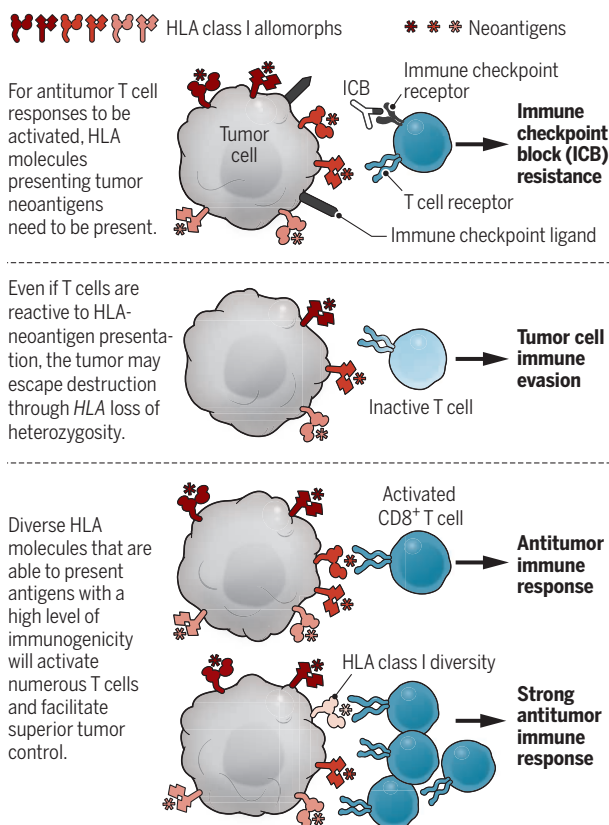
tion. This is particularly important if there is somatic loss of an *HLA* allele by which the majority of antitumor T cells are activated. The *HLA-A* and *HLA-B* genes present in the cohort were divided into six supergene families based on the immunopeptidomes they present (6). Further implicating a critical role of HLA class I in ICB responses, Chowell *et al.* found a strongly positive effect on survival of *HLA-B44* supertype alleles in melanoma patients receiving ICBs, but no association with survival in patients not receiving ICBs. Intriguingly, the correlation did not extend to ICBs of lung cancer, which may reflect the low number of lung cancer patients in the study. Chowell *et al.* also found that *HLA-B62* supertype expression was associated with decreased effectiveness of ICBs in melanoma.

These findings affect the selection of patients who are likely to benefit from ICBs, which have potentially life-threatening side effects, and are sufficiently costly to strain national health care budgets. More generally, they provide important clues about the mechanisms of antitumor immunity. Tumors must walk a tightrope in escaping CD8<sup>+</sup> T cells: loss of HLA class I molecules triggers natural killer (NK) cells, which also play an important role in tumor control. Indeed, this may be key to the observation by Chowell *et al.* that *HLA-C* alleles had the strongest positive effect on ICB responses because *HLA-C* molecules play a critical role in activating NK cell inhibitory receptors and present a more limited immunopeptidome compared with *HLA-A* and *HLA-B* molecules.

The effect of *HLA* polymorphism on immunosurveillance is well preceded by studies of antiviral CD8<sup>+</sup> T cells. The most striking example is fourfold overrepresentation of *HLA-B57* in the 0.5% of patients who can control HIV infections without antiviral therapy (7). Despite intense study, the underlying mechanisms for these effects remain obscure. Chowell *et al.* propose that the generally higher complexity of the *HLA-B44* supertype-associated immunopeptidome increases the number of cancer neoantigens presented.

## Tumor cell killing by T cells

T cell activation and tumor cell killing depends on multiple factors involving antigen presentation and T cell activation.



<sup>1</sup>Division of Molecular Oncology and Immunology, The Netherlands Cancer Institute, Amsterdam, Netherlands. <sup>2</sup>Cellular Biology Section, Laboratory of Viral Diseases, National Institute of Allergy and Infectious Diseases, Bethesda, MD 20892, USA. Email: p.kvistborg@nki.nl

Conversely, they propose that a more rigid structure of the HLA-B62 molecule somehow reduces the capacity of T cell receptors to recognize neoantigens. This would not, however, explain the negative correlation of HLA-B62 with survival, which implies interference with antitumor responses.

Too little is known about the effects of polymorphism on HLA function to have any confidence in interpreting these important observations. It is likely that only a minor fraction of potentially immunogenic neoantigens have been identified by traditional exome sequencing, which ignores peptides generated from noncanonical translation products. These include CUG-initiated proteins, small frameshifted proteins, proteins from “untranslated” reading frames, and intron-encoded proteins (8). Given the plethora of abnormalities in cancer cells, noncanonical translation is likely to increase. This might occur in parallel with mutation load, contributing to the correlation with ICB responsiveness.

Additionally, HLA class I allomorphs exhibit astonishing differences, including 10-fold or more variation in their export rate from the endoplasmic reticulum (9), where they are loaded with peptides. Furthermore, little is known about how allelism affects the pathways used by dendritic cells and macrophages to acquire tumor antigens from, for example, dying tumor cells. This can influence the quality of both T cell activation and antitumor immunity (10).

There are large gaps in our understanding of the astonishing diversity of the HLA molecules (HLA class II molecules, which present extracellular peptides, are similarly polymorphic). As HLA molecules are involved in all aspects of immunity and autoimmunity, as well as neuronal development, and likely other biological phenomena, the findings of Chowell *et al.* encourage increasing understanding of the impact of HLA polymorphism on vertebrate biology and human health. ■

#### REFERENCES AND NOTES

1. J. D. Wolchok, R. Rollin, J. Larkin, *New Engl. J. Med.* **377**, 2503 (2017).
2. D. Chowell *et al.*, *Science* **359**, 582 (2018).
3. J. Robinson *et al.*, *PLOS Genet.* **13**, e1006862 (2017).
4. T. N. Schumacher, R. D. Schreiber, *Science* **348**, 69 (2015).
5. N. McGranahan *et al.*, *Cell* **171**, 1259 (2017).
6. J. Sidney, B. Peters, N. Frahm, C. Brander, A. Sette, *BMC Immunol.* **9**, 1 (2008).
7. I. Messaoudi, J. A. Guevara Patiño, R. Dyall, J. LeMaout, J. Nikolich-Zugich, *Science* **298**, 1797 (2002).
8. S. R. Starck, N. Shastri, *Immunol. Rev.* **272**, 8 (2016).
9. D. B. Williams, S. J. Swiedler, G. W. Hart, *J. Cell Biol.* **101**, 725 (1985).
10. B. Zhang *et al.*, *J. Exp. Med.* **204**, 49 (2007).

#### ACKNOWLEDGMENTS

J.W.Y. is supported by the Division of Intramural Research, National Institute of Allergy and Infectious Diseases.

10.1126/science.aar6574



The bacteria, *D. acetivorans*, that carry out the rTCA cycle are found in volcanic caldera in Kamchatka, Russia.

#### MICROBIOLOGY

## Stealth reactions driving carbon fixation

New twists to bacterial metabolic pathways that contribute to the global carbon cycle

By Stephen W. Ragsdale

Organisms live in an interconnected dynamic web in which they make, degrade, and interconvert compounds containing carbon, hydrogen, nitrogen, oxygen, phosphorus, and sulfur. The carbon cycle involves the oxidation of organic compounds to produce CO<sub>2</sub> by heterotrophic organisms and the incorporation (“fixation”) of CO<sub>2</sub> from the environment into living tissue by autotrophic organisms. Heterotrophic organisms (most animals) obtain the energy for life by conserving the energy obtained by oxidizing organic molecules to CO<sub>2</sub> in the form of reducing equivalents (electrons) and adenosine triphosphate (ATP). Autotrophic plants, bacteria, and archaea fix CO<sub>2</sub> by a process in which the energy of electrons and ATP is used to produce biomolecules, such as sugars, amino acids, and lipids, thereby replenishing these essential organic molecules in the ecosystem. Cumulatively, autotrophy occurs on the huge scale of 7 × 10<sup>16</sup> g of carbon fixed annually (1). Six CO<sub>2</sub> fixation pathways differing in

their ATP requirements are known to exist. Fixing CO<sub>2</sub> using the least ATP possible is key for anaerobes because their metabolism generates much less ATP than does growth on oxygen. The reductive tricarboxylic acid (rTCA) cycle is one of the most evolutionarily ancient and least ATP-demanding autotrophic pathways. On pages 563 and 559 of this issue, Mall *et al.* (2) and Nunoura *et al.* (3), respectively, uncover an unexpected ATP-conserving mechanism, and Pachiadaki *et al.* (4) report a surprising source of reducing equivalents in the rTCA cycle.

According to our understanding of the rTCA cycle, three enzymes that catalyze reactions in the oxidative TCA (oTCA) cycle that would be highly thermodynamically unfavorable if performed in the reverse direction are substituted by parallel energetically favorable reactions (5) (see the figure). In the rTCA cycle, citrate synthase (CS), which catalyzes conversion of oxaloacetate and acetyl-CoA (coenzyme A) to citrate and CoA, the most irreversible reaction in the oTCA cycle, is replaced by ATP-dependent citrate lyase (ACL) or homologous enzymes that catalyze this reaction in two steps (5, 6). The existence of ACL is the typical indicator for an operational rTCA cycle in an organism (2). The use of ATP in the ACL reaction converts a notoriously thermodynamically unfavorable

Department of Biological Chemistry, University of Michigan Medical School, 1150 West Medical Center Drive, Ann Arbor, MI 48109-0606, USA. Email: sragdsal@umich.edu

reaction into a slightly uphill one. Thus, the direct physiological reversal of CS has been thought to be impossible.

It is time to change these views of CS and of the rTCA cycle. Mall *et al.* and Nunoura *et al.* showed that two thermophilic sulfur-reducing anaerobic bacteria, *Desulfurella acetivorans* and *Thermosulfidibacter takaii*, lack the gene encoding ACL, but retain all other rTCA cycle enzymes. Both organisms also lack the genes required for other known CO<sub>2</sub> fixation pathways. Surprisingly, despite the high thermodynamic barrier, they found that CS functioned in reverse. To distinguish this pathway from the typical rTCA cycle, this was termed the reversed oxidative TCA (roTCA) cycle.

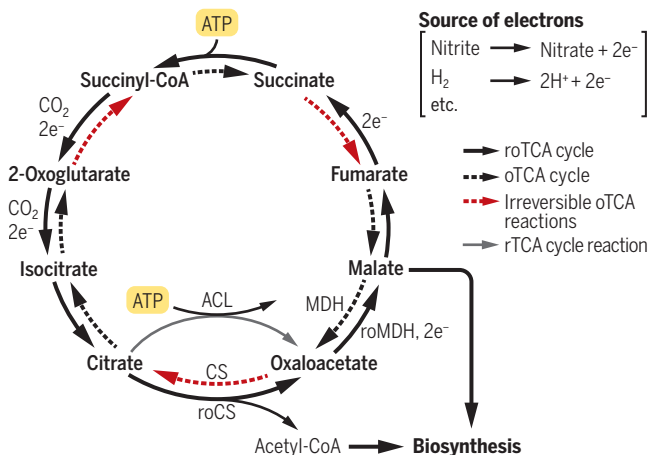
How do these organisms overcome the huge thermodynamic barrier for reversing CS? Apparently, they rapidly and efficiently couple roCS with the succeeding reaction(s) to keep the cellular roCS substrate (citrate and CoA)/product (oxaloacetate and acetyl-CoA) ratio extremely high. Mall *et al.* detected very high activities of both roCS and reverse oxidative malate dehydrogenase (roMDH), which would promote rapid conversion of oxaloacetate (from roCS activity) to malate. However, enhancing the activity of these enzymes will increase the rates of both oxidative and reverse reactions. The key is flux. These organisms must efficiently convert the roCS and roMDH products malate and acetyl-CoA into cellular biomolecules, thus maintaining very low concentrations of oxaloacetate, effectively pulling flux through the roTCA cycle. Even by bypassing the ATP provided by the ACL-catalyzed reaction in the rTCA cycle, the overall roTCA cycle is thermodynamically favorable; thus, another trick may involve bifurcation, driving the unfavorable half of the cycle by coupling it to the favorable half, yielding a net spontaneous process. Nunoura *et al.* suggest that abundant input of redox equivalents via electron bifurcation may provide the needed thermodynamic push. However, roCS is not an electron transfer reaction, so it would require an unknown mechanism to make such a bifurcation work. To add to the mystery, the CSs of both organisms have rather typical kinetic parameters and are not specifically adapted to run in reverse.

The essential point is that reversing CS spares one ATP for the organism during autotrophic growth, which is potentially important as ATP is of premium value for anaerobes. Thus, these fastidious microbes

require a single ATP to fix two molecules of CO<sub>2</sub>, matching the stoichiometry, efficiency, and free energy (a measure of thermodynamics) of the Wood-Ljungdahl pathway (6). By contrast, the prevalent CO<sub>2</sub>-fixing pathway, the Calvin cycle, uses seven ATPs. Not surprisingly, the Calvin cycle has not been found to occur in anaerobes. In aerobes, O<sub>2</sub> is used to generate an ample supply of ATP. Furthermore, because ACL is the signature enzyme for detecting the rTCA cycle, the roTCA cycle could be a widespread mode of CO<sub>2</sub> fixation, previously hidden below our genomic radar.

## The autotrophic roTCA cycle

In the rTCA cycle, two CO<sub>2</sub> molecules are fixed to generate acetyl-CoA, which is used for biosynthesis. Because three reactions in the oTCA cycle are irreversible, they are substituted by ATP-dependent enzymes, such as ACL, in the rTCA cycle. Surprisingly, in the roTCA cycle, citrate synthase (CS) is retained. The discovery of nitrite oxidation as a source of electrons in the rTCA cycle is another unexpected turn of events in carbon fixation.



Mall *et al.* asked if there are other examples of unexpected reactions running in reverse. Anaerobic methane oxidation is one example. This is an unusual pathway that is important to the carbon cycle in which the entire reaction sequence of CO<sub>2</sub> reduction to methane runs in reverse (7). This appears to occur in a microbial community by coupling the reducing equivalents generated by the methane oxidizer to the highly favorable reduction of electron acceptors—for example, sulfate—by a sulfate-reducing organism (8).

A mode of CO<sub>2</sub> fixation by rTCA based on nitrite oxidation has also been lurking unnoticed in the dark ocean where light is too dim to produce ATP using photosynthesis and fix CO<sub>2</sub> by the Calvin cycle. Using metagenomic analyses of nitrite-oxidizing bacteria, Pachiadaki *et al.* identified a highly represented phylum, Nitrospinae, which contains genes encoding nitrite oxidoreductase and rTCA cycle enzymes, yet lacks the genes for other energy-generating pathways such as ammonium or sulfur oxi-

dation. Thus, CO<sub>2</sub> fixation in the rTCA cycle by these bacteria appears to rely on nitrite oxidation as a source of electrons. The authors found that nitrite oxidation appears to be responsible for up to 43% of the total carbon fixed at some ocean depths. This process has thus far been underappreciated because nitrite is found at very low concentrations in the ocean. Correspondingly, the nitrite oxidoreductases in these bacteria have very high affinity for their substrates.

Coming back full circle to the roTCA cycle, the genome of the most frequently encountered marine nitrite oxidizer contains coding sequences with similarity to those of known ACLs (9), making it likely that Nitrospinae use the rTCA cycle to fix carbon, using nitrite as an electron source. However, given the precedent described by Mall *et al.* and Nunoura *et al.*, the use of roCS in this phylum of bacteria cannot be excluded.

It is exciting to alter our conceptions about the rTCA cycle, a key pathway in the global carbon cycle. rTCA has been studied for 50 years. This pathway and the Wood-Ljungdahl pathway (6) have the lowest energetic costs and are considered to be the most evolutionarily ancient modes of CO<sub>2</sub> fixation. The new findings highlighted here reveal an unusual mechanism (roCS) that microbes use to lower the ATP requirement for the roTCA cycle and the use of a surprising source of electrons (nitrite) to drive the

rTCA cycle. Future studies could reveal the details of how such efficient flux through the roCS reaction is accomplished. In addition, we can look forward to results that clearly quantify the relative global contributions of the various pathways and electron sources that drive CO<sub>2</sub> fixation and to the discovery of other autotrophic metabolic systems. ■

## REFERENCES AND NOTES

1. I. A. Berg, *Appl. Environ. Microbiol.* **77**, 1925 (2011).
2. A. Mall *et al.*, *Science* **359**, 563 (2018).
3. T. Nunoura *et al.*, *Science* **359**, 559 (2018).
4. M. G. Pachiadaki *et al.*, *Science* **358**, 1046 (2017).
5. M. C. W. Evans, B. B. Buchanan, D. I. Arnon, *Proc. Natl. Acad. Sci. U.S.A.* **55**, 928 (1966).
6. G. Fuchs, *Annu. Rev. Microbiol.* **65**, 631 (2011).
7. K. Knittel, A. Boetius, *Annu. Rev. Microbiol.* **63**, 311 (2009).
8. S. E. McGlynn *et al.*, *Nature* **526**, 531 (2015).
9. S. Lucker *et al.*, *Front. Microbiol.* **4**, 27 (2013).

## ACKNOWLEDGMENTS

Work on the carbon cycle in the author's lab is supported by the Department of Energy (DE-FG02-08ER15931) and National Institute of General Medical Sciences (R37-GM39451).

10.1126/science.aar6329

# Lighting up superconducting stripes

Terahertz radiation reveals pair density waves in underdoped cuprate superconductors

By **Emre Ergeçen** and **Nuh Gedik**

Cuprate superconductors display a plethora of complex phases as a function of temperature and carrier concentration, the understanding of which could provide clues into the mechanism of superconductivity. For example, when about one-eighth of the conduction electrons are removed from the copper oxygen planes in cuprates such as  $\text{La}_{2-x}\text{Ba}_x\text{CuO}_4$  (LBCO), the doped holes (missing electrons) organize into one-dimensional stripes (1). The bulk superconducting transition temperature ( $T_c$ ) is greatly reduced, and just above  $T_c$ , electrical transport perpendicular to the planes (along the  $c$  axis) becomes resistive, but parallel to the copper oxygen planes, resistivity remains zero for a range of temperatures (2). It was proposed a decade ago (3) that this anisotropic behavior is caused by pair density waves (PDWs); superconducting Cooper pairs exist along the stripes within the planes but cannot tunnel to the adjacent layers. On page 575 of this issue, Rajasekaran *et al.* (4) now report detection of this state in LBCO using nonlinear reflection of high-intensity terahertz (THz) light.

In LBCO, charge and spin ordering coexist in the stripe phase. Within the plane, stripes are separated by regions with insulating antiferromagnetic spin order. In the adjacent plane, this pattern is rotated by  $90^\circ$ . The PDW state predicted by Berg *et al.* (3) is based on this intricate packing of the stripes that cancels the interlayer coupling confining the superconductivity to two dimensions. Hamidian *et al.* (5) captured this elusive order using a scanning Josephson tunneling microscope in  $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8+x}$  below  $T_c$ . Whether this state also exists in LBCO above  $T_c$  has been an open question.

The dynamics of paired electrons in a superconductor can be described by the superconducting order parameter. Any spatial change of phase will result in a current without dissipation. A current will flow between two superconducting regions with different phases if separated spatially by a thin layer of insulator (a Josephson junction), and the current through the thin insulating layer is a nonlinear function of the phase differ-

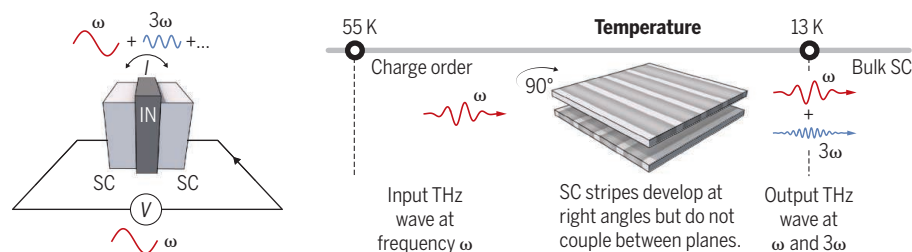
ence (see the figure, left panel). In the case of cuprates below  $T_c$ , superconducting  $\text{CuO}_2$  planes are separated by insulating regions that allow the interlayer tunneling of Cooper pairs by means of the Josephson effect. Below  $T_c$ , a plasma resonance appears in the  $c$ -axis optical properties because of this Josephson coupling. Such optical properties are not seen in the normal state of underdoped cuprates because  $c$ -axis transport is resistive.

For the proposed PDW state in the stripe phase, Josephson coupling is highly sup-

pressed between the planes and reduces bulk  $T_c$  and the plasma resonance. Rajasekaran *et al.* used THz light pulses to study the  $c$ -axis reflectivity of LBCO in samples with dopings of 9.5% ( $T_c = 34$  K) and 15.5% ( $T_c = 32$  K), where superconductivity coexists with a weak charge order that also vanishes near  $T_c$  (6). They observed the plasma edge in the  $c$ -axis reflectivity in the superconducting state. As the intensity of THz light increased, another feature arose at a frequency three times higher than that of the incident light from the inherent nonlinearity of the Josephson effect. Both of these features disappeared at  $T_c$ . For the sample with 11.5% doping, in which strong charge order persisted up to the charge-ordering temperature  $T_{co} = 55$  K and  $T_c$  was suppressed down to 13 K (see the figure, right panel), a third harmonic peak was again seen, but it persisted up to  $T_{co}$ .

## Josephson junctions and the pair density wave (PDW)

Current-voltage ( $I$ - $V$ ) nonlinearity in the Josephson effect enables detection of stripe-phase PDW.



### Current nonlinearities in Josephson junctions

A thin insulator (IN) between two superconductors (SC) creates odd current harmonics.

### Pair density wave in underdoped LBCO

Superconductivity develops along the stripes. However, special ordering of the stripes decouples neighboring planes up to linear order, preventing bulk superconductivity. A high-intensity THz pulse can probe the nonlinearities in interlayer Josephson coupling.

pressed between the planes and reduces bulk  $T_c$  and the plasma resonance. Rajasekaran *et al.* used THz light pulses to study the  $c$ -axis reflectivity of LBCO in samples with dopings of 9.5% ( $T_c = 34$  K) and 15.5% ( $T_c = 32$  K), where superconductivity coexists with a weak charge order that also vanishes near  $T_c$  (6). They observed the plasma edge in the  $c$ -axis reflectivity in the superconducting state. As the intensity of THz light increased, another feature arose at a frequency three times higher than that of the incident light from the inherent nonlinearity of the Josephson effect. Both of these features disappeared at  $T_c$ . For the sample with 11.5% doping, in which strong charge order persisted up to the charge-ordering temperature  $T_{co} = 55$  K and  $T_c$  was suppressed down to 13 K (see the figure, right panel), a third harmonic peak was again seen, but it persisted up to  $T_{co}$ .

No bulk superconductivity exists above  $T_c$ , so another phenomenon must be responsible for this feature. Rajasekaran *et al.* show that the PDW state proposed by Berg *et al.*

field. Through this mechanism, the third-harmonic signal survived up to  $T_{co}$  without bulk superconductivity.

Nonlinear THz measurements performed with ultrafast laser pulses reveal the PDW order in LBCO as predicted by theory a decade ago. However, how these ultrafast measurements should be compared to the static measurements is still an open question. A next step is to use this technique to study charge orders seen in other cuprates (7). These results also reaffirm the usefulness of nonlinear light-matter interaction in revealing hidden orders in quantum materials. ■

### REFERENCES

1. J. M. Tranquada, B. J. Sternlieb, J. D. Axe, Y. Nakamura, S. Uchida, *Nature* **375**, 561 (1995).
2. Q. Li, M. Hücker, G. D. Gu, A. M. Tsvelik, J. M. Tranquada, *Phys. Rev. Lett.* **99**, 067001 (2007).
3. E. Berg *et al.*, *Phys. Rev. Lett.* **99**, 127003 (2007).
4. S. Rajasekaran *et al.*, *Science* **359**, 575 (2018).
5. M. H. Hamidian *et al.*, *Nature* **532**, 343 (2016).
6. M. Hücker *et al.*, *Phys. Rev. B* **83**, 104506 (2011).
7. G. Ghiringhelli *et al.*, *Science* **337**, 821 (2012).

10.1126/science.aar5254

Department of Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. Email: gedik@mit.edu

## ASTRONOMY

# Galaxy motions cause trouble for cosmology

Satellite galaxy motions around a nearby galaxy seem to be at odds with dark matter models

By Michael Boylan-Kolchin

**A**ccording to the widely accepted dark energy plus cold dark matter ( $\Lambda$ CDM) model, dark matter is responsible for both the growth of cosmological structures and the motions of galaxies relative to the expansion of the universe. The dynamics of small galaxies orbiting larger ones provides a crucial window into this mysterious dark matter, which leaves its gravitational mark throughout the universe but has not yet been detected directly. On page 534 of this issue, Müller *et al.* (1) describe observations of satellite galaxies around Centaurus A, the largest galaxy system in the vicinity of the Milky Way. The results may lead to either a better understanding of galaxy formation within the  $\Lambda$ CDM model or a push to overthrow its underlying assumptions.

Although the fundamental nature of both dark matter and dark energy remains elusive, the  $\Lambda$ CDM model has proved to be a remarkably effective framework for understanding the large-scale distribution of matter and energy from the earliest cosmic epochs to the present day (2, 3).  $\Lambda$ CDM is

also the basis of modern theories of galaxy formation, and cosmological simulations using the model can produce virtual universes that match the observed universe both broadly and in detail (4, 5).

However,  $\Lambda$ CDM model predictions do not agree with observations of the motions of small satellite galaxies around the Milky Way and its nearby companion, the Andromeda galaxy (6, 7). According to  $\Lambda$ CDM, the dark matter surrounding each of these galaxies should be roughly spherically distributed, yet both the Milky Way and Andromeda appear to be surrounded by planar structures of satellites. In the Milky Way, where measurements of the full, three-dimensional velocity of a satellite can be made in many cases, several satellites also share a common sense of rotation. Cosmological simulations can produce thin planes of satellite galaxies, but these planes are chance alignments rather than dynamically stable configurations. Reproducing stable, rotating planes has proved elusive in  $\Lambda$ CDM (8).

Is this discrepancy a reason to abandon  $\Lambda$ CDM and search for an alternative model? Some scientists have pointed to sat-

ellite planes as evidence that the very idea of dark matter is fundamentally incorrect and that effects currently ascribed to dark matter are instead the result of a modification to gravity (9).

Most astrophysicists remain unconvinced by this argument, however. The evidence for cold dark matter—or something that behaves very much like it—is very strong on various astrophysical scales and over a range of cosmic times. Modifying the laws of gravity is a straightforward idea but turns out to be extremely difficult in practice when constrained by the wealth of data in support of dark matter and dark energy. Furthermore, many astronomers have been concerned about drawing conclusions from the nearest galaxy systems: The census of Milky Way satellite galaxies might be affected by the gas and stars in the Galaxy's disk, and it is not currently possible to measure motions perpendicular to the plane of satellites in Andromeda, meaning its long-term stability remains unknown.

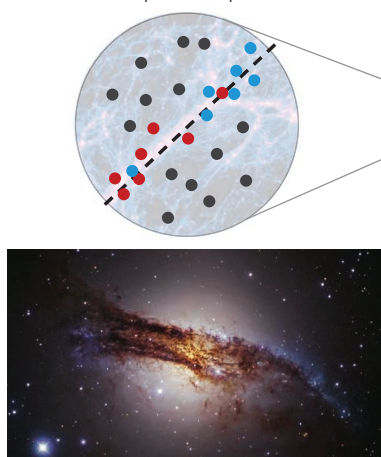
The system that Müller *et al.* report on—Centaurus A—is dynamically isolated from both the Milky Way and Andromeda, allowing them to test for planar structures in the galaxy distribution around a system that is not affected by either of these two galaxies. Their results strongly point to the existence of correlated structure in satellite galaxies around Centaurus A (see the figure). These findings make it less plausible that the observations in all three galaxy systems—the Milky Way, Andromeda, and Centaurus A—arise from statistically rare configurations.

Nevertheless, there are good reasons to be cautious when interpreting the data. In  $\Lambda$ CDM, the distribution of structures in nearby galaxies is not truly independent: It also depends on the distribution of matter on larger scales of approximately 10 to 20 times the distance between Earth and Centaurus A. The large-scale distribution of matter that is specific to our cosmic neighborhood might induce features similar to those observed. Cosmological  $\Lambda$ CDM simulations typically would not capture these system-specific effects. The planes in Andromeda and Centaurus A appear to have a common cosmic orientation, providing support for the idea that large-scale struc-

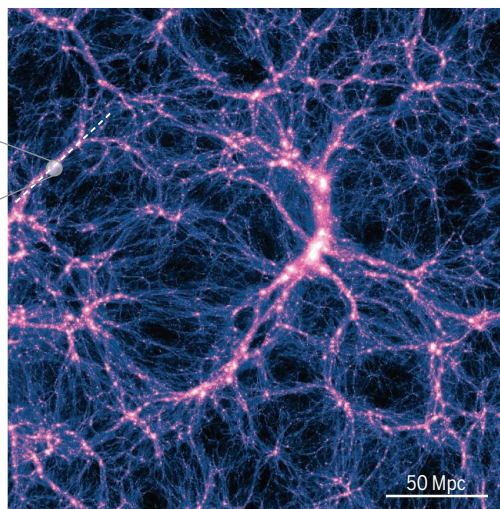
## Do planes of satellites defy the standard cosmological model?

The Centaurus A galaxy is surrounded by a plane of satellites, which defies naïve interpretations of the standard cosmological model. A possible reconciliation of theory and observation lies in correlations induced by the large-scale distribution of matter. Galaxies trace high-density filaments of dark matter; these filaments might be responsible for the observed planes of satellites in the nearby universe.

● Approaching galaxies ● Receding galaxies  
● Not part of the plane



Centaurus A (half-light radius: 25,000 light-years), appears to host a plane of satellite galaxies, much like the Milky Way and Andromeda.



The large-scale distribution of matter in the TNG300 numerical simulation is shown. Light (dark) colors denote regions of high (low) dark matter density. 1 Mpc =  $3.26 \times 10^6$  light-years.

Department of Astronomy, University of Texas at Austin, Austin, TX 78712, USA. Email: mbk@astro.as.utexas.edu

ture is responsible for the existence of the observed satellite planes.

Furthermore, only a subset of the satellites in each of the three systems (the Milky Way, Andromeda, and Centaurus A) lie in planes. Could the observations of satellite planes be a result of astronomers looking too hard for any sign of nonuniform structure in the distribution of satellite galaxies and declaring success upon finding it?

Müller *et al.* argue that this is not the case. Fortunately, a definitive resolution may be in sight. The major uncertainty in assessing the origin of the observed planar structures is the perpendicular (out-of-plane) motion of the galaxies: True corotation, indicating a dynamically stable plane, would require these velocities to be minimal. If the perpendicular velocities are instead comparable to the in-plane velocities, then the observed configurations are likely to be chance alignments. Obtaining the requisite tangential velocities is observationally prohibitive for galaxies in Centaurus A, but detecting (or excluding) large tangential velocities in the Andromeda system is firmly within the grasp of the Hubble and James Webb Space Telescopes over the next few years.

Steven Weinberg once noted that “Physics thrives on crisis” (10). In this light, the current problem for astronomers and physicists is arguably the lack of a major crisis. They are in the uncomfortable position of having a theory— $\Lambda$ CDM—that explains most astronomical observations by invoking a postulated material that has eluded increasingly sophisticated attempts at detection.

The distribution of satellites around nearby galaxies does not yet rise to the level of a crisis driving us to a new understanding, but the results of Müller *et al.* raise the stakes. Perhaps most excitingly, any potential resolution of the puzzle of satellite planes is interesting: At worst, we improve our understanding of galaxy formation; at best, we are led to a deeper understanding of the laws of physics. If this and other issues related to small galaxies (11) do not constitute a crisis, they may have to suffice until the crisis gets here. ■

#### REFERENCES

1. O. Müller, M. S. Pawłowski, H. Jerjen, F. Lelli, *Science* **359**, 534 (2018).
2. S. White, M. Rees, *Mon. Not. R. Astron. Soc.* **183**, 341 (1978).
3. V. Springel, C. Frenk, S. White, *Nature* **440**, 1137 (2006).
4. M. Vogelsberger *et al.*, *Mon. Not. R. Astron. Soc.* **444**, 1518 (2014).
5. J. Schaye *et al.*, *Mon. Not. R. Astron. Soc.* **446**, 521 (2015).
6. D. Lynden-Bell, *Mon. Not. R. Astron. Soc.* **174**, 695 (1976).
7. R. A. Ibata *et al.*, *Nature* **493**, 62 (2013).
8. T. Buck, A. Dutton, A. Maccio, *Mon. Not. R. Astron. Soc.* **460**, 4348 (2016).
9. P. Kroupa, C. Theis, C. Boily, *Astron. Astrophys.* **431**, 517 (2005).
10. S. Weinberg, *Rev. Mod. Phys.* **61**, 1 (1989).
11. J. Bullock, M. Boylan-Kolchin, *Annu. Rev. Astron. Astrophys.* **55**, 343 (2017).

10.1126/science.aar7375

#### GENOMICS

## Tracing single-cell histories

Origins of mutations in single cells during human brain development and aging are revealed

By Je H. Lee

**D**NA mutations accumulate at a steady pace across the human genome, passing from one generation to another. On the basis of the degree of shared mutations, a genealogical relationship can be reconstructed from ancient and modern individuals, allowing one to go back hundreds of thousands of years in human evolutionary history (1). Instead of comparing individuals, on pages 550 and 555 of this issue, Bae *et al.* (2) and Lodato *et al.* (3), respectively, assessed the rate of DNA mutation in single cells from developing and aging human brains, revealing mutational histories in neurodevelopment, aging, and neurodegeneration. These approaches also have implications for understanding complex diseases that could result from somatic mutations that arise later in life, such as cancer.

De novo mutations in the DNA of egg or sperm can be associated with devastating disorders affecting young individuals (4). These germline mutations are widely detected because all cells in the body inherit them. By contrast, somatic DNA mutations sporadically occur throughout the life of an organism (that is, postzygotically) as a result of DNA damage and errors in DNA replication or repair. When somatic mutations occur early in life in dividing cells, they are found in a large number of cellular descendants. If mutations occur in dividing cells as humans age, they are found in only a limited number of cells, resulting in tissue mutational mosaicism (see the figure). The inheritance pattern of mutations in cells within a tissue can be used to establish a temporal or genealogical relationship of mutations to better understand the role of mutational mosaicism in human diseases. Bae *et al.* analyzed the rate and origin of somatic mutations in the brain occurring before birth that could predispose to functional alterations, including neuropsychiatric or neurodevelopmental disorders. Lodato *et al.* address whether mutations continue to occur later in life and whether DNA mutagenesis in

nondividing cells is associated with neurodegenerative disorders.

To study somatic mutations that arise in development or during aging, an accurate single-cell whole-genome sequencing (WGS) method is necessary. This method requires sensitive whole-genome amplification techniques (5); however, DNA nucleotide bases are susceptible to damage (which can lead to mutation), and enzymes used in amplification introduce additional errors. Indeed, the number of false-positive single-nucleotide mutations [also referred to as single-nucleotide variations (SNVs)] can be

**“...advances in single-cell WGS [whole-genome sequencing] allow for examining recent, uncommon mutational events specific to each cell...”**

as high as  $10^4$  in single-cell WGS (5), vastly outnumbering naturally occurring SNVs ( $10^2$  to  $10^3$  per cell). Even in cancer or population genome-sequencing projects, mutagenic DNA damage can be a major source of sequencing errors for rare SNVs (6).

Building on previous methodology (7), Bae *et al.* addressed this challenge by adopting a single-cell cloning method using cultured neuronal precursors from three human fetal brains at 20 weeks of gestation. If a cell contains a true heterozygous mutation, the wild-type and mutant alleles in the genome should each comprise ~50% of the cell population, even after multiple cell divisions. If mutations occur during cell culture or DNA amplification, the frequency of the heterozygous mutation drops to less than 25%. With this criterion, Bae *et al.* sequenced single cells from 31 clonal cultures, identifying somatic mutations by performing clone-to-clone and clone-to-tissue comparisons. This reduced the rate of false positives to 5% and false negatives to 17% in single-cell clones.

Subsequently, they found 200 to 400 mosaic SNVs in cultures of each neuronal precursor, with the most common substitutions

Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring Harbor, NY 11724, USA. Email: jlee@cshl.edu

being C-to-A transversions (a common result of oxidative damage) and 5-methylcytosine-to-thymine transitions (common in germline SNVs). Assuming a linear increase in somatic mutations over time, they estimate 8.6 mutations per cell division, a significant increase compared to mutation rate estimates in early postzygotic cell divisions (1.3 mutations per cell division) (8, 9). Extending these estimates to regions where neurogenesis continues throughout life (for example, in the hippocampus), Bae *et al.* hypothesize that mutational mosaicism arising from dividing neuronal precursors over time may alter neuronal function in the mature brain for processes such as memory and may thus contribute to neurodegeneration.

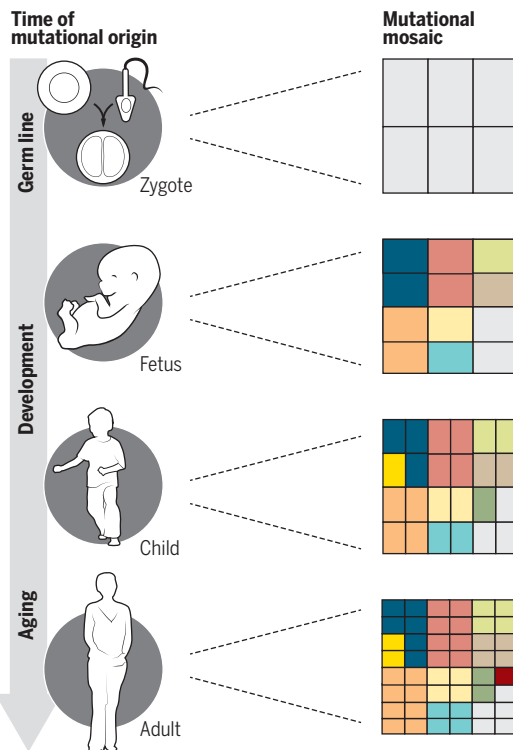
Lodato *et al.* took a different approach. Examining postmitotic neurons, they used an algorithm to extrapolate the frequency of somatic mutation using nearby known germline SNVs. With this method, mutations caused by DNA damage or amplification errors occur on one DNA strand but not the other, whereas naturally occurring somatic mutations are present on both strands. Although only mutations near germline SNVs can be discovered this way, it does not require cell culture or clonal expansion of individual cells, as in other approaches (2, 7).

Lodato *et al.* carried out single-cell WGS on 93 normal postmitotic prefrontal cortex (PFC) neurons in individuals ranging from 4 months to 82 years of age, 26 normal hippocampal dentate gyrus neurons, and 42 PFC neurons from patients with DNA damage repair-deficient neurodegenerative disorders. Notably, they found 300 to 900 SNVs per cell in PFC neurons within 1 year of birth, consistent with the estimate of Bae *et al.* In addition, they observed a twofold higher rate of somatic mutations in hippocampal dentate gyrus neurons (with 40 mutations occurring per cell each year) than in PFC neurons, as was also proposed by Bae *et al.* Furthermore, patients with DNA damage repair-deficient neurodegenerative disorders showed a twofold increase in somatic mutations in postmitotic neurons across different age groups. They also detected three distinct mutational signatures in all of the cells studied, suggesting multiple mechanisms of somatic mutation associated with brain development, aging, and neurodegeneration.

Although these studies do not provide definite mechanisms of how somatic mutations occur or establish their functional

## Reconstructing cellular mutation history

Mutations can be present in most cells or in just one cell within a tissue, depending on when they occur. Emerging sequencing technologies can accurately profile genome-wide somatic mutations, even in the tiniest mosaic piece—a single cell.



consequences, it is exciting that Bae *et al.* and Lodato *et al.* utilized readily accessible methods to accurately sequence single-cell whole genomes and reconstructed mutation histories (akin to a “mutational clock”) with single-cell resolution. Because recent mutations are restricted to the tiniest mosaic pieces (see the figure), mutational clocks had previously been limited to the early mutation events shared by many cells. These technological advances in single-cell WGS allow for examining recent, uncommon mutational events specific to each cell, considerably widening the temporal breadth of mutation-history reconstruction.

In so doing, fundamental biological questions that had been inaccessible can now be addressed, especially regarding the onset and progression of relatively recent somatic mutations in development or aging. Indeed, these studies now firmly establish the presence of ongoing DNA damage later in life that accelerates in neurodegeneration (3). They also demonstrate the occurrence of coding and noncoding mutations that arise later during neurogenesis (2), potentially dividing the brain into smaller mutational territories, with implications for brain connectivity, function, and predisposition to neurological disorders.

The accuracy of single-cell WGS continues to improve, including new amplification methods with a lower error rate and more uniform coverage of the genome (5). Unfortunately, single-cell WGS is limited to a small number of cells because of throughput and cost bottlenecks. To accurately reconstruct mutation history and mosaicism throughout life, it is necessary to trace cell lineages comprehensively. Recently, somatic alterations induced by members of the CRISPR-associated (Cas) family of DNA nucleases have been used to reconstruct cell lineages in model organisms and cell lines (10, 11). In this method, Cas9-induced alterations within a single synthetic locus can be traced. Such reporters are transcriptionally active, making them compatible with high-throughput single-cell RNA sequencing (12) to “read” the reporter and identify cell type-specific RNA signatures. Indeed, reporters can even be programmed to record sequential events in cells (13), potentially revealing environmental interactions that shape the tissue hierarchy and cell lineages within.

One day, it might be possible to engineer cell lineage and activity recorders throughout an entire organism, followed by comprehensive single-cell sequencing of recorders containing induced mutations. It might then be possible to compare naturally occurring somatic mutations in a subset of single cells using WGS across specific developmental stages. This could reveal biological processes that affect the mutational clock across different tissues or cellular activities and expand our understanding of mutational signatures (14, 15), especially if such signatures are conserved across species. Precisely answering when, how, and where mutational events occur in cells in different tissues will help clarify important questions in development, aging, and disease research. ■

## REFERENCES

1. M. Kuhlmann *et al.*, *Nature* **530**, 429 (2016).
2. T. Bae *et al.*, *Science* **359**, 550 (2018).
3. M. A. Lodato *et al.*, *Science* **359**, 555 (2018).
4. A. Poduri *et al.*, *Science* **341**, 1237758 (2013).
5. C. Chen *et al.*, *Science* **356**, 189 (2017).
6. L. Chen *et al.*, *Science* **355**, 752 (2017).
7. S. Behjati *et al.*, *Nature* **513**, 422 (2014).
8. R. Rahbari *et al.*, *Nat. Genet.* **48**, 126 (2016).
9. Y. S. Ju *et al.*, *Nature* **543**, 714 (2017).
10. A. McKenna *et al.*, *Science* **353**, aaf7907 (2016).
11. K. L. Frieda *et al.*, *Nature* **541**, 107 (2017).
12. J. Cao *et al.*, *Science* **357**, 661 (2017).
13. S. L. Shipman *et al.*, *Nature* **547**, 345 (2017).
14. L. B. Alexandrov *et al.*, *Nat. Genet.* **47**, 1402 (2015).
15. I. Martincorena, P. J. Campbell, *Science* **349**, 1483 (2015).

10.1126/science.aar6335

## CORPORATE SUSTAINABILITY

# Scientifically assess impacts of sustainable investments

Metrics can inform investors wary of “green washing”

By C. J. Vörösmarty,<sup>1,2</sup> V. Rodríguez Osuna,<sup>1</sup> D. A. Koehler,<sup>3</sup> P. Klop,<sup>4</sup> J. D. Spengler,<sup>5</sup> J. J. Buonocore,<sup>5</sup> A. D. Cak,<sup>1</sup> Z. D. Tessler,<sup>1</sup> F. Corsi,<sup>1,2</sup> P. A. Green,<sup>1</sup> R. Sánchez<sup>5</sup>

**T**he practice of selecting and managing financial assets based on their social and environmental performance is undergoing rapid growth and fundamental change. Investors are increasingly pressed by asset owners to prove how one company's practices are materially more or less sustainable than those of another. Yet, the basic information that companies declare is hardly standardized and is difficult to verify, with unreliable assertions (1) that are widely criticized as “green washing.” Metrics are mainly restricted to documenting changes to internal business practices but offer limited guidance on whether a company's actions, products, and services promote human well-being or preserve environmental integrity in the external, real-world domain, fueling reluctance on the part of otherwise enthusiastic investors (2, 3). It is here where science can play an important role. Our consortium of an asset owner, an asset manager, and two research universities is designing a next generation of traceable indicators to quantify external context and impact of investments and place these into a decision-making framework useful to investors. Tests of these science-based sustainability metrics are under way on a \$2.1 billion portfolio of public equities invested on behalf of a large European pension fund. Estimated growth in assets under professional management for sustainable investment is well above market growth rates

worldwide (2, 3). In the United States, sustainable investments grew 33% from 2014 to 2016 and now total nearly \$9 trillion, or about one-fifth of all professionally managed assets (4). Several initiatives have emerged to help guide sustainable investing, including the Principles for Responsible Investment, Sustainability Accounting Standards Board, and Global Reporting Initiative. While such efforts' primary focus on minimizing a company's operational and supply chain risks is important—e.g., via emission reductions, efficiency gains, or improved worker health and safety—such evaluation procedures fall far short of what is needed for true verification of company-level sustainability performance and may produce assessments and investment decisions that are at odds with sustainable development.

***“...a more comprehensive approach to impact measurement could be realized if the financial sector joins forces with the scientific community.”***

Although the idea of linking actions to environmental and social impacts is not new (e.g., in economics and policy analysis), its broader value to financial decision-making, in a way that makes sense to investors, has yet to be fully embraced. Many large, publicly traded companies are actively developing and putting to work technologies that create business opportunities going well beyond compliance per se, as with the explosion of energy- and life-saving innovations in the 21st-century automobile, whose genesis can be traced to enforcement of the U.S. Clean Air Act of 1970 (5). Nevertheless, the positive social and environmental impacts of individual companies and their business activities have proven difficult to systematically evaluate.

## SEEKING CONTEXT

One active test bed for assessing corporate sustainability performance is in carbon

emissions management, with various protocols in place (e.g., Carbon Disclosure Project, Science-Based Targets). Given the well-mixed nature of the atmosphere, emitting one less ton of carbon should yield an identical positive impact regardless of the source. Thus, estimating climate change mitigation benefits becomes the relatively straightforward task of crediting the efficiency of a company's operations in avoiding emissions, often normalized to revenue, but with little attention paid to how those operations contribute to other environmental risks or how its products and services ultimately generate positive or negative impact (6).

Carbon accounting will have limited usefulness in the broader development agenda, like that represented by the United Nations Sustainable Development Goals (SDGs), where climate change mitigation is but one of 17 focal points. The SDGs require additional metrics on human health, access to clean water, poverty alleviation, biodiversity conservation, and many others. Positive contributions can no longer be considered to operate in a well-mixed box, as for climate. Given the social and economic disparities that the SDGs seek to redress, investment strategies need to target particular human beneficiaries or environmental systems. What is needed is the notion of context, by which a company's business model or manufacturing practices generate real-world impacts.

Context-oriented metrics move the investment process one step closer to identifying actual solutions, which tabulations of modified corporate operations alone do not necessarily uncover. The approach, then, is to go beyond documenting changes in facility-scale outputs and combine them with what we term outcomes and impacts, which by their very nature evaluate corporate products and services within broader environmental or human beneficiary settings. Such an expanded view of business practices, services, or products consumed begins to uncover counterintuitive results that are often masked by a singular focus on outputs.

For example, for publicly held companies selling renewable energy technologies, reductions in CO<sub>2</sub> and other air pollutant emissions at the smokestack (outputs) produce ambient air quality improvement (outcomes) and associated health benefits (impacts). However, the tabulated impacts are often poorly correlated with the originating outputs, especially if the investment choices embody large differences in plant technologies, the local electrical grid, and potential beneficiary populations living downwind (7). This means that a small renewable energy company in a re-

<sup>1</sup>Environmental Sciences Initiative, Advanced Science Research Center at the Graduate Center, City University of New York (CUNY), New York, USA. <sup>2</sup>Civil Engineering Department, The City College of New York, CUNY, New York, USA. <sup>3</sup>UBS Asset Management, New York, NY, USA.

<sup>4</sup>Responsible Investment at PGGM Investments, Zeist, Netherlands. <sup>5</sup>Center for Health and the Global Environment, Harvard T. H. Chan School of Public Health, Boston, MA, USA. Email: Charles.Vorosmarty@asrc.cuny.edu

gion with high baseline levels of air pollution could generate greater positive impact on human health than a larger company operating in a region where the air is cleaner and the electrical grid emits less pollution per megawatt-hour generated. Apart from its financial performance, that smaller company may be passed over as a good candidate for inclusion within a sustainability portfolio if that judgment were based solely on its much lower outputs of clean energy at the plant level.

Similarly, we have been designing human health impact scores for pharmaceutical companies. The monotherapy drug effectiveness from U.S. Food and Drug Administration–approved clinical studies is combined with the number of prescriptions

to an impact have been well enough established in many cases for us to adopt some basic assumptions to compute corporate contributions to sustainability—for example, knowing that CO<sub>2</sub> and other emissions from power plants contribute to climate change and bear negative human health effects and that renewables reduce both emissions and their climate and health burdens. Until there is more comprehensive disclosure of information, however, estimating the benefits that a company’s products and services accrue thus requires some pragmatic, simplifying assumptions (see the table).

In our pilot work, we use relatively abundant revenue data as a proxy for business activity at the country level, based on available

but also on data from recent advances in Earth observation and modeling, epidemiology, public health, nutrition science, demographic mapping, and life-cycle analysis of resource production and consumption, which all have the advantage of being, for the most part, transparent, reproducible, and peer-reviewed. Although these technical capabilities are supporting breakthroughs in understanding the complexity and causal links that bind human-environment systems, we repurpose the science to define context. In some of our work on air and water pollution abatement, we have been successful in “draping” corporate output data onto geospatial Earth system and human dimension data sets to compute outcome and impact indicators that explicitly link corporate manufacturing activities, products, and services to environmental and human beneficiaries (7, 8).

Moving a step further, combining such measures with reported financial data would create intuitive, company-level metrics of value to sustainable investment decision-making—for example, using our computed impacts per dollar of revenue as an indicator of sustainable investment “efficiency” to facilitate comparisons across companies. This mixture of quantitative information brings with it the critical issue of data harmonization in terms of thematic, spatial, and temporal coherency, but frameworks from other domains, such as the Earth system modeling framework community (9), provide useful guideposts. Protocols being developed under the next generation of the Global Earth Observation System of Systems (NextGEOSS) to accelerate the adoption of Earth observations into business practices are another such resource.

Context-based metrics must also be able to evaluate cross-sectoral synergies, trade-offs, and unintended consequences associated with multiple sustainability end points (e.g., improved water management, human health, and biodiversity conservation). A company that is highly rated for reducing air pollution could simultaneously present material risks to occupational health and safety, or at the expense of clean water. Context also helps investors to more clearly define a company’s potential reputational risk. For example, portfolio managers might automatically consider selling off holdings of a pharmaceutical company that is depleting water in a dry region and risking a major public backlash, but that decision might sensibly be reconsidered should the facilities serve large numbers of people in a developing country, where the positive health benefits per dollar invested in drugs are far higher than in wealthier nations (10).

A first and fundamental step is to establish a formal dialogue space for information

Context-based metrics for investment decisions

Examples of metrics applied to holdings from the pension fund portfolio tested. Current data limitations necessitate practical simplifications that affect accuracy and precision. Access to more precise data on corporate products and services would enable the impact metrics to more fully capture environmental and public health context. The framework can adopt additional modules, representing alternative state-of-the-art impact assessment models and/or their outputs.

ASSESSMENT MEASURES	WASTEWATER TREATMENT	RENEWABLE ENERGY	PHARMACEUTICALS
<b>Impact metrics</b> <i>Data source: Combined business activity/technology/sociobiogeophysical impact metrics, as proposed here</i>	<b>Outputs:</b> Volumes of wastewater treated (e.g., cubic meters per day) <b>Outcomes:</b> Pollutant concentration reductions in receiving waters (e.g., milligrams per liter) <b>Impacts:</b> Beneficiary populations drawing drinking water from downstream (1000s), stream length of improved fish habitat (kilometers below threshold)	<b>Outputs:</b> Emissions reduced (CO <sub>2</sub> , PM <sub>2.5</sub> , SO <sub>2</sub> , NO <sub>x</sub> ) (e.g., metric tons per year) <b>Outcomes:</b> Ambient air pollution improvement (e.g., parts per million; number of smog alerts) <b>Impacts:</b> Lives extended, hospitalizations and sick days avoided, reduced health-care spending	<b>Outputs:</b> Spatially distributed drug sales <b>Outcomes:</b> Target populations reached (e.g., millions; % of total) <b>Impacts:</b> Lives extended, hospitalizations and sick days avoided, reduced health-care spending
<b>Financial (standards-based)</b> <i>Data source: Mandated corporate financial statements</i>	Valuation Corporate strategy and relative competitive position		
<b>Environmental, social, and governmental norms (nonfinancial performance)</b> <i>Data source: Voluntary or mandated corporate self-disclosure, sustainability reports</i>	Facility worker health and safety Environmental compliance Governance and ethical policies		

per drug, epidemiological information from peer-reviewed articles, and official occupational health reports for related diseases to estimate life spans extended, hospitalizations averted, and sick days prevented by every drug. This information is then aggregated to the national scale, by company, to estimate overall health impact.

Because companies typically fail to disclose many of the details of what products they sell, where, and in what quantity, we are not yet able to explicitly link an investment dollar to a specific yield of impacts across an entire portfolio of large global companies. However, the causal chains linking an action

corporate disclosures. We link these to estimates made by impact models operating on a geographical scale that captures the particular relevant health and environmental context. For example, the wastewater treatment impact model operates regionally with product deployments that are assessed with respect to contrasting dilution potentials in receiving waters. Health benefits are computed using country-level revenue information prorated by population.

DESIGN AND PRACTICE

We see it as essential that metrics rely not only on self-disclosed corporate reports

exchange and best practices, organized as a strategic alliance of asset owners and managers, companies, auditors, nongovernmental organizations, and scientists. The partnership would emphasize codesign of unified output-outcome-impact assessment frameworks and catalyze rational debate on the specific kinds of corporate disclosure and scientific data, their reporting frequencies, and spatial characteristics. In addition, the consortium would create momentum for third-party sustainability reporting and auditing standards, setting the stage for verifiable broad-scale indices. Established efforts like the Global Impact Investing Network, UN Global Compact, Sustainable Development Solutions Network, and World Business Council for Sustainable Development provide ready-made forums but would need to expand the current dialogue to consider new approaches that uncover and measure context.

## NEW BUSINESS MODELS

To itself be sustainable, sustainable investing must ultimately yield financial returns. In a global survey of sustainable investment organizations (3), 98% of respondents reported that their investments either met or exceeded expectations for social and/or environmental performance, whereas 91% reported satisfaction with financial performance. And, despite the willingness of some investors to accept below-market performance on some of their managed assets in the name of public good or to seed new investment frontiers (3), the assessment model being tested here assumes that sound business practices based on durable social or environmental benefits translate into profit. This investment strategy rewards a long time horizon—the essence of sustainability—in lieu of short-range financial performance, which preoccupies much of the current market.

Although our own impact investment experiment has just begun, we see early promise that the fund being tested could ultimately generate returns competitive with a broad benchmark while producing measurable environmental and health benefits (11). Before initiating our study, stocks had already been selected on the basis of financial valuation; positive performance on environmental, social, and environmental risks in their operations and supply chains; and sales of products and services in four critical challenge arenas—water, climate change mitigation, human health, and food security. These impact categories, selected by the pension fund trustees themselves, are each indispensable to the global economy and present considerable environmental and social risks. Representative companies

were also chosen based on how well their products and services potentially contribute to the SDGs. We are testing our metrics on this subset of pension fund assets.

Our explicit assumption is that financial and impact metrics can be combined in a portfolio without adversely affecting financial returns. Most important, well-designed impact metrics for public equity portfolios can be used effectively to inform asset owners, influence stakeholders, and shape investment decisions. The global effect of this transformation would be consequential and would reasonably be expected to save trillions of dollars by mitigating climate change, the growing water crisis, and widespread loss of ecosystem services (12–14).

In the end, a more comprehensive approach to impact measurement could be realized if the financial sector joins forces with the scientific community. The dialogue space proposed here could be used to negotiate the terms of engagement and confront difficult carrot-and-stick issues such as that regarding corporate data disclosure. If one agrees with the wisdom of mandated financial disclosures such as under the U.S.

## ***“The approach...is to... evaluate corporate products and services within broader environmental or human beneficiary settings.”***

Securities and Exchange Commission, built on well-established accounting principles and standards, then it is not unreasonable to warrant authentication of nonfinancial information that investors would find useful. In fact, more than 50 stock exchanges worldwide (e.g., Securities and Exchange Board of India) encourage or mandate such disclosures today (15). Might inclusion of a particular company in a sustainable asset portfolio be made contingent upon authentication of mandatory self disclosures? If so, the reward to that company, beyond direct revenue generated, would be a bona fide declaration that it is promoting sustainable business practices, especially if the impact measures show the company outperforming its peers or broader benchmarks. The willingness of corporations to share detailed information in such an alliance—as increasingly demanded by their investors—will be as much a shift in corporate culture as in how companies report on sustainability issues. If such a change is realized, we foresee a huge investment space opening, based on verifiable net positive impacts.

In the 12 years remaining to achieve the UN's Sustainable Development Goals, we see an opportunity to road test this and similar context-based frameworks. If sustainable investors can adopt suitable impact metrics to identify sustainable investment opportunities, the community of nations will have in the business community a new and reliable ally, thereby increasing the likelihood that the goals will actually be realized within this short, ambitious time frame. ■

## REFERENCES AND NOTES

1. The Economist Intelligence Unit, The future of climate-related disclosure (2016) (available at <https://perspectives.eiu.com/sustainability/future-climate-related-disclosure>).
2. E. Lewis, A. Pinchot, G. Christianson, “Navigating the Sustainable Investment Landscape” [World Resources Institute (WRI), Washington, DC, 2016] (available at [https://www.wri.org/sites/default/files/Navigating\\_the\\_Sustainable\\_Investment\\_Landscape.pdf](https://www.wri.org/sites/default/files/Navigating_the_Sustainable_Investment_Landscape.pdf)).
3. Global Impact Investing Network (GIIN), “Annual Impact Investor Survey: 2017” (New York, 2017), (available at [https://theGIIN.org/assets/GIIN\\_AnnualImpactInvestorSurvey\\_2017\\_Web\\_Final.pdf](https://theGIIN.org/assets/GIIN_AnnualImpactInvestorSurvey_2017_Web_Final.pdf)).
4. US SIF Foundation, “Report on US Sustainable, Responsible and Impact Investing Trends, 2016” (The Forum for Sustainable and Responsible Investment, Bloomberg, New York, 2016).
5. A. Johnson, “Environmental regulation and technological development in the U.S. auto industry: The causes and consequences for sustained economic development” (Washington Center for Equitable Growth, Washington, DC, 2016).
6. B. L. Caldecott, L. Kruitwagen, “Guest Opinion: How asset-level data can improve the assessment of environmental risk in credit analysis” (S&P Global Market Intelligence, New York, 2016) (available at [https://www.globalcreditportal.com/ratingsdirect/renderArticle.do?articleId=1728982&SctArtId=402968&from=CM&nsf\\_code=LIME&sourceObjectID=9796605&sourceRevid=1&fee\\_ind=N&exp\\_date=20261003-21:55:05](https://www.globalcreditportal.com/ratingsdirect/renderArticle.do?articleId=1728982&SctArtId=402968&from=CM&nsf_code=LIME&sourceObjectID=9796605&sourceRevid=1&fee_ind=N&exp_date=20261003-21:55:05)).
7. J. J. Buonocore et al., *Nat. Clim. Chang.* **6**, 100 (2016).
8. P. A. Green et al., *Global Environ. Change* **34**, 108 (2015).
9. Earth System Modeling Framework (ESMF) (2017) (available at [www.earthsystemcog.org/projects/esmf/](http://www.earthsystemcog.org/projects/esmf/)).
10. P. M. Danzon et al., *Health Aff.* **30**, 1529 (2011).
11. PGGM, Responsible Investment Reports (2017) (available at <https://www.pggm.nl/english/what-we-do/Pages/Responsible-investing-reports.aspx>).
12. The Economist Intelligence Unit, “The Cost of Inaction: Recognising the value at risk from climate change” (The Economist Intelligence Unit Limited, London, New York, Hong Kong, Geneva, 2015) (available at [www.eiuperspectives.economist.com/sustainability/cost-inaction](http://www.eiuperspectives.economist.com/sustainability/cost-inaction)).
13. World Economic Forum (WEF), “From the Margins to the Mainstream Assessment of the Impact Investment Sector and Opportunities to Engage Mainstream Investors” (World Economic Forum, Geneva, 2013) (available at [http://www3.weforum.org/docs/WEF\\_IL\\_FromMarginsMainstream\\_Report\\_2013.pdf](http://www3.weforum.org/docs/WEF_IL_FromMarginsMainstream_Report_2013.pdf)).
14. R. Costanza et al., *Global Environ. Chang.* **26**, 152 (2014).
15. E. J. Hespenheide and D. A. Koehler, “Disclosure of Long-term Business Value: What Matters?” (Deloitte University Press, 2013) (available at [https://www2.deloitte.com/content/dam/insights/us/articles/disclosure-of-long-term-business-value/DUP150\\_Reporting\\_What\\_Matters.pdf](https://www2.deloitte.com/content/dam/insights/us/articles/disclosure-of-long-term-business-value/DUP150_Reporting_What_Matters.pdf)).

## ACKNOWLEDGMENTS

Main sponsorship for this work was provided by UBS Asset Management, with additional support from the Advanced Science Research Center at the Graduate Center, CUNY and Grove School of Engineering at The City College of New York, CUNY. We also thank three anonymous reviewers for their helpful inputs. The views expressed herein reflect those of the authors alone, are subject to change, and do not necessarily represent the views, policies, or positions of their employers. Industry and related data contained herein has been compiled in good faith from sources believed to be reliable.

10.1126/science.aao3895

BOOKS *et al.*

## SCIENCE COMMUNICATION

# Picturing the future

A new film celebrates a mid-century illustrator who stoked imaginations with prescient comics

By **Dov Greenbaum**

**D**irected by Brett Ryan Bonowicz (*The Perfect 46*, 2014), *Closer Than We Think* is an eponymous documentary about a mid-20th-century comic strip and its creator, Arthur Radebaugh. The film provides important insight into what made the comic a foundation for other popular science fiction endeavors. More important, however, it shows how the strip directly encouraged nascent engineers and scientists to actualize the future described therein. To this end, Radebaugh's efforts succeeded where so many other science communication efforts continue to fail.

Published weekly from 1958 through 1963 in the Sunday funnies of more than 200 newspapers, Radebaugh's educational futurism provided playful, tongue-in-cheek visions of a technological utopia that the public

could readily understand and appreciate. His predictions were based on information that he obtained from researchers, which he dutifully recorded in each strip. In one panel, for example, a colorful rendering of a solar-powered car accompanies a pithy interview with the vice president of the Chrysler Corporation, James Zeder. "If we continue to increase the efficiency of these [silicon] converters, and if we are able to develop small, efficient energy storage cells, solar powered cars will be feasible," Zeder explains. In the future, translates Radebaugh, "[y]our solar sedan will take energy from sunrays and store it in accumulators that work like a battery."

Bonowicz's documentary stresses that Radebaugh's relationships with researchers were central to the success of his predictions, which included the VCR ("[W]hen a worthwhile program comes over the air ... you'll be able to preserve both the picture and sound on tape for replaying at any time"), Skype and the globalization of the job market, and driverless vehicles ("Future trains will be fully automatic—robots that can regulate their own

speed and control their own movements").

The film suggests that Radebaugh's ability to connect to the lay public can also be attributed to his technological optimism. By rejecting the dystopian tropes common to science fiction, he enticed a younger generation to become interested in science and technology, encouraging them to envision fantastical technological futures. This mirrored a growing feeling in society: Even during the chaos and fear of the Cold War—*Closer Than We Think* was bookended by Sputnik and Kennedy's assassination—it was also a time of growing hope that technology could be harnessed for the betterment of mankind.

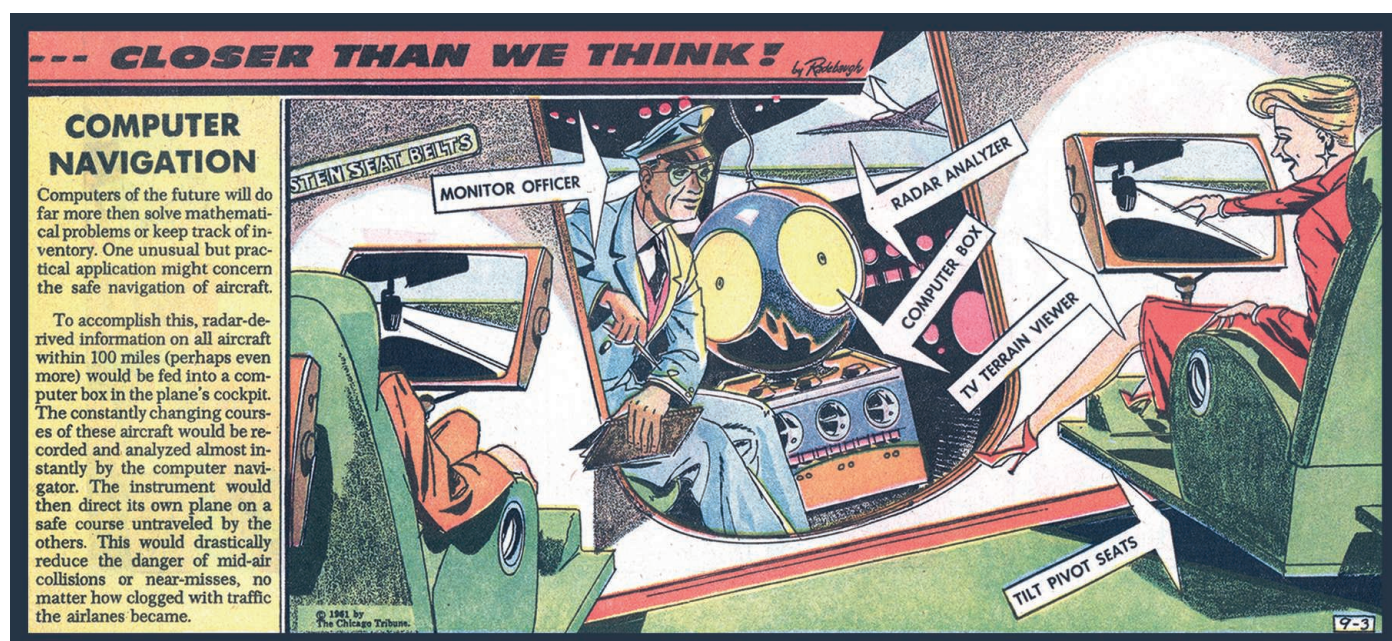
Before *Closer Than We Think*, Radebaugh was a well-regarded commercial illustrator who saw the demand for his work plummet as photographs began to take the place of drawings in print publications and television advertising began to obtain prominence. He moved to Grand Rapids, Michigan, in 1967, ostensibly to retire with his wife, Nancy. There, he fell on hard times and lived out his days in relative obscurity.

Bonowicz's film, although an exhaustive examination of Radebaugh's art and life, would have benefited from a more thorough exploration of both his scientific muses and his sources. A more in-depth exploration of those who were influenced by his work would also have been worthwhile.

Radebaugh's cartoons have an important take-home message for scientists: Learn to communicate with the public. They are listening, and they want to learn. You never know who you might influence. ■

10.1126/science.aar6193

The reviewer is at the Zvi Meitar Institute for Legal Implications of Emerging Technologies, Herzliya, Israel, and the Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT, USA. Email: dov.greenbaum@idc.ac.il



Radebaugh's vivid depictions of soon-to-be technologies resonated with the American public in the mid-20th century.



## MEASUREMENT

# The fetishization of quantification

A historian explores the dark side of metric-based performance evaluation

By Theodore M. Porter

**T**he notion of a “metric” as a performance measure became familiar in the 1970s and 1980s as a tool of business management. Almost immediately, its use was extended to the assessment of a range of activities and institutions, from medical outcomes and educational programs to military projects. The credibility of metrics rests in part on an affiliation with ideals of business efficiency and in part on the supposition that measurement is tantamount to science.

Although the numbers whose “tyranny” forms the subject of Jerry Muller’s timely book share some of the attributes of scientific measurement, their purposes are primarily administrative and political. They are designed to be incorporated into systems of what might be called “data-ocracy,” often for the sake of public accountability: Schools, hospitals, and corporate divisions whose numbers meet or exceed their goals are to be rewarded, whereas poor numbers, taken to imply underperformance, may bring penalties or even annihilation. In *The Tyranny of Metrics*, Muller shows how teachers, doctors, researchers, and managers are driven to sacrifice the professional goals they value in order to improve their numbers.

Scientists these days sometimes contrast “data-driven” science with institutional

routines—and especially with the fake news that seems now to run rampant in public debate. Yet metrics, when used to ease administration, have become graveyards of such naïve idealism. Data are almost never, as the etymology implies, “given,” but have to be actively made or taken. The appropriate uses of data are subject to intense discussion and debate and depend in most cases on the framing and interpretation of models.

Metrics are typically designed for routine administrative use, often in a context of political disagreement. As Muller shows, the expectation that they should be transparent leaves little space for expert interpreters to adjust or explain misleading measures of therapeutic effectiveness, school quality, research impacts, or even corporate profitability. Because none of these are concrete in a way that would permit rigorous measurement, such metrics can only function as indicators and can lead to terrible consequences when they are imposed reflexively.

The American obsession with metrics of school performance, which took off in the early 1980s and is now increasingly globalized, has probably done more harm than good, suggests Muller. Standardized tests are unable to capture what is really valuable in an education, and schools have no control over many of the variables that enter into measures of student performance. In one specific sense, the metrics are counterproductive

Overvaluing standardized test scores creates perverse incentives for teachers, argues Muller.

because they encourage a focus on test performance rather than on the knowledge that a proper curriculum aims to advance. Metrics of this sort privilege symbols over reality.

There are, likewise, many well-documented instances of needy patients refused treatment for the sake of statistics. (“V.A. doctors say rating push hurts patient care,” proclaimed *The New York Times* on the front page on 1 January 2018.) When doctors or hospitals are judged by the success of operations on patients with a particular diagnosis, they face a strong incentive to avoid treating the neediest patients, because these are ordinarily the least likely to recover.

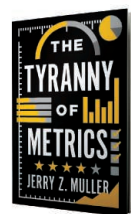
Not infrequently, metrics produce almost the opposite results for which they were intended. Muller gives the example of hospital emergency rooms that improved their metric for timely admission of patients by allowing a line of ambulances to form outside.

Muller’s basic arguments, as he acknowledges, are not new. Thoroughgoing critiques of performance metrics began appearing in the 1970s in response to a wave of extravagant ambitions for numbers-based management. A particularly noteworthy target of such criticism was Robert McNamara, who became infamous for his fetishization of body counts during the Vietnam War.

There were at least two fundamental problems with this metric: first, that the numbers could be fudged—for example, by including noncombatants in the death counts—and second, that the number of deaths did not seem to contribute much to winning the war. We must add the moral point that this metric encouraged the slaughter of innocents.

In 1975, the American social psychologist Donald Campbell and the British economist C. A. E. Goodhart articulated independently the principle that reliance on measurement to incentivize behaviors leads almost inevitably to a corruption of the measures. Muller explains the logic of this corruption and defends, in place of indiscriminate numbers, an ideal of professional knowledge and experience.

Measurement, he concludes, can contribute to better performance, but only if the measures are designed to function in alliance with professional values rather than as an alternative to them. Good metrics cannot be detached from customs and practices but must depend on a willingness to immerse oneself in the work of these institutions. ■



### The Tyranny of Metrics

Jerry Z. Muller  
Princeton University Press, 2018. 232 pp.

The reviewer is at the Department of History, University of California at Los Angeles, Los Angeles, CA 90095, USA. Email: tporter@history.ucla.edu

PHOTO: PEOPLEIMAGES/ISTOCKPHOTO

## LETTERS

Edited by Jennifer Sills

## Malaria in Venezuela requires response

The Americas have witnessed a substantial decline in malaria-related morbidity (62%) and mortality (61%) during the past 15 years as part of the implementation of the Global Malaria Action Plan (1). Venezuela, the first World Health Organization (WHO)-certified country to eradicate malaria in 1961 (2), has been the alarming exception in the region, displaying an unprecedented 365% increase in malaria cases between 2000 and 2015 (1). In 2016 alone, 240,613 malaria cases were officially reported (3). Worryingly, 2017 witnessed an increase of 68% in the cumulative number of cases compared to the previous year, totaling 319,765 malaria cases by 21 October (4). The disease has spread to areas where malaria was eradicated previously (such as near the capital, Caracas), prompting alarm in the health sector. Moreover, malaria cases have overloaded frontier health care infrastructure in Brazil and Colombia, where 78 and 81%, respectively, of imported malaria cases in 2016 originated in Venezuela (5).

Economic and political mismanagement have precipitated a general collapse of Venezuela's health system, creating an ongoing humanitarian crisis with severe social consequences (6, 7). The malaria epidemic has been fueled by financial constraints for procurement of malaria commodities (such as insecticides, drugs, diagnostic supplies, and mosquito nets) and surveillance activities, internal human migration associated with illegal gold mining, and lack of provision and implementation of services (2, 3). The continued upsurge of malaria in Venezuela threatens to become uncontrollable, and it is jeopardizing the hard-won gains in the Americas' elimination agenda and global malaria targets. There is an urgent need for support and action from WHO, United Nations agencies, the U.S. Agency for International Development, the Global Fund, and other international organizations to control this epidemic. Failure to address this dangerous situation may result in one of the worst malaria epidemics in the history of the Americas.

**Maria Eugenia Grillet,<sup>1\*</sup> Leopoldo Villegas,<sup>2</sup> José Felix Oletta,<sup>3</sup> Adriana Tami,<sup>4,5</sup> Jan E. Conn<sup>6</sup>**

<sup>1</sup>Laboratorio de Biología de Vectores y Parásitos,



Venezuelans wait outside a health center for malaria treatment in November, 2017.

Instituto de Zoología y Ecología Tropical, Facultad de Ciencias, Universidad Central de Venezuela, Caracas, Venezuela. <sup>2</sup>Asociación Civil Impacto Social, Tumeremo, Venezuela. <sup>3</sup>Sociedad Venezolana de Salud Pública, Caracas, Venezuela. <sup>4</sup>University of Groningen, University Medical Center Groningen, Department of Medical Microbiology, Groningen, Netherlands. <sup>5</sup>Departamento de Parasitología, Facultad de Ciencias de la Salud, Universidad de Carabobo, Valencia, Venezuela. <sup>6</sup>Griffin Laboratory, Wadsworth Center, New York State Department of Health, Slingerlands, NY 12159, USA. \*Corresponding author. Email: mariaeugenia.grillet@gmail.com

### REFERENCES

1. WHO, *World Malaria Report 2016* (Geneva, World Health Organization, 2016); <http://apps.who.int/iris/bitstream/10665/252038/1/9789241511711-eng.pdf>.
2. P. J. Hotez et al., *PLOS Negl. Trop. Dis.* **11**, e0005423 (2017).
3. WHO, *World Malaria Report 2017* (Geneva, World Health Organization, 2017); <http://apps.who.int/iris/bitstream/10665/259492/1/9789241565523-eng.pdf>.
4. Sociedad Venezolana de Salud Pública (SVSP) y Red Defendamos la Epidemiología Nacional (RDEN), "Reporte sumario de la compleja situación de malaria en Venezuela" (2017); [www.svinfectologia.org/index.php/component/content/article/11-noticias/editorial/631-paludismo-rde-svsp.html](http://www.svinfectologia.org/index.php/component/content/article/11-noticias/editorial/631-paludismo-rde-svsp.html) [in Spanish].
5. ICASO, "Triple threat: Resurging epidemics, a broken health system, and global indifference to Venezuela's crisis" (2017); [www.icaso.org/new-report-details-health-crisis-venezuela/](http://www.icaso.org/new-report-details-health-crisis-venezuela/).
6. A. Tami, *Lancet* **383**, 1968 (2014).
7. J. F. Oletta, R. A. Orihuela, P. Pulido, C. Walter, *Lancet* **383**, 1967 (2014).

10.1126/science.aar5440

## Integrated approach to malaria control

To sustain reductions in child mortality due to malaria, J. Hemingway argues that new public health insecticides are required ("The way forward for vector control," Perspectives, 24 November 2017, p. 998). However, because of the heavy selection pressure placed on mosquito populations,

the use of new compounds will likely lead to the development of new mechanisms to resist them. Integrated vector management programs should therefore devote equal attention to noninsecticide methods.

Numerous biological control agents, from viruses to predatory fish, have been evaluated for effectiveness in controlling malaria mosquitoes. Foremost, the use of the bacterium *Bacillus thuringiensis* var. *israelensis* has substantially diminished malaria when applied to aquatic habitats, such as in urban areas of Tanzania (1) and highland villages in Kenya (2). Adult mosquitoes can also be controlled with entomopathogenic fungi (3). Research has focused on the development of innovative delivery platforms of fungal spores (4) as well as genetic modification of the fungus to increase effectiveness (5).

People affected by malaria often reside in poorly constructed houses with opportunities for mosquito entry. Improved housing and living conditions protect against malaria (6), as well as providing other health benefits, such as reducing exposure to *Culex* mosquitoes, vectors of the parasitic disease filariasis (7).

Lastly, manipulating mosquito behavior through exploitation of volatile infochemicals presents opportunities for pesticide-free vector control. For example, a large intervention trial using odor-baited mosquito traps removed up to 70% of the *Anopheles funestus* population, resulting in a 30% decrease in malaria (8).

Integration of a variety of tools for malaria control is central to the Global Vector Control Response 2017–2030 (9). The plan emphasizes integrated vector management as a pillar of sustainable malaria control, in addition

to engagement of communities, investments in research, and capacity building. Pesticide resistance needs to be addressed in this context, not by a narrow focus on new pesticides alone. Only with an integrated approach can we make malaria elimination a realistic target.

**Constantianus J. M. Koenraadt\***

**and Willem Takken**

Laboratory of Entomology, Wageningen University,  
6700 AA, Netherlands.

\*Corresponding author.

Email: sander.koenraadt@wur.nl

## REFERENCES

1. Y. Geissbühler et al., *PLOS One* **4**, e5107 (2009).
2. U. Fillinger et al., *Bull. World Health Org.* **87**, 655 (2009).
3. S. Blanford et al., *Science* **308**, 1638 (2005).
4. L. L. Mnyone et al., *Malar. J.* **11**, 87 (2012).
5. W. Fang et al., *Science* **331**, 1074 (2011).
6. M. J. Kirby et al., *Lancet* **374**, 998 (2009).
7. H. Van den Berg, L. A. Kelly-Hope, S. W. Lindsay, *Lancet Infect. Dis.* **13**, 89 (2013).
8. T. Homan et al., *Lancet* **388**, 1193 (2016).
9. P. Alonso, D. Engels, J. Reeder, *Lancet* **389**, 2270 (2017).

10.1126/science.aar7554

## Response

Koenraadt and Takken argue that, in addition to finding new mosquito vector control insecticides to replace those that are currently failing due to resistance, we devote equal attention to alternative methods. We agree with the principle of reducing selection pressure on new insecticides by combining them with other robust, cost-effective, operationally viable vector control activities. The challenge lies in finding properly evaluated interventions that can be recommended for use and deployed cost-effectively in a variety of transmission settings. The alternative interventions cited by Koenraadt and Takken have never been successfully used operationally on a large scale. Although some interventions have been shown to be effective in small-scale studies, the vector control literature is awash with poorly designed intervention “trials” that are riddled with bias, lack robust controls, and have insufficient statistical power (1–3). Financial support for malaria control is already well below the investment needed for interventions that work (1–3). Reducing coverage even further by diverting resources to unproven alternative interventions is ethically unacceptable.

In 2004, the Cochrane infectious disease group began commissioning a series of systematic reviews to assess the evidence for various interventions. One review concluded that electronic buzzer devices, sold directly to the public as mosquito repellents, do not work and should not be used to reduce mosquito bites or malaria infection rates (1). The 2013 review on larval source management, which assessed all

published and unpublished data from 1900 to 2012, located only 13 studies of sufficient quality to include in the assessment. The authors concluded that larval source management is a valid policy option in limited situations where sufficient breeding sites can be targeted. However, they pointed out that further research is still needed to evaluate larval source management feasibility in rural Africa (2).

The review on larvivorous fish (2013) concluded that there were no reliable studies demonstrating an impact on malaria infection rates in nearby communities and that all the studies reporting reductions in mosquito numbers had a high risk of bias. The evidence for use of fish was insufficient to recommend their use (3). The review on larvicides, which will include *Bacillus thuringiensis* var. *israelensis*, was commissioned in 2017, and the protocol for the study search has been published (4). It remains to be seen how many of the publications on larval interventions are sufficiently robust and lacking in bias to warrant inclusion.

Other potential interventions, such as entomopathogenic fungi, are not yet at the point where the evidence can be sensibly evaluated. The fatal flaw in this intervention, which has yet to be overcome despite the efforts of several groups, is the inability of the spore formulation to survive for any significant periods of time at tropical storage temperatures, severely limiting its shelf life.

A few interventions, such as odor-baited traps and house eave tubes, are entering large-scale trials, most with substantial funding from the Bill and Melinda Gates Foundation (5). These trials take several years and will need to be completed before they generate the evidence that may allow them to be recommended for wide-scale operational use.

Reviews on long-lasting nets (2004) and indoor residual spraying (2010) supported the use of these interventions (6, 7). Therefore, limited resources should be directed to these solutions.

**Janet Hemingway**

Liverpool School of Tropical Medicine, Liverpool L3  
5QA, UK. Email: janet.hemingway@lstm.ac.uk

## REFERENCES

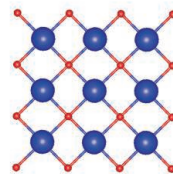
1. A. Enayati et al., *Cochrane Database Syst. Rev.* **2007**, CD005434 (2007).
2. L. S. Tusting et al., *Cochrane Database Syst. Rev.* **2013**, CD008923 (2013).
3. D. P. Walsh et al., *Cochrane Database Syst. Rev.* **2013**, CD008090 (2013).
4. L. Choi, A. Wilson, *Cochrane Database Syst. Rev.* **2017**, CD012736 (2017).
5. In2Care, Projects ([www.in2care.org/projects/](http://www.in2care.org/projects/)).
6. C. Lengeler, *Cochrane Database Syst. Rev.* **2004**, CD000363 (2004).
7. B. Pluess et al., *Cochrane Database Syst. Rev.* **2010**, CD006657 (2010).

10.1126/science.aar8094

# RESEARCH

Avoiding polar catastrophe  
on perovskite surfaces

Setvin et al., p. 572



## IN SCIENCE JOURNALS

Edited by **Stella Hurtley**



### GENETICS

#### Myoediting for Duchenne muscular dystrophy

**D**uchenne muscular dystrophy (DMD) is an X-linked recessive disorder that results in progressive weakening and muscle loss. Long *et al.* used CRISPR-Cas9 genome editing to correct the hotspot mutations that disrupt expression of the dystrophin gene in DMD. So-called myoediting abolished RNA splice sites and allowed the removal of the most common mutations. The approach successfully restored dystrophin expression in patient-derived induced pluripotent stem cells and restored mechanical force contraction in derived cardiomyocytes to near normal. —ASH

*Sci. Adv.* 10.1126/sciadv.aap9004 (2018).

A new technology brings hope  
in the fight against DMD.

### NEAR-FIELD COSMOLOGY

#### Dwarf galaxies move in unexpected ways

Massive galaxies like our Milky Way are orbited by satellite dwarf galaxies. Standard cosmological simulations of galaxy formation predict that these satellites should move randomly around their host. Müller *et al.* examined the satellites of the nearby elliptical galaxy Centaurus A (see the Perspective by Boylan-Kolchin). They found that the satellites are distributed in a planar arrangement, and the members of the plane are orbiting in a coherent direction. This is inconsistent with more than 99% of comparable galaxies in simulations. Centaurus A, the Milky Way, and Andromeda all have highly

statistically unlikely satellite systems. This observational evidence suggests that something is wrong with standard cosmological simulations. —KTS

*Science*, this issue p. 534;  
see also p. 520

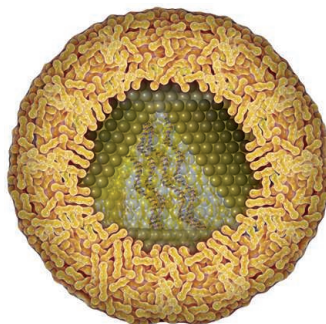
### EMERGING INFECTIONS

#### Fearsome flaviviruses

Although Zika virus was identified more than half a century ago, its impact on the unborn children of infected pregnant women was only detected in the recent devastating epidemic. Unfortunately, Zika virus may not be unique in its ability to inflict fetal damage. To investigate this possibility, Platt *et al.* infected pregnant immunocompetent mice with related viruses. Two other flaviviruses, West Nile virus and Powassan

virus, caused fetal demise. These viruses could also replicate in human maternal and fetal explant tissue. If these or other neurotropic flaviviruses take off in the human population, we may again experience congenital infections with devastating effects. —LP

*Sci. Transl. Med.* 10, eaao7090 (2018).



West Nile virus, like its cousin Zika, may damage fetuses.

### TOPOLOGICAL MATTER

#### A phonon merry-go-round

Chirality is associated with the breaking of symmetry, often described as left- or right-handed behavior. Such asymmetry can be seen, for example, in the electronic responses of particular materials or the reactions between particular chemical species. Zhu *et al.* observed a chiral phonon mode in a monolayer of the transition metal dichalcogenide  $\text{WSe}_2$ , detected spectroscopically as the circular dichroism of the phonon-assisted transition of holes. Phonon chirality could be used to control the electron-phonon coupling and/or the phonon-driven topological states of solids. —ISO

*Science*, this issue p. 579

## BIOCHEMISTRY

### Remember the sugar when making proteins

Eukaryotes have an elaborate trafficking and quality-control system for secreted glycoproteins. The glycosylation pathway begins in the endoplasmic reticulum with the enzyme oligosaccharyl-transferase (OST), which attaches a long chain of sugars to asparagine residues of target proteins. Wild *et al.* report a cryo-electron microscopy structure of yeast OST, which includes eight separate membrane proteins. The central catalytic subunit contains binding sites for substrates and is flanked by accessory subunits that may facilitate delivery of newly translocated proteins for glycosylation. —MAF

*Science*, this issue p. 545

## STRUCTURAL BIOLOGY

### Structure of the human spliceosome

Catalyzed by the spliceosome, precursor mRNA splicing proceeds in two steps: branching and exon ligation. Transition from the C (catalytic post-branching spliceosome) to the C\* (catalytic pre-exon ligation spliceosome) complex is driven by the adenosine triphosphatase/helicase Prp16. Zhan *et al.* report the cryo-electron microscopy structure of the human C complex, showing that two step I splicing factors stabilize the active site and link it to Prp16. —SYM

*Science*, this issue p. 537

## ECOPHYSIOLOGY

### A demanding lifestyle

Polar bears appear to be well adapted to the extreme conditions of their Arctic habitat. Pagano *et al.*, however, show that the energy balance in this harsh environment is narrower than we might expect (see the Perspective by Whiteman). They monitored the behavior and metabolic rates of nine free-ranging polar bears over 2 years. They found that high energy demands required consumption of high-fat prey, such as seals, which are easy to come by

on sea ice but nearly unavailable in ice-free conditions. Thus, as sea ice becomes increasingly short-lived annually, polar bears are likely to experience increasingly stressful conditions and higher mortality rates. —SNV

*Science*, this issue p. 568;

see also p. 514

## COLON CANCER

### Biofilms provide refuge for cancerous bacteria

Familial adenomatous polyposis (FAP) causes benign polyps along the colon. If left untreated, FAP leads to a high incidence of colon cancer. To understand how polyps influence tumor formation, Dejea *et al.* examined the colonic mucosa of FAP patients. They discovered biofilms containing the carcinogenic versions of the bacterial species *Escherichia coli* and *Bacteroides fragilis*. Colon tissue from FAP patients exhibited greater expression of two bacterial genes that produce secreted oncotoxins. Studies in mice showed that specific bacteria could work together to induce colon inflammation and tumor formation. —PNK

*Science*, this issue p. 592

## MOLECULAR BIOLOGY

### Self-defense by avoiding self-targeting

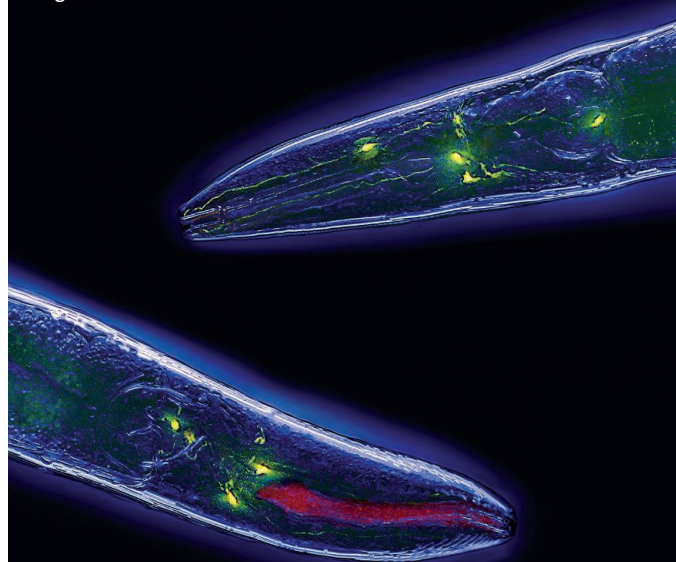
By silencing transposons, Piwi-interacting RNAs (piRNAs) protect the stability of animal genomes in germ lines. However, many piRNAs do not map to transposons, and their functions have remained undefined. Zhang *et al.* described the piRNA targeting logic in *Caenorhabditis elegans* and identified an intrinsic sequence signal in endogenous germline genes that confer resistance to piRNA silencing. Thus, diverse piRNAs silence foreign nucleic acids but spare self genes to defend the *C. elegans* genome. In addition, multiple foreign transgenes can be engineered to escape piRNA targeting, allowing successful expression in the germline. —SYM

*Science*, this issue p. 587

## IN OTHER JOURNALS

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

Confocal micrograph  
of *Caenorhabditis  
elegans* worms



## CELL BIOLOGY

### Targeting the demise of male mitochondria

In the nematode worm *Caenorhabditis elegans*, as in humans, offspring inherit mitochondria and mitochondrial DNA only from their mother. Sato *et al.* identified key molecular components of the machinery that recognizes male mitochondria and targets them for degradation. In a screen for kinases that might be required for the process, they identified IKKE-1, a protein kinase with similarity to the mammalian protein kinases TBK1 and IKKε. IKKE-1 interacted with a protein that they named ALLO-1, which had sequence similarity to autophagy receptors and was required for clearance of paternal organelles. The similarity of the role of IKKE-1 in removing male mitochondria to that of mammalian TBK1, which functions in the innate immune response, hints that the two processes might have a similar evolutionary origin. —LBR

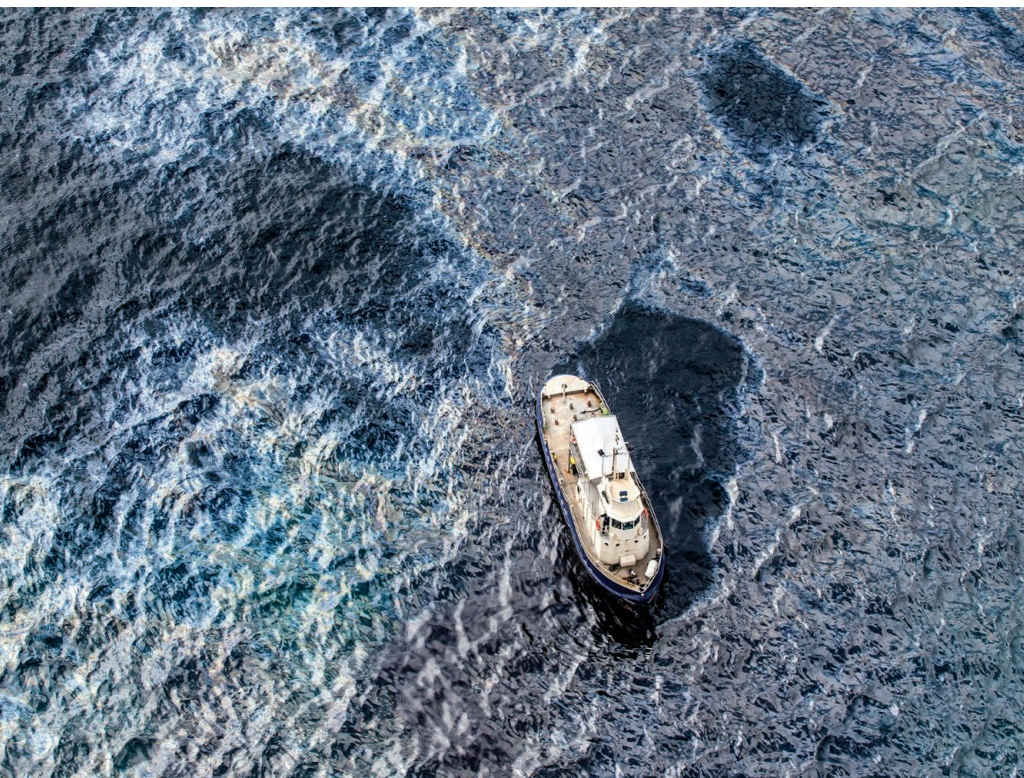
*Nat. Cell Biol.* 10.1038/s41556-017-0008-9 (2018).

## IMMUNOLOGY

### High caloric intake induces inflammation

The consumption of high-calorie diets has become prevalent in industrialized nations and is associated with increased body mass, inflammation, and resulting pathologies. What are the underlying mechanisms

connecting diets of high-energy, processed food to inflammation? Christ *et al.* show that mice fed on a high-calorie Western diet developed systemic inflammation and functionally reprogrammed granulocyte monocyte precursor cells. After a standard chow diet was restored to the animals, inflammation persisted, indicating that



## OCEAN PHYSICS

## Shear by the centimeter

**P**redicting the transport of oil and floating debris in the ocean is very difficult because it is challenging to measure differences in water motion very close to the surface. Laxague *et al.* present measurements of the ocean's current vector profile made in the Gulf of Mexico, showing that the velocity of the uppermost centimeter of the water column was four times that at a depth of 10 m and displaced by nearly 90°. Therefore, incorporating these kinds of dynamics into forecasting efforts is essential for better understanding the movement and fates of plastic particles and spilled oil in the oceans. —HJS

*Geophys. Res. Lett.* 10.1002/2017GL075891 (2017).

Oil from the 2011 MV *Rena* spill surrounds a service vessel.

components in the Western diet induced innate immune memory. This study identifies possible targets for preventing diet-induced inflammation. —GKA

*Cell* **172**, 162 (2018).

## COMPARATIVE GENOMICS

## Using phylogeny to test evolution

Sequencing of whole genomes and obtaining information about gene transcription offer the promise to compare changes and variation in gene expression and function across species. However, Dunn *et al.* show that analyses of gene expression data that use pairwise comparisons across multiple species may not be able to distinguish between differing models of evolution. By examining published studies of how gene expression may change across species, they show that it is necessary to consider the mode of evolution—i.e., duplication generating paralogs or orthologous evolution through speciation—as well as the time since the species diverged. Thus, phylogenetic information is needed to discern

commonalities and differences across species and among genes that share an ancestor but may have undergone different evolutionary trajectories. —LMZ

*Proc. Natl. Acad. Sci. U.S.A.* 10.1073/pnas.1707515115 (2018).

## STAR FORMATION

## Many stars don't form in clusters

Most stars are thought to form in dense, gravitationally bound molecular clouds, which should produce a bound cluster of stars. As unused gas is expelled from the system by stellar feedback, the cluster becomes gravitationally unbound to form an association, which gradually drifts apart. Associations of young stars (no more than a few million years old) should therefore be slowly expanding. Ward and Kruijssen use astrometric data to test this idea by looking for evidence of expansion in several nearby young associations. They do not find any; the associations show no sign of ever having been gravitationally bound. The authors suggest that this indicates that star formation

is dominated by turbulent fragmentation, not monolithic collapse. —KTS

*Mon. Not. R. Astron. Soc.* 10.1093/mnras/sty117 (2018).

## ANTIBIOTICS

## Recycling antibiotic sensitivity

Cystic fibrosis (CF) is accompanied by chronic lung infections, requiring lifelong consumption of many antibiotics to maintain airway function. Antibiotic resistance and clinical deterioration are apparently inevitable. In a multifactorial study, Imamovic *et al.* discovered that mutations in *Pseudomonas aeruginosa* that endow resistance to some classes of antibiotic concomitantly result in sensitivity to others. For example, in any strain of *P. aeruginosa*, resistance to the fluoroquinolone ciprofloxacin was consistently associated with sensitivity to aminoglycoside antibiotics, such as amikacin. In this case, resistance was mediated by mutations in the transcriptional regulator *nfxB* of the MexC transporter. Using biomarkers such as *nfxB*, it

appears to be possible to design antibiotic regimens for individual CF sufferers that flip resistant bacteria between predictable drug-sensitive states. —CA

*Cell* 10.1016/j.cell.2017.12.012 (2018).

## MICROBIOLOGY

## Cutting up in the deep

Life in the deep ocean is cut off from sunlight as an energy source. Fixed carbon from the surface—namely, dead microbes and larger organisms—therefore serves as a source of energy at depths below the reach of sunlight. Bergauer *et al.* analyzed genomic and proteomic data from microorganisms collected at depths down to 4 to 5 km. Microbial transport proteins varied only slightly in abundance with depth, reflecting changes in the abundance of organisms and in lifestyle. Organic acid transporters were found throughout the water column and were more abundant in the deep ocean, suggesting that uptake of dilute dissolved organic matter is crucial for microbial metabolism at these depths. —MAF

*Proc. Natl. Acad. Sci. U.S.A.* 10.1073/pnas.1708779115 (2018).

ALSO IN *SCIENCE* JOURNALSEdited by **Stella Hurtley****EARTH SYSTEMS****Integrating the biosphere into climate models**

High-quality climate predictions are crucial for understanding the impacts of different greenhouse gas emission scenarios and for mitigating and adapting to the resulting climatic changes. Bonan and Doney review advances in Earth system models that include the terrestrial and marine biosphere. Such models capture interactions between physical and biological aspects of the Earth system. This provides insight into climate impacts of societal importance, such as altered crop yields, wildfire risk, and water availability. Further research is needed to better understand model uncertainties, some of which may be unavoidable, and to better translate observations into abstract model representations. —JFU

*Science*, this issue p. 533

**NEURODEVELOPMENT****Brain mutations, young and old**

Most neurons that make up the human brain are postmitotic, living and functioning for a very long time without renewal (see the Perspective by Lee). Bae *et al.* examined the genomes of single neurons from the prenatal developing human brain. Both the type of mutation and the rates of accumulation changed between gastrulation and neurogenesis. These early mutations could be generating useful neuronal diversity or could predispose individuals to later dysfunction. Lodato *et al.* also found that neurons take on somatic mutations as they age by sequencing single neurons from subjects aged 4 months to 82 years. Somatic mutations accumulated with increasing age and accumulated faster in individuals affected by inborn errors in DNA repair. Postmitotic

mutations might only affect one neuron, but the accumulated divergence of genomes across the brain could affect function. —PJH

*Science*, this issue p. 550, p. 555;  
see also p. 521

**SURFACE SCIENCE****Compensating a polar surface**

An ionic crystal surface can be electrostatically unstable, and the surface must reconstruct in some way to avoid this “polar catastrophe.” Setvin *et al.* used scanning probe microscopies and density functional theory to study the changes in the polar surface of the perovskite  $\text{KTaO}_3$ . They observed several structural reconstructions as the surface cleaved in vacuum was heated to higher temperatures. These ranged from surface distortions to the formation of oxygen vacancies to the development of  $\text{KO}$  and  $\text{TaO}_2$  stripes. Hydroxylation after exposure to water vapor also stabilized the surface. —PDS

*Science*, this issue p. 572

**CARBON FIXATION****About-face for citrate synthase**

Classically, it is thought that citrate synthase only works in one direction: to catalyze the production of citrate from acetyl coenzyme A and oxaloacetate in the tricarboxylic acid (TCA) cycle. The TCA cycle can run in reverse to cleave citrate and fix carbon dioxide autotrophically, but this was thought to occur only with alternative enzymes, such as citrate lyase. Now Nunoura *et al.* and Mall *et al.* have discovered thermophilic bacteria with highly efficient and reversible citrate synthase that requires reduced ferredoxin (see the Perspective by Ragsdale). This function is undetectable by metagenomics, but classical

biochemistry filled in the gaps seen between the genome sequences and the phenotypes of the organisms. The direction of catalysis depends on the availability of organic versus inorganic carbon and reflects a flexible bet-hedging strategy for survival in fluctuating environments. In evolutionary terms, this capacity might predate the classical TCA cycle and is likely to occur in a wide range of anaerobic microorganisms. —CA

*Science*, this issue p. 559, p. 563;  
see also p. 517

**CANCER IMMUNOTHERAPY****HLA genotype affects response**

Immunotherapy works by activating the patient's own immune system to fight cancer. For effective tumor killing,  $\text{CD8}^+$  T cells recognize tumor peptides presented by human leukocyte antigen class I (HLA-I) molecules. In humans, there are three major HLA-I genes (*HLA-A*, *HLA-B*, and *HLA-C*). Chowell *et al.* asked whether germline HLA-I genotype influences how T cells recognize tumor peptides and respond to checkpoint inhibitor immunotherapies (see the Perspective by Kvistborg and Yewdell). They examined more than 1500 patients and found that heterozygosity at HLA-I loci was associated with better survival than homozygosity for one or more HLA-I genes. Thus, specific HLA-I mutations could have implications for immune recognition and for the design of epitopes for cancer vaccines and immunotherapies. —PNK

*Science*, this issue p. 582;  
see also p. 516

**AUTOIMMUNE DISEASE****The X chromosome link to lupus**

Nine out of 10 people who develop systemic lupus erythematosus (SLE) are women.

Furthermore, individuals with Klinefelter syndrome (47,XXY) also have increased incidence of SLE. This suggests that X chromosome dosage could be an important risk factor in SLE. Using sensitive quantification methods, Souyris *et al.* demonstrate that Toll-like receptor 7 (TLR7), which is encoded on the X chromosome, escapes X inactivation in B cells and myeloid cells in females and Klinefelter individuals. TLR7 binds single-stranded RNA and activates type I interferon signaling, a pathway that is also activated in SLE patients. Thus, biallelic expression of TLR7 appears to contribute to greater SLE risk. —AB

*Sci. Immunol.* **3**, eaap8855 (2018).

**CANCER****The cytoskeleton-EGFR connection**

Some cancers are caused by aberrant epidermal growth factor receptor (EGFR) signaling. Two papers now reveal connections between this receptor and cytoskeletal remodeling (see the Focus by Chiasson-MacKenzie and McClatchey). Pike *et al.* delineated a cytoskeletal pathway that delayed receptor internalization, thereby increasing EGFR signal duration from the membrane. Roth *et al.* identified a cytoskeleton-associated effector of EGFR signaling that promoted migration and proliferation of mammary epithelial cells and was associated with poor prognosis in breast cancer patients. —LKF

*Sci. Signal.* **10**, eaan0949, eaaq1060;  
see also eaas9473 (2018).

**SUPERCONDUCTIVITY****A nonlinear peek into stripes**

In many theoretical models of high-temperature superconductors, remnants of superconductivity persist to

---

temperatures higher than the transition temperature,  $T_c$ . Rajasekaran *et al.* used nonlinear terahertz spectroscopy to probe this region of the phase diagram of a cuprate superconductor that is well known for a stripe phase that appears for certain doping levels (see the Perspective by Ergeçen and Gedik). For a sample deep in the stripe phase, a large nonlinear signal persisted from the superconducting region up to temperatures much higher than  $T_c$ . The findings suggest the formation of a peculiar spatially modulated superconducting state called the pair-density wave. —JS

*Science*, this issue p. 575;  
see also p. 519

## REVIEW SUMMARY

## EARTH SYSTEMS

# Climate, ecosystems, and planetary futures: The challenge to predict life in Earth system models

Gordon B. Bonan\* and Scott C. Doney\*

**BACKGROUND:** Earth system models (ESMs) simulate physical, chemical, and biological processes that underlie climate and are the most complex in a hierarchy of models of Earth's interacting atmosphere–land–ocean–sea ice system. As terrestrial and marine ecosystems have been added to ESMs, the distinction between the physical basis for climate change, mitigation, and vulnerability, impacts, and adaptation (VIA) no longer necessarily holds. The same global change stresses that affect terrestrial and marine ecosystems are critical processes that determine the magnitude and trajectory of climate change, and many of the interventions that might lessen anthropogenic climate change pertain to the biosphere. Here we describe environmental changes that are stressing terrestrial and marine ecosystems. We discuss how these stressors are being included in ESMs, initially with an emphasis on climate processes, but also show their emerging utility for VIA analyses and examine them in the context of Earth system prediction.

**ADVANCES:** Terrestrial ecosystems face stresses from changing climate and atmospheric composition that alter phenology, growing season length, and community composition; these stresses enhance productivity and water-use efficiency in some regions, but also lead to mortality and increased disturbances from wildfires, insects, and extreme events in other regions. The addition of reactive nitrogen, elevated levels of tropospheric O<sub>3</sub>, and anthropogenic land-use and land-cover change stress ecosystems as well. The terrestrial biosphere models included in ESMs simulate the ecological impacts of these stresses and their effects on Earth system functioning. Ocean ecosystems and living marine resources face threats from ocean warming, changing large-scale circulation, increased vertical stratification, declining oxygen, and acidification, which alter nutrient supply, the light environment, and phytoplankton productivity; result in coral bleaching; and produce novel marine communities. Three-dimensional ocean models simulate the carbon cycle and

associated biogeochemistry. Plankton ecosystem models both drive biogeochemistry models and characterize marine ecological dynamics.

**OUTLOOK:** The untapped potential of ESMs is to bring dispersed terrestrial and marine ecosystem research related to climate processes, VIA, and mitigation into a common framework. ESMs offer an opportunity to move beyond physical descriptors of atmospheric and oceanic states to societally relevant quantities such as habitat loss, water availability, wildfire risk, air quality, and crop, fishery, and timber yields. To do so, the science of climate prediction has to

## ON OUR WEBSITE

Read the full article at <http://dx.doi.org/10.1126/science.aam8328>

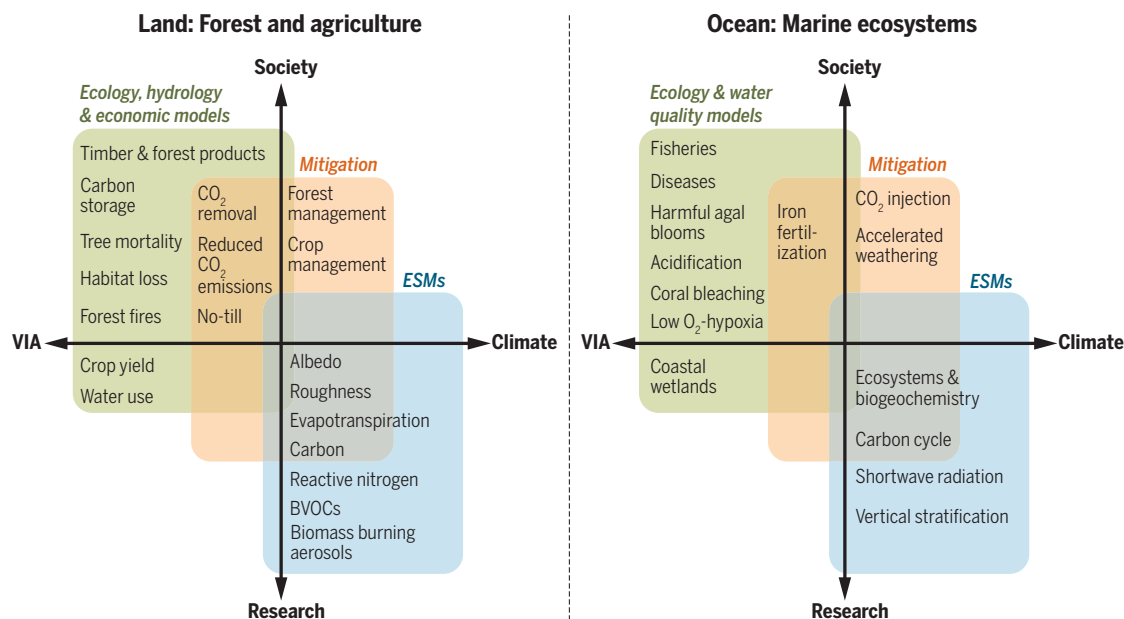
be extended to a more multifaceted Earth system prediction, including the biosphere and its resources. ESMs provide the means not just to assess the potential for future global change stresses, but also to determine the outcome of those stresses on the biosphere. Such Earth system prediction is necessary to inform sound policy that maintains a healthy biosphere and provides the food, energy, and fresh water needed for a growing global population without further exacerbating climate change. Substantial impediments that must be overcome include advancing our knowledge of biosphere-related climate processes; reducing model uncertainty; and effectively communicating among, rather than across, the disparate science communities of climate prediction, global biosphere modeling, VIA analyses, and climate change mitigation. ■

The list of author affiliations is available in the full article online.

\*Corresponding author. Email: [bonan@ucar.edu](mailto:bonan@ucar.edu); (G.B.B.); [sdoney@virginia.edu](mailto:sdoney@virginia.edu) (S.C.D.)

Cite this article as G. B. Bonan, S. C. Doney, *Science* 359, eaam8328 (2018). DOI: 10.1126/science.aam8328

The various models used for climate projections and mitigation and VIA analyses overlap in scope and would benefit from a broad perspective of Earth system prediction. Shown are the domains of ESMs, mitigation models, and VIA models along axes from VIA to climate processes (horizontal) and from primarily serving the research community to informing societal needs (vertical). Panels show forests and agriculture (left) and marine ecosystems (right) as represented across modeling domains.



## REVIEW

## EARTH SYSTEMS

# Climate, ecosystems, and planetary futures: The challenge to predict life in Earth system models

Gordon B. Bonan<sup>1\*</sup> and Scott C. Doney<sup>2\*</sup>

Many global change stresses on terrestrial and marine ecosystems affect not only ecosystem services that are essential to humankind, but also the trajectory of future climate by altering energy and mass exchanges with the atmosphere. Earth system models, which simulate terrestrial and marine ecosystems and biogeochemical cycles, offer a common framework for ecological research related to climate processes; analyses of vulnerability, impacts, and adaptation; and climate change mitigation. They provide an opportunity to move beyond physical descriptors of atmospheric and oceanic states to societally relevant quantities such as wildfire risk, habitat loss, water availability, and crop, fishery, and timber yields. To achieve this, the science of climate prediction must be extended to a more multifaceted Earth system prediction that includes the biosphere and its resources.

Human activities are transforming Earth's atmosphere, ocean, and land surfaces at a scale and magnitude not previously seen during the past several thousand years of human history. These changes threaten healthy planetary functions and socioeconomic well-being (1, 2). Fossil fuel combustion, industrialized agriculture, urbanization, and other facets of modern human societies are changing climate and atmospheric composition; melting permafrost, glaciers, ice sheets, and Arctic sea ice; raising sea levels; warming and acidifying the oceans; polluting air, water, and soils; altering biogeochemical cycles and freshwater availability; increasing the cycling of reactive nitrogen; reducing forest cover and degrading land; and destroying habitats and reducing biodiversity (3–5). The ecological consequences of these changes are apparent in individual organisms, the communities they inhabit, and the ecosystems in which they function (6–8).

The interconnectedness and global scope of this changing environment have transformed the scientific study of Earth as a system. It is now understood that climate change must be studied in terms of a myriad of interrelated physical, chemical, biological, and socioeconomic processes. This broadening basis for climate change research underlies the transformation from global climate models to Earth system models (ESMs). These models have shown that the biosphere not only responds to climate change, but also directly influences the direction and magnitude of climate change. Terrestrial and marine ecosystems, and their uses by humans, are fundamental to

addressing the climate change problem. How do we provide the food, energy, and fresh water needed for a growing global population without further exacerbating climate change? Can terrestrial and marine ecosystems be managed to reduce greenhouse gas emissions? With the advent of ESMs, climate science is no longer limited to the physical basis for climate projections, but also includes projections of the biosphere—for example, regarding carbon storage on land and in the ocean, forest dieback, wildfires, crop yield, and fisheries and marine resources.

However, the study of climate change is still often parsed into separate activities of observing changes and deducing causes (3); assessing the vulnerability, impacts, and adaptation (VIA) of natural and human systems to these changes (6, 7); and determining the socioeconomic transformations needed to mitigate them (9). The untapped potential of ESMs is to bring these dispersed activities into a common framework. There has been success, for example, in coordinating climate projections with the integrated assessment models that identify the societal transformations needed to mitigate climate change (10) and even some initial attempts at directly coupling ESMs and integrated assessment models (11). As terrestrial and marine ecosystems have been added to ESMs, the distinction between the physical basis for climate change, VIA, and mitigation no longer necessarily holds. Land-use and land-cover change, for example, is driven by socioeconomic needs for food, fiber, and fuel, but is also an ecological problem that alters habitat and biodiversity and a means to mitigate anthropogenic CO<sub>2</sub> emissions (4).

In this Review, we discuss the treatment of the biosphere in ESMs, considering terrestrial and marine ecosystems as they are now represented in the models, exploring how ESMs can be used

to study the biosphere, and highlighting opportunities for future research. We then describe environmental changes that are occurring globally and that are stressing terrestrial and marine ecosystems and show how these stresses are included in ESMs, in the past primarily with an emphasis on climate processes, but now with additional utility for VIA and mitigation research. Last, we examine these stresses in the context of Earth system prediction. Our list of stressors is not meant to be exhaustive. Rather, we highlight several key stressors and their coincidence among climate processes, VIA, and mitigation with the goal of initiating a dialog among the scientific communities that study climate change. This Review is timely because it identifies synergies across the climate and ESM research communities involved in the next Coupled Model Intercomparison Project (CMIP6) (12), which provides an unparalleled opportunity to model and analyze the Earth system.

## Earth system models

ESMs simulate physical, chemical, and biological processes that underlie climate. They are the most complex in the ongoing evolution of global models of Earth's atmosphere, ocean, cryosphere, and land (Fig. 1). Climate models focus on the physical climate system, as represented by atmosphere, ocean, and sea ice physics and dynamics and land surface hydrometeorology. In climate models, land and ocean are coupled with the atmosphere through energy and momentum fluxes and the hydrologic cycle. ESMs have the same representation of the physical climate system but additionally include the carbon cycle, terrestrial and marine ecosystems and biogeochemistry, atmospheric chemistry, and natural and human disturbances. ESMs typically couple distinct component modules for land, atmosphere, and ocean physics, and ecosystem dynamics and biogeochemistry are embedded into these modules.

A prominent feature of ESMs is their inclusion of the biosphere and abiotic interactions that together make up an ecosystem. On land, terrestrial ecosystems are represented in ESMs by the type of vegetation, the amount of leaf area, the stomata on leaves, and carbon and nitrogen pools (13). Similarly, ESMs simulate ocean phytoplankton production of chlorophyll that influences the vertical profile of light absorption in the upper ocean, which in turn affects model sea surface temperature and mixed layer dynamics, as well as large-scale ocean circulation, heat transport, and climate variability (14, 15).

Biogeochemical cycles were added to ESMs because of the potential for large climate feedbacks arising from the carbon cycle. Terrestrial ecosystems and the ocean together absorb about one-half of the annual anthropogenic CO<sub>2</sub> emissions (16), but the future efficacy of these sinks is uncertain (17). Biogeochemical processes on land encompass spatial scales from leaves to plant canopies and from ecosystems to landscapes to biomes (13). Temporal scales include near-instantaneous physiological responses (e.g., stomatal conductance, photosynthesis, and respiration) to prevailing environmental conditions; the seasonal emergence

<sup>1</sup>National Center for Atmospheric Research (NCAR), Boulder, CO 80307, USA. <sup>2</sup>Department of Environmental Sciences, University of Virginia, Charlottesville, VA 22904, USA.

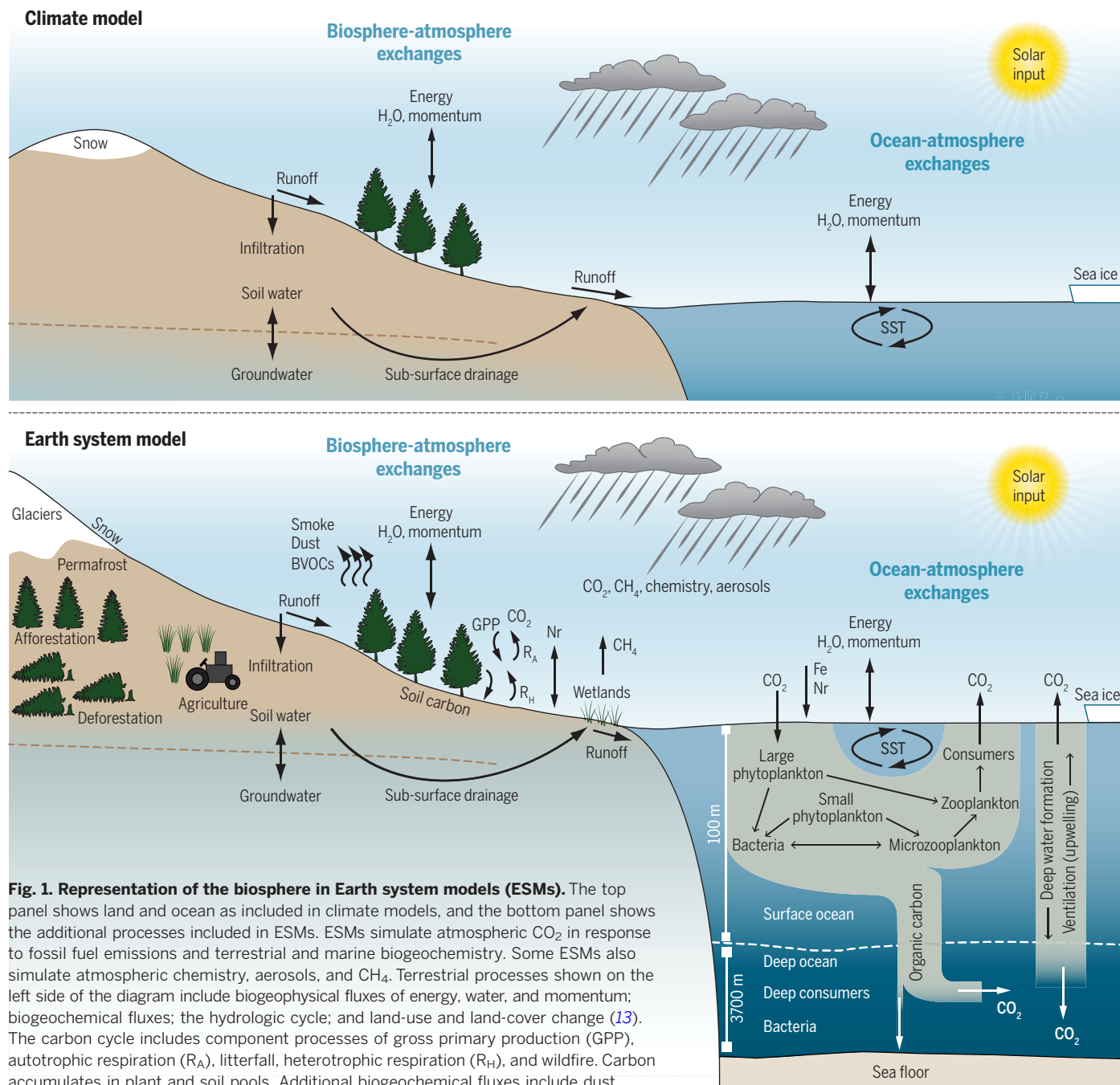
\*Corresponding author. Email: bonan@ucar.edu (G.B.B.); sdoney@virginia.edu (S.C.D.)

and senescence of leaves; and changes in ecosystem structure and biogeography over decades and centuries in response to natural disturbances (e.g., wildfires), anthropogenic disturbances (e.g., land-use transitions), and climate change. Ongoing model development aims to more authentically represent plant demography and life history char-

acteristics using cohorts of individuals of similar functional traits in vertically structured plant canopies (18).

The three-dimensional carbon cycle models used to estimate ocean uptake of anthropogenic  $\text{CO}_2$  evolved from model tracer studies of ocean physical circulation. Biogeochemical models ad-

ditionally track natural cycling of inorganic carbon, alkalinity, macronutrients (nitrogen, phosphorus, and silicon), and often  $\text{O}_2$ ; net organic matter and  $\text{CaCO}_3$  production and export from the surface ocean; particle sinking and respiration and remineralization at depth; and air-sea  $\text{CO}_2$  (and  $\text{O}_2$ ) gas exchange (19). Plankton ecosystem models



**Fig. 1. Representation of the biosphere in Earth system models (ESMs).** The top panel shows land and ocean as included in climate models, and the bottom panel shows the additional processes included in ESMs. ESMs simulate atmospheric  $\text{CO}_2$  in response to fossil fuel emissions and terrestrial and marine biogeochemistry. Some ESMs also simulate atmospheric chemistry, aerosols, and  $\text{CH}_4$ . Terrestrial processes shown on the left side of the diagram include biogeophysical fluxes of energy, water, and momentum; biogeochemical fluxes; the hydrologic cycle; and land-use and land-cover change (13). The carbon cycle includes component processes of gross primary production (GPP), autotrophic respiration ( $R_A$ ), litterfall, heterotrophic respiration ( $R_H$ ), and wildfire. Carbon accumulates in plant and soil pools. Additional biogeochemical fluxes include dust entrainment, wildfire chemical emissions, biogenic volatile organic compounds (BVOCs), the reactive nitrogen cycle (Nr), and  $\text{CH}_4$  emissions from wetlands. Ocean processes are shown on the right side of the diagram. Physical processes include sea ice dynamics, ocean mixing and circulation, changes in sea surface temperature (SST), and ocean-atmosphere fluxes. The gray shaded area depicts the marine carbon cycle, consisting of the phytoplankton-based food web in the upper ocean, export and remineralization in the deep sea and sediments, and the physiochemical solubility pump controlled by surface-deep ocean exchange (100).

that simulate interactions of phytoplankton, zooplankton, nutrients, and detrital pools arose both to drive biogeochemistry models and to characterize marine ecological dynamics, such as seasonal phytoplankton blooms. Recent biogeochemical developments include incorporation of iron and other trace elements (20), iron limitation being a major controlling factor for phytoplankton growth in much of the ocean; more sophisticated treatment of marine biological nitrogen fixation and denitrification (21); coastal inputs of nutrients; and ocean acidification. Major model advances under way involve expansion of plankton biological complexity to incorporate functional groups, trait-based dynamics, and biodiversity (22–24) and efforts to integrate simulated plankton productivity with fisheries catch (25).

An active research frontier for ESMs is incorporating more extensive chemistry-climate interactions. Additional reactive nitrogen affects climate through enhanced terrestrial carbon storage, emissions of N<sub>2</sub>O, and chemical reactions that determine the amount of tropospheric O<sub>3</sub>, CH<sub>4</sub>, and aerosols (5, 26). Atmospheric deposition of nitrogen to the surface ocean can enhance biological productivity in low-nutrient subtropical regions; globally, however, marine biogeochemistry may be more sensitive to anthropogenic iron deposition (27, 28). Increased concentrations of tropospheric O<sub>3</sub> decrease plant productivity and reduce the terrestrial carbon sink, but the appropriate way to parameterize this in models is uncertain (29, 30). Emissions of biogenic volatile organic compounds from terrestrial ecosystems influence atmospheric concentrations of O<sub>3</sub>, CH<sub>4</sub>, and secondary organic aerosols (31). Wetlands are an important source of CH<sub>4</sub>, as are permafrost soils and hydrates. Global models of wetlands and CH<sub>4</sub> emissions are being developed (32), and some ESMs include methane chemistry in their climate projections (33). In the ocean, more diverse biogeochemistry is needed (e.g., trace gases such as dimethyl sulfide) in models to link ocean-atmosphere chemistry.

Natural and human disturbances are a continuing research priority for ESMs. Wildfires affect climate and air quality through emissions of long-lived greenhouse gases, short-lived reactive gases, and aerosols and by altering surface albedo (34). Wildfires are included in ESMs, but our ability to model the precise details of fire regimes is limited (35). The mountain pine beetle epidemic in western North American forests has reduced terrestrial carbon uptake (36), increased surface albedo (37), and warmed the surface by reducing evapotranspiration (38). Efforts to represent insect outbreaks in ESMs are promising but still nascent (39). Human disturbances include land-use transitions (e.g., deforestation, reforestation, farm abandonment, and shifting cultivation) and wood harvest (40). Global models of crop growth are included in ESMs, but specific cultivars, time of planting, crop rotation, and other management practices are lacking (41, 42). Nor is forest management and commercial timber production included, despite having an effect on temperature similar to that of land-cover change (43).

Planetary stresses and climate feedbacks

The inclusion of terrestrial and marine ecosystems in ESMs enables study of the global change stresses on the biosphere and feedbacks with climate change (Table 1). A prominent signal on land is a “greening” of the biosphere, although this is partially countered by increased tree mortality and disturbance. Novel community assemblages are likely to emerge that depend on the magnitude and rate of climate change and the ability of species to adjust to these changes through dispersal. Marine ecosystems also face numerous threats from climate change (44–46). Surface ocean waters are warming globally and freshening at high latitudes; together, these trends act to increase vertical physical stratification, resulting in altered regional patterns of nutrient supply, light environment, and phytoplankton productivity. Ocean warming leads to shifts in plankton seasonal phenology and poleward migration of plankton, invertebrate, and fish species; coral bleaching; and sea ice loss in polar marine ecosystems. Climate change is projected to alter the spatial patterns and size of marine wild-caught fisheries (47) and may potentially change marine disease outbreaks (48). Marine communities and ecosystems also may be reorganizing into novel assemblages, requiring more sophisticated ESMs that can track in more detail the effects of warming on plankton community composition and trophic interactions (49).

Human activities imperil ecosystems and biota in ways other than direct climate change (Table 1). Additional reactive nitrogen alters biodiversity; terrestrial, freshwater, and marine biogeochemistry;

and water and air quality. Anthropogenic aerosols increase the amount of diffuse solar radiation, which can enhance terrestrial productivity. High concentrations of tropospheric O<sub>3</sub> can cause stomata to close and thus decrease plant productivity and transpiration. Vast areas of forests have been cleared over the industrial era, and many of the remaining forests are managed or secondary rather than old-growth primary forests. About one-third of the ice-free land is covered by cropland or pastureland, and much of the terrestrial productivity has been appropriated for human uses.

Marine impacts arise from ocean acidification owing to increasing atmospheric CO<sub>2</sub> and deoxygenation from climate-related circulation changes. Regional stresses, particularly on continental shelves and some parts of the open ocean, are occurring from overfishing, anthropogenic noise, seabed habitat destruction, pollution, and coastal eutrophication, as well as loss of coastal wetlands, mangrove forests, and seagrass beds owing to development and sea level rise (50).

Vulnerability, impacts, and adaptation

Agriculture and food security, forest and water resources, terrestrial ecosystems, and fisheries and marine ecosystems are facets of the biosphere that sustain socioeconomic well-being. Assessing the impact of climate change on these goods and services, their vulnerability to disruption in a changing climate, and the adaptations needed to maintain their future availability is critical for informing sound climate policies (6, 7). Such assessments are commonly obtained by using climate projections

Table 1. Planetary stresses faced by terrestrial and marine ecosystems.	
Stress	Reference
<b>Terrestrial ecosystems</b>	
Greening of the biosphere	
Earlier springtime and longer growing season	(101)
Higher leaf area index	(55)
Greater productivity	(56, 91)
Higher water-use efficiency	(102)
Increased nitrogen deposition	(5, 26, 28)
Diffuse radiation	(103)
Browning of the biosphere	
Tree mortality	(104)
Extreme events	(105)
Wildfire and insects outbreaks	(36, 106)
Ozone damage	(29, 30)
Community assemblages	(107)
Land-use and land-cover change	(4, 40)
Human appropriation of net primary production	(108)
<b>Marine ecosystems</b>	
Vertical stratification, nutrient supply and phytoplankton productivity	(86, 109)
Plankton seasonal phenology	(46)
Coral bleaching	(110)
Polar marine ecosystems and sea ice loss	(111, 112)
Community assemblages	(46, 49, 113)
Acidification	(114)
Deoxygenation	(57)
Aerosol deposition	(28)

to drive models of terrestrial ecosystems (51), crop yield (52), water availability (53), and fisheries and marine ecosystems (54). This indirect, two-step approach has deficiencies because archived climate model output may not capture variable types or temporal resolution needed for some VIA models and restricts the ability to study feedbacks on climate and biogeochemistry.

With inclusion of the biosphere in ESMs, VIA can be investigated directly. For example, the ESMs used to quantify future carbon-climate feedbacks can also be used in retrospective studies to assess responses of terrestrial and marine ecosystems to historical anthropogenic forcings and climate variability (16, 55, 56). The ocean component of ESMs provides a tool for reconstructing the variability, trends, and mechanisms of historical ocean biogeochemistry (57) and plankton dynamics (58). ESMs with high-resolution ocean circulation models can be used to track larval dispersal and connectivity among coral reefs subject to bleaching (59).

ESMs provide further opportunity to move beyond physical descriptors of atmospheric and oceanic states (such as temperature and precipitation) to societally relevant quantities related to food, energy, and fresh water. For example, the croplands in ESMs allow direct study of the impacts of climate change on agricultural production, the vulnerability of the food supply to future climate disruption, and adaptations to make food production more resilient (42). Some ESMs include urban land cover, which allows study of extreme heat waves and facilitates assessment of heat-stress mortality in cities (60). ESMs are particularly useful for assessing air quality and human health issues because of the interactions among agriculture, wildfire, nitrogen gaseous fluxes, and biogenic volatile organic compounds that affect regional air quality.

Achieving this potential requires effective communication between the scientists developing ESMs and those using climate projections to study VIA (61). One result of better collaboration would be to identify and reconcile discrepancies between ESMs and VIA models, such as are evident in their assessments of water availability in a future climate. ESMs account for the effects of elevated atmospheric CO<sub>2</sub> on stomatal conductance and evapotranspiration, but many VIA models do not, resulting in an inconsistency in projections of water availability (62). Other examples of processes included in ESMs but not VIA models are the effect of O<sub>3</sub> on stomata (29, 30) and the effect of vegetation greening and land use on runoff (63). Closer collaboration between the communities would help to identify capabilities relevant to VIA, define impact-relevant metrics that ESMs should produce, and develop data sets and protocols for validation of simulated impacts (61).

ESMs remain just one of several means to study VIA. A suite of specialized research tools including statistical models and process-based crop, ecosystem, and hydrology models is required (6, 7). These models have the advantage that they can be run at the fine-scale spatial resolution needed to inform decision making. Also, they are less computationally expensive than ESMs and

can therefore be used in an ensemble of simulations to assess uncertainty.

ESMs cannot yet represent the rich ecological detail needed to capture spatial heterogeneity at local scales. Similarly, the ocean ecosystem models used in ESMs typically incorporate the lowest trophic levels of the marine food web (phytoplankton, herbivorous zooplankton) and have only a limited representation of biodiversity. Often, ESMs lack the ecological complexity required to predict outcomes in higher trophic levels and fisheries. The spatial resolution of global models is too coarse to capture regional dynamics of highly productive coastal ecosystems and coral reefs, and models are just beginning to incorporate adequate land-ocean connectivity to assess nutrient eutrophication, water quality, and harmful algal blooms (64). Variable-resolution global models with a horizontal resolution that refines from a 1° global grid to a regional 0.125° (14-km) grid help to bridge the gap between coarse-scale ESMs and the finer scales needed for VIA research (65, 66).

### Climate change mitigation

Reducing the sources and enhancing the sinks of long-lived greenhouse gases are the most direct means to mitigate anthropogenic climate change (67). However, many interventions that might reduce greenhouse gas emissions affect the biosphere and have other effects on climate and ecosystem services. Afforestation, reforestation, or avoided deforestation, for example, enhance the terrestrial carbon sink, but also warm climate annually by decreasing surface albedo, cool climate through evapotranspiration and turbulent mixing with the atmosphere, and have additional effects through atmospheric chemistry and aerosols (13, 68). These biogeophysical effects can counter the carbon mitigation benefits of forests so that even more extensive forested land may be required to achieve climate stabilization at a target that avoids dangerous climate change (e.g., 2°C). ESMs are an imperfect but necessary tool to study the net climate effects of forests (68).

Agriculture is another example of the need to consider mitigation in an ESM context. Efficient application of nitrogen fertilizer, tillage, and other management can enhance carbon storage and reduce N<sub>2</sub>O emissions (69). Crops also affect climate through biogeophysical coupling with the atmosphere; it is likely that expansion of agricultural lands over the industrial era has cooled climate because of these changes (70). Intensification of agriculture is thought to have cooled summer temperatures in the Midwest United States (71). No-till agriculture can increase surface albedo and cool climate (72), and other increases in surface albedo may have geoengineering potential (73). Production of bioenergy for carbon capture and storage (BECCS) can also mitigate climate change, but land use for BECCS must be balanced by arable land for food production (74). ESMs provide a necessary tool to investigate the multidisciplinary outcomes of BECCS for climate, food, energy, and fresh water. ESMs are also being used to determine the effects on ecosystems of geoengineering techniques involving solar radiation modification

such as stratospheric aerosol injection, cloud brightening, and surface albedo manipulation (75, 76).

At present, the ocean removes roughly a quarter of anthropogenic CO<sub>2</sub> emissions to the atmosphere, with the magnitude modulated by chemical dissolution into surface seawater and the physical rate of exchange between surface and deep waters (16). Over the centuries-long time scales of ocean-overturning circulation, an increasing fraction of anthropogenic CO<sub>2</sub> will be stored in the deep ocean reservoir. Several geoengineering approaches have been proposed to increase ocean carbon uptake by injecting CO<sub>2</sub> directly at depth, fertilizing phytoplankton to speed up the marine biological pump that transports organic carbon from the surface layer into the deep ocean, and accelerating weathering processes to add alkalinity to seawater (67). A number of studies with ocean-only models and full ESMs have explored the possible efficacy of these approaches, as well as the potential for ecological impacts and biogeochemical feedbacks (75). Substantial alterations to marine ecosystems could also arise from solar radiation modification.

### Earth system prediction

Atmospheric science has long embraced models to make predictions of near-term weather and long-term climate. ESMs enable predictions of the biosphere, but the atmospheric-centric view of prediction needs to be extended to the biological realm. In this section, we introduce terminology and concepts specific to weather forecasts and climate prediction and then show extension to the biosphere.

Forecasting the weather on time scales of a few hours to 2 weeks is a classic prediction problem in which a model that describes the atmosphere-land-ocean-sea ice system is stepped forward in time from initial conditions. The predictability—the capability to make a skillful forecast—is limited by uncertainty in the exact initial conditions, imperfections in the model and understanding of the underlying physics and dynamics, and the degree of randomness or chaotic behavior in the system. The same concept applies to predicting climate at seasonal, interannual, or decadal time scales (77–79). Climate projection over several decades must consider additional long-term Earth system processes such as ocean circulation, ice sheet melting, and changes in vegetation, terrestrial and marine biogeochemistry, and human behavior. The lattermost is particularly poorly known and is imposed using anthropogenic forcing scenarios. At multidecadal time scales, the exact initial state is less critical. Instead, uncertainty in climate projections is largely dominated by the choice of an anthropogenic forcing scenario, although uncertainties also remain with regard to climate sensitivity and feedback processes (Fig. 2).

The term Earth system prediction is used to capture this spectrum of temporal scales from subseasonal to multidecadal, mostly in the context of weather and climate (77–81). In a broader perspective, however, the scope of Earth system prediction can be expanded to include other facets

of the Earth system. The predictability of Arctic sea ice loss is a prominent such example (82). As climate models have evolved into more complex ESMs, the predictability of biosphere states and processes needs to be considered jointly with that of weather and climate.

The predictive capability of a model depends on the sources of errors in the forecast. For climate, these errors are broadly classified into initial conditions, boundary conditions, and model uncertainty, including both model structure and parameters (83). Uncertainty in exact initial conditions manifests in unforced variability internal to the climate system (also known as natural variability), in which small differences in initial conditions produce different climate trajectories. The importance of natural variability can be assessed through a multimember ensemble of simulations with a single model initialized with different states.

The second source of uncertainty is model error, seen in the model response to the imposed forcing scenarios. Models are imperfect and differ in their forced response owing to their spatial resolution and imprecision in their parameterization of the various physical, chemical, and biological processes. Model uncertainty is assessed through multimodel ensemble studies.

The third source of uncertainty is the forcing scenarios and their depiction of the time evolution of greenhouse gases, land use, and other anthropogenic forcings of climate, which are imposed as boundary conditions to the models. For temperature projections at the global scale, model uncertainty and natural variability dominate at near-term decadal time scales (10 to 30 years) (83). Scenarios are the major source of uncertainty at multidecadal lead times. Total uncertainty is larger

at regional scales, mostly from natural variability and model structure.

A related concept is time of emergence. Determining the time at which the forced climate change signal emerges from the noise of natural variability is a necessary requirement in assessing when an expected change can be detected or whether observed changes can be attributed to anthropogenic forcings (84, 85). Time of emergence has been mostly studied for temperature and can vary from a few decades in mid-latitudes with low natural variability to several decades or longer in regions with larger natural variability.

Can these concepts of predictability, uncertainty, and time of emergence be extended to study the biosphere in the Earth system? ESMs predict prominent changes in terrestrial and marine biogeochemistry, but only recently have the necessary large, multimodel and multimember ensembles become available to distinguish the forced response from natural variability and model uncertainty. Such analyses give important insights into the use of ESMs to understand changes in Earth system biogeochemistry. In addition to warmer temperatures, the ocean has trends of increased carbon uptake, acidification, lower O<sub>2</sub>, and reduced net primary production over the next several decades in the absence of mitigation (86). The forced trend in air-sea CO<sub>2</sub> flux emerges rapidly in some ocean regions, but large natural variability precludes detection of changes in the rate of carbon uptake until mid-century or later in many regions (87). Other aspects of ocean biogeochemistry such as pH, O<sub>2</sub> concentration, and net primary production also have large, regionally dependent natural variability (88–90). Ocean acidification has a rapid time of emergence driven

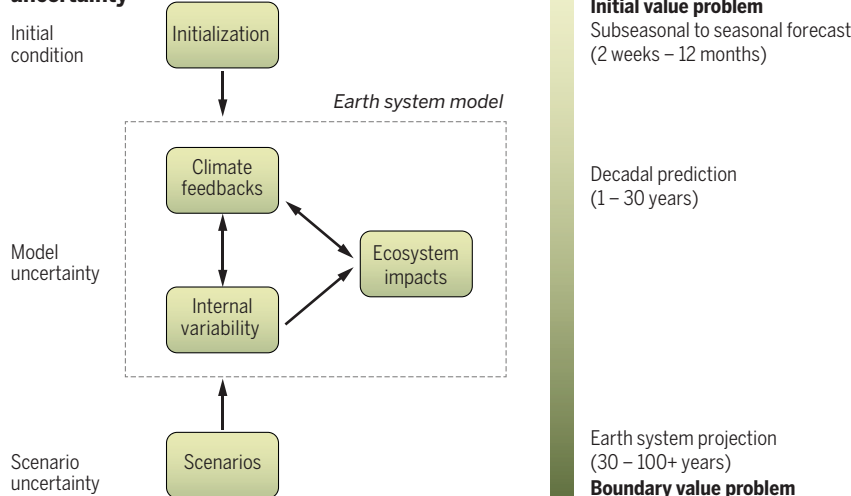
by the accumulation of anthropogenic CO<sub>2</sub> in the surface layer, and the sea surface temperature warming signal also emerges within a few decades in many regions, but forced changes in O<sub>2</sub> concentration and net primary production do not emerge from natural variability until mid- to late-century, if at all (Fig. 3, A to C).

Time of emergence has been less studied in the terrestrial biosphere. Observational and modeling analyses support an enhanced terrestrial carbon sink arising from global change (17, 56, 91). Unforced variability in the land-atmosphere CO<sub>2</sub> flux is large and precludes detection of change at decadal time scales (92). There is considerable variability within and among models, and the forced response statistically emerges only after several decades in many regions of the world. The HadGEM2-ES model, for example, shows the forced signal of terrestrial carbon gain as emerging from internal variability in many regions by 2030, but other models show a weaker signal that has yet to statistically emerge, and even carbon loss rather than gain (Fig. 3, D to F).

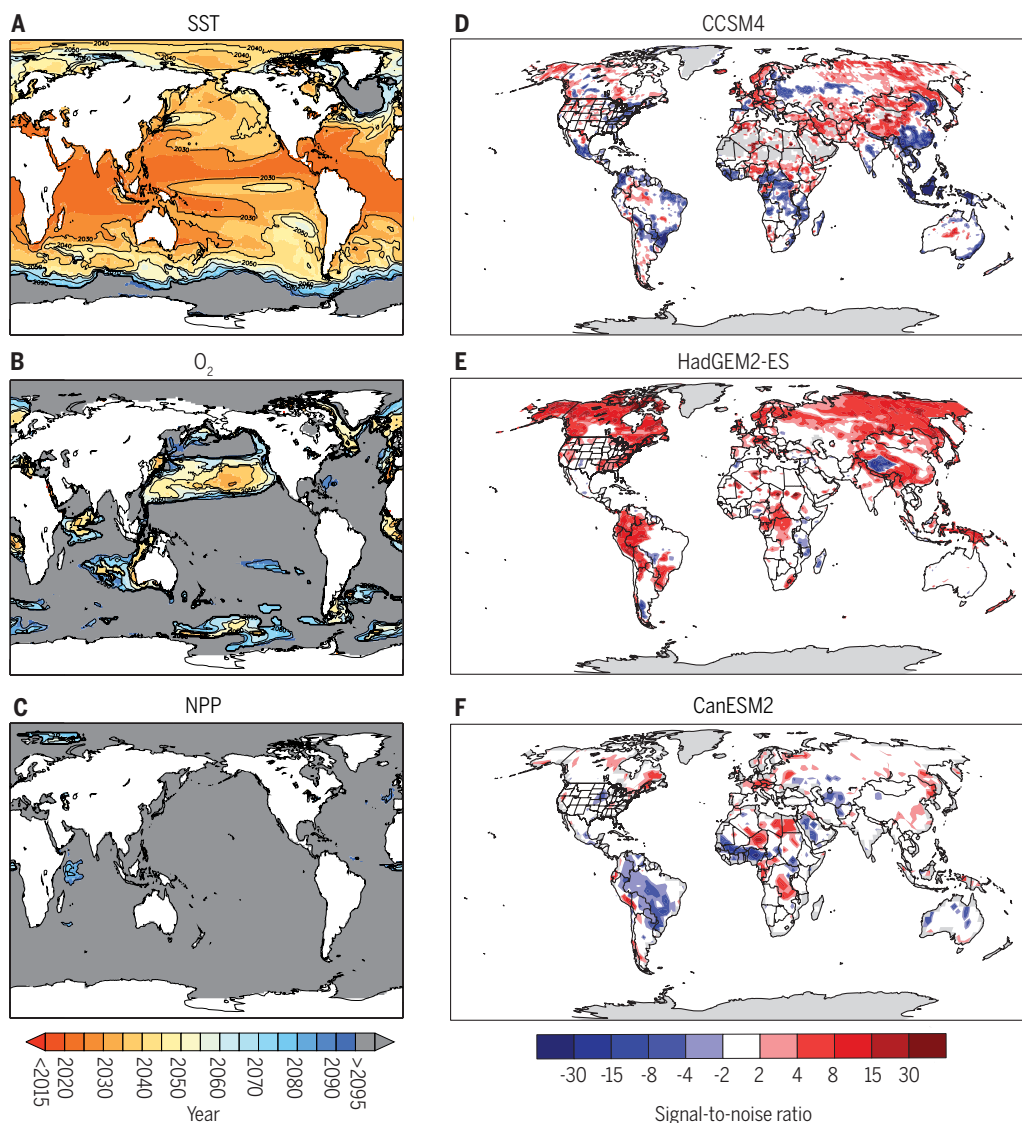
The various contributions to uncertainty differ depending on the quantity of interest, lead time, and spatial scale. The uncertainty from natural variability is particularly large at small spatial scales and short lead times for pH, O<sub>2</sub> concentration, and net primary production in the ocean (89) and air-sea carbon flux (93). By the end of the 21st century, scenario uncertainty dominates total uncertainty for these states and fluxes at the global scale (except for net primary production), but natural variability and model uncertainty remain large at the regional or biome scale. Simulations of the terrestrial carbon cycle are much more variable among models and largely dominated by model uncertainty (94). Comparisons of ocean and land carbon cycle projections point to a markedly different assessment of uncertainty (Fig. 4). For ocean carbon uptake, model uncertainty is initially large, but scenario uncertainty dominates during the latter part of the 21st century. In contrast, model structure contributes 80% of total uncertainty throughout the 21st century for the terrestrial carbon cycle.

Much of the study of Earth system prediction is focused on climate, and decadal climate prediction is particularly focused on the dynamics and thermal characteristics of the ocean because of its prominent role in climate variability. When soil moisture and vegetation are considered by atmospheric modelers, it is often in the context of how these affect climate predictability, rather than whether they can be predicted (79). In a global change perspective, ecological predictions are as essential as those of the physical climate system. ESMs provide the means not just to assess the potential for future stresses engendered by a changing climate, but also to determine the outcome of those stresses on crop yield, tree mortality, fisheries, and other aspects of the biosphere. For example, characterizing when and where wildfires might occur would be valuable to aid governmental agencies charged with wildfire protection. Particular fire events are nearly impossible to forecast, especially because so many

## Sources of uncertainty



**Fig. 2. Schematic depiction of Earth system prediction of the biosphere.** The synergies between climate feedback processes, internal climate variability, and ecosystem impacts determine model outcomes. Subseasonal to seasonal forecasts and decadal climate prediction are initial value problems. Earth system projections are a boundary value problem driven by anthropogenic forcing scenarios. Uncertainty arises from inexactness of initial conditions, model imperfections, and scenarios.



**Fig. 3. Ocean and land forced trends relative to internal variability and model uncertainty.**

Data are from a range of ESMs contributed to CMIP5. (A to C) Multimodel time of emergence for SST,  $O_2$ , and net primary production (NPP) (89). Time of emergence is defined as the year at which the signal exceeds the noise, which, as used here, includes both internal variability and model uncertainty. The forced SST signal emerges rapidly in many locations,  $O_2$  time of emergence is regionally variable, and the forced NPP signal does not statistically emerge by 2100. (D to F) Signal-to-noise ratio for cumulative land carbon uptake in a business-as-usual scenario at 2030 for three different ESMs (92). Positive (negative) values indicate carbon gain (loss). In these panels, the noise is strictly internal variability, and a ratio greater than 2 or less than  $-2$  indicates that the signal has emerged from the internal variability. There are considerable differences among models in the sign of the terrestrial carbon flux and whether the change has emerged from natural variability by 2030. CCSM4, Community Climate System Model version 4; HadGEM2-ES, Hadley Centre Global Environmental Model version 2; CanESM2, second-generation Canadian Earth System Model.

fires are caused by people, but fire behavior on seasonal time scales can be forecast on the basis of relationships with sea surface temperatures (95). Prediction of future fire behavior requires models that accurately depict fire occurrence and severity, but wildfire prediction will also require an understanding of the predictability of the climate that drives fire behavior. A similar argument pertains to crop yield, marine resources, and dust emissions simulated in the current generation of ESMs and to forest mortality and habit loss that will be simulated by the next generation of models. Such forecasts may be particularly relevant at the subseasonal to seasonal time scale (79).

### Research needs

As this Review highlights, the biosphere is central to understanding why and how the Earth system is changing and to adapting to and mitigating future changes. Many of the global change stressors that terrestrial and marine ecosystems face

need to be understood not only for their impacts on ecosystem services that are essential to humankind, but also as processes that affect the magnitude and trajectory of climate change. A strategy is needed to extend the study of subseasonal and seasonal forecasts and decadal climate prediction to a more multifaceted Earth system prediction, including the biosphere and its resources. The extension of seasonal to decadal climate forecasts to living marine resources, for example, has considerable potential to aid marine management (96). A similar extension to terrestrial ecosystems would aid land resource management.

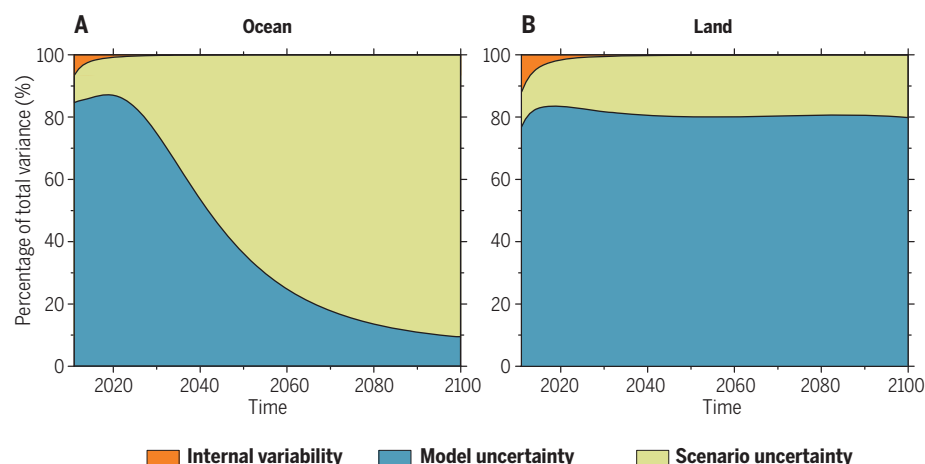
Toward this goal, this Review has presented several pathways to further define Earth system prediction. First is continued advancement of terrestrial and marine science in light of climate processes and the many ways in which the biosphere influences climate. A prominent example is the carbon cycle, its feedback with climate change, and whether terrestrial and marine eco-

systems can be purposely managed to mitigate anthropogenic  $CO_2$  emissions. There is considerable uncertainty in ocean carbon cycle projections, particularly at regional or biome scales (89, 93), and land carbon model uncertainty precludes distinguishing among various alternative scenarios (94). Moreover, if planting forests and biofuels are essential to maintaining atmospheric  $CO_2$  concentrations within some planetary warming target, how confident are we in our ability to know the net climate outcome of these policies (68)?

The current generation of ecosystem models are abstractions of complex systems. Many ecological and biogeochemical processes are represented, but the challenge of representing the biosphere—with its rich diversity of life forms, their assemblage into communities and ecosystems, and the complexity of ecological systems—is daunting, as is evident in large model uncertainty in terrestrial carbon cycle projections. Theoretical

**Fig. 4. Ocean and land carbon cycle**

**uncertainty.** The percentage of total variance attributed to internal variability, model uncertainty, and scenario uncertainty in projections of cumulative global carbon uptake from 2006 to 2100 differs widely between (A) ocean and (B) land. The ocean carbon cycle is dominated by scenario uncertainty by the middle of the century, but uncertainty in the land carbon cycle is mostly from model structure. Data are from 12 ESMs using four different scenarios (94).



advances are needed, but there may be a limit to how much model uncertainty can be reduced (94). More complexity does not necessarily lead to better predictions or reduce uncertainty.

A second pathway is to better integrate ESMs and VIA models. The gap between models arises from disciplinary expertise (atmospheric and ocean sciences for ESMs and hydrology, ecology, biogeochemistry, agronomy, forestry, and marine sciences for VIA models), but effective communication among, rather than across, disciplines is not trivial. There are also pragmatic considerations, particularly with regard to spatial scale and process complexity, that limit collaboration between global ESMs and VIA models with a more local to regional domain. However, just as the science of Earth system prediction is seen as a means to unite the weather and climate modeling communities (80, 81), so, too, can the broadening of Earth system prediction to include the biosphere stimulate collaborations with the VIA community.

A third promising research pathway is to expand the concepts and methodology of seasonal to decadal climate prediction to include terrestrial and marine ecosystems and to quantify prediction uncertainty at spatial and temporal scales relevant to stakeholders. The predictability of the terrestrial carbon cycle can be considered from an ecological perspective (97), but only recently has it been considered in an Earth system perspective of natural climate variability, the forced climate response, and model uncertainty (92, 94). Analysis of natural variability, model uncertainty, and scenario uncertainty is similarly informing marine biogeochemistry (87–90, 93). Whether the biosphere is a source of climate predictability is not necessarily the right question to pose. A more fruitful research pathway may be to investigate how to predict the biosphere and its resources in a changing environment, as identified specifically for marine living resources (96) and considered also for atmospheric CO<sub>2</sub> (98). Initial condition uncertainty and the difficulty in separating natural variability from the forced trend likely produces irreducible uncertainty in climate prediction (99). At the regional or biome scale, natural variability is large for the ocean and land

carbon cycles (89, 92, 93). Whether a similar irreducible uncertainty manifests in terrestrial and marine ecosystems remains to be explored.

With their terrestrial and marine ecosystems, biogeochemical cycles, and simulation of plants, microbes, and marine life, ESMs challenge terrestrial and marine ecologists and biogeochemists to think in terms of broad generalizations and to find the mathematical equations to describe the biosphere, its functioning, and its response to global change. ESMs similarly challenge geoscientists to think beyond a physical understanding of climate to include biology. The models show much promise to advance our understanding of global change but must move from the synthetic world of an ESM toward the real world. Bridging the gap between observations and theory as atmospheric CO<sub>2</sub> rises, climate changes, more nitrogen is added to the system, forests are cleared, grasslands are plowed or converted to pastures, coastal wetlands and coral reefs are degraded or lost, and oceans warm and are increasingly polluted poses challenging opportunities for the next generation of scientists to advance planetary ecology and climate science.

**REFERENCES AND NOTES**

1. J. Rockström *et al.*, A safe operating space for humanity. *Nature* **461**, 472–475 (2009). doi: [10.1038/461472a](https://doi.org/10.1038/461472a); pmid: [19779433](https://pubmed.ncbi.nlm.nih.gov/19779433/)
2. W. Steffen *et al.*, Planetary boundaries: Guiding human development on a changing planet. *Science* **347**, 1259855 (2015). doi: [10.1126/science.1259855](https://doi.org/10.1126/science.1259855); pmid: [25592418](https://pubmed.ncbi.nlm.nih.gov/25592418/)
3. Intergovernmental Panel on Climate Change (IPCC), *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, T. F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, P. M. Midgley, Eds. (Cambridge Univ. Press, 2013).
4. J. A. Foley *et al.*, Global consequences of land use. *Science* **309**, 570–574 (2005). doi: [10.1126/science.1111772](https://doi.org/10.1126/science.1111772); pmid: [16040698](https://pubmed.ncbi.nlm.nih.gov/16040698/)
5. J. N. Galloway *et al.*, The nitrogen cascade. *Bioscience* **53**, 341–356 (2003). doi: [10.1641/0006-3568\(2003\)053\[0341:TNC\]2.0.CO;2](https://doi.org/10.1641/0006-3568(2003)053[0341:TNC]2.0.CO;2)
6. IPCC, *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, C. B. Field, V. R. Barros, D. J. Dokken, K. J. Mach, M. D. Mastrandrea, T. E. Bilir, M. Chatterjee, K. L. Ebi,

- Y. O. Estrada, R. C. Genova, B. Girma, E. S. Kissel, A. N. Levy, S. MacCracken, P. R. Mastrandrea, L. L. White, Eds. (Cambridge Univ. Press, 2014).
7. IPCC, *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part B: Regional Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, V. R. Barros, C. B. Field, D. J. Dokken, M. D. Mastrandrea, K. J. Mach, T. E. Bilir, M. Chatterjee, K. L. Ebi, Y. O. Estrada, R. C. Genova, B. Girma, E. S. Kissel, A. N. Levy, S. MacCracken, P. R. Mastrandrea, L. L. White, Eds. (Cambridge Univ. Press, 2014).
8. B. R. Scheffers *et al.*, The broad footprint of climate change from genes to biomes to people. *Science* **354**, aaf7671 (2016). doi: [10.1126/science.aaf7671](https://doi.org/10.1126/science.aaf7671); pmid: [27846577](https://pubmed.ncbi.nlm.nih.gov/27846577/)
9. IPCC, *Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, O. Edenhofer, R. Pichs-Madruga, Y. Sokona, E. Farahani, S. Kadner, K. Seyboth, A. Adler, I. Baum, S. Brunner, P. Eickemeier, B. Kriemann, J. Savolainen, S. Schlömer, C. von Stechow, T. Zwickel, J. C. Minx, Eds. (Cambridge Univ. Press, 2014).
10. R. H. Moss *et al.*, The next generation of scenarios for climate change research and assessment. *Nature* **463**, 747–756 (2010). doi: [10.1038/nature08823](https://doi.org/10.1038/nature08823); pmid: [20148028](https://pubmed.ncbi.nlm.nih.gov/20148028/)
11. W. D. Collins *et al.*, The integrated Earth system model version 1: Formulation and functionality. *Geosci. Model Dev.* **8**, 2203–2219 (2015). doi: [10.5194/gmd-8-2203-2015](https://doi.org/10.5194/gmd-8-2203-2015)
12. V. Eyring *et al.*, Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geosci. Model Dev.* **9**, 1937–1958 (2016). doi: [10.5194/gmd-9-1937-2016](https://doi.org/10.5194/gmd-9-1937-2016)
13. G. B. Bonan, *Ecological Climatology: Concepts and Applications* (Cambridge Univ. Press, ed. 3, 2016).
14. C. Sweeney *et al.*, Impacts of shortwave penetration depth on large-scale ocean circulation and heat transport. *J. Phys. Oceanogr.* **35**, 1103–1119 (2005). doi: [10.1175/JPO2740.1](https://doi.org/10.1175/JPO2740.1)
15. M. Jochum, S. Yeager, K. Lindsay, K. Moore, R. Murtugudde, Quantification of the feedback between phytoplankton and ENSO in the Community Climate System Model. *J. Clim.* **23**, 2916–2925 (2010). doi: [10.1175/2010JCLI3254.1](https://doi.org/10.1175/2010JCLI3254.1)
16. C. Le Quéré *et al.*, Global carbon budget 2016. *Earth Syst. Sci. Data* **8**, 605–649 (2016). doi: [10.5194/essd-8-605-2016](https://doi.org/10.5194/essd-8-605-2016)
17. P. Friedlingstein *et al.*, Uncertainties in CMIP5 climate projections due to carbon cycle feedbacks. *J. Clim.* **27**, 511–526 (2014). doi: [10.1175/JCLI-D-12-00579.1](https://doi.org/10.1175/JCLI-D-12-00579.1)
18. R. A. Fisher *et al.*, Taking off the training wheels: The properties of a dynamic vegetation model without climate envelopes, CLM4.5(ED). *Geosci. Model Dev.* **8**, 3593–3619 (2015). doi: [10.5194/gmd-8-3593-2015](https://doi.org/10.5194/gmd-8-3593-2015)
19. J. L. Sarmiento, N. Gruber, *Ocean Biogeochemical Dynamics* (Princeton Univ. Press, 2006).
20. A. Tagliabue *et al.*, How well do global ocean biogeochemistry models simulate dissolved iron distributions? *Global Biogeochem. Cycles* **30**, 149–174 (2016). doi: [10.1002/2015GB005289](https://doi.org/10.1002/2015GB005289)

21. J. K. Moore, K. Lindsay, S. C. Doney, M. C. Long, K. Misumi, Marine ecosystem dynamics and biogeochemical cycling in the Community Earth System Model [CESM1(BGC)]: Comparison of the 1990s with the 2090s under the RCP4.5 and RCP8.5 scenarios. *J. Clim.* **26**, 9291–9312 (2013). doi: [10.1175/JCLI-D-12-00566.1](https://doi.org/10.1175/JCLI-D-12-00566.1)
22. R. R. Hood *et al.*, Pelagic functional group modeling: Progress, challenges and prospects. *Deep Sea Res. Part II Top. Stud. Oceanogr.* **53**, 459–512 (2006). doi: [10.1016/j.dsr.2.2006.01.025](https://doi.org/10.1016/j.dsr.2.2006.01.025)
23. E. Litchman, C. A. Klausmeier, Trait-based community ecology of phytoplankton. *Annu. Rev. Ecol. Evol. Syst.* **39**, 615–639 (2008). doi: [10.1146/annurev.ecolsys.39.110707.173549](https://doi.org/10.1146/annurev.ecolsys.39.110707.173549)
24. M. J. Follows, S. Dutkiewicz, Modeling diverse communities of marine microbes. *Annu. Rev. Mar. Sci.* **3**, 427–451 (2011). doi: [10.1146/annurev-marine-120709-142848](https://doi.org/10.1146/annurev-marine-120709-142848); pmid: [21329212](https://pubmed.ncbi.nlm.nih.gov/21329212/)
25. C. A. Stock *et al.*, Reconciling fisheries catch and ocean productivity. *Proc. Natl. Acad. Sci. U.S.A.* **114**, E1441–E1449 (2017). doi: [10.1073/pnas.1610238114](https://doi.org/10.1073/pnas.1610238114); pmid: [28115722](https://pubmed.ncbi.nlm.nih.gov/28115722/)
26. R. W. Pinder *et al.*, Climate change impacts of US reactive nitrogen. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 7671–7675 (2012). doi: [10.1073/pnas.114243109](https://doi.org/10.1073/pnas.114243109); pmid: [22547815](https://pubmed.ncbi.nlm.nih.gov/22547815/)
27. A. Krishnamurthy *et al.*, Impacts of increasing anthropogenic soluble iron and nitrogen deposition on ocean biogeochemistry. *Global Biogeochem. Cycles* **23**, GB3016 (2009). doi: [10.1029/2008GB003440](https://doi.org/10.1029/2008GB003440)
28. N. M. Mahowald *et al.*, Aerosol deposition impacts on land and ocean carbon cycles. *Curr. Clim. Change Rep.* **3**, 16–31 (2017). doi: [10.1007/s40641-017-0056-z](https://doi.org/10.1007/s40641-017-0056-z)
29. S. Sith, P. M. Cox, W. J. Collins, C. Huntingford, Indirect radiative forcing of climate change through ozone effects on the land-carbon sink. *Nature* **448**, 791–794 (2007). doi: [10.1038/nature06059](https://doi.org/10.1038/nature06059); pmid: [17653194](https://pubmed.ncbi.nlm.nih.gov/17653194/)
30. D. Lombardozzi, S. Levis, G. Bonan, P. G. Hess, J. P. Sparks, The influence of chronic ozone exposure on global carbon and water cycles. *J. Clim.* **28**, 292–305 (2015). doi: [10.1175/JCLI-D-14-00223.1](https://doi.org/10.1175/JCLI-D-14-00223.1)
31. N. Unger, Human land-use-driven reduction of forest volatiles cools global climate. *Nat. Clim. Change* **4**, 907–910 (2014). doi: [10.1038/nclimate2347](https://doi.org/10.1038/nclimate2347)
32. J. R. Melton *et al.*, Present state of global wetland extent and wetland methane modelling: Conclusions from a model inter-comparison project (WETCHIMP). *Biogeosciences* **10**, 753–788 (2013). doi: [10.5194/bg-10-753-2013](https://doi.org/10.5194/bg-10-753-2013)
33. D. T. Shindell *et al.*, Interactive ozone and methane chemistry in GISS-E2 historical and future climate simulations. *Atmos. Chem. Phys.* **13**, 2653–2689 (2013). doi: [10.5194/acp-13-2653-2013](https://doi.org/10.5194/acp-13-2653-2013)
34. D. S. Ward *et al.*, The changing radiative forcing of fires: Global model estimates for past, present and future. *Atmos. Chem. Phys.* **12**, 10857–10886 (2012). doi: [10.5194/acp-12-10857-2012](https://doi.org/10.5194/acp-12-10857-2012)
35. S. Hantson *et al.*, The status and challenge of global fire modelling. *Biogeosciences* **13**, 3359–3375 (2016). doi: [10.5194/bg-13-3359-2016](https://doi.org/10.5194/bg-13-3359-2016)
36. W. A. Kurz *et al.*, Mountain pine beetle and forest carbon feedback to climate change. *Nature* **452**, 987–990 (2008). doi: [10.1038/nature06777](https://doi.org/10.1038/nature06777); pmid: [18432244](https://pubmed.ncbi.nlm.nih.gov/18432244/)
37. B. C. Bright, J. A. Hicke, A. J. H. Meddens, Effects of bark beetle-caused tree mortality on biogeochemical and biogeophysical MODIS products. *J. Geophys. Res. Biogeosci.* **118**, 974–982 (2013). doi: [10.1002/jgrg.20078](https://doi.org/10.1002/jgrg.20078)
38. H. Maness, P. J. Kushner, I. Fung, Summertime climate response to mountain pine beetle disturbance in British Columbia. *Nat. Geosci.* **6**, 65–70 (2013). doi: [10.1038/ngeo1642](https://doi.org/10.1038/ngeo1642)
39. S. L. Edburg, J. A. Hicke, D. M. Lawrence, P. E. Thornton, Simulating coupled carbon and nitrogen dynamics following mountain pine beetle outbreaks in the western United States. *J. Geophys. Res.* **116**, G04033 (2011). doi: [10.1029/2011JG001786](https://doi.org/10.1029/2011JG001786)
40. G. C. Hurtt *et al.*, Harmonization of land-use scenarios for the period 1500–2100: 600 years of global gridded annual land-use transitions, wood harvest, and resulting secondary lands. *Clim. Change* **109**, 117–161 (2011). doi: [10.1007/s10584-011-0153-2](https://doi.org/10.1007/s10584-011-0153-2)
41. S. Levis *et al.*, Interactive crop management in the Community Earth System Model (CESM1): Seasonal influences on land-atmosphere fluxes. *J. Clim.* **25**, 4839–4859 (2012). doi: [10.1175/JCLI-D-11-00446.1](https://doi.org/10.1175/JCLI-D-11-00446.1)
42. S. Levis, A. Badger, B. Drewniak, C. Nevison, X. Ren, CLMcrop yields and water requirements: Avoided impacts by choosing RCP 4.5 over 8.5. *Clim. Change* (2016). doi: [10.1007/s10584-016-1654-9](https://doi.org/10.1007/s10584-016-1654-9)
43. S. Luyssaert *et al.*, Land management and land-cover change have impacts of similar magnitude on surface temperature. *Nat. Clim. Change* **4**, 389–393 (2014). doi: [10.1038/nclimate2196](https://doi.org/10.1038/nclimate2196)
44. S. C. Doney *et al.*, Climate change impacts on marine ecosystems. *Annu. Rev. Mar. Sci.* **4**, 11–37 (2012). doi: [10.1146/annurev-marine-041911-111611](https://doi.org/10.1146/annurev-marine-041911-111611); pmid: [22457967](https://pubmed.ncbi.nlm.nih.gov/22457967/)
45. O. Hoegh-Guldberg, J. F. Bruno, The impact of climate change on the world's marine ecosystems. *Science* **328**, 1523–1528 (2010). doi: [10.1126/science.1189930](https://doi.org/10.1126/science.1189930); pmid: [20558709](https://pubmed.ncbi.nlm.nih.gov/20558709/)
46. E. S. Poloczanska *et al.*, Global imprint of climate change on marine life. *Nat. Clim. Change* **3**, 919–925 (2013). doi: [10.1038/nclimate1958](https://doi.org/10.1038/nclimate1958)
47. A. B. Hollowed *et al.*, Projected impacts of climate change on marine fish and fisheries. *ICES J. Mar. Sci.* **70**, 1023–1037 (2013). doi: [10.1093/icesjms/fst081](https://doi.org/10.1093/icesjms/fst081)
48. C. A. Burge *et al.*, Climate change influences on marine infectious diseases: Implications for management and society. *Annu. Rev. Mar. Sci.* **6**, 249–277 (2014). doi: [10.1146/annurev-marine-010213-135029](https://doi.org/10.1146/annurev-marine-010213-135029); pmid: [23808894](https://pubmed.ncbi.nlm.nih.gov/23808894/)
49. J. G. Molinos *et al.*, Climate velocity and the future global redistribution of marine biodiversity. *Nat. Clim. Change* **6**, 83–88 (2016). doi: [10.1038/nclimate2769](https://doi.org/10.1038/nclimate2769)
50. B. S. Halpern *et al.*, Spatial and temporal changes in cumulative human impacts on the world's ocean. *Nat. Commun.* **6**, 7615 (2015). doi: [10.1038/ncomms8615](https://doi.org/10.1038/ncomms8615); pmid: [26172980](https://pubmed.ncbi.nlm.nih.gov/26172980/)
51. J. M. Melillo *et al.*, Vegetation/ecosystem modeling and analysis project: Comparing biogeography and biogeochemistry models in a continental-scale study of terrestrial ecosystem responses to climate change and CO<sub>2</sub> doubling. *Global Biogeochem. Cycles* **9**, 407–437 (1995). doi: [10.1029/95GB02746](https://doi.org/10.1029/95GB02746)
52. C. Rosenzweig *et al.*, Assessing agricultural risks of climate change in the 21st century in a global gridded crop model intercomparison. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 3268–3273 (2014). doi: [10.1073/pnas.1222463110](https://doi.org/10.1073/pnas.1222463110); pmid: [24344314](https://pubmed.ncbi.nlm.nih.gov/24344314/)
53. J. Schewe *et al.*, Multimodel assessment of water scarcity under climate change. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 3245–3250 (2014). doi: [10.1073/pnas.1222460110](https://doi.org/10.1073/pnas.1222460110); pmid: [24344289](https://pubmed.ncbi.nlm.nih.gov/24344289/)
54. C. A. Stock *et al.*, On the use of IPCC-class models to assess the impact of climate on Living Marine Resources. *Prog. Oceanogr.* **88**, 1–27 (2011). doi: [10.1016/j.pocean.2010.09.001](https://doi.org/10.1016/j.pocean.2010.09.001)
55. Z. Zhu *et al.*, Greening of the Earth and its drivers. *Nat. Clim. Change* **6**, 791–795 (2016). doi: [10.1038/nclimate3004](https://doi.org/10.1038/nclimate3004)
56. S. Sith *et al.*, Recent trends and drivers of regional sources and sinks of carbon dioxide. *Biogeosciences* **12**, 653–679 (2015). doi: [10.5194/bg-12-653-2015](https://doi.org/10.5194/bg-12-653-2015)
57. T. Ito, S. Minobe, M. C. Long, C. Deutsch, Upper ocean O<sub>2</sub> trends: 1958–2015. *Geophys. Res. Lett.* **44**, 4214–4223 (2017). doi: [10.1002/2017GL073613](https://doi.org/10.1002/2017GL073613)
58. S. A. Henson, J. P. Dunne, J. L. Sarmiento, Decadal variability in North Atlantic phytoplankton blooms. *J. Geophys. Res.* **114**, C04013 (2009). doi: [10.1029/2008JC005139](https://doi.org/10.1029/2008JC005139)
59. J. A. Kleypas *et al.*, Larval connectivity across temperature gradients and its potential effect on heat tolerance in coral populations. *Glob. Change Biol.* **22**, 3539–3549 (2016). doi: [10.1111/gcb.13347](https://doi.org/10.1111/gcb.13347); pmid: [27154763](https://pubmed.ncbi.nlm.nih.gov/27154763/)
60. G. B. Anderson, K. W. Oleson, B. Jones, R. D. Peng, Projected trends in high-mortality heatwaves under different scenarios of climate, population, and adaptation in 82 US communities. *Clim. Change* (2016). doi: [10.1007/s10584-016-1779-x](https://doi.org/10.1007/s10584-016-1779-x)
61. A. C. Ruane *et al.*, The Vulnerability, Impacts, Adaptation and Climate Services Advisory Board (VIACS AB v1.0) contribution to CMIP6. *Geosci. Model Dev.* **9**, 3493–3515 (2016). doi: [10.5194/gmd-9-3493-2016](https://doi.org/10.5194/gmd-9-3493-2016)
62. A. L. S. Swann, F. M. Hoffman, C. D. Koven, J. T. Randerson, Plant responses to increasing CO<sub>2</sub> reduce estimates of climate impacts on drought severity. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 10019–10024 (2016). doi: [10.1073/pnas.1604581113](https://doi.org/10.1073/pnas.1604581113); pmid: [27573831](https://pubmed.ncbi.nlm.nih.gov/27573831/)
63. S. Piao *et al.*, Changes in climate and land use have a larger direct impact than rising CO<sub>2</sub> on global river runoff trends. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 15242–15247 (2007). doi: [10.1073/pnas.0707213104](https://doi.org/10.1073/pnas.0707213104); pmid: [17878298](https://pubmed.ncbi.nlm.nih.gov/17878298/)
64. Q. Sun, M. M. Whitney, F. O. Bryan, Y. Tseng, A box model for representing estuarine physical processes in Earth system models. *Ocean Model.* **112**, 139–153 (2017). doi: [10.1016/j.ocemod.2017.03.004](https://doi.org/10.1016/j.ocemod.2017.03.004)
65. X. Huang, A. M. Rhoades, P. A. Ullrich, C. M. Zarzycki, An evaluation of the variable-resolution CESM for modeling California's climate. *J. Adv. Model. Earth Syst.* **8**, 345–369 (2016). doi: [10.1002/2015MS000559](https://doi.org/10.1002/2015MS000559)
66. A. M. Rhoades, X. Huang, P. A. Ullrich, C. M. Zarzycki, Characterizing Sierra Nevada snowpack using variable-resolution CESM. *J. Appl. Meteorol. Climatol.* **55**, 173–196 (2016). doi: [10.1175/JAMC-D-15-0156.1](https://doi.org/10.1175/JAMC-D-15-0156.1)
67. National Research Council, *Climate Intervention: Carbon Dioxide Removal and Reliable Sequestration* (National Academies Press, 2015).
68. G. B. Bonan, Forests, climate, and public policy: A 500-year interdisciplinary odyssey. *Annu. Rev. Ecol. Evol. Syst.* **47**, 97–121 (2016). doi: [10.1146/annurev-ecolsys-121415-032359](https://doi.org/10.1146/annurev-ecolsys-121415-032359)
69. P. Smith *et al.*, "Agriculture, forestry and other land use (AFOLU)," in *Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, O. Edenhofer *et al.*, Eds. (Cambridge Univ. Press, 2014), pp. 811–922.
70. N. de Noblet-Ducoudré *et al.*, Determining robust impacts of land-use-induced land cover changes on surface climate over North America and Eurasia: Results from the first set of LUCID experiments. *J. Clim.* **25**, 3261–3281 (2012). doi: [10.1175/JCLI-D-11-00338.1](https://doi.org/10.1175/JCLI-D-11-00338.1)
71. N. D. Mueller *et al.*, Cooling of US Midwest summer temperature extremes from cropland intensification. *Nat. Clim. Change* **6**, 317–322 (2016). doi: [10.1038/nclimate2825](https://doi.org/10.1038/nclimate2825)
72. E. L. Davin, S. I. Seneviratne, P. Ciais, A. Olioso, T. Wang, Preferential cooling of hot extremes from cropland albedo management. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 9757–9761 (2014). doi: [10.1073/pnas.1317323111](https://doi.org/10.1073/pnas.1317323111); pmid: [24958872](https://pubmed.ncbi.nlm.nih.gov/24958872/)
73. K. Caldeira, G. Bala, L. Cao, The science of geoengineering. *Annu. Rev. Earth Planet. Sci.* **41**, 231–256 (2013). doi: [10.1146/annurev-earth-042711-105548](https://doi.org/10.1146/annurev-earth-042711-105548)
74. P. Smith *et al.*, Biophysical and economic limits to negative CO<sub>2</sub> emissions. *Nat. Clim. Change* **6**, 42–50 (2016). doi: [10.1038/nclimate2870](https://doi.org/10.1038/nclimate2870)
75. L. M. Russell *et al.*, Ecosystem impacts of geoengineering: A review for developing a science plan. *Ambio* **41**, 350–369 (2012). doi: [10.1007/s13280-012-0258-5](https://doi.org/10.1007/s13280-012-0258-5); pmid: [22430307](https://pubmed.ncbi.nlm.nih.gov/22430307/)
76. P. J. Irvine *et al.*, Towards a comprehensive climate impacts assessment of solar geoengineering. *Earth's Future* **5**, 93–106 (2017). doi: [10.1002/2016EF000389](https://doi.org/10.1002/2016EF000389)
77. G. A. Meehl *et al.*, Decadal prediction: Can it be skillful? *Bull. Am. Meteorol. Soc.* **90**, 1467–1485 (2009). doi: [10.1175/2009BAMS2778.1](https://doi.org/10.1175/2009BAMS2778.1)
78. G. A. Meehl *et al.*, Decadal climate prediction: An update from the trenches. *Bull. Am. Meteorol. Soc.* **95**, 243–267 (2014). doi: [10.1175/BAMS-D-12-00241.1](https://doi.org/10.1175/BAMS-D-12-00241.1)
79. National Academies of Sciences, Engineering, and Medicine, *Next Generation Earth System Prediction: Strategies for Subseasonal to Seasonal Forecasts* (National Academies Press, 2016).
80. W. Hazeleger *et al.*, EC-Earth: A seamless Earth-system prediction approach in action. *Bull. Am. Meteorol. Soc.* **91**, 1357–1363 (2010). doi: [10.1175/2010BAMS2877.1](https://doi.org/10.1175/2010BAMS2877.1)
81. J. C. Carman *et al.*, The national Earth system prediction capability: Coordinating the giant. *Bull. Am. Meteorol. Soc.* **98**, 239–252 (2017). doi: [10.1175/BAMS-D-16-0002.1](https://doi.org/10.1175/BAMS-D-16-0002.1)
82. A. Jahn, J. E. Kay, M. M. Holland, D. M. Hall, How predictable is the timing of a summer ice-free Arctic? *Geophys. Res. Lett.* **43**, 9113–9120 (2016). doi: [10.1002/2016GL070067](https://doi.org/10.1002/2016GL070067)
83. E. Hawkins, R. Sutton, The potential to narrow uncertainty in regional climate predictions. *Bull. Am. Meteorol. Soc.* **90**, 1095–1107 (2009). doi: [10.1175/2009BAMS2607.1](https://doi.org/10.1175/2009BAMS2607.1)
84. C. Deser, R. Knutti, S. Solomon, A. S. Phillips, Communication of the role of natural variability in future North American climate. *Nat. Clim. Change* **2**, 775–779 (2012). doi: [10.1038/nclimate1562](https://doi.org/10.1038/nclimate1562)
85. E. Hawkins, R. Sutton, Time of emergence of climate signals. *Geophys. Res. Lett.* **39**, L01702 (2012). doi: [10.1029/2011GL005087](https://doi.org/10.1029/2011GL005087)
86. L. Bopp *et al.*, Multiple stressors of ocean ecosystems in the 21st century: Projections with CMIP5 models. *Biogeosciences* **10**, 6225–6245 (2013). doi: [10.5194/bg-10-6225-2013](https://doi.org/10.5194/bg-10-6225-2013)
87. G. A. McKinley *et al.*, Timescales for detection of trends in the ocean carbon sink. *Nature* **530**, 469–472 (2016). doi: [10.1038/nature16958](https://doi.org/10.1038/nature16958); pmid: [26911782](https://pubmed.ncbi.nlm.nih.gov/26911782/)

88. K. B. Rodgers, J. Lin, T. L. Frölicher, Emergence of multiple ocean ecosystem drivers in a large ensemble suite with an Earth system model. *Biogeosciences* **12**, 3301–3320 (2015). doi: [10.5194/bg-12-3301-2015](https://doi.org/10.5194/bg-12-3301-2015)
89. T. L. Frölicher, K. B. Rodgers, C. A. Stock, W. W. L. Cheung, Sources of uncertainties in 21st century projections of potential ocean ecosystem stressors. *Global Biogeochem. Cycles* **30**, 1224–1243 (2016). doi: [10.1002/2015GB005338](https://doi.org/10.1002/2015GB005338)
90. S. A. Henson *et al.*, Rapid emergence of climate change in environmental drivers of marine ecosystems. *Nat. Commun.* **8**, 14682 (2017). doi: [10.1038/ncomms14682](https://doi.org/10.1038/ncomms14682); pmid: [28267144](https://pubmed.ncbi.nlm.nih.gov/28267144/)
91. J. E. Campbell *et al.*, Large historical growth in global terrestrial gross primary production. *Nature* **544**, 84–87 (2017). doi: [10.1038/nature22030](https://doi.org/10.1038/nature22030); pmid: [28382993](https://pubmed.ncbi.nlm.nih.gov/28382993/)
92. D. Lombardozi, G. B. Bonan, D. W. Nychka, The emerging anthropogenic signal in land–atmosphere carbon-cycle coupling. *Nat. Clim. Change* **4**, 796–800 (2014). doi: [10.1038/nclimate2323](https://doi.org/10.1038/nclimate2323)
93. N. S. Lovenduski, G. A. McKinley, A. R. Fay, K. Lindsay, M. C. Long, Partitioning uncertainty in ocean carbon uptake projections: Internal variability, emission scenario, and model structure. *Global Biogeochem. Cycles* **30**, 1276–1287 (2016). doi: [10.1002/2016GB005426](https://doi.org/10.1002/2016GB005426)
94. N. S. Lovenduski, G. B. Bonan, Reducing uncertainty in projections of terrestrial carbon uptake. *Environ. Res. Lett.* **12**, 044020 (2017). doi: [10.1088/1748-9326/aa66b8](https://doi.org/10.1088/1748-9326/aa66b8)
95. Y. Chen, D. C. Morton, N. Andela, L. Giglio, J. T. Randerson, How much global burned area can be forecast on seasonal time scales using sea surface temperatures? *Environ. Res. Lett.* **11**, 045001 (2016). doi: [10.1088/1748-9326/11/4/045001](https://doi.org/10.1088/1748-9326/11/4/045001)
96. D. Tommasi *et al.*, Managing living marine resources in a dynamic environment: The role of seasonal to decadal climate forecasts. *Prog. Oceanogr.* **152**, 15–49 (2017). doi: [10.1016/j.pocean.2016.12.011](https://doi.org/10.1016/j.pocean.2016.12.011)
97. Y. Luo, T. F. Keenan, M. Smith, Predictability of the terrestrial carbon cycle. *Glob. Change Biol.* **21**, 1737–1751 (2015). doi: [10.1111/gcb.12766](https://doi.org/10.1111/gcb.12766); pmid: [25327167](https://pubmed.ncbi.nlm.nih.gov/25327167/)
98. S. M. Polavarapu *et al.*, Greenhouse gas simulations with a coupled meteorological and transport model: The predictability of CO<sub>2</sub>. *Atmos. Chem. Phys.* **16**, 12005–12038 (2016). doi: [10.5194/acp-16-12005-2016](https://doi.org/10.5194/acp-16-12005-2016)
99. E. Hawkins, R. S. Smith, J. M. Gregory, D. A. Stainforth, Irreducible uncertainty in near-term climate projections. *Clim. Dyn.* **46**, 3807–3819 (2016). doi: [10.1007/s00382-015-2806-8](https://doi.org/10.1007/s00382-015-2806-8)
100. S. W. Chisholm, Stirring times in the Southern Ocean. *Nature* **407**, 685–687 (2000). doi: [10.1038/35037696](https://doi.org/10.1038/35037696); pmid: [11048702](https://pubmed.ncbi.nlm.nih.gov/11048702/)
101. R. Buitenwerf, L. Rose, S. I. Higgins, Three decades of multi-dimensional change in global leaf phenology. *Nat. Clim. Change* **5**, 364–368 (2015). doi: [10.1038/nclimate2533](https://doi.org/10.1038/nclimate2533)
102. T. F. Keenan *et al.*, Increase in forest water-use efficiency as atmospheric carbon dioxide concentrations rise. *Nature* **499**, 324–327 (2013). doi: [10.1038/nature12291](https://doi.org/10.1038/nature12291); pmid: [23842499](https://pubmed.ncbi.nlm.nih.gov/23842499/)
103. L. M. Mercado *et al.*, Impact of changes in diffuse radiation on the global land carbon sink. *Nature* **458**, 1014–1017 (2009). doi: [10.1038/nature07949](https://doi.org/10.1038/nature07949); pmid: [19396143](https://pubmed.ncbi.nlm.nih.gov/19396143/)
104. C. D. Allen *et al.*, A global overview of drought and heat-induced tree mortality reveals emerging climate change risks for forests. *For. Ecol. Manage.* **259**, 660–684 (2010). doi: [10.1016/j.foreco.2009.09.001](https://doi.org/10.1016/j.foreco.2009.09.001)
105. D. Frank *et al.*, Effects of climate extremes on the terrestrial carbon cycle: Concepts, processes and potential future impacts. *Glob. Change Biol.* **21**, 2861–2880 (2015). doi: [10.1111/gcb.12916](https://doi.org/10.1111/gcb.12916); pmid: [25752680](https://pubmed.ncbi.nlm.nih.gov/25752680/)
106. W. M. Jolly *et al.*, Climate-induced variations in global wildfire danger from 1979 to 2013. *Nat. Commun.* **6**, 7537 (2015). doi: [10.1038/ncomms8537](https://doi.org/10.1038/ncomms8537); pmid: [26172867](https://pubmed.ncbi.nlm.nih.gov/26172867/)
107. A. Ordóñez, J. W. Williams, J.-C. Svenning, Mapping climatic mechanisms likely to favour the emergence of novel communities. *Nat. Clim. Change* **6**, 1104–1109 (2016). doi: [10.1038/nclimate3127](https://doi.org/10.1038/nclimate3127)
108. H. Haberl *et al.*, Quantifying and mapping the human appropriation of net primary production in Earth's terrestrial ecosystems. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 12942–12947 (2007). doi: [10.1073/pnas.0704243104](https://doi.org/10.1073/pnas.0704243104); pmid: [17616580](https://pubmed.ncbi.nlm.nih.gov/17616580/)
109. C. Laufkötter *et al.*, Drivers and uncertainties of future global marine primary production in marine ecosystem models. *Biogeosciences* **12**, 6955–6984 (2015). doi: [10.5194/bg-12-6955-2015](https://doi.org/10.5194/bg-12-6955-2015)
110. T. P. Hughes *et al.*, Global warming and recurrent mass bleaching of corals. *Nature* **543**, 373–377 (2017). doi: [10.1038/nature21707](https://doi.org/10.1038/nature21707); pmid: [28300113](https://pubmed.ncbi.nlm.nih.gov/28300113/)
111. P. Wassmann, C. M. Duarte, S. Agustí, M. K. Sejr, Footprints of climate change in the Arctic marine ecosystem. *Glob. Change Biol.* **17**, 1235–1249 (2011). doi: [10.1111/j.1365-2486.2010.02311.x](https://doi.org/10.1111/j.1365-2486.2010.02311.x)
112. A. J. Constable *et al.*, Climate change and Southern Ocean ecosystems I: How changes in physical habitats directly affect marine biota. *Glob. Change Biol.* **20**, 3004–3025 (2014). doi: [10.1111/gcb.12623](https://doi.org/10.1111/gcb.12623); pmid: [24802817](https://pubmed.ncbi.nlm.nih.gov/24802817/)
113. T. Wernberg *et al.*, Climate-driven regime shift of a temperate marine ecosystem. *Science* **353**, 169–172 (2016). doi: [10.1126/science.aad8745](https://doi.org/10.1126/science.aad8745); pmid: [27387951](https://pubmed.ncbi.nlm.nih.gov/27387951/)
114. J. P. Gattuso, L. Hansson, Ed., *Ocean Acidification* (Oxford Univ. Press, 2011).

# ACKNOWLEDGMENTS

We acknowledge funding from the National Institute of Food and Agriculture/U.S. Department of Agriculture (2015-67003-23485) and the NASA Ocean Biology and Biogeochemistry Program (NNX14AL86G). We thank C. Tebaldi and J. Kleypas (NCAR) for comments on the manuscript and figures. NCAR is sponsored by the National Science Foundation.

10.1126/science.aam8328

## RESEARCH ARTICLE

## NEAR-FIELD COSMOLOGY

# A whirling plane of satellite galaxies around Centaurus A challenges cold dark matter cosmology

Oliver Müller,<sup>1\*</sup> Marcel S. Pawlowski,<sup>2</sup> Helmut Jerjen,<sup>3</sup> Federico Lelli<sup>4</sup>

The Milky Way and Andromeda galaxies are each surrounded by a thin plane of satellite dwarf galaxies that may be corotating. Cosmological simulations predict that most satellite galaxy systems are close to isotropic with random motions, so those two well-studied systems are often interpreted as rare statistical outliers. We test this assumption using the kinematics of satellite galaxies around the Centaurus A galaxy. Our statistical analysis reveals evidence for corotation in a narrow plane: Of the 16 Centaurus A satellites with kinematic data, 14 follow a coherent velocity pattern aligned with the long axis of their spatial distribution. In standard cosmological simulations, <0.5% of Centaurus A-like systems show such behavior. Corotating satellite systems may be common in the universe, challenging small-scale structure formation in the prevailing cosmological paradigm.

**T**he presence of planes of satellite dwarf galaxies around the Milky Way (1–4) and Andromeda (5–7) galaxies has challenged our understanding of structure formation on galactic and subgalactic scales. Similar structures are rare in galaxy formation simulations based on the standard Lambda cold dark matter (ΛCDM) cosmological model, which predicts close to isotropic distributions and random kinematics for satellite systems (8). The existence of planes of satellite galaxies around these two largest galaxies in the Local Group is difficult to explain within the ΛCDM framework. Some authors have argued that preferential accretion of satellites along filaments may explain such flattened structures (9). Others suggest that the Local Group should be considered a rare exception in an otherwise successful cosmological model (10–12). This interpretation, however, has been challenged by emerging evidence for anisotropic satellite distributions around massive galaxies beyond the Local Group (13, 14).

The cosmic expansion of the Local Void (a vast, empty region of space adjacent to the Local Group) has been suggested as a possible origin for the formation of these planar structures (15). An issue that is mostly ignored in this context is the coherent kinematics of the satellite galaxies, which are likely corotating around their host. This is clear for the Milky Way (16, 17), where accurate proper motions are available for several satellites, but it remains more uncertain for Andromeda

(17) because only velocities projected along the line of sight (LoS) are measurable. Such orderly kinematic motions are extremely rare in high-resolution cosmological *N*-body simulations (18) and statistically should not be observed in typical galaxy groups. It remains unclear whether such planes of satellites are unique to the Local Group or ubiquitous in the nearby universe.

In this Research Article, we study the galaxy group in the constellation Centaurus. The Centaurus Group is the richest assembly of galaxies within a distance of 10 megaparsecs (Mpc) from the Milky Way, the so-called Local Volume (19, 20). It comprises two concentrations: the Cen A subgroup dominated by a radio-active elliptical galaxy Centaurus A (Cen A, NGC 5128) at a distance of 3.8 Mpc, and the M 83 subgroup dominated by a late-type spiral galaxy M 83 (NGC 5236) at a distance of 4.9 Mpc (19, 20). The galaxies, which are gravitationally bound to Cen A, were claimed to be distributed in two parallel planes (21). The discovery of additional satellite galaxies in the group weakened the case for a double-planar structure, and a single-plane interpretation has become more statistically significant (22–24). This plane has a small-scale height with a root-mean-square (RMS) thickness of 69 kiloparsecs (kpc) and a major-axis RMS length of 309 kpc (25). We investigate the kinematics of this planar structure and compare it with galaxy formation simulations in ΛCDM cosmology.

## Dynamics of the Cen A satellite system

From Earth, the satellite plane around Cen A is seen nearly edge-on at an inclination of 14.6° (25). This coincidental geometrical alignment allows us to scrutinize the kinematics of the plane. We use all available heliocentric velocities for the Cen A satellites, taken from the Local Volume catalog (19, 20). The vast majority of satellites have

accurate distances derived from the tip magnitude of the red giant branch (TRGB) method with a typical uncertainty of ~5%. There are 31 confirmed satellites of Cen A with accurate distance measurements. Half of them have measured LoS velocities. One sample galaxy (KKs 59) has a measured velocity but lacks a TRGB distance: We adopt the same distance as for Cen A; excluding this galaxy does not change our results. The adopted data are listed in table S1.

The on-sky distribution of the satellites is plotted in Fig. 1 together with their motions relative to Cen A. Figure 1 also shows the positions and kinematic information for 1239 planetary nebulae (PNe) (26) and the three-dimensional (3D) distribution of the satellites with measured velocities. The mean velocity of the Cen A satellite system ( $555 \text{ km s}^{-1}$ ) is equal to the recession velocity of Cen A ( $556 \pm 10 \text{ km s}^{-1}$ ) within the measurement uncertainties. Hereafter, the recession velocity of Cen A is used as a zero-point reference, and the terms “approaching” and “receding” are intended with respect to this velocity. The dust lane of Cen A serves as a natural dividing line: Its position angle ( $PA = 110^\circ$ ) roughly coincides with the geometrical minor axis of the satellite plane (25). Clearly, approaching and receding satellites tend to lie to the southwest and northeast of the dividing line, respectively, indicating a kinematically coherent structure.

To determine the statistical significance of the kinematic coherence, we compare the velocities of Cen A satellites to a random phase-space distribution. Every galaxy has a 50% chance of approaching or receding along the LoS. The probability of finding at least 14 out of 16 galaxies with coherent velocity movement is 0.42%. Consequently, the observed velocity pattern of the Cen A satellites is statistically different from a random phase-space distribution at the  $2.6\sigma$  confidence level.

Figure 2 shows the heliocentric velocities of the satellite galaxies as a function of their distances to Cen A. The geometrical minor axis of the plane (or equivalently the dust lane) is used to assign a positive or negative sign to the distance between satellite galaxies and Cen A. Figure 2 shows a clear trend: Galaxies to the south of Cen A are approaching, whereas galaxies to the north are receding. This is to be expected if the satellites are rotating around Cen A. Only two satellite galaxies (KK 221 and ESO 269-058) deviate from this trend and may potentially be counter-rotating, analogous to the Sculptor dwarf in the Milky Way halo (27). An inspection of their properties and alignment inside the plane does not reveal any peculiar characteristics (e.g., they are not more massive or luminous than other satellites). The velocity field of the planetary nebulae within Cen A follows a similar trend: Planetary nebulae in the northern and southern hemisphere are (on average) systematically blue and red shifted, respectively.

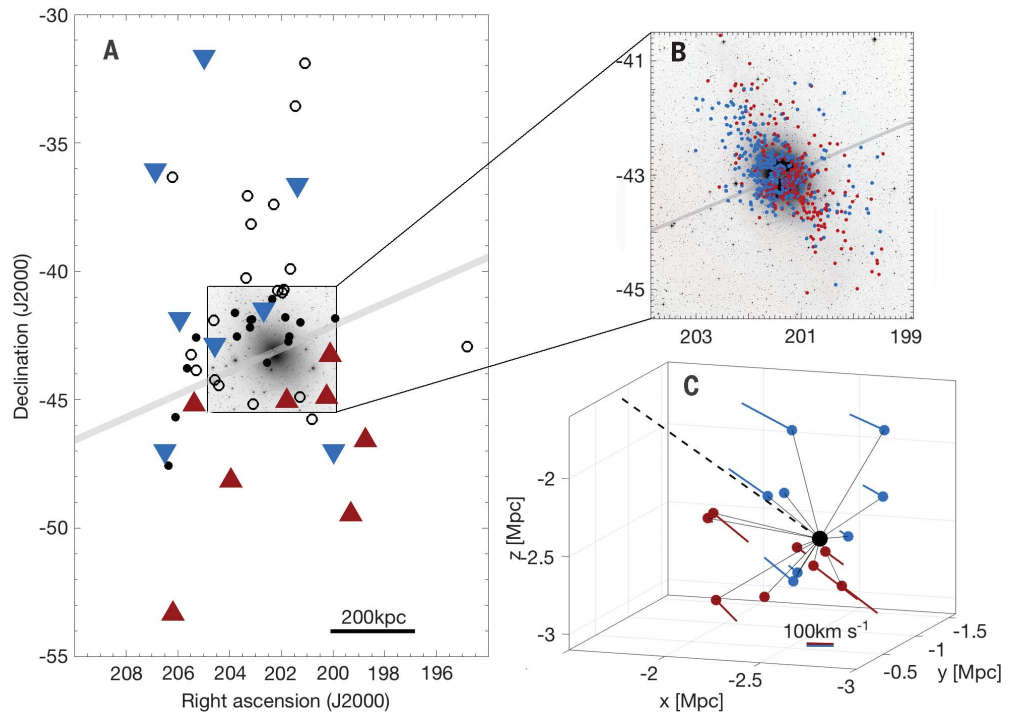
To explore the observed velocity pattern for the satellite galaxies, we ran three statistical tests: namely, Pearson's *R*, Spearman's Rho, and Kendall's Tau. These are standard methods to test correlations between independent variables. Whereas

<sup>1</sup>Departement Physik, Universität Basel, Klingelbergstrasse 82, CH-4056 Basel, Switzerland. <sup>2</sup>Department of Physics and Astronomy, University of California, Irvine, CA 92697, USA. <sup>3</sup>Research School of Astronomy and Astrophysics, Australian National University, Canberra, ACT 2611, Australia. <sup>4</sup>European Southern Observatory (ESO), Karl-Schwarzschild-Strasse 1, 85748 Garching bei München, Germany.

\*Corresponding author. Email: [oliver89.mueller@unibas.ch](mailto:oliver89.mueller@unibas.ch)

**Fig. 1. On-sky and 3D distribution of the satellite system.**

**(A)** The on-sky distribution of the Cen A subgroup. The central image of Cen A has been scaled up by a factor of 5 to illustrate the features of the host galaxy. Blue downwards- and red upwards-pointing triangles show approaching and receding satellite galaxies with respect to Cen A velocity, respectively. Open circles are group member candidates; filled circles are confirmed satellites without velocity measurements. The line that optimally separates the approaching and receding satellites is indicated with the wide gray band; it coincides with the dust lane of Cen A. **(B)** The kinematic distribution of 1239 planetary nebulae (26). Blue PNe are approaching, and red PNe are receding relative to Cen A. **(C)** A 3D representation of the spatial distribution of the Cen A satellite galaxies in equatorial Cartesian coordinates (with the origin at Earth). The length of the colored lines is proportional to the observed velocity; the dashed line is our line of sight toward Cen A.



the Pearson's method tests for a strictly linear correlation, the Spearman's Rho and Kendall's Tau methods test for a general correlation between the variables. The null hypothesis is that velocities and separations are uncorrelated. The velocity pattern is significant within a confidence interval of  $2\sigma$  ( $P$  value  $< 0.03$ ) for the projected separation and  $3\sigma$  ( $P$  value  $< 0.01$ ) for the 3D separation (28). These low  $P$  values lead us to reject the null hypothesis, implying a small chance of finding such a correlation in random, normal distributed data. We further consider how much more likely the hypothesis of correlated data is with respect to the hypothesis of uncorrelated data. We applied a Bayesian correlation test (29) and found that the scenario of coherently moving satellites is 4.5 times more likely using the projected separation and 16.5 times more likely using the full 3D information than uncorrelated satellite movements (28). Projected separations consistently give lower statistical significance than 3D distances because they contain less physical information: This highlights the importance of having TRGB distance measurements for dwarf galaxies in Centaurus.

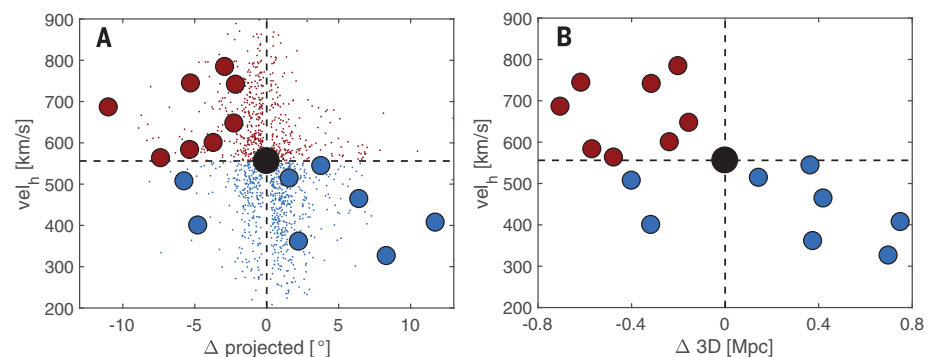
### Implications for galaxy formation

The satellite galaxies in the Cen A subgroup collectively form a coherent kinematical structure. Comparable structures have been discovered in the Milky Way halo, where the majority of the 11 classical satellites share a coherent orbital motion (established with proper motion measurements of individual stars from the satellites) (16), and for the Andromeda galaxy, for which 13 out of 15 satellites follow a coherent LoS velocity trend (7).

Although we find that the kinematics of the Cen A satellites are unlikely to occur by chance, this does not immediately allow us to draw conclusions about its agreement with predictions from  $\Lambda$ CDM cosmology. Satellite galaxy systems in cosmological simulations generally exhibit some degree of phase-space coherence, owing to the accretion of subhalos from preferred directions, along filaments and in groups (9). To judge whether this effect is sufficient to explain the observed coherence in the Cen A satellites, we determined the occurrence of such extreme structures in two cosmological simulations: Millennium II (30) and Illustris (37). Millennium II is a dark-matter-only  $N$ -body simulation that includes gravitational effects such as subhalo accre-

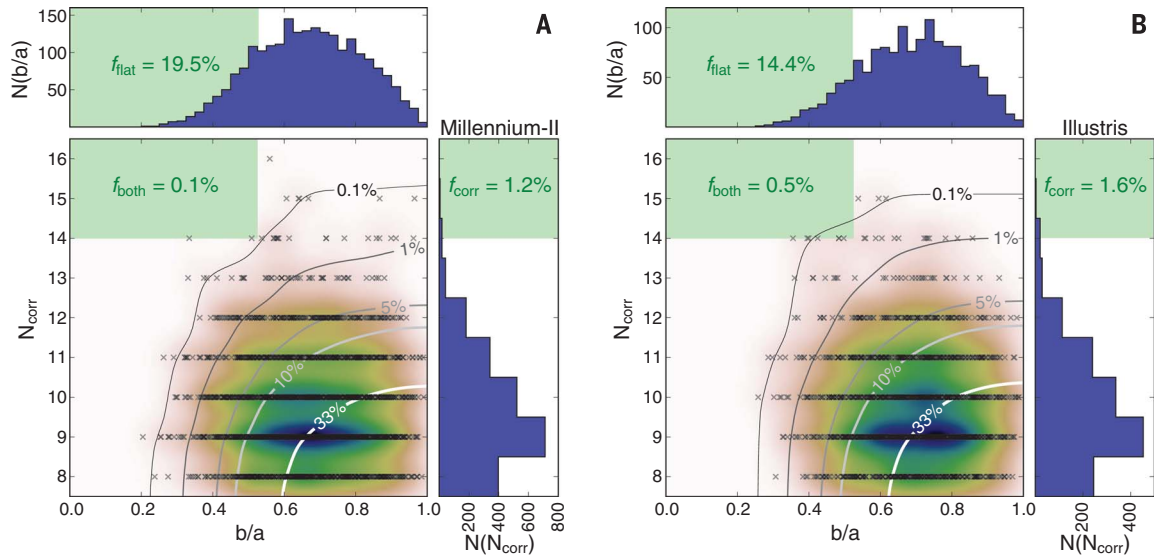
tion from filaments, but neglects baryonic effects such as stellar and black-hole feedback and possible destruction of satellite galaxies due to the enhanced tidal effects from the baryonic disk (32). The relative importance of these effects is highly debated (33–36). Hence, we also analyze the hydrodynamical Illustris simulation (37), which additionally includes gas physics, star formation, and feedback processes.

Our approach is analogous to recent studies of the frequency of the satellite planes around the Milky Way and the Andromeda galaxy (18, 37). We identify Cen A analogs within the simulations by selecting dark matter halos with masses between  $4 \times 10^{12}$  and  $12 \times 10^{12}$  solar masses ( $M_{\odot}$ ) and by rejecting any candidate hosts that have a



**Fig. 2. Velocities and separations to Cen A.** Heliocentric velocities versus angular **(A)** and 3D **(B)** distances  $\Delta$  from Cen A (black sphere), in the north (positive  $\Delta$ ) or south (negative  $\Delta$ ) of the dust lane. Large and small spheres show, respectively, satellite galaxies and planetary nebulae. Blue and red colors indicate, respectively, approaching and receding objects with respect to the Cen A velocity. The angular distances of the PNe are scaled up by a factor of 10.

**Fig. 3. Comparison to  $\Lambda$ CDM cosmological simulations.** The number of kinematically correlated satellites  $N_{\text{corr}}$  and the on-sky axis-ratio flattening  $a$  are plotted for Cen A analogs from the Millennium II (A) and Illustris (B) simulations. The density of simulated systems is indicated by the color map. The contours show the frequency of more extreme realizations, i.e., the number of points to the upper left of each position divided by the total number of realizations.



The top and right histograms show the number of realizations with a given axis ratio  $N(b/a)$  and a given number of correlated velocities  $N(N_{\text{corr}})$ , respectively. The green areas delineate the regions in which systems are as extreme as or more extreme than the observed one:  $f_{\text{flat}}$ ,  $f_{\text{corr}}$ , and  $f_{\text{both}}$

give the frequency of realizations that are, respectively, at least as flattened as the observed system, at least as kinematically correlated, or both. A system must fulfill both conditions simultaneously to reproduce the observed properties of the Cen A satellite system.

companion galaxy with dark matter halo mass  $\geq 1 \times 10^{12} M_{\odot}$  within 1.4-Mpc distance. We require a simulated galaxy-satellite system to fulfill two simplified criteria to be considered similar to the observed system: (i) The projected on-sky axis ratio of the system must be  $b/a \leq 0.52$ , where  $a$  and  $b$  are the semimajor and semiminor axes, respectively; and (ii) the kinematic coherence along the long axis is at least 14 out of 16 satellites. We find that the occurrence of arrangements similar to those of Cen A in the cosmological simulations is 0.1% for Millennium II and 0.5% for Illustris (Fig. 3). These estimates must be considered upper limits, because we do not take into account the full 3D distribution of satellite galaxies. Even though the hydrodynamical Illustris simulation does contain a higher frequency of systems analogous to Cen A than the dark-matter-only Millennium II simulation, they are rare cases in both. The observed Cen A satellite system is thus in serious tension with the expectations from these  $\Lambda$ CDM simulations, to a similar degree as the satellite planes in the Local Group.

Could the coherent motion be the result of cosmic expansion? If that were the case, a correlation between the velocities of the satellites and their distances to the Milky Way would be expected. This is not found for the sample of Cen A satellite galaxies (Fig. 1C). We thus can rule out that the cosmic expansion is responsible for the observed velocity field. Another possible origin of a velocity gradient is a perspective effect. For angular offsets  $\delta$  along the direction of motion, a fraction  $\sin(\delta)$  of a system's bulk tangential velocity is projected along the LoS (38, 39). The velocity gradient found for the Cen A system in Fig. 2 implies a tangential velocity on the order of  $1000 \text{ km s}^{-1}$ , comparable to what would be required for the Andromeda satellite plane (7). This

is unphysically high given that it exceeds the cosmic expansion at the distance of Cen A by a factor of 3.7. Such an interpretation of the velocity gradient would imply that the Cen A group moves in the direction defined by the satellite plane, which is unlikely. Another potential systematic issue is the contribution to the LoS velocity by the motion of the Sun around the Galactic center. However, we found that this contribution is negligible:  $-2$  to  $-4 \text{ km s}^{-1}$ , depending on the sky position, well within the uncertainties of the heliocentric velocity measurements (table S1).

The coherent kinematics of the Cen A satellites is, instead, best explained by corotation within the plane. We explored a toy model with purely circular orbits and tried to deproject the LoS velocities into circular velocities (28). The results were unsatisfactory as many satellites would have unrealistic circular velocities, which randomly vary with the distance from Cen A. This suggests that the satellites must be on elliptical orbits, as expected for collisionless objects. Two galaxies do not follow the general trend: They may be counterrotating, or on highly elliptical orbits, or simply unrelated to the planar structure. Planetary nebulae provide additional evidence (40): They also show coherent motion, although this is less pronounced than for the dwarf satellites: Only 65% of PNe partake in the common motion (Fig. 1). Because the same trend is present in two independent populations of objects with different orbital times, we can expect this correlation to be long-lived and thus indicative of corotation within the planes.

Corotation outside the Local Group has been investigated with satellite galaxy pairs on opposite sides of their hosts (41). The LoS velocities of satellite pairs are preferentially anticorrelated, suggesting a high incidence ( $>50\%$ ) of corotating satellite pairs in the universe (41), although that

result remains controversial (42–44). For the Cen A subgroup, the presence of a plane of satellite galaxies is known independently of velocity information and is established by using multiple group members. This is unlike previous studies, which could not determine whether specific pairs of satellites actually lie in a plane (41).

In alternative frameworks for the formation of dwarf galaxies, corotating planes of satellites could be a consequence of past interactions and mergers between disk galaxies (16). During galaxy mergers, tidal tails form from disk material as a result of angular momentum conservation and can collapse into tidal dwarf galaxies (45–47). Hydrodynamical simulations show that these may survive the interaction and begin orbiting around the central merger remnant as dwarf satellites (48, 49). In the Local Group, a major merger forming the Andromeda galaxy has been proposed as a possible origin of the observed satellite galaxy planes around both the Milky Way (50) and the Andromeda galaxy (51). The recent finding of a correlation between the size of spiral galaxy bulges (thought to form via major mergers) and the number of satellites is in agreement with this picture (52). Even the existence of some counterrotating satellites can be understood in this framework (27).

Here, we have provided evidence for a kinematically coherent plane of satellite galaxies around Cen A, demonstrating that the phenomenon is not restricted to the Milky Way and Andromeda galaxies. The kinematic coherence can be understood if the satellites are corotating within the plane, as seen around the Milky Way. Considering that the likelihood of finding a single kinematically coherent plane is  $\leq 0.5\%$  in cosmological  $\Lambda$ CDM simulations, finding three such systems in the nearby universe seems extremely unlikely.

## REFERENCES AND NOTES

- W. E. Kunkel, S. Demers, *Royal Greenwich Observatory Bull.* **182**, 241 (1976).
- D. Lynden-Bell, *Mon. Not. R. Astron. Soc.* **174**, 695–710 (1976).
- M. S. Pawlowski, J. Pflamm-Altenburg, P. Kroupa, *Mon. Not. R. Astron. Soc.* **423**, 1109–1126 (2012).
- M. S. Pawlowski, P. Kroupa, H. Jerjen, *Mon. Not. R. Astron. Soc.* **435**, 1928–1957 (2013).
- A. W. McConnachie, M. J. Irwin, *Mon. Not. R. Astron. Soc.* **365**, 902–914 (2006).
- A. Koch, E. K. Grebel, *Astron. J.* **131**, 1405–1415 (2006).
- R. A. Ibata *et al.*, *Nature* **493**, 62–65 (2013).
- P. Kroupa, C. Theis, C. M. Boily, *Astron. Astrophys.* **431**, 517–521 (2005).
- A. R. Zentner, A. V. Kravtsov, O. Y. Gnedin, A. A. Klypin, *Astrophys. J.* **629**, 219–232 (2005).
- J. P. Ostriker, P. J. Steinhardt, *Nature* **377**, 600–602 (1995).
- S. D. M. White, J. F. Navarro, A. E. Evrard, C. S. Frenk, *Nature* **366**, 429–433 (1993).
- D. J. Eisenstein *et al.*, *Astrophys. J.* **633**, 560–574 (2005).
- O. Müller, R. Scalera, B. Binggeli, H. Jerjen, *Astron. Astrophys.* **602**, A119 (2017).
- K. Chiboucas, B. A. Jacobs, R. B. Tully, I. D. Karachentsev, *Astron. J.* **146**, 126 (2013).
- N. I. Libeskind *et al.*, *Mon. Not. R. Astron. Soc.* **452**, 1052–1059 (2015).
- M. S. Pawlowski, P. Kroupa, *Mon. Not. R. Astron. Soc.* **435**, 2116–2131 (2013).
- M. S. Pawlowski, S. S. McGaugh, H. Jerjen, *Mon. Not. R. Astron. Soc.* **453**, 1047–1061 (2015).
- M. S. Pawlowski *et al.*, *Mon. Not. R. Astron. Soc.* **442**, 2362–2380 (2014).
- I. D. Karachentsev, V. E. Karachentseva, W. K. Huchtmeier, D. I. Makarov, *Astron. J.* **127**, 2031–2068 (2004).
- I. D. Karachentsev, D. I. Makarov, E. I. Kaisina, *Astron. J.* **145**, 101 (2013).
- R. B. Tully *et al.*, *Astrophys. J.* **802**, L25 (2015).
- O. Müller, H. Jerjen, B. Binggeli, *Astron. Astrophys.* **583**, A79 (2015).
- O. Müller, H. Jerjen, B. Binggeli, *Astron. Astrophys.* **597**, A7 (2017).
- D. Crnojević *et al.*, *Astrophys. J.* **823**, 19 (2016).
- O. Müller, H. Jerjen, M. S. Pawlowski, B. Binggeli, *Astron. Astrophys.* **595**, A119 (2016).
- J. R. Walsh, M. Rejkuba, N. A. Walton, *Astron. Astrophys.* **574**, A109 (2015).
- M. S. Pawlowski, P. Kroupa, K. S. de Boer, *Astron. Astrophys.* **532**, A118 (2011).
- Materials and methods are available as supplementary materials.
- R. Wetzels, E.-J. Wagenmakers, *Psychon. Bull. Rev.* **19**, 1057–1064 (2012).
- M. Boylan-Kolchin, V. Springel, S. D. M. White, A. Jenkins, G. Lemson, *Mon. Not. R. Astron. Soc.* **398**, 1150–1164 (2009).
- M. Vogelsberger *et al.*, *Nature* **509**, 177–182 (2014).
- S. Garrison-Kimmel *et al.*, *Mon. Not. R. Astron. Soc.* **471**, 1709–1727 (2017).
- J. S. Bullock, M. Boylan-Kolchin, *Annu. Rev. Astron. Astrophys.* **55**, 343–387 (2017).
- T. Buck, A. A. Dutton, A. V. Macciò, *Mon. Not. R. Astron. Soc.* **460**, 4348–4365 (2016).
- N. Gillet *et al.*, *Astrophys. J.* **800**, 34 (2015).
- H. Bahl, H. Baumgardt, *Mon. Not. R. Astron. Soc.* **438**, 2916–2923 (2014).
- M. Cautun *et al.*, *Mon. Not. R. Astron. Soc.* **452**, 3838–3852 (2015).
- R. P. van der Marel, P. Guhathakurta, *Astrophys. J.* **678**, 187–199 (2008).
- J.-B. Salomon, R. A. Ibata, B. Famaey, N. F. Martin, G. F. Lewis, *Mon. Not. R. Astron. Soc.* **456**, 4432–4440 (2016).
- E. W. Peng, H. C. Ford, K. C. Freeman, *Astrophys. J.* **602**, 685–704 (2004).
- N. G. Ibata, R. A. Ibata, B. Famaey, G. F. Lewis, *Nature* **511**, 563–566 (2014).
- J. I. Phillips, M. C. Cooper, J. S. Bullock, M. Boylan-Kolchin, *Mon. Not. R. Astron. Soc.* **453**, 3839 (2015).
- M. Cautun, W. Wang, C. S. Frenk, T. Sawala, *Mon. Not. R. Astron. Soc.* **449**, 2576–2587 (2015).
- R. A. Ibata, B. Famaey, G. F. Lewis, N. G. Ibata, N. Martin, *Astrophys. J.* **805**, 67 (2015).
- J. E. Hibbard, J. M. van der Hulst, J. E. Barnes, R. M. Rich, *Astron. J.* **122**, 2969–2992 (2001).
- F. Bournaud *et al.*, *Science* **316**, 1166–1169 (2007).
- F. Lelli *et al.*, *Astron. Astrophys.* **584**, A113 (2015).
- F. Bournaud, P.-A. Duc, *Astron. Astrophys.* **456**, 481–492 (2006).
- S. Ploekinger, S. Recchi, G. Hensler, P. Kroupa, *Mon. Not. R. Astron. Soc.* **447**, 2512–2525 (2015).
- S. Fouquet, F. Hammer, Y. Yang, M. Puech, H. Flores, *Mon. Not. R. Astron. Soc.* **427**, 1769–1783 (2012).
- F. Hammer *et al.*, *Mon. Not. R. Astron. Soc.* **431**, 3543–3549 (2013).
- M. López-Corredoira, P. Kroupa, *Astrophys. J.* **817**, 75 (2016).

## ACKNOWLEDGMENTS

We thank B. Binggeli and M. Rejkuba for interesting discussions and helpful inputs. O.M. thanks E. Schnider for inputs concerning the figures. O.M. is grateful to the Swiss National Science Foundation for financial support. M.S.P. acknowledges support for this work provided by NASA through Hubble Fellowship grant HST-HF2-51379.001-A awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA, under contract NAS5-26555. H.J. acknowledges the support of the Australian Research Council through Discovery Project DP150100862. The work of F.L. is supported by an ESO postdoctoral fellowship. The Millennium-II Simulation databases used in this paper and the web application providing online access to them were constructed as part of the activities of the German Astrophysical Virtual Observatory. The Illustris Simulation databases are provided by the Illustris Collaboration. The observational data we used are given in table S1. The software for our Monte Carlo analysis is provided in data file S1, and the catalogs of Cen A analogs in the simulations are in data files S2 and S3. We thank the anonymous referees for helpful comments that improved the paper.

## SUPPLEMENTARY MATERIALS

[www.sciencemag.org/content/359/6375/534/suppl/DC1](http://www.sciencemag.org/content/359/6375/534/suppl/DC1)  
 Materials and Methods  
 Supplementary Text  
 Figs. S1 and S2  
 Tables S1 and S2  
 References (53–70)  
 Movie S1  
 Data Files S1 to S3  
 24 June 2017; accepted 20 December 2017  
 10.1126/science.aal1858

Cite as: X. Zhan *et al.*, *Science*  
10.1126/science.aar6401 (2018).

# Structure of a human catalytic step I spliceosome

Xiechao Zhan,<sup>1\*</sup> Chuangye Yan,<sup>1\*†</sup> Xiaofeng Zhang,<sup>1\*</sup> Jianlin Lei,<sup>1,2</sup> Yigong Shi<sup>1,3†</sup>

<sup>1</sup>Beijing Advanced Innovation Center for Structural Biology, Tsinghua-Peking Joint Center for Life Sciences, School of Life Sciences and School of Medicine, Tsinghua University, Beijing 100084, China. <sup>2</sup>Technology Center for Protein Sciences, Ministry of Education Key Laboratory of Protein Sciences, School of Life Sciences, Tsinghua University, Beijing 100084, China. <sup>3</sup>Institute of Biology, Westlake Institute for Advanced Study, Westlake University, Shilongshan Road No. 18, Xihu District, Hangzhou 310064, Zhejiang Province, China.

\*These authors contributed equally to this work.

†Corresponding author. Email: yancy05@mails.tsinghua.edu.cn (C.Y.); shi-lab@tsinghua.edu.cn (Y.S.)

Splicing by the spliceosome involves branching and exon ligation. The branching reaction leads to the formation of the catalytic step I spliceosome (C complex). Here we report the cryo-EM structure of the human C complex at an average resolution of 4.1 Å. Compared to the structure of the *S. cerevisiae* C complex, the human complex contains 11 additional proteins. The step I splicing factors CCDC49 and CCDC94 (Cwc25 and Yju2 in *S. cerevisiae*, respectively) closely interact with the DEAH-family ATPase/helicase Prp16 and bridge the gap between Prp16 and the active site RNA elements. These features, together with structural comparison between the human C and C\* complexes, reveal mechanistic insights into ribonucleoprotein remodeling and allow proposition of a working mechanism for the C-to-C\* transition.

Each cycle of pre-mRNA splicing, executed by the spliceosome, proceeds in two sequential steps of transesterification: branching and exon ligation (1, 2). The products of the branching reaction are an intron lariat-3'-exon intermediate and a 5'-exon, both of which remain bound to the catalytic step I spliceosome (C complex). The ribonucleoprotein remodeling of the C-to-C\* transition is driven by the RNA-dependent DEAH-family ATPase/helicase Prp16 (3), which is thought to bind and pull the single-stranded RNA sequences in a 3'-to-5' direction (4).

The first atomic model of an intact spliceosome was generated in 2015 from the cryo-electron microscopy (cryo-EM) structure of the *S. pombe* intron lariat spliceosome (ILS) at 3.6 Å resolution (5, 6). Since 2016, a burst of structural advances has allowed atomic visualization of the *S. cerevisiae* and human spliceosomes (7–20). Here we report the cryo-EM structure of the human C complex at an average resolution of 4.1 Å. Compared to the yeast C complex (10, 11), the structure of the human C complex contains 11 additional proteins including 4 peptidyl prolyl isomerases (PPIs), the exon junction complex (EJC), U5-40K, RBM22, and Aquarius. The structural features and comparison with the human C\* complex (19, 20) reveal mechanistic insights into the transition from the C to the C\* complex.

## Cryo-EM analysis

The human spliceosomes were assembled using HeLa S3 nuclear extract and a synthetic pre-mRNA. Following a pilot experiment (fig. S1A), we chose to use the small molecule BN82685, known to inhibit exon ligation (21), at a final concentration of 250 μM in the splicing assay. The spliceosomes

were purified and chemically crosslinked. The spliceosome-containing fractions were pooled on the basis of RNA analysis (fig. S1B), and the sample was examined by EM under negative staining (fig. S1C) and cryogenic condition (fig. S1D). Micrographs were collected on a Titan Krios electron microscope, yielding 1,464,033 auto-picked particles.

A subset of 157,388 particles was used to generate the initial references for the human spliceosomal complexes (fig. S2), which were applied to the entire dataset. Following three-dimensional classifications, 95,064 particles gave a reconstruction of the human C complex at 4.5 Å resolution (fig. S3). Further selection of 53,633 particles yielded a reconstruction at an average resolution of 4.1 Å (figs. S3 to S7 and tables S1 and S2). The local resolutions vary greatly in the human C complex (fig. S4), reaching 3.8–4.2 Å in the core but considerably lower in the peripheral regions (table S2). A soft mask on the Brr2 region improved the local EM map to an average resolution of 6.5 Å (figs. S3 and S7 and table S2).

## Overall structure

The refined model of the human C complex contains 15,479 amino acids from 47 proteins and 414 nucleotides from three snRNAs (U2, U5, and U6) and the pre-mRNA (Fig. 1A and tables S1 and S2), with a combined molecular mass of about 1.8 mega-Daltons. Approximately two-thirds of the 15,479 amino acids have side chains and the rest were built as Ala. The poly-Ala sequences are mostly assigned to proteins in the peripheral regions, including Prp16 and a few components (Prp19 tetramer, Spf27, Syf1 and part of Cdc5) of the NineTeen complex (NTC, also known as the Prp19-CDC5L complex). The 47 proteins in the atomic model include 11 from U5 snRNP, 9

from U2 snRNP, 7 from the NTC, 6 from the NTC-related (NTR), 4 from the EJC, 5 splicing factors (SRm300, Cwc22, CCDC49, CCDC94, and Prp17), 4 PPIs (tentatively assigned as PPIL1, CypE, PPIG, and PPWD1), and Prp16 (Fig. 1A).

Compared to the yeast C complex (10, 11), the human C complex adopts a similarly extended, highly asymmetric organization but is considerably larger, with 11 additional protein components that are scattered in discrete regions (Fig. 1B). These 11 proteins include the EJC (eIF4AIII, MAGOH, Y14, and MLN51), the U5 snRNP component U5-40K, the NTR proteins Aquarius and RBM22, and 4 PPIs. In contrast to the yeast C complex structure (10), Prp16 and Brr2 are unambiguously identified in the human C complex. The human protein RBM22 bears significant sequence homology to the yeast NTR proteins Cwc2 and Ecm2 (22), likely reflecting convergent evolution. Other than Cwc2 and Ecm2, all other proteins in the structure of the yeast C complex have corresponding functional orthologs in the human C complex (Fig. 1B).

Previously unknown structural features have also been observed for a number of the human proteins that have functional orthologs in *S. cerevisiae*. Most notably, compared to their *S. cerevisiae* counterparts Cwc25 and Yju2, a number of previously unrecognized structural features of CCDC49 and CCDC94 have been identified. These structural elements play a crucial role in connecting the active site RNA elements to Prp16. Because the human and yeast C complexes share nearly identical structural features in the core of the spliceosome, we focus our discussion on the new structural elements and the human-unique proteins.

### The RNA map and the active site

The RNA sequences in the core regions of the human C complex display clear features in the EM density map (fig. S5). Of the 414 modeled nucleotides, 174, 84, and 97 are assigned to U2, U5, and U6 snRNA, respectively (table S2). The remaining 59 nucleotides belong to pre-mRNA, with 13 nucleotides in the 5'-exon and 46 in the intron lariat-3'-exon intermediate. The overall structure of the three snRNAs, the 5'-exon, and the intron lariat-3'-exon intermediate in the human C complex is quite similar to that of the yeast C complex (10, 11) (Fig. 2, A and B).

The active site of the human C complex comprises the intramolecular stem loop (ISL) of U6 snRNA, the catalytic triplex between U2 and U6, the loop I of U5 snRNA, and five metal ions. Three of the five metals, presumably  $Mg^{2+}$ , appear to stabilize the ISL by neutralizing the negative charges of the RNA backbone phosphates (Fig. 2B, left panel). All three metals are identically coordinated in the yeast C complex (10, 11) (Fig. 2B, right panel).

Two catalytic metal ions known as M1 and M2 are thought to play indispensable but reciprocal roles in the branching reaction and exon ligation (23, 24). Compared to the yeast C complex (10), the position of M2 is shifted by approximately

3 Å in the human C complex (Fig. 2B, right panel). Coordination of M1 in the human C complex is different from that in the yeast complex where M1 is specifically bound by the 3'-OH at the 3'-end of the 5'-exon (10, 11). In the human C complex, the ribose at the 3'-end of the 5'-exon is flipped by approximately 10 Å away from M1 (Fig. 2C). Consequently, M1 is only coordinated by three ligands: the phosphate of the nucleophile-containing adenine nucleotide in the branch point sequence (BPS) and the phosphates of G72 and U74 of U6 snRNA (Fig. 2B).

### The step I splicing factors and Prp16

Similar to the yeast C complex (10, 11), the conformations of the active site RNA elements in the human C complex are sustained by 15 surrounding protein components, particularly CCDC49 (Cwc25 in *S. cerevisiae*) and CCDC94 (Yju2 in *S. cerevisiae*), the NTC components Isy1 and the N-terminal fragment of Cdc5, and the RNaseH-like domain of Prp8 (Fig. 3, A and B, and fig. S6). The ATPase/helicases Brr2 and Prp16 in the peripheral regions are connected to the core of the spliceosome mainly through CCDC49 and CCDC94.

The structurally identified fragment of CCDC49 comprises an extended  $\alpha$ -helix (N-helix, residues 17-61), which is capped by an N-terminal plug (residues 2-16) and a C-terminal hook (residues 62-80) (Fig. 3C). Three Trp residues (Trp12, Trp24, and Trp72) are invariant among the yeast orthologs. The N-terminal plug is placed into the active site (Fig. 3, B and D), with a conserved Gly-rich motif (Gly2-Gly3-Gly4 in CCDC49) penetrating into a small crevice generated by the U2/BPS duplex and the helix Ia of the U2/U6 duplex (Fig. 3E and fig. S6). Residues 5-16 of the N-terminal plug interact with the major groove of the U2/BPS duplex and the zinc-binding domain of CCDC94. Trp12 stacks against Phe76 of CCDC94. The N- and C-terminal ends of the N-helix interact with the 1585-loop of Prp8 and Prp16, respectively (Fig. 3, D and F). Trp24 directly contacts the tip of the 1585-loop. The hook associates with the RNaseH-like domain (Fig. 3G). Notably, the interactions of CCDC49 with Prp16 and the RNaseH-like domain are unknown in the *S. cerevisiae* C complex due to the lack of the C-terminal hook and half of the N-helix in the structurally identified portion of Cwc25 (10).

Although both CCDC94 and CCDC130 share sequence homology with the *S. cerevisiae* Yju2 (Fig. 4A), the fine features of the EM density map are only consistent with CCDC94, not CCDC130. The structure of CCDC94 consists of an extended N terminus, a Cys4-type zinc-binding motif, and three  $\alpha$ -helices (H1/H2/H3) at the C terminus (Fig. 4B). Similar to CCDC49, the N terminus is inserted deeply into the active site, interacting with the 5'-exon, the U2/BPS duplex, Isy1, and Prp8 (Fig. 4C and fig. S6). The zinc-binding motif contacts the ISL, the U2/BPS duplex, the N-helix of CCDC49, and the N-terminal portion of Cdc5 (Fig. 4D and fig. S6). The  $\alpha$ -helices H1/H2/H3 interact with Prp16, Prp8, the N-terminal

portion of Cdc5, and the C-terminal HAT repeats of Syf1 (Fig. 4E). Notably, the interactions involving the CCDC94 helices H1/H2/H3 are previously uncharacterized in the *S. cerevisiae* C complex because these helices remain structurally unidentified in Yju2 (10, 11).

The elongated helices of CCDC49 and CCDC94 directly associate with Prp16 (Fig. 4F). Specifically, the N-helix of CCDC49 binds the RecA2 domain of Prp16, whereas the helices H2/H3 of CCDC94 interact with the OB and Ratchet-like domains. Prp16, in turn, interacts with the Jab1/MPN domain of Prp8 and the N-terminal Plug (residues 107-179) and the PWI-like domain (residues 1030-1181) of Brr2 (fig. S7A). The Jab1/MPN domain also stably associates with Brr2 (25). The RNA-binding site of Prp16 remains unoccupied (fig. S7, B and C), consistent with the structural comparison between Prp16 of the human C complex and Prp43 bound to RNA (26) (fig. S7D). Prp16 is thought to remodel the C complex by pulling the 3'-end sequences of the intron lariat-3'-exon intermediate (3, 4, 27, 28). In the human C complex, the putative RNA-binding site of Prp16 is located approximately 60 Å away from the last ordered nucleotide at the 3'-end of the intron lariat-3'-exon intermediate (Fig. 4F). 9-10 RNA nucleotides are minimally required to cover the distance, consistent with the observation that shortening of the 3'-exon led to the inhibition of exon ligation (4, 29).

### Peptidyl prolyl isomerases (PPIs)

PPI is absent in the structure of the *S. cerevisiae* C complex (10, 11). Eight PPIs are known to be present in the human spliceosome (30) (fig. S8 and table S3), of which four have been identified in the structure of the human C complex. CypE contains an RNA recognition motif (RRM, residues 6-99) and a canonical PPI domain (residues 108-266) (fig. S8). In the human C complex, both the RRM and the PPI of CypE associate with the intron lariat (Fig. 5A). The RRM also binds the U2 snRNP component U2-A' (Lea1 in yeast) and the superhelical protein Syf1; the PPI interacts with U2-A', Isy1, and the heptameric U2 Sm ring. The C-terminal sequences of SmB contain a conserved motif GLXG**P**VRGVGG**P** (Fig. 5A), which is spatially close to the PPI domain of CypE and may serve as a substrate.

PPIL1 (Cyp1 in *S. pombe*) closely interacts with the extended sequences of both SKIP (Prp45 in yeast) (31, 32) and the step II factor Prp17 (Fig. 5B). PPIL1 also binds both the N- and C-domains of RBM22. The PPIL1-bound sequences of Prp17 contain the motif NPX<sub>7</sub>PXXG**P**, which likely represents the substrate. The remaining two PPIs associate with the RNaseH-like domain or the RT Finger/Palm domain of Prp8. One PPI is located close to the C-terminal sequences of CCDC49, which harbor the sequence motif FAPX<sub>15</sub>D**P** but lack a conserved Gly residue that is usually required for substrate specificity (33, 34) (Fig. 5C). The other PPI may catalyze proline isomerization of Ad-002 (Cwc15 in yeast) (Fig. 5D).

These two PPIs may represent PPWD1 and PPIG, although the current EM density map does not allow conclusive assignment.

All four PPIs bind the extended sequence elements of the protein components in the human spliceosome. The switch of the proline cis-trans configuration usually results in alteration of the local structure, which is known to trigger signaling and impact on a range of important cellular processes (34, 35). These PPIs are thought to join spliceosomal complexes at distinct stages of the splicing reaction to regulate pre-mRNA splicing (30, 36). The structural changes caused by proline isomerization likely lead to alteration of the interactions between the substrate protein and surrounding components of the spliceosome.

### Structural comparison with the human C\* complex

Structural resolution of the human C complex allows comparison with the human C\* complex (19, 20) and visualization of the compositional and conformational changes. These changes are concentrated in two aspects. First, the protein components at the peripheral regions on one side of the spliceosome have undergone major reshuffling, exemplified by the large-scale movement of the U2 snRNP (Fig. 6A, left panels). The overall translocation of the U2 snRNP exceeds 100 Å (fig. S9 and movies S1 and S2). The RRM domain of CypE binds the intron sequences upstream of the BPS in the human C complex but instead interacts with the helix Iia/b of U2 snRNA in the human C\* complex (19) (fig. S9). In contrast, the RRM domain binds Syf1 identically in both the C and C\* complexes. Such changes may be facilitated by proline isomerization in the U2 snRNP component SmB. During the C-to-C\* transition, the protein PRKRIP1 is recruited into the C\* complex to stabilize the new position of the U2 snRNP. Brr2 is translocated by about 150 Å in the C-to-C\* transition (Fig. 6A, middle panels, and fig. S10). Prp16 is dissociated in the transition and Prp22 is recruited into the C\* complex (19). Second, the protein components at or near the active site exhibit major changes. During the transition, CCDC49 and CCDC94, along with Isy1, are dissociated (Fig. 6A, right panels). The step II splicing factor Slu7 is recruited into the C\* complex; Prp17 and the RNaseH-like domain of Prp8 undergo major positional changes (movie S3). During the C-to-C\* transition, the WD40 domain of Prp17 is translocated by approximately 70 Å, the RNaseH-like domain of Prp8 is rotated by about 70 degrees, and the β-finger is inserted into the gap between the U6/5'SS duplex and the U2/BPS duplex (fig. S10).

### Discussion

Our structural analysis reveals a working model for the C-to-C\* transition (Fig. 6B). In the C complex, the active site RNA elements are stabilized by CCDC49 and CCDC94 together with Isy1, Prp17, and the RNaseH-like domain of Prp8. The ATPase/helicase Prp16 is framed adjacent to the active site

mainly through interactions with the two step I factors, and Brr2 binds to the other side of Prp16 from the periphery of the spliceosome. Given their direct interactions with the U2/BPS at the active site, CCDC49 and CCDC94 may be immediately dissociated by the action of Prp16 (27). This is likely followed by the dissociation of Prp16 because it is bound to the spliceosome mainly through the two step I factors. These changes allow the recruitment of Slu7, PRKRIP1, and the ATPase/helicase Prp22, ultimately forming the C\* complex.

The step I factors CCDC49 and CCDC94 stabilize the active site RNA conformation to prime the branching reaction. This structural finding is fully consistent with available biochemical and genetic observations (37–39). The structure of the human C complex also reveals notable features of CCDC49 and CCDC94 that were previously unknown in the *S. cerevisiae* C complex. These structural features – elongated N-helix and the hook in CCDC49 and three  $\alpha$ -helices H1/H2/H3 – directly interact with the ATPase/helicase Prp16 and the central protein Prp8 (Figs. 3 and 4), linking Prp16 to the active site RNA elements.

Structural analysis suggests a key role for the 1585-loop in the upload of the 3'-exon into the active site (fig. S11). In the human C-to-C\* transition, the location of the 1585-loop flips around the sequence element at the 3'-side of the lariat junction. In the yeast P complex (14–16), the 1585-loop is located in a position that resembles that in the human C\* complex (19) but is sandwiched by the 3'SS of the intron lariat and the ligated exon (fig. S11). During these processes, the conformation of the 1585-loop undergoes drastic changes.

In this study, the splicing inhibitor BN82685 was used to enrich the human C complex. The EM density map in the core of the spliceosome does not support the presence of the inhibitor. Although direct binding of BN82685 cannot be ruled out, a more likely scenario is that the inhibitor regulates the kinase or phosphatase activity that is essential for pre-mRNA splicing (21). Structure of the human C complex contains 11 additional proteins compared to that of the *S. cerevisiae* C complex (10, 11). The majority of these additional proteins in the human spliceosome play a regulatory role and are a testimony of the more sophisticated nature of pre-mRNA splicing in mammals.

## REFERENCES AND NOTES

- M. C. Wahl, C. L. Will, R. Lührmann, The spliceosome: Design principles of a dynamic RNP machine. *Cell* **136**, 701–718 (2009). [doi:10.1016/j.cell.2009.02.009](https://doi.org/10.1016/j.cell.2009.02.009) [Medline](#)
- Y. Shi, Mechanistic insights into precursor messenger RNA splicing by the spliceosome. *Nat. Rev. Mol. Cell Biol.* **18**, 655–670 (2017). [doi:10.1038/nrm.2017.86](https://doi.org/10.1038/nrm.2017.86) [Medline](#)
- B. Schwer, C. Guthrie, A conformational rearrangement in the spliceosome is dependent on PRP16 and ATP hydrolysis. *EMBO J.* **11**, 5033–5039 (1992). [Medline](#)
- D. R. Semlow, M. R. Blanco, N. G. Walter, J. P. Staley, Spliceosomal DEAH-box ATPases remodel pre-mRNA to activate alternative splice sites. *Cell* **164**, 985–998 (2016). [doi:10.1016/j.cell.2016.01.025](https://doi.org/10.1016/j.cell.2016.01.025) [Medline](#)
- C. Yan, J. Hang, R. Wan, M. Huang, C. C. L. Wong, Y. Shi, Structure of a yeast spliceosome at 3.6-angstrom resolution. *Science* **349**, 1182–1191 (2015). [doi:10.1126/science.aac7629](https://doi.org/10.1126/science.aac7629) [Medline](#)
- J. Hang, R. Wan, C. Yan, Y. Shi, Structural basis of pre-mRNA splicing. *Science* **349**, 1191–1198 (2015). [doi:10.1126/science.aac8159](https://doi.org/10.1126/science.aac8159) [Medline](#)
- C. Plaschka, P. C. Lin, K. Nagai, Structure of a pre-catalytic spliceosome. *Nature* **546**, 617–621 (2017). [Medline](#)
- C. Yan, R. Wan, R. Bai, G. Huang, Y. Shi, Structure of a yeast activated spliceosome at 3.5 Å resolution. *Science* **353**, 904–911 (2016). [doi:10.1126/science.aag0291](https://doi.org/10.1126/science.aag0291) [Medline](#)
- R. Rauhut, P. Fabrizio, O. Dybkov, K. Hartmuth, V. Pena, A. Chari, V. Kumar, C.-T. Lee, H. Urlaub, B. Kastner, H. Stark, R. Lührmann, Molecular architecture of the *Saccharomyces cerevisiae* activated spliceosome. *Science* **353**, 1399–1405 (2016). [doi:10.1126/science.aag1906](https://doi.org/10.1126/science.aag1906) [Medline](#)
- R. Wan, C. Yan, R. Bai, G. Huang, Y. Shi, Structure of a yeast catalytic step I spliceosome at 3.4 Å resolution. *Science* **353**, 895–904 (2016). [doi:10.1126/science.aag2235](https://doi.org/10.1126/science.aag2235) [Medline](#)
- W. P. Galej, M. E. Wilkinson, S. M. Fica, C. Oubridge, A. J. Newman, K. Nagai, Cryo-EM structure of the spliceosome immediately after branching. *Nature* **537**, 197–201 (2016). [doi:10.1038/nature19316](https://doi.org/10.1038/nature19316) [Medline](#)
- C. Yan, R. Wan, R. Bai, G. Huang, Y. Shi, Structure of a yeast step II catalytically activated spliceosome. *Science* **355**, 149–155 (2017). [doi:10.1126/science.aag9979](https://doi.org/10.1126/science.aag9979) [Medline](#)
- S. M. Fica, C. Oubridge, W. P. Galej, M. E. Wilkinson, X.-C. Bai, A. J. Newman, K. Nagai, Structure of a spliceosome remodelled for exon ligation. *Nature* **542**, 377–380 (2017). [doi:10.1038/nature21078](https://doi.org/10.1038/nature21078) [Medline](#)
- R. Bai, C. Yan, R. Wan, J. Lei, Y. Shi, Structure of the post-catalytic spliceosome from *Saccharomyces cerevisiae*. *Cell* **171**, 1589–1598.e8 (2017). [doi:10.1016/j.cell.2017.10.038](https://doi.org/10.1016/j.cell.2017.10.038) [Medline](#)
- S. Liu, X. Li, L. Zhang, J. Jiang, R. C. Hill, Y. Cui, K. C. Hansen, Z. H. Zhou, R. Zhao, Structure of the yeast spliceosomal postcatalytic P complex. *Science* **358**, 1278–1283 (2017). [doi:10.1126/science.aar3462](https://doi.org/10.1126/science.aar3462) [Medline](#)
- M. E. Wilkinson, S. M. Fica, W. P. Galej, C. M. Norman, A. J. Newman, K. Nagai, Postcatalytic spliceosome structure reveals mechanism of 3'-splice site selection. *Science* **358**, 1283–1288 (2017). [doi:10.1126/science.aar3729](https://doi.org/10.1126/science.aar3729) [Medline](#)
- R. Wan, C. Yan, R. Bai, J. Lei, Y. Shi, Structure of an intron lariat spliceosome from *Saccharomyces cerevisiae*. *Cell* **171**, 120–132.e12 (2017). [doi:10.1016/j.cell.2017.08.029](https://doi.org/10.1016/j.cell.2017.08.029) [Medline](#)
- K. Bertram, D. E. Agafonov, O. Dybkov, D. Haselbach, M. N. Leelaram, C. L. Will, H. Urlaub, B. Kastner, R. Lührmann, H. Stark, Cryo-EM structure of a pre-catalytic human spliceosome primed for activation. *Cell* **170**, 701–713.e11 (2017). [doi:10.1016/j.cell.2017.07.011](https://doi.org/10.1016/j.cell.2017.07.011) [Medline](#)
- X. Zhang, C. Yan, J. Hang, L. I. Finci, J. Lei, Y. Shi, An atomic structure of the human spliceosome. *Cell* **169**, 918–929.e14 (2017). [doi:10.1016/j.cell.2017.04.033](https://doi.org/10.1016/j.cell.2017.04.033) [Medline](#)
- K. Bertram, D. E. Agafonov, W.-T. Liu, O. Dybkov, C. L. Will, K. Hartmuth, H. Urlaub, B. Kastner, H. Stark, R. Lührmann, Cryo-EM structure of a human spliceosome activated for step 2 of splicing. *Nature* **542**, 318–323 (2017). [doi:10.1038/nature21079](https://doi.org/10.1038/nature21079) [Medline](#)
- M. G. Berg, L. Wan, I. Younis, M. D. Diem, M. Soo, C. Wang, G. Dreyfuss, A quantitative high-throughput in vitro splicing assay identifies inhibitors of spliceosome catalysis. *Mol. Cell. Biol.* **32**, 1271–1283 (2012). [doi:10.1128/MCB.05788-11](https://doi.org/10.1128/MCB.05788-11) [Medline](#)
- N. Rasche, O. Dybkov, J. Schmitzová, B. Akyildiz, P. Fabrizio, R. Lührmann, Cwc2 and its human homologue RBM22 promote an active conformation of the spliceosome catalytic centre. *EMBO J.* **31**, 1591–1604 (2012). [doi:10.1038/emboj.2011.502](https://doi.org/10.1038/emboj.2011.502) [Medline](#)
- T. A. Steitz, J. A. Steitz, A general two-metal-ion mechanism for catalytic RNA. *Proc. Natl. Acad. Sci. U.S.A.* **90**, 6498–6502 (1993). [doi:10.1073/pnas.90.14.6498](https://doi.org/10.1073/pnas.90.14.6498) [Medline](#)
- K. S. Keating, N. Toor, P. S. Perlman, A. M. Pyle, A structural analysis of the group II intron active site and implications for the spliceosome. *RNA* **16**, 1–9 (2010). [doi:10.1261/rna.1791310](https://doi.org/10.1261/rna.1791310) [Medline](#)
- T. H. Nguyen, J. Li, W. P. Galej, H. Oshikane, A. J. Newman, K. Nagai, Structural basis of Brr2-Prp8 interactions and implications for U5 snRNP biogenesis and the spliceosome active site. *Structure* **21**, 910–919 (2013). [doi:10.1016/j.str.2013.04.017](https://doi.org/10.1016/j.str.2013.04.017) [Medline](#)
- M. J. Tauchert, J. B. Fourmann, R. Lührmann, R. Ficner, Structural insights into the mechanism of the DEAH-box RNA helicase Prp43. *eLife* **6**, e21510 (2017).

- [doi:10.7554/eLife.21510](https://doi.org/10.7554/eLife.21510) Medline
27. C. K. Tseng, H. L. Liu, S. C. Cheng, DEAH-box ATPase Prp16 has dual roles in remodeling of the spliceosome in catalytic steps. *RNA* **17**, 145–154 (2011). [doi:10.1261/rna.2459611](https://doi.org/10.1261/rna.2459611) Medline
  28. T. Ohrt, P. Odenwälder, J. Dannenberg, M. Prior, Z. Warkocki, J. Schmitzová, R. Karaduman, I. Gregor, J. Enderlein, P. Fabrizio, R. Lührmann, Molecular dissection of step 2 catalysis of yeast pre-mRNA splicing investigated in a purified system. *RNA* **19**, 902–915 (2013). [doi:10.1261/rna.039024.113](https://doi.org/10.1261/rna.039024.113) Medline
  29. S. C. Cheng, Formation of the yeast splicing complex A1 and association of the splicing factor PRP19 with the pre-mRNA are independent of the 3' region of the intron. *Nucleic Acids Res.* **22**, 1548–1554 (1994). [doi:10.1093/nar/22.9.1548](https://doi.org/10.1093/nar/22.9.1548) Medline
  30. B. M. Adams, M. N. Coates, S. R. Jackson, M. S. Jurica, T. L. Davis, Nuclear cyclophilins affect spliceosome assembly and function in vitro. *Biochem. J.* **469**, 223–233 (2015). [doi:10.1042/BJ20150396](https://doi.org/10.1042/BJ20150396) Medline
  31. C. M. Stegmann, R. Lührmann, M. C. Wahl, The crystal structure of PPI1 bound to cyclosporine A suggests a binding mode for a linear epitope of the SKIP protein. *PLOS ONE* **5**, e10013 (2010). [doi:10.1371/journal.pone.0010013](https://doi.org/10.1371/journal.pone.0010013) Medline
  32. C. Xu, J. Zhang, X. Huang, J. Sun, Y. Xu, Y. Tang, J. Wu, Y. Shi, Q. Huang, Q. Zhang, Solution structure of human peptidyl prolyl isomerase-like protein 1 and insights into its interaction with SKIP. *J. Biol. Chem.* **281**, 15900–15908 (2006). [doi:10.1074/jbc.M51155200](https://doi.org/10.1074/jbc.M51155200) Medline
  33. R. K. Harrison, R. L. Stein, Substrate specificities of the peptidyl prolyl cis-trans isomerase activities of cyclophilin and FK-506 binding protein: Evidence for the existence of a family of distinct enzymes. *Biochemistry* **29**, 3813–3816 (1990). [doi:10.1021/bi00468a001](https://doi.org/10.1021/bi00468a001) Medline
  34. T. L. Davis, J. R. Walker, V. Campagna-Slater, P. J. Finerty, R. Paramanathan, G. Bernstein, F. MacKenzie, W. Tempel, H. Ouyang, W. H. Lee, E. Z. Eisenmesser, S. Dhe-Paganon, Structural and biochemical characterization of the human cyclophilin family of peptidyl-prolyl isomerases. *PLOS Biol.* **8**, e1000439 (2010). [doi:10.1371/journal.pbio.1000439](https://doi.org/10.1371/journal.pbio.1000439) Medline
  35. P. Wang, J. Heitman, The cyclophilins. *Genome Biol.* **6**, 226 (2005). [doi:10.1186/gb-2005-6-7-226](https://doi.org/10.1186/gb-2005-6-7-226) Medline
  36. D. S. Horowitz, E. J. Lee, S. A. Mabon, T. Misteli, A cyclophilin functions in pre-mRNA splicing. *EMBO J.* **21**, 470–480 (2002). [doi:10.1093/emboj/21.3.470](https://doi.org/10.1093/emboj/21.3.470) Medline
  37. Y. F. Chiu, Y.-C. Liu, T.-W. Chiang, T.-C. Yeh, C.-K. Tseng, N.-Y. Wu, S.-C. Cheng, Cwc25 is a novel splicing factor required after Prp2 and Yju2 to facilitate the first catalytic reaction. *Mol. Cell. Biol.* **29**, 5671–5678 (2009). [doi:10.1128/MCB.00773-09](https://doi.org/10.1128/MCB.00773-09) Medline
  38. Y. C. Liu, H. C. Chen, N. Y. Wu, S. C. Cheng, A novel splicing factor, Yju2, is associated with NTC and acts after Prp2 in promoting the first catalytic reaction of pre-mRNA splicing. *Mol. Cell. Biol.* **27**, 5403–5413 (2007). [doi:10.1128/MCB.00346-07](https://doi.org/10.1128/MCB.00346-07) Medline
  39. C. K. Tseng, C. S. Chung, H. C. Chen, S. C. Cheng, A central role of Cwc25 in spliceosome dynamics during the catalytic phase of pre-mRNA splicing. *RNA* **23**, 546–556 (2017). [doi:10.1261/rna.059204.116](https://doi.org/10.1261/rna.059204.116) Medline
  40. W. L. DeLano, The PyMOL Molecular Graphics System (Schrödinger, 2002); [www.pymol.org](http://www.pymol.org).
  41. J. D. Dignam, R. M. Lebovitz, R. G. Roeder, Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res.* **11**, 1475–1489 (1983). [doi:10.1093/nar/11.5.1475](https://doi.org/10.1093/nar/11.5.1475) Medline
  42. B. Kastner, N. Fischer, M. M. Golas, B. Sander, P. Dube, D. Boehringer, K. Hartmuth, J. Deckert, F. Hauer, E. Wolf, H. Uchtenhagen, H. Urlaub, F. Herzog, J. M. Peters, D. Poerschke, R. Lührmann, H. Stark, GraFix: Sample preparation for single-particle electron cryomicroscopy. *Nat. Methods* **5**, 53–55 (2008). [doi:10.1038/nmeth1139](https://doi.org/10.1038/nmeth1139) Medline
  43. J. Lei, J. Frank, Automated acquisition of cryo-electron micrographs for single particle reconstruction on an FEI Tecnai electron microscope. *J. Struct. Biol.* **150**, 69–80 (2005). [doi:10.1016/j.jsb.2005.01.002](https://doi.org/10.1016/j.jsb.2005.01.002) Medline
  44. S. Q. Zheng, E. Palovcak, J.-P. Armache, K. A. Verba, Y. Cheng, D. A. Agard, MotionCor2: Anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017). [doi:10.1038/nmeth.4193](https://doi.org/10.1038/nmeth.4193) Medline
  45. K. Zhang, Gctf: Real-time CTF determination and correction. *J. Struct. Biol.* **193**, 1–12 (2016). [doi:10.1016/j.jsb.2015.11.003](https://doi.org/10.1016/j.jsb.2015.11.003) Medline
  46. F. Wang, H. Gong, G. Liu, M. Li, C. Yan, T. Xia, X. Li, J. Zeng, DeepPicker: A deep learning approach for fully automated particle picking in cryo-EM. *J. Struct. Biol.* **195**, 325–336 (2016). [doi:10.1016/j.jsb.2016.07.006](https://doi.org/10.1016/j.jsb.2016.07.006) Medline
  47. D. Kimanius, B. O. Forsberg, S. H. Scheres, E. Lindahl, Accelerated cryo-EM structure determination with parallelisation using GPUs in RELION-2. *eLife* **5**, e18722 (2016). [doi:10.7554/eLife.18722](https://doi.org/10.7554/eLife.18722) Medline
  48. S. H. Scheres, RELION: Implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* **180**, 519–530 (2012). [doi:10.1016/j.jsb.2012.09.006](https://doi.org/10.1016/j.jsb.2012.09.006) Medline
  49. S. Chen, G. McMullan, A. R. Faruqi, G. N. Murshudov, J. M. Short, S. H. W. Scheres, R. Henderson, High-resolution noise substitution to measure overfitting and validate resolution in 3D structure determination by single particle electron cryomicroscopy. *Ultramicroscopy* **135**, 24–35 (2013). [doi:10.1016/j.ultramic.2013.06.004](https://doi.org/10.1016/j.ultramic.2013.06.004) Medline
  50. P. B. Rosenthal, R. Henderson, Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *J. Mol. Biol.* **333**, 721–745 (2003). [doi:10.1016/j.jmb.2003.07.013](https://doi.org/10.1016/j.jmb.2003.07.013) Medline
  51. A. Kucukelbir, F. J. Sigworth, H. D. Tagare, Quantifying the local resolution of cryo-EM density maps. *Nat. Methods* **11**, 63–65 (2014). [doi:10.1038/nmeth.2727](https://doi.org/10.1038/nmeth.2727) Medline
  52. P. Emsley, K. Cowtan, Coot: Model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004). [doi:10.1107/S0907444904019158](https://doi.org/10.1107/S0907444904019158) Medline
  53. E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin, UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004). [doi:10.1002/jcc.20084](https://doi.org/10.1002/jcc.20084) Medline
  54. N. Stein, CHAINSAW: A program for mutating pdb files used as templates in molecular replacement. *J. Appl. Cryst.* **41**, 641–643 (2008). [doi:10.1107/S0021889808006985](https://doi.org/10.1107/S0021889808006985)
  55. F. DiMaio, M. D. Tyka, M. L. Baker, W. Chiu, D. Baker, Refinement of protein structures into low-resolution density maps using Rosetta. *J. Mol. Biol.* **392**, 181–190 (2009). [doi:10.1016/j.jmb.2009.07.008](https://doi.org/10.1016/j.jmb.2009.07.008) Medline
  56. F. DiMaio, Y. Song, X. Li, M. J. Brunner, C. Xu, V. Conticello, E. Egelman, T. Marlovits, Y. Cheng, D. Baker, Atomic-accuracy models from 4.5-Å cryo-electron microscopy data with density-guided iterative local refinement. *Nat. Methods* **12**, 361–365 (2015). [doi:10.1038/nmeth.3286](https://doi.org/10.1038/nmeth.3286) Medline
  57. Y. Song, F. DiMaio, R. Y.-R. Wang, D. Kim, C. Miles, T. Brunette, J. Thompson, D. Baker, High-resolution comparative modeling with RosettaCM. *Structure* **21**, 1735–1742 (2013). [doi:10.1016/j.str.2013.08.005](https://doi.org/10.1016/j.str.2013.08.005) Medline
  58. S. Bessonov, M. Anokhina, A. Krasauskas, M. M. Golas, B. Sander, C. L. Will, H. Urlaub, H. Stark, R. Lührmann, Characterization of purified human Bact spliceosomal complexes reveals compositional and morphological changes during spliceosome activation and first step catalysis. *RNA* **16**, 2384–2403 (2010). [doi:10.1261/rna.2456210](https://doi.org/10.1261/rna.2456210) Medline
  59. R. A. Hom, P.-Y. Chang, S. Roy, C. A. Musselman, K. C. Glass, A. I. Selezneva, O. Gozani, R. F. Ismagilov, M. L. Cleary, T. G. Kutateladze, Molecular mechanism of MLL PHD3 and RNA recognition by the Cyp33 RRM domain. *J. Mol. Biol.* **400**, 145–154 (2010). [doi:10.1016/j.jmb.2010.04.067](https://doi.org/10.1016/j.jmb.2010.04.067) Medline
  60. F. C. Chou, P. Sripakdeevong, S. M. Dibrov, T. Hermann, R. Das, Correcting pervasive errors in RNA crystallography through enumerative structure prediction. *Nat. Methods* **10**, 74–76 (2013). [doi:10.1038/nmeth.2262](https://doi.org/10.1038/nmeth.2262) Medline
  61. G. N. Murshudov, A. A. Vagin, E. J. Dodson, Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.* **53**, 240–255 (1997). [doi:10.1107/S0907444996012255](https://doi.org/10.1107/S0907444996012255) Medline
  62. R. A. Nicholls, M. Fischer, S. McNicholas, G. N. Murshudov, Conformation-independent structural comparison of macromolecules with ProSMART. *Acta Crystallogr. D Biol. Crystallogr.* **70**, 2487–2499 (2014). [doi:10.1107/S1399004714016241](https://doi.org/10.1107/S1399004714016241) Medline
  63. A. Amunts, A. Brown, X. C. Bai, J. L. Llácer, T. Hussain, P. Emsley, F. Long, G. Murshudov, S. H. W. Scheres, V. Ramakrishnan, Structure of the yeast mitochondrial large ribosomal subunit. *Science* **343**, 1485–1489 (2014). [doi:10.1126/science.1249410](https://doi.org/10.1126/science.1249410) Medline
  64. I. W. Davis, A. Leaver-Fay, V. B. Chen, J. N. Block, G. J. Kapral, X. Wang, L. W. Murray, W. B. Arendall 3rd, J. Snoeyink, J. S. Richardson, D. C. Richardson, MolProbity: All-atom contacts and structure validation for proteins and nucleic

- acids. *Nucleic Acids Res.* **35**, W375–W383 (2007). [doi:10.1093/nar/gkm216](https://doi.org/10.1093/nar/gkm216) [Medline](#)
65. A. T. Brünger, P. D. Adams, G. M. Clore, W. L. DeLano, P. Gros, R. W. Grosse-Kunstleve, J. S. Jiang, J. Kuszewski, M. Nilges, N. S. Pannu, R. J. Read, L. M. Rice, T. Simonson, G. L. Warren, Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D Biol. Crystallogr.* **54**, 905–921 (1998). [doi:10.1107/S0907444498003254](https://doi.org/10.1107/S0907444498003254) [Medline](#)
66. N. Echols, D. Milburn, M. Gerstein, MolMovDB: Analysis and visualization of conformational change and structural flexibility. *Nucleic Acids Res.* **31**, 478–482 (2003). [doi:10.1093/nar/gkg104](https://doi.org/10.1093/nar/gkg104) [Medline](#)
67. W. G. Krebs, M. Gerstein, The morph server: A standardized system for analyzing and visualizing macromolecular motions in a database framework. *Nucleic Acids Res.* **28**, 1665–1675 (2000). [doi:10.1093/nar/28.8.1665](https://doi.org/10.1093/nar/28.8.1665) [Medline](#)

## ACKNOWLEDGMENTS

We thank Dr. Liansuo Zu for synthesizing the compound BN82685, the Tsinghua University Branch of China National Center for Protein Sciences (Beijing) for the cryo-EM facility, and the computational facility support on the cluster of Bio-Computing Platform. This work was supported by funds from the National Natural Science Foundation of China (31621092 and 31430020) and the Ministry of Science and Technology (2016YFA0501100 to J.L.). The atomic coordinates have been deposited in the Protein Data Bank with the accession code 5YZG, and the EM maps have been deposited in EMDb with the accession code EMD-6864.

## SUPPLEMENTARY MATERIALS

Materials and Methods

Figs. S1 to S11

Tables S1 to S3

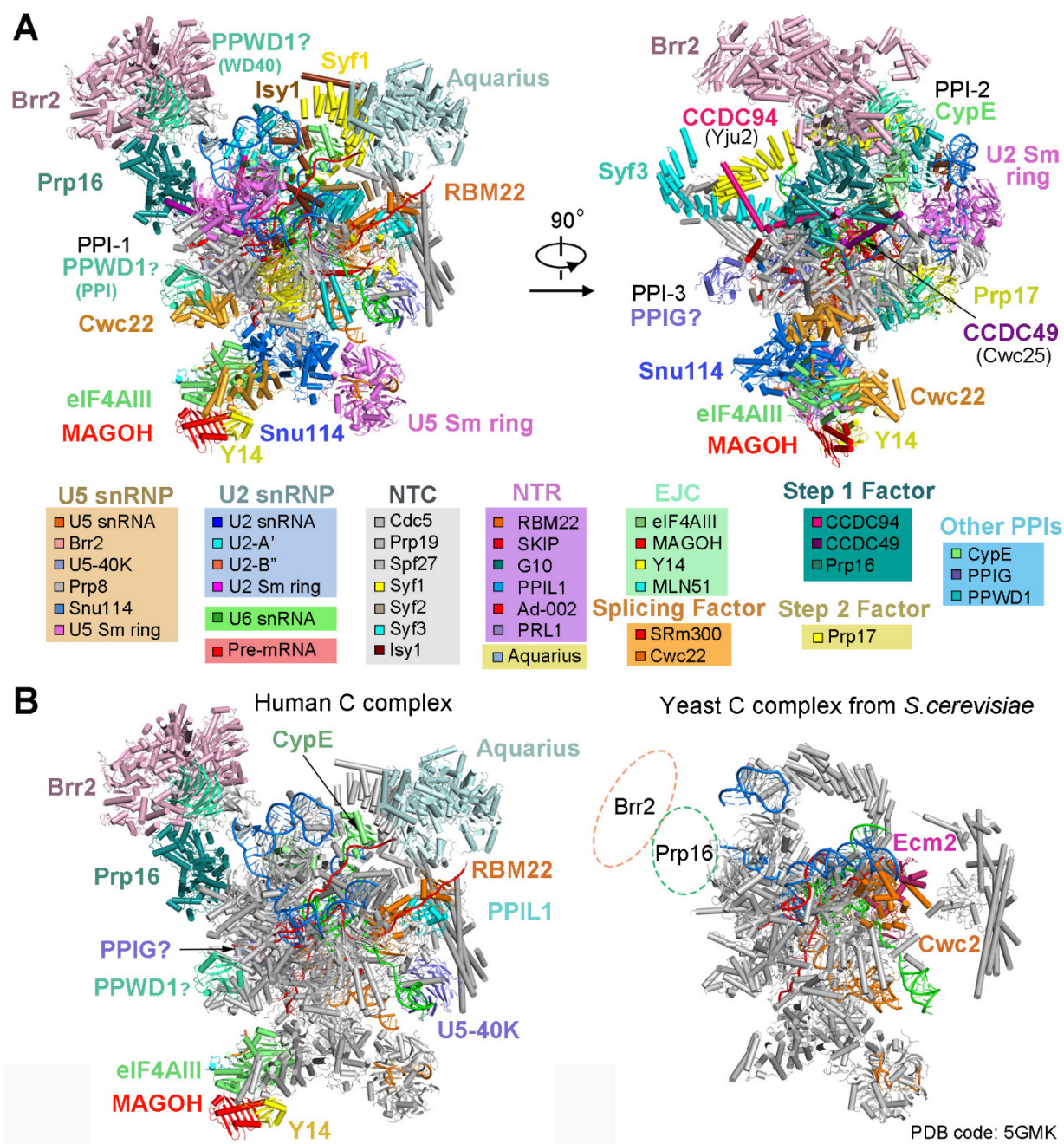
References (41–67)

Movies S1 to S3

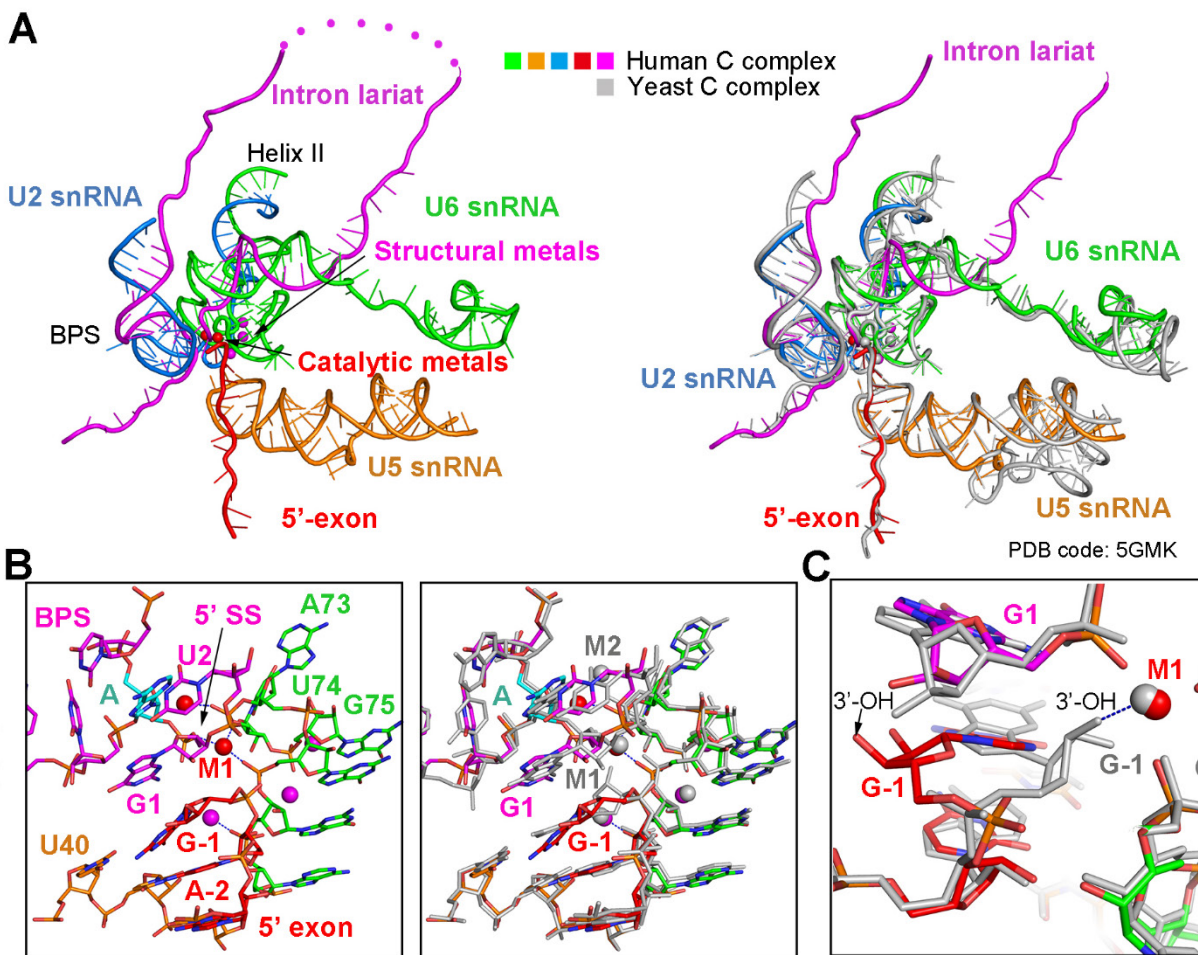
1 December 2017; accepted 15 December 2017

Published online 4 January 2018

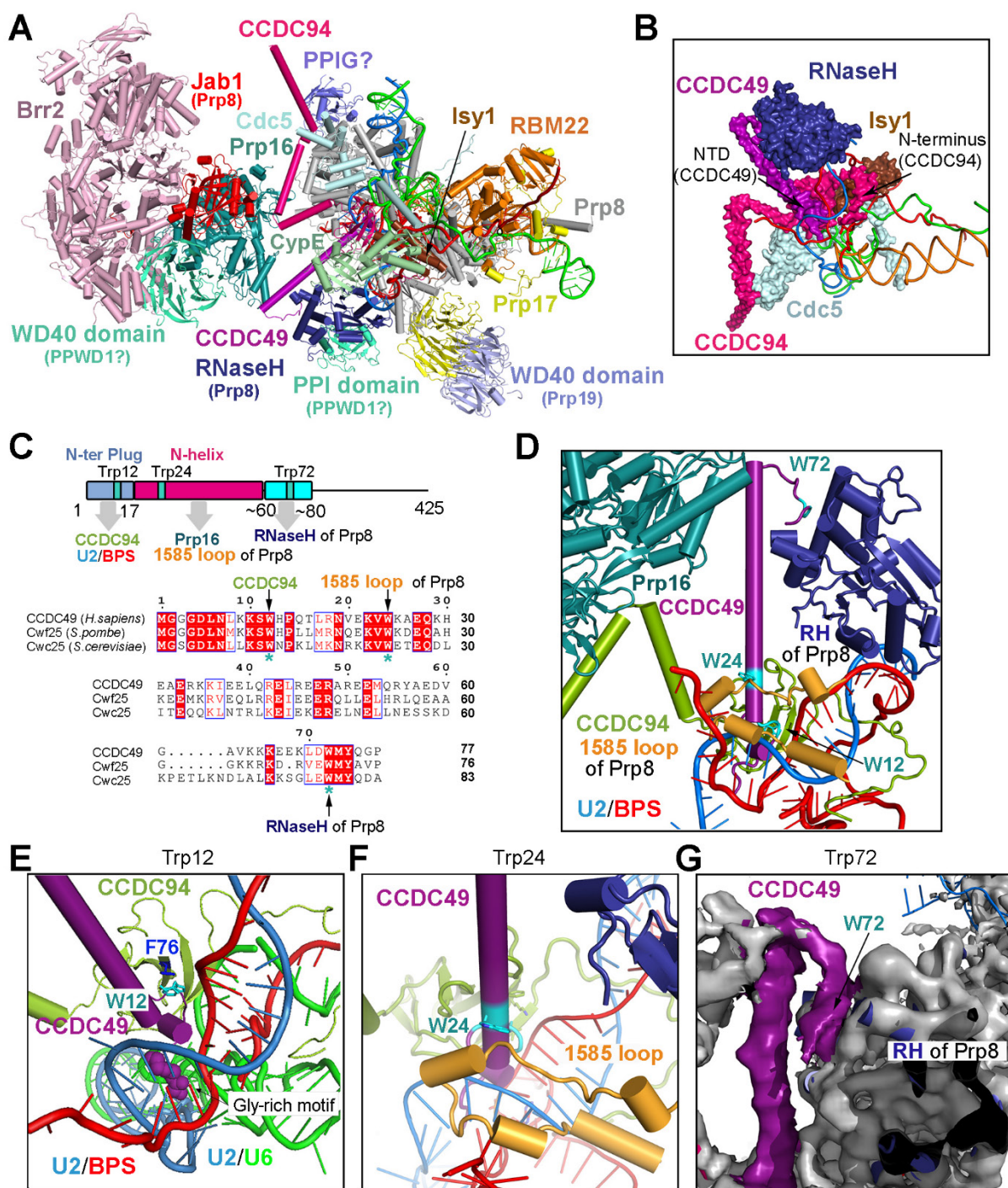
10.1126/science.aar6401



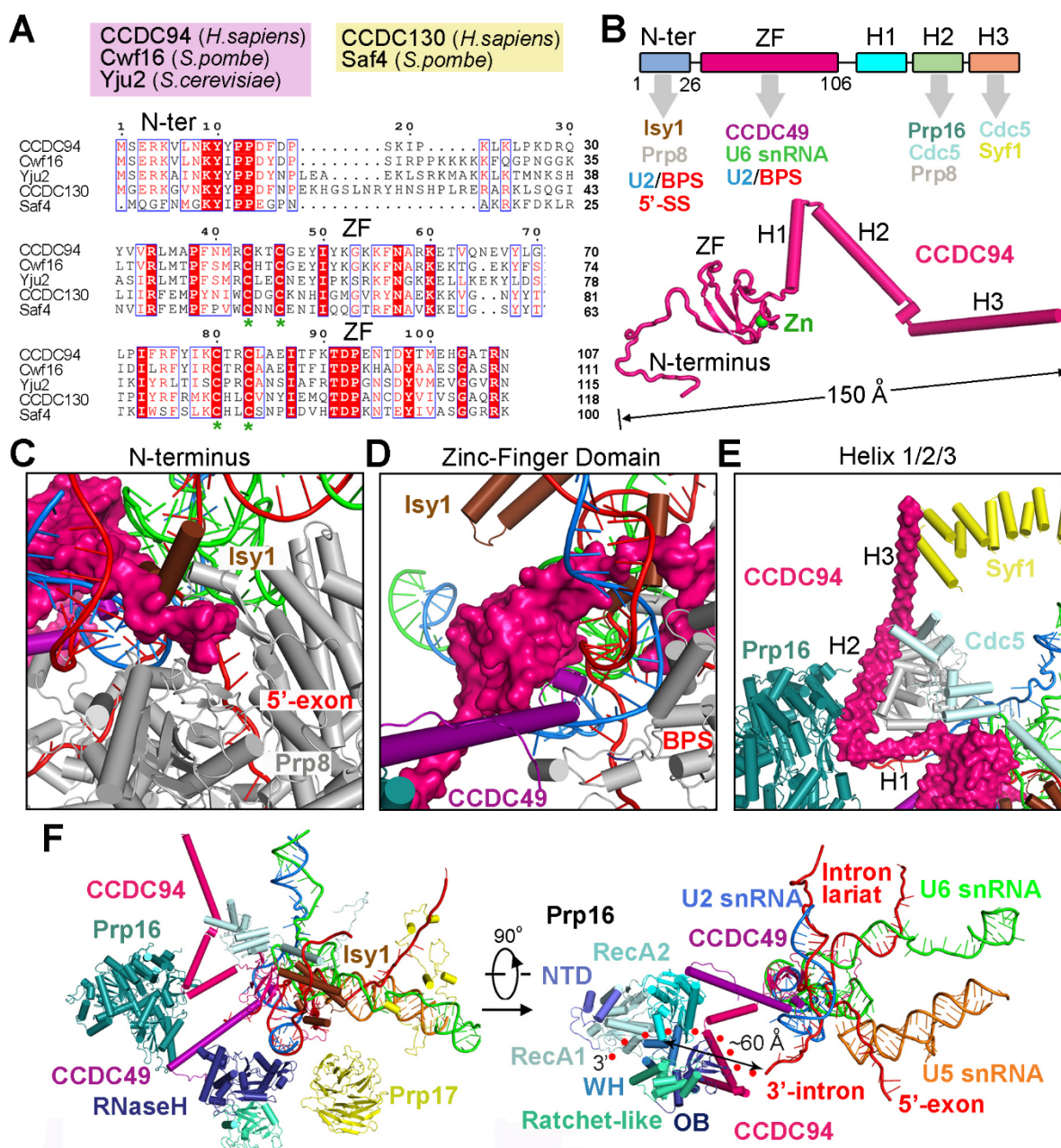
**Fig. 1. Cryo-EM structure of the human catalytic step I spliceosome (C complex) at 4.1 Å resolution.** (A) Overall structure of the human C complex. Two perpendicular views are shown. The atomic model includes 47 proteins, three snRNAs (U2, U5, and U6), a 5'-exon, and a 44-nucleotide intron lariat, with a molecular weight of approximately 1.8 MDa. Among the modeled 15,479 amino acids, about two-thirds have side chains. All structural images in the figures were created using PyMol (40). (B) An overall structural comparison between the C complexes from human (left) and *S. cerevisiae* (right). In both structures, the U2, U5, and U6 snRNAs are colored marine, orange, and green, respectively; the pre-mRNA is highlighted in red; and the shared protein components are shaded gray. Protein components that are uniquely present in either structure are color-coded. Compared to the structure of the yeast C complex, 11 additional protein components have been identified in the human C complex. These include 4 peptidyl prolyl isomerases (PPIs), 4 proteins of the exon junction complex (EJC) (eIF4AIII, MAGOH, Y14, and MLN51), U5-40K, RBM22, Aquarius. The EM density map also allows unambiguous identification of the ATPase/helicases Prp16 and Brr2, which were absent in our own preparation of the yeast C complex (10). Notably, the human protein RBM22 appears to be derived from the yeast Ecm2 and Cwc2 (22).



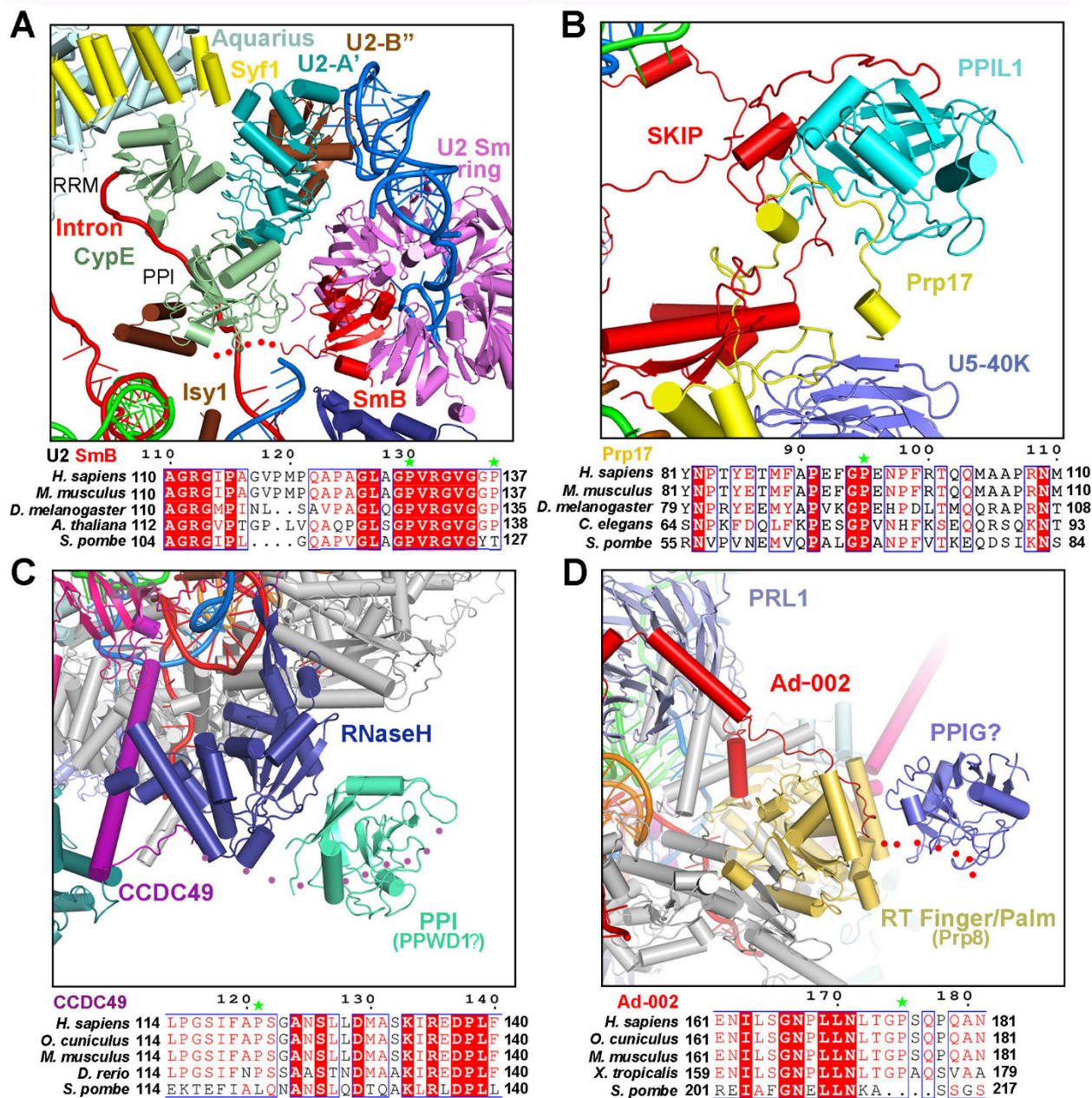
**Fig. 2. The RNA elements and the active site metals in the human C complex.** (A) Overall structure of the core RNA elements. The color-coded RNA structure of the human C complex is shown in the left panel, and its comparison to that of the *S. cerevisiae* C complex (colored gray) is displayed in the right panel. In the C complex, the 5'-exon remains anchored to the loop I of U5 snRNA. The nucleophile-containing adenine nucleotide from the branch point sequence (BPS) is covalently joined to the guanine nucleotide at the 5'-end of 5'SS. The disordered RNA sequences of the intron lariat are indicated by the dotted lines. (B) Coordination of the metal ions in the active site of the human C complex (left panel) and its comparison to the *S. cerevisiae* C complex (right panel). All structural elements of the *S. cerevisiae* C complex are shown in gray. Both the location and the coordination of the three structural metals (colored magenta) in the active site of the human C complex are nearly identical to those in the *S. cerevisiae* C complex. However, despite a similar location for the catalytic M1 metal (colored red) in the human C complex, its coordination has changed compared to the yeast C complex. In addition, compared to the yeast C complex, the catalytic M2 in the human complex has been translocated by about 3 Å. (C) A close-up view on the coordination of the M1 metal. Due to a positional shift of the ribose of the guanine nucleotide (G<sub>-1</sub>) at the 3'-end of the 5'-exon, the M1 metal is coordinated by the 3'-OH of G<sub>-1</sub> only in the *S. cerevisiae* C complex, but not in the human C complex.



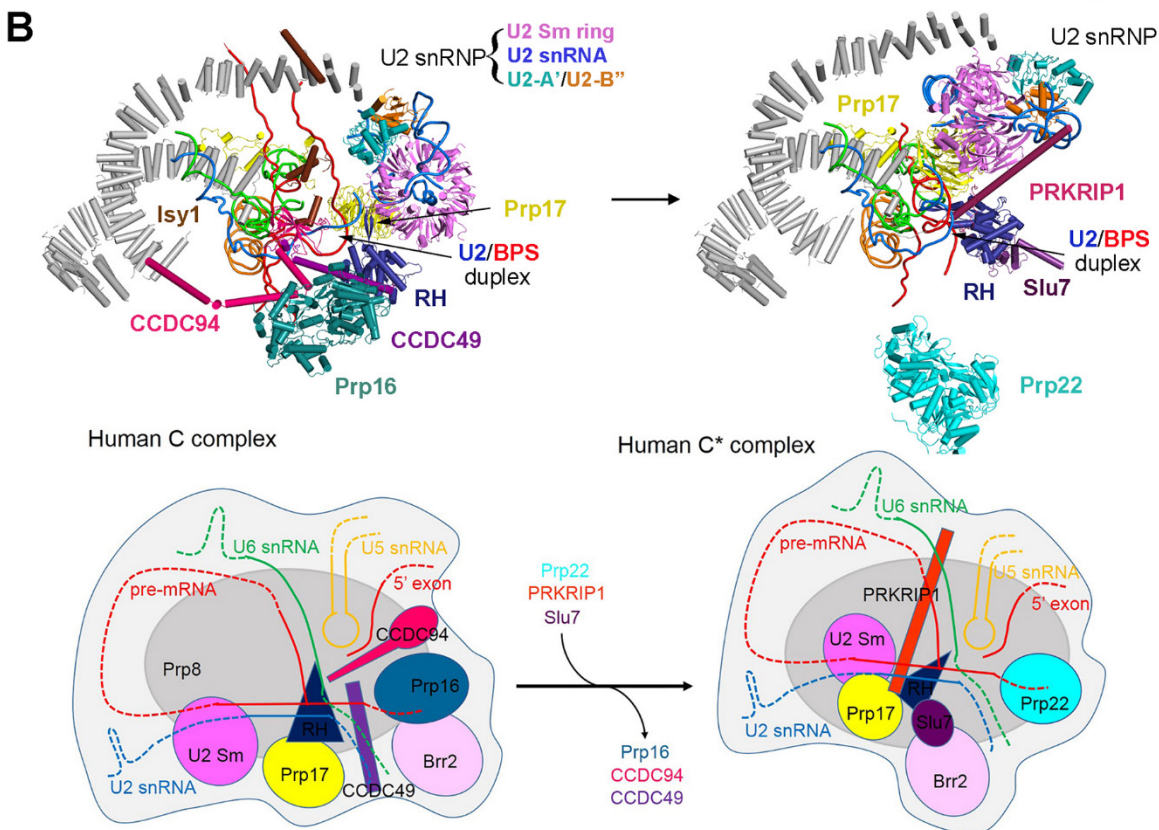
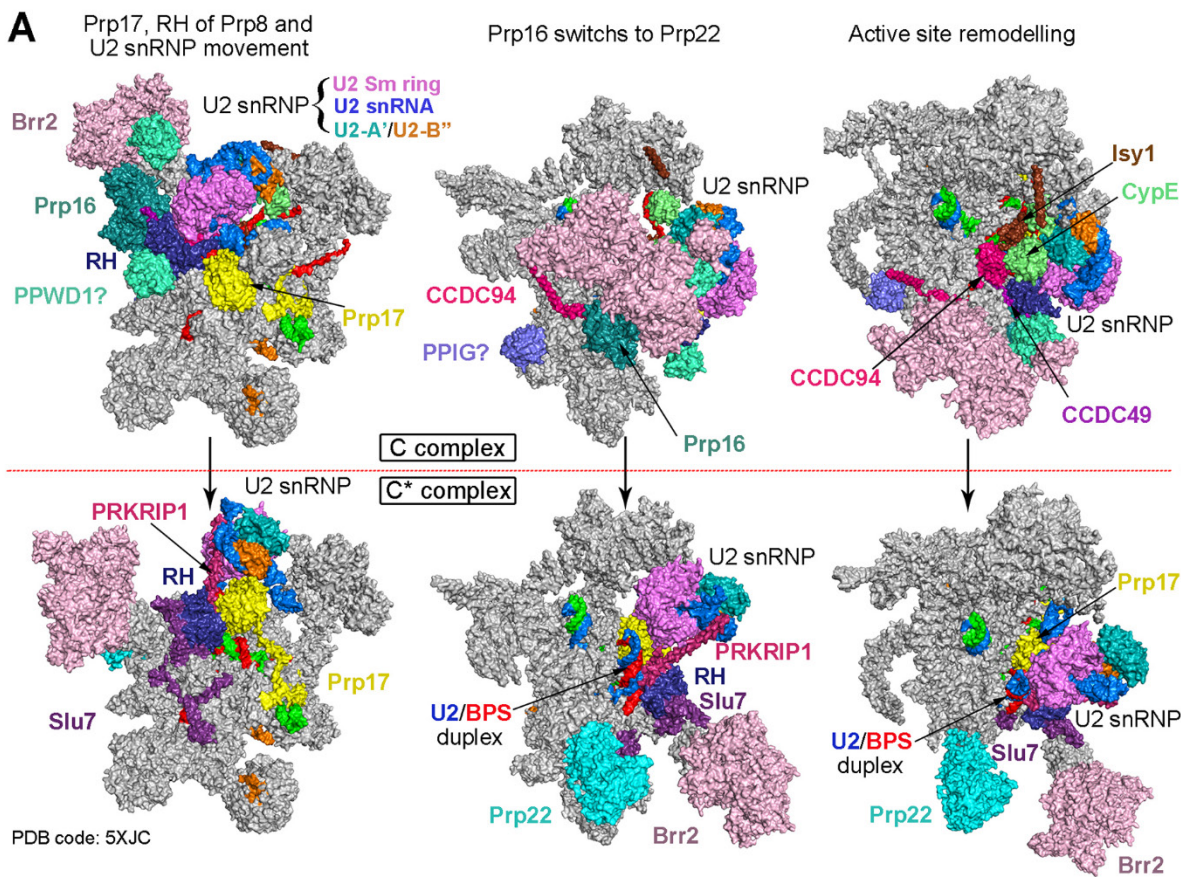
**Fig. 3. Structure and functional implications of the step I splicing factor CCDC49 in the human C complex.** (A) An overall view on the protein components surrounding the active site. The protein components are color-coded. The ATPase/helicase Prp16 is placed between Brr2 and the core of the spliceosome. The step I splicing factors CCDC49 and CCDC94 correspond to the yeast proteins Cwc25 and Yju2, respectively. (B) A surface view on the key proteins that stabilize the active site conformation. The N-termini of both CCDC49 and CCDC94 insert deeply into the active site to interact with the RNA elements. The RNaseH-like domain of Prp8, the NTC component Isy1, and the N-terminal portion of Cdc5 also stabilize the RNA elements at the active site. (C) Domain structure of CCDC49 and its sequence alignment with Cwf25 (*S. pombe*) and Cwc25 (*S. cerevisiae*). Three invariant Trp residues of CCDC49 – Trp12 in the N-terminal plug, Trp24 in the N-helix, and Trp72 in the C-terminal hook – interact with CCDC94, the 1585-loop of Prp8, and the RNaseH-like domain of Prp8, respectively. (D) A close-up view on CCDC49. The three invariant Trp residues are indicated. (E) A close-up view on the N-terminal plug of CCDC49 and its interactions with surrounding components. The N-terminal Gly2-Gly3-Gly4 motif is inserted deeply between the U2/BPS duplex and the U2/U6 duplex at the active site. Trp12 of CCDC49 interacts with hydrophobic amino acids from CCDC94. (F) A close-up view on the N-terminal portion of the N-helix. Trp24 of CCDC49 directly contacts the 1585-loop of Prp8. (G) A close-up view on the C-terminal hook of CCDC49. Shown here is the EM density map in this region. The C-terminal portion of the N-helix interacts with Prp16 whereas the hook binds to the RNaseH-like domain of Prp8, with Trp72 of CCDC49 playing a critical role.



**Fig. 4. The step I splicing factor CCDC94 and the ATPase/helicase Prp16 in the human C complex.** (A) Sequence alignment involving the N-terminal fragment and the zinc-binding domain of CCDC94. The corresponding sequences from CCDC94 and CCDC130 (human), Cwf16 and Saf4 (*S. pombe*), and Yju2 (*S. cerevisiae*) are aligned. Compared to CCDC130 and Saf4, CCDC94 and Cwf16 shares a higher degree of sequence similarity with Yju2. (B) The structure of CCDC94 is shown below its schematic domain representation. CCDC94 sequentially consists of an N-terminal fragment, a Cys4-type zinc-binding motif, and three  $\alpha$ -helices known as H1/H2/H3. The protein and RNA components that interact with these domains are indicated below the schematic domain representation. (C) The N terminus of CCDC94 inserts deeply into the active site, making direct interactions with Isy1, Prp8, U2/BPS duplex, and 5'SS. (D) The zinc-binding motif binds the N terminus of CCDC49, U6 snRNA and the U2/BPS duplex. (E) The helix H2 interacts with Prp16, Prp8 and the N-terminal domain of Cdc5, whereas the helix H3 binds to the N-terminal domain of Cdc5 and the C-terminal region of Syf1. (F) Prp16 is connected to the active site RNA elements through its interactions with the two step I factors CCDC49 and CCDC94. Two perpendicular views are shown. Prp16 is thought to pull the 3'-exon sequences in a 3'-to-5' direction, dissociating CCDC49 and CCDC94 and eventually forming the step II activated spliceosome (the C\* complex).



**Fig. 5. The location and putative mechanism of the peptidyl prolyl isomerases (PPIs) in the human C complex.** (A) CypE associates with the intron lariat through both the N-terminal RRM domain and the C-terminal PPI domain. The PPI domain of CypE may catalyze proline isomerization in the C-terminal fragment of the SmB protein. Sequence alignment of SmB from five organisms is shown below the structure. Conserved residues are boxed and invariant residues are shared red. (B) PPIL1 associates with SKIP, Prp17, and RBM22. Sequence alignment of Prp17 from five organisms is shown below the structure. Prp17 may serve as the substrate for PPIL1. (C) A PPI domain associates with the RNaseH-like domain of Prp8. Sequence alignment of CCDC49 from five organisms is shown below the structure. This PPI might come from PPWD1. (D) A PPI domain interacts with the RT Finger/Palm domain of Prp8. The C-terminal sequences of Ad-002 may serve as the substrate for this PPI, which is tentatively assigned to PPIG.



**Fig. 6. A working model on the transition from the human C complex to the C\* complex.** (A) Structural comparison between the human C and C\* complexes. Three surface views of the human C and C\* complexes are shown in the upper and lower panels, respectively. The protein components that are unique to either complex or undergo marked positional movement are color-coded. (B) A working model on the transition from the human C complex to the C\* complex. An overall structural comparison between the human C and C\* complexes is shown in the upper panels. The protein components that remain unchanged in the transition are removed to better display the key players. A cartoon diagram of the working model is shown in the lower panels. In this model, the transition is triggered by Prp16, which, through ATP binding and hydrolysis, pulls the 3'-end single-stranded sequences of the intron lariat-3'-exon intermediate. This leads to the dissociation of the step I factors CCDC49 and CCDC94 and recruitment of step II factors Prp18 and Slu7 and the protein PRKRIP1. Prp16 is dissociated presumably due to the loss of interactions with the step I factors and Prp22 is recruited into the C\* complex.

Cite as: R. Wild *et al.*, *Science*  
10.1126/science.aar5140 (2018).

# Structure of the yeast oligosaccharyltransferase complex gives insight into eukaryotic *N*-glycosylation

Rebekka Wild,<sup>1\*</sup> Julia Kowal,<sup>1\*</sup> Jillianne Eyring,<sup>2\*</sup> Elsy M. Ngwa,<sup>2</sup> Markus Aebi,<sup>2†</sup> Kaspar P. Locher<sup>1†</sup>

<sup>1</sup>Institute of Molecular Biology and Biophysics, Department of Biology, ETH Zurich, CH-8093 Zurich, Switzerland. <sup>2</sup>Institute of Microbiology, Department of Biology, ETH Zurich, CH-8093 Zurich, Switzerland.

\*These authors contributed equally to this work.

†Corresponding author. Email: markus.aebi@micro.biol.ethz.ch (M.A.); locher@mol.biol.ethz.ch (K.P.L.)

Oligosaccharyltransferase (OST) is an essential membrane protein complex in the Endoplasmic Reticulum, where it transfers an oligosaccharide from a dolichol-pyrophosphate-activated donor to glycosylation sites of secretory proteins. We here describe the atomic structure of yeast OST determined by cryo-EM, revealing a conserved subunit arrangement. The active site of the catalytic STT3 subunit points away from the center of the complex, allowing unhindered access to substrates. The dolichol-pyrophosphate moiety binds to a lipid-exposed groove of STT3, while two non-catalytic subunits and an ordered *N*-glycan form a membrane-proximal pocket for the oligosaccharide. The acceptor polypeptide site faces an oxidoreductase domain in standalone OST complexes or is immediately adjacent to the translocon, suggesting how eukaryotic OSTs efficiently glycosylate a large number of polypeptides prior to their folding.

*N*-linked glycosylation is a post-translational modification of asparagine residues found in all domains of life (1). The covalently attached glycans are essential for correct protein folding, sorting and secretion, or for modulating specific cell surface interactions (2–4). The central enzyme in the *N*-glycosylation pathway is the oligosaccharyltransferase (OST), which catalyzes the initial transfer of a defined glycan (Glc<sub>3</sub>Man<sub>9</sub>GlcNAc<sub>2</sub> in higher eukaryotes) from the lipid carrier dolichol-pyrophosphate to polypeptide chains entering the secretory pathway in the Endoplasmic Reticulum (ER) (5, 6). OST specifically recognizes the Asn-Xaa-Ser/Thr motif, where Xaa can be any amino acid except proline (7). Whereas in bacteria and some lower eukaryotes, this process is carried out by a single-subunit oligosaccharyltransferase (ssOST) (8–10), most eukaryotes encode large, multi-subunit, membrane-embedded OST complexes that contain eight subunits in yeast and possibly more in multicellular organisms (11, 12). In *Saccharomyces cerevisiae* five subunits (STT3, SWP1, WBP1, OST1 and OST2) were shown to be essential for cell survival, while deletion of the remaining three subunits was found to reduce complex stability and glycosylation activity (11, 13).

Eukaryotic OST enzymes have been visualized using single particle cryo-electron microscopy or cryo-electron tomography (14–16), but the resolution of these studies, ranging from 20–9 Å, did not allow unambiguous assignment of OST subunits. X-ray structures of ssOSTs from bacteria and archaea (PglB and AglB proteins, homologous to the catalytic STT3 protein of multi-subunit OSTs) have provided insight

into the substrate recognition and glycan transfer mechanisms of prokaryotic enzymes (8, 9, 17, 18). X-ray structures of the isolated luminal domains of yeast OST6 and of the human Tusc3 (also named N33) have revealed the disulfide formation and cleavage mechanism of this redox chaperone (19, 20). However, in the absence of a high-resolution structure of a eukaryotic OST complex, it is unknown how the non-catalytic subunits extend the range of acceptor polypeptides over that of ssOST enzymes. To reveal the architecture of eukaryotic OSTs and to understand how they recognize and process a striking number of acceptor proteins (21), we determined a high-resolution structure of the yeast OST complex using single-particle cryo-EM.

## Preparation of OST complex for cryo-EM analysis

Yeast OST exists in two isoforms, containing either OST3 or the homologous OST6 subunit. To avoid potential heterogeneity in our purified samples, we generated a *S. cerevisiae* strain lacking the OST6 gene and overexpressed OST3 from a plasmid to compensate for the absence of OST6. A 1D4-tag fused to the C terminus of the OST4 subunit facilitated efficient affinity purification and sufficient yields of the OST complex despite low endogenous expression levels. The resulting strain ( $\Delta$ ost6 pOST3 OST4-1D4) showed almost complete *N*-glycosylation of OST1 and WBP1 proteins, whereas a strain expressing only an OST6-containing complex ( $\Delta$ ost3 pOST6 OST4-1D4) exhibited hypoglycosylation (Fig. 1A). SILAC-based mass spectrometric analyses demonstrated that for 6 out of 8 subunits, the detergent-purified, OST3-

containing complex has the same subunit composition and stoichiometry as OST in wild-type yeast cells (Fig. 1B and table S1). OST4-derived peptides could not be detected, and the observed ratios for OST3 and OST6 showed that the two paralogs are mutually exclusive in the assembled OST complex. The OST3-containing complex reconstituted into lipidic nanodiscs was fully functional as shown by an in vitro glycosylation assay developed earlier (10, 22), using a synthetic lipid-linked oligosaccharide (LLO) analog and a fluorescently labeled acceptor peptide (Fig. 1C). The measured glycosylation activity of 3.5 peptides per minute per OST matches previously reported rates of a eukaryotic ssOST (10) (Fig. 1D).

The resolution of the 3D reconstruction of nanodisc-reconstituted OST was 3.3 Å based on the FSC = 0.143 criterion (figs. S1 and S2). The EM map was of excellent quality in the transmembrane (TM) region and in most of the luminal regions. The luminal domain of SWP1, most distant to the membrane, featured lower map quality, probably indicating higher domain flexibility, but still displayed secondary structure features (Fig. 2A and figs. S1, D to F, and S3). Missing parts in the model include the luminal domain and TM1 of OST3 and the external loop EL5 of STT3, for which no density was observed. Because we visualized OST in an apo state, these segments are likely mobile (see below).

### Architecture and subunit structure

The membrane-embedded part of yeast OST contains a total of 28 transmembrane helices, with each subunit contributing at least one TM segment (Fig. 2, B to D). The membrane topologies of the subunits agree with previous predictions (11) except for OST5, whose N- and C-termini are located in the ER lumen. Previous in vivo experiments suggested that OST assembly occurs through the formation of three sub-complexes (13). Our structure revealed that the spatial arrangement of the OST subunits agrees with this subdivision: Sub-complex 1 contains OST1 and OST5, sub-complex 2 contains STT3, OST3, and OST4, and sub-complex 3 contains OST2, WBP1, and SWP1 (Fig. 2D). At the membrane-embedded interfaces between sub-complexes, several ordered phospholipids could be identified, in particular in the lipid layer facing the ER lumen (fig. S4, F to I).

In sub-complex I, the N-terminal helix of OST5 is in close contact with the OST1 subunit. OST1 has a single TM helix and a luminal domain containing two sub-domains that have similar folds and superimpose well on each other (r.m.s.d = 2.2 Å for 153 out of 191 residues, fig. S5A). The fold features two stacked  $\beta$ -sheets and no  $\alpha$ -helices and was previously observed in other multi-domain proteins, e.g., in aminopeptidases and leukotriene hydrolases (23, 24) (fig. S5B). Because it does not contain the catalytic residues in these proteins, its function within OST1 cannot be deduced.

Sub-complex II includes the catalytic subunit STT3,

which is homologous to the functionally and structurally related ssOST enzymes PglB from *Campylobacter lari* and AglB from *Archaeoglobus fulgidus* (8, 9, 18). Our structure confirms that yeast STT3 contains a similar transmembrane topology with 13 TM helices (Fig. 2C) (25, 26), of which TM9 is poorly resolved in the EM map and is probably flexible in the absence of bound substrate (fig. S6). The luminal domain of STT3 also resembles the PglB and AglB proteins, although the folds are not identical (fig. S7). Importantly, we found the external loop EL5 of STT3 to be disordered, in line with previous observations that EL5 of PglB only becomes fully ordered when substrates are bound (18). The TM helices of STT3 are tightly interacting with OST4 (which consists of a single TM helix) and with TM2-TM4 of OST3 (Fig. 2C). Although the presence of full-length OST3 in the OST complex was confirmed by SDS-PAGE and mass spectrometric analysis (Fig. 1A and table S1), no clear density for TM1 and the N-terminal luminal domain of OST3 was visible in the EM map. We conclude that in the absence of a peptide substrate, the luminal domain of OST3 is highly flexible. Both TM1 of OST3 and TM9 of STT3 are in close proximity to the likely LLO-binding site, suggesting that their flexibility might be associated with the absence of bound LLO.

Sub-complex III contains the OST2 subunit, whose TM helices mediate contacts between STT3 and the TM helices of WBP1 and SWP1. OST2 contains an N-terminal  $\alpha$ -helix (aa21-aa38) that is located at the cytosolic membrane boundary and parallel to the membrane plane, where it forms contacts to TM8 and TM9 of STT3. The WBP1 subunit contains two luminal domains. The N-terminal domain shares structural homology with various proteins of distinct functions including an intraflagellar transport protein or glutamine amidotransferase (27, 28) (fig. S5C), whereas the central domain of WBP1 features a fold found in proteins of the complement system (29, 30) (fig. S5D). However, none of these structural homologs unambiguously identify a potential function of WBP1. The SWP1 subunit contains a single luminal domain, which is most distant to the membrane of all OST domains and connected to a single transmembrane helix TM1 by a long linker.

OST1, WBP1 and SWP1 were previously suggested to be involved in substrate recognition and to act as chaperones, which would coordinate protein folding and glycosylation (31–33). Our structural data show that none of their luminal domains adopt a chaperone-like fold. However, these domains might serve as docking platforms for interaction partners including chaperones or enzymes acting on nascent glycoproteins. For example, the human OST1 homolog Ribophorin I was shown to interact with the carbohydrate-binding protein malectin (34, 35).

To analyze whether eukaryotic OST complexes share common structural features we plotted the degree of sequence conservation between different species onto our yeast

OST structure. We found that the active site groove in STT3 is highly conserved (Fig. 3A and fig. S7D). In addition, regions of high sequence conservation were found at the interfaces between subunits OST1 and STT3 (1129 Å<sup>2</sup> buried surface area) as well as between WBP1 and SWP1 (2369 Å<sup>2</sup> buried surface area) (Fig. 3B). This suggests that both the arrangement of the luminal domains and the active site are conserved in eukaryotic, multi-subunit OST.

Previous genetic and mass spectrometric analyses suggested that yeast OST contains seven *N*-glycosylation sites (36). We observed EM density for five of these glycans (at N539 of STT3, N336 and N400 of OST1, N60 and N332 of WBP1), whereas the remaining two glycans (at N99 and N217 of OST1) are located in flexible loop regions and are disordered. The best-ordered *N*-glycan is attached to the strictly conserved STT3 residue N539, where we observed density covering eight glycan moieties (Man<sub>6</sub>GlcNAc<sub>2</sub>). This *N*-glycan forms interactions with WBP1 and SWP1 (fig. S4, A to E) and is in the immediate vicinity of the proposed binding site for the 14-unit-saccharide of bound LLO (see below).

### Active site and substrate-binding pockets

The STT3 subunit contains the active site and the acceptor peptide and donor LLO binding pockets. Although we visualized yeast OST in the apo state with no substrates bound, significant mechanistic insight could be deduced by comparing our structure to bacterial and archaeal ssOSTs (8, 9, 37). A superposition of yeast STT3 and substrate-bound PglB (18) (fig. S7, A and B) revealed that functionally important, conserved residues and motifs have a similar spatial arrangement (Fig. 4A). These include the D-X-D/E motif involved in coordinating the catalytic metal ion, an aspartate (D47 in yeast STT3 corresponding to D56 in PglB) that binds both the metal ion and the carboxamide group of the acceptor asparagine, the WWD motif providing hydrogen bond contacts to the β-hydroxyl group of the +2 serine/threonine of the acceptor sequon, and the K586 residue of the so-called “DK motif” that contributes additional contacts to the +2 serine/threonine of the acceptor peptide (Fig. 4A). Residues shown to be involved in LLO binding to PglB (18) are also conserved. These include an essential arginine (R404 in yeast STT3) that interacts with the pyrophosphate group of LLO, and a tyrosine residue (Y521 in yeast STT3) that forms a hydrogen bond with the *N*-acetyl group of the reducing-end GlcNAc moiety (Fig. 4B).

Comparing the electrostatic surface potential maps of the peptide-binding pocket of PglB with the predicted binding pocket of yeast STT3 provides a molecular explanation for some of the differences in acceptor peptide specificities of bacterial and eukaryotic OSTs (Fig. 4C) (38, 39). In PglB, an arginine residue (R331) was found to interact with a negatively charged side chain (Asp or Glu) at the -2 position of the

sequon. This arginine is conserved in bacterial ssOST enzymes and correlates with an extended sequon requirement (DxNxS/T) for bacterial *N*-glycosylation (8, 38) (Fig. 4C). In contrast, no such requirement is present at the -2 position of acceptor sequons in eukaryotes, and no positively charged residue is present in yeast STT3 where the -2 side chain is expected to bind (Fig. 4C). Instead, a larger cavity providing space for more voluminous side chains at the -2 position of sequons is observed, which is in line with previous findings that human OSTs display an increased glycosylation efficiency for substrates with aromatic residues at the -2 position (39). At the expected binding site of the -1 residue of the sequon, we found a “knob” in the yeast STT3 structure formed by the side chain of E45 (Fig. 4C). The reduced space may result in a more efficient binding of sequons with smaller side chains at the -1 position.

A groove between TM6 and TM11 of STT3 indicates where the dolichol tail of the LLO will likely bind (Fig. 3A). A similar groove was demonstrated to represent the polyprenyl-binding site in PglB, with four prenyl units ordered in the X-ray structure (18). Given their locations, it is conceivable that in yeast OST, the partially ordered TM9 of STT3 and the disordered TM1 of OST3 might interact with bound LLO. Our results suggest that PglB and yeast STT3 share a common substrate recognition and glycan transfer mechanism despite differing in their substrate specificities.

Because glycosylation is an essential process in yeast, abolishing OST function prevents cell growth (40). To validate the importance of the residues identified in the yeast OST structure, we therefore used a tester strain that expressed the ssOST enzyme *LmSTT3D* from *Leishmania major* (41) and generated six chromosomal mutations in the *STT3* locus (D47A, D166A, E168Q, E350A, R404A, K586A). These mutations were predicted to affect metal binding, peptide binding, or LLO binding, but did not affect OST complex stability (fig. S7C). With the exception of K586A, which resulted in a temperature-sensitive phenotype, all generated mutations prevented growth in the absence of *LmSTT3D* (Fig. 4D). We conclude that these mutations impaired the catalytic activity of the STT3 subunit and thus of OST function in vivo (40).

### Acceptor polypeptide delivery by the redox chaperone OST3 or the translocon

Higher eukaryotes express two paralogs of the catalytic STT3 subunit, termed STT3A and STT3B, which are part of distinct OST complexes. STT3A-containing OST associates with the translocon, thereby bringing native peptide chains entering the ER into close proximity of the glycosylation machinery (14, 15, 42–46). In contrast, STT3B-containing OST complexes (including the yeast enzyme) are standalone units that contain either an OST3 or OST6 subunit (homologous to

the mammalian Tusc3 and MagT1, respectively), neither of which is present in translocon-associated OST. The luminal domains of OST6 and Tusc3 have oxidoreductase activity and feature a thioredoxin fold. Both can form disulfide bonds with acceptor proteins, but OST3- or OST6-containing complexes process different subsets of disulfide-forming acceptor polypeptides in vivo (19, 47, 48). In our apo structure of yeast OST, the oxidoreductase domain of OST3 appears disordered and is likely flexible, which may allow diverse and transiently bound OST substrates to be efficiently glycosylated.

Recent studies reported a mammalian ribosome-translocon-OST supercomplex visualized by electron cryo-tomography (14, 15). We docked our high-resolution OST structure into this tomography map and found a remarkably good fit with the portion of the density previously assigned to OST (14, 15) (Fig. 5A). This provides additional evidence that yeast and mammalian OST complexes share a common architecture. Additional density in the tomography map near the N terminus of SWP1 (Fig. 5A) probably corresponds to an N-terminal extension of ~300 residues of the mammalian SWP1 homolog Ribophorin II compared to yeast SWP1 (11).

Our docking analysis suggests that the active site of STT3A faces the heterotrimeric translocon complex (49) (Fig. 5, A and B). The contact point between OST and the translocon comprises the region corresponding to TM2-TM4 of the OST3 subunit in yeast OST (Fig. 5B). However, OST3 is not present in the mammalian, translocon-associated OST (43). Instead, recent studies suggest that the OST-translocon interaction is mediated by the DC2 protein (50). A sequence comparison reveals that the 3 TM helices of DC2 share a striking similarity with TM2-TM4 of OST3 (fig. S8), lending further weight to the docking model shown in Fig. 5. Multi-subunit OSTs may thus be described as a modular assembly of an OST core complex formed by seven subunits (STT3, OST1, OST2, OST4, OST5, WBP1 and SWP1), which associates either with oxidoreductases (OST3/6 or Tusc3/MagT1) or with the translocon.

The arrangement shown in Fig. 5C allows unhindered access of the polypeptide substrate from the luminal exit of the translocon to the catalytic center of OST. The observed separation of ~40 Å is in agreement with earlier distance estimates based on the minimal polypeptide length of 65 residues between the peptidyltransferase center of the ribosome and the first possible glycosylation site (51). It is also in line with the previously reported minimal distance of 30-40 Å between the OST active site and the ER membrane deduced from evaluating the glycosylation of integral membrane proteins (52).

## Conclusions: Functions of OST subunits

Organisms encoding multi-subunit OSTs are found to

glycosylate a significantly expanded range of protein substrates compared to single-subunit OST enzymes (53). This implies that the auxiliary subunits increase the efficiency of the catalytic STT3 core by contributing to substrate acquisition or by affecting the folding of acceptor proteins. Our yeast OST structure suggests that OST2, OST4 and OST5, which contain mostly TM helices, have a scaffolding function and are thus important for complex stability, but without directly contacting the substrates (11, 54). The luminal domains of OST1, SWP1, and WBP1 may also have structural roles by stabilizing the STT3 subunit conformation. In addition, some of them may directly interact with the substrates. Our structure revealed a cavity ranging from the active site of STT3 to the WBP1 and SWP1 subunits, just above the membrane boundary (Fig. 5C). The cavity is lined by the highly ordered N-glycan attached to N539 of STT3 and is sufficiently large to accommodate the glycan moiety of bound LLO substrate. It is conceivable that WBP1, SWP1, and possibly even the ordered N-glycan contribute to the recognition of the lipid-linked  $\text{Glc}_3\text{Man}_6\text{GlcNAc}_2$  moiety and thus help define the preference for an LLO substrate containing terminal,  $\alpha$ -1,2-linked glucose (55). For OST1, it is worth speculating that given its proximity to the peptide-binding pocket of STT3, it might interact with acceptor proteins and influence their folding. The active site of the OST complex faces the peptide-binding OST3/6 subunits or the translocon, both of which present polypeptide substrates in an unfolded state. This arrangement favors N-glycosylation over the competing folding reactions and thus extends the substrate range of OST.

## REFERENCES AND NOTES

1. F. Schwarz, M. Aebi, Mechanisms and principles of N-linked protein glycosylation. *Curr. Opin. Struct. Biol.* **21**, 576–582 (2011). [doi:10.1016/j.sbi.2011.08.005](https://doi.org/10.1016/j.sbi.2011.08.005) Medline
2. C. G. Gahmberg, M. Tolvanen, Why mammalian cell surface proteins are glycoproteins. *Trends Biochem. Sci.* **21**, 308–311 (1996). [doi:10.1016/S0968-0004\(96\)10034-7](https://doi.org/10.1016/S0968-0004(96)10034-7) Medline
3. A. Grigorian, H. Mkhikian, M. Demetriou, Interleukin-2, interleukin-7, T cell-mediated autoimmunity, and N-glycosylation. *Ann. N.Y. Acad. Sci.* **1253**, 49–57 (2012). [doi:10.1111/j.1749-6632.2011.06391.x](https://doi.org/10.1111/j.1749-6632.2011.06391.x) Medline
4. A. Helenius, M. Aebi, Roles of N-linked glycans in the endoplasmic reticulum. *Annu. Rev. Biochem.* **73**, 1019–1049 (2004). [doi:10.1146/annurev.biochem.73.011303.073752](https://doi.org/10.1146/annurev.biochem.73.011303.073752) Medline
5. M. Chavan, W. Lennarz, The molecular basis of coupling of translocation and N-glycosylation. *Trends Biochem. Sci.* **31**, 17–20 (2006). [doi:10.1016/j.tibs.2005.11.010](https://doi.org/10.1016/j.tibs.2005.11.010) Medline
6. R. Kornfeld, S. Kornfeld, Assembly of asparagine-linked oligosaccharides. *Annu. Rev. Biochem.* **54**, 631–664 (1985). [doi:10.1146/annurev.bi.54.070185.003215](https://doi.org/10.1146/annurev.bi.54.070185.003215) Medline
7. R. D. Marshall, Glycoproteins. *Annu. Rev. Biochem.* **41**, 673–702 (1972). [doi:10.1146/annurev.bi.41.070172.003325](https://doi.org/10.1146/annurev.bi.41.070172.003325) Medline

8. C. Lizak, S. Gerber, S. Numao, M. Aebi, K. P. Locher, X-ray structure of a bacterial oligosaccharyltransferase. *Nature* **474**, 350–355 (2011). [doi:10.1038/nature10151](https://doi.org/10.1038/nature10151) [Medline](#)
9. S. Matsumoto, A. Shimada, J. Nyirenda, M. Igura, Y. Kawano, D. Kohda, Crystal structures of an archaeal oligosaccharyltransferase provide insights into the catalytic cycle of N-linked protein glycosylation. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 17868–17873 (2013). [doi:10.1073/pnas.1309777110](https://doi.org/10.1073/pnas.1309777110) [Medline](#)
10. A. S. Ramírez, J. Boilevin, R. Biswas, B. H. Gan, D. Janser, M. Aebi, T. Darbre, J.-L. Reymond, K. P. Locher, Characterization of the single-subunit oligosaccharyltransferase STT3A from *Trypanosoma brucei* using synthetic peptides and lipid-linked oligosaccharide analogs. *Glycobiology* **27**, 525–535 (2017). [doi:10.1093/glycob/cwx017](https://doi.org/10.1093/glycob/cwx017) [Medline](#)
11. D. J. Kelleher, R. Gilmore, An evolving view of the eukaryotic oligosaccharyltransferase. *Glycobiology* **16**, 47R–62R (2006). [doi:10.1093/glycob/cwj066](https://doi.org/10.1093/glycob/cwj066) [Medline](#)
12. N. A. Cherepanova, R. Gilmore, Mammalian cells lacking either the cotranslational or posttranslational oligosaccharyltransferase complex display substrate-dependent defects in asparagine linked glycosylation. *Sci. Rep.* **6**, 20946 (2016). [doi:10.1038/srep20946](https://doi.org/10.1038/srep20946) [Medline](#)
13. S. Mueller, A. Wahlander, N. Selevsek, C. Otto, E. M. Ngwa, K. Poljak, A. D. Frey, M. Aebi, R. Gauss, Protein degradation corrects for imbalanced subunit stoichiometry in OST complex assembly. *Mol. Biol. Cell* **26**, 2596–2608 (2015). [doi:10.1091/mbc.F15-03-0168](https://doi.org/10.1091/mbc.F15-03-0168) [Medline](#)
14. S. Pfeffer, L. Burbaum, P. Unverdorben, M. Pech, Y. Chen, R. Zimmermann, R. Beckmann, F. Förster, Structure of the native Sec61 protein-conducting channel. *Nat. Commun.* **6**, 8403 (2015). [doi:10.1038/ncomms9403](https://doi.org/10.1038/ncomms9403) [Medline](#)
15. S. Pfeffer, J. Dudek, M. Schaffer, B. G. Ng, S. Albert, J. M. Plitzko, W. Baumeister, R. Zimmermann, H. H. Freeze, B. D. Engel, F. Förster, Dissecting the molecular organization of the translocon-associated protein complex. *Nat. Commun.* **8**, 14516 (2017). [doi:10.1038/ncomms14516](https://doi.org/10.1038/ncomms14516) [Medline](#)
16. H. Li, M. Chavan, H. Schindelin, W. J. Lennarz, H. Li, Structure of the oligosaccharyl transferase complex at 12 Å resolution. *Structure* **16**, 432–440 (2008). [doi:10.1016/j.str.2007.12.013](https://doi.org/10.1016/j.str.2007.12.013) [Medline](#)
17. M. Igura, N. Maita, J. Kamishikiryo, M. Yamada, T. Obita, K. Maenaka, D. Kohda, Structure-guided identification of a new catalytic motif of oligosaccharyltransferase. *EMBO J.* **27**, 234–243 (2008). [doi:10.1038/sj.emboj.7601940](https://doi.org/10.1038/sj.emboj.7601940) [Medline](#)
18. M. Napiórkowska, J. Boilevin, T. Sovdat, T. Darbre, J.-L. Reymond, M. Aebi, K. P. Locher, Molecular basis of lipid-linked oligosaccharide recognition and processing by bacterial oligosaccharyltransferase. *Nat. Struct. Mol. Biol.* **24**, 1100–1106 (2017). [doi:10.1038/nsmb.3491](https://doi.org/10.1038/nsmb.3491) [Medline](#)
19. B. L. Schulz, C. U. Stirnimann, J. P. A. Grimshaw, M. S. Brozzo, F. Fritsch, E. Mohorko, G. Capitani, R. Glockshuber, M. G. Grütter, M. Aebi, Oxidoreductase activity of oligosaccharyltransferase subunits Ost3p and Ost6p defines site-specific glycosylation efficiency. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 11061–11066 (2009). [doi:10.1073/pnas.0812515106](https://doi.org/10.1073/pnas.0812515106) [Medline](#)
20. E. Mohorko, R. L. Owen, G. Malojić, M. S. Brozzo, M. Aebi, R. Glockshuber, Structural basis of substrate specificity of human oligosaccharyl transferase subunit N33/Tusc3 and its role in regulating protein N-glycosylation. *Structure* **22**, 590–601 (2014). [doi:10.1016/j.str.2014.02.013](https://doi.org/10.1016/j.str.2014.02.013) [Medline](#)
21. D. F. Zielinska, F. Gnad, J. R. Wiśniewski, M. Mann, Precision mapping of an in vivo N-glycoproteome reveals rigid topological and sequence constraints. *Cell* **141**, 897–907 (2010). [doi:10.1016/j.cell.2010.04.012](https://doi.org/10.1016/j.cell.2010.04.012) [Medline](#)
22. S. Gerber, C. Lizak, G. Michaud, M. Bucher, T. Darbre, M. Aebi, J.-L. Reymond, K. P. Locher, Mechanism of bacterial oligosaccharyltransferase: In vitro quantification of sequon binding and catalysis. *J. Biol. Chem.* **288**, 8849–8861 (2013). [doi:10.1074/jbc.M112.445940](https://doi.org/10.1074/jbc.M112.445940) [Medline](#)
23. M. M. Thunnissen, P. Nordlund, J. Z. Haeggström, Crystal structure of human leukotriene A(4) hydrolase, a bifunctional enzyme in inflammation. *Nat. Struct. Biol.* **8**, 131–135 (2001). [doi:10.1038/84117](https://doi.org/10.1038/84117) [Medline](#)
24. A. Mpakali, E. Saridakis, K. Harlos, Y. Zhao, P. Kokkala, D. Georgiadis, P. Giastas, A. Papakyriakou, E. Stratikos, Ligand-induced conformational change of insulin-regulated aminopeptidase: Insights on catalytic mechanism and active site plasticity. *J. Med. Chem.* **60**, 2963–2972 (2017). [doi:10.1021/acs.jmedchem.6b01890](https://doi.org/10.1021/acs.jmedchem.6b01890) [Medline](#)
25. H. Kim, G. von Heijne, I. Nilsson, Membrane topology of the STT3 subunit of the oligosaccharyl transferase complex. *J. Biol. Chem.* **280**, 20261–20267 (2005). [doi:10.1074/jbc.M412213200](https://doi.org/10.1074/jbc.M412213200) [Medline](#)
26. P. Lara, K. Öjemalm, J. Reithinger, A. Holgado, Y. Maojun, A. Hammed, D. Mattle, H. Kim, I. Nilsson, Refined topology model of the STT3/Stt3 protein subunit of the oligosaccharyltransferase complex. *J. Biol. Chem.* **292**, 11349–11360 (2017). [doi:10.1074/jbc.M117.779421](https://doi.org/10.1074/jbc.M117.779421) [Medline](#)
27. S. Korolev, T. Skarina, E. Evdokimova, S. Beasley, A. Edwards, A. Joachimiak, A. Savchenko, Crystal structure of glutamine amidotransferase from *Thermotoga maritima*. *Proteins* **49**, 420–422 (2002). [doi:10.1002/prot.10161](https://doi.org/10.1002/prot.10161) [Medline](#)
28. M. Taschner, K. Weber, A. Mourão, M. Vetter, M. Awasthi, M. Stiegler, S. Bhogaraju, E. Lorentzen, Intraflagellar transport proteins 172, 80, 57, 54, 38, and 20 form a stable tubulin-binding IFT-B2 complex. *EMBO J.* **35**, 773–790 (2016). [doi:10.15252/emboj.201593164](https://doi.org/10.15252/emboj.201593164) [Medline](#)
29. V. Krishnan, K. Ponnuraj, Y. Xu, K. Macon, J. E. Volanakis, S. V. L. Narayana, The crystal structure of cobra venom factor, a cofactor for C3- and C5-convertase C5b. *Structure* **17**, 611–619 (2009). [doi:10.1016/j.str.2009.01.015](https://doi.org/10.1016/j.str.2009.01.015) [Medline](#)
30. N. S. Laursen, K. R. Andersen, I. Braren, E. Spillner, L. Sottrup-Jensen, G. R. Andersen, Substrate recognition by complement convertases revealed in the C5-cobra venom factor complex. *EMBO J.* **30**, 606–616 (2011). [doi:10.1038/emboj.2010.341](https://doi.org/10.1038/emboj.2010.341) [Medline](#)
31. R. Pathak, T. L. Hendrickson, B. Imperiali, Sulfhydryl modification of the yeast Wbp1p inhibits oligosaccharyl transferase activity. *Biochemistry* **34**, 4179–4185 (1995). [doi:10.1021/bi00013a005](https://doi.org/10.1021/bi00013a005) [Medline](#)
32. Q. Yan, G. D. Prestwich, W. J. Lennarz, The Ost1p subunit of yeast oligosaccharyl transferase recognizes the peptide glycosylation site sequence, -Asn-X-Ser-/Thr-. *J. Biol. Chem.* **274**, 5021–5025 (1999). [doi:10.1074/jbc.274.8.5021](https://doi.org/10.1074/jbc.274.8.5021) [Medline](#)
33. C. M. Wilson, C. Kraft, C. Duggan, N. Ismail, S. G. Crawshaw, S. High, Ribophorin I associates with a subset of membrane proteins after their integration at the sec61 translocon. *J. Biol. Chem.* **280**, 4195–4206 (2005). [doi:10.1074/jbc.M410329200](https://doi.org/10.1074/jbc.M410329200) [Medline](#)
34. T. Schallus, C. Jaech, K. Fehér, A. S. Palma, Y. Liu, J. C. Simpson, M. Mackeen, G. Stier, T. J. Gibson, T. Feizi, T. Pieler, C. Muhle-Goll, Malectin: A novel carbohydrate-binding protein of the endoplasmic reticulum and a candidate player in the early steps of protein N-glycosylation. *Mol. Biol. Cell* **19**, 3404–3414 (2008). [doi:10.1091/mbc.F08-04-0354](https://doi.org/10.1091/mbc.F08-04-0354) [Medline](#)
35. S.-Y. Qin, D. Hu, K. Matsumoto, K. Takeda, N. Matsumoto, Y. Yamaguchi, K. Yamamoto, Malectin forms a complex with ribophorin I for enhanced association with misfolded glycoproteins. *J. Biol. Chem.* **287**, 38080–38089 (2012). [doi:10.1074/jbc.M112.394288](https://doi.org/10.1074/jbc.M112.394288) [Medline](#)
36. G. Li, Q. Yan, A. Nita-Lazar, R. S. Haltiwanger, W. J. Lennarz, Studies on the N-glycosylation of the subunits of oligosaccharyl transferase in *Saccharomyces cerevisiae*. *J. Biol. Chem.* **280**, 1864–1871 (2005). [doi:10.1074/jbc.M410969200](https://doi.org/10.1074/jbc.M410969200) [Medline](#)
37. C. Lizak, S. Gerber, G. Michaud, M. Schubert, Y.-Y. Fan, M. Bucher, T. Darbre, M. Aebi, J.-L. Reymond, K. P. Locher, Unexpected reactivity and mechanism of carboxamide activation in bacterial N-linked protein glycosylation. *Nat.*

- Commun.* **4**, 2627 (2013). [doi:10.1038/ncomms3627](https://doi.org/10.1038/ncomms3627) [Medline](#)
38. M. Kowarik, N. M. Young, S. Numao, B. L. Schulz, I. Hug, N. Callewaert, D. C. Mills, D. C. Watson, M. Hernandez, J. F. Kelly, M. Wacker, M. Aebi, Definition of the bacterial N-glycosylation site consensus sequence. *EMBO J.* **25**, 1957–1966 (2006). [doi:10.1038/sj.emboj.7601087](https://doi.org/10.1038/sj.emboj.7601087) [Medline](#)
  39. A. N. Murray, W. Chen, A. Antonopoulos, S. R. Hanson, R. L. Wiseman, A. Dell, S. M. Haslam, D. L. Powers, E. T. Powers, J. W. Kelly, Enhanced aromatic sequons increase oligosaccharyltransferase glycosylation efficiency and glycan homogeneity. *Chem. Biol.* **22**, 1052–1062 (2015). [doi:10.1016/j.chembiol.2015.06.017](https://doi.org/10.1016/j.chembiol.2015.06.017) [Medline](#)
  40. Q. Yan, W. J. Lennarz, Studies on the function of oligosaccharyl transferase subunits. Stt3p is directly involved in the glycosylation process. *J. Biol. Chem.* **277**, 47692–47700 (2002). [doi:10.1074/jbc.M208136200](https://doi.org/10.1074/jbc.M208136200) [Medline](#)
  41. F. P. Nasab, B. L. Schulz, F. Gamarró, A. J. Parodi, M. Aebi, All in one: *Leishmania major* STT3 proteins substitute for the whole oligosaccharyltransferase complex in *Saccharomyces cerevisiae*. *Mol. Biol. Cell* **19**, 3758–3768 (2008). [doi:10.1091/mbc.F08-05-0467](https://doi.org/10.1091/mbc.F08-05-0467) [Medline](#)
  42. M. Chavan, A. Yan, W. J. Lennarz, Subunits of the translocon interact with components of the oligosaccharyl transferase complex. *J. Biol. Chem.* **280**, 22917–22924 (2005). [doi:10.1074/jbc.M502858200](https://doi.org/10.1074/jbc.M502858200) [Medline](#)
  43. N. A. Cherepanova, S. Shrimal, R. Gilmore, Oxidoreductase activity is necessary for N-glycosylation of cysteine-proximal acceptor sites in glycoproteins. *J. Cell Biol.* **206**, 525–539 (2014). [doi:10.1083/jcb.201404083](https://doi.org/10.1083/jcb.201404083) [Medline](#)
  44. Y. Harada, H. Li, H. Li, W. J. Lennarz, Oligosaccharyltransferase directly binds to ribosome at a location near the translocon-binding site. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 6945–6949 (2009). [doi:10.1073/pnas.0812489106](https://doi.org/10.1073/pnas.0812489106) [Medline](#)
  45. T. Shibatani, L. L. David, A. L. McCormack, K. Frueh, W. R. Skach, Proteomic analysis of mammalian oligosaccharyltransferase reveals multiple subcomplexes that contain Sec61, TRAP, and two potential new subunits. *Biochemistry* **44**, 5982–5992 (2005). [doi:10.1021/bi047328f](https://doi.org/10.1021/bi047328f) [Medline](#)
  46. T. A. Rapoport, L. Li, E. Park, Structural and mechanistic insights into protein translocation. *Annu. Rev. Cell Dev. Biol.* **33**, 369–390 (2017). [doi:10.1146/annurev-cellbio-100616-060439](https://doi.org/10.1146/annurev-cellbio-100616-060439) [Medline](#)
  47. L. F. Zacchi, B. L. Schulz, SWATH-MS glycoproteomics reveals consequences of defects in the glycosylation machinery. *Mol. Cell. Proteomics* **15**, 2435–2447 (2016). [doi:10.1074/mcp.M115.056366](https://doi.org/10.1074/mcp.M115.056366) [Medline](#)
  48. K. Poljak, N. Selevsek, E. Ngwa, J. Grossmann, M. E. Losfeld, M. Aebi, Quantitative profiling of N-linked glycosylation machinery in yeast *Saccharomyces cerevisiae*. *Mol. Cell. Proteomics* mcp.RA117.000096 (2017). [doi:10.1074/mcp.RA117.000096](https://doi.org/10.1074/mcp.RA117.000096) [Medline](#)
  49. B. Van den Berg, W. M. Clemons Jr., I. Collinson, Y. Modis, E. Hartmann, S. C. Harrison, T. A. Rapoport, X-ray structure of a protein-conducting channel. *Nature* **427**, 36–44 (2004). [doi:10.1038/nature02218](https://doi.org/10.1038/nature02218) [Medline](#)
  50. S. Shrimal, N. A. Cherepanova, R. Gilmore, DC2 and KCP2 mediate the interaction between the oligosaccharyltransferase and the ER translocon. *J. Cell Biol.* **216**, 3625–3638 (2017). [doi:10.1083/jcb.201702159](https://doi.org/10.1083/jcb.201702159) [Medline](#)
  51. M. Kowarik, S. Küng, B. Martoglio, A. Helenius, Protein folding during cotranslational translocation in the endoplasmic reticulum. *Mol. Cell* **10**, 769–778 (2002). [doi:10.1016/S1097-2765\(02\)00685-8](https://doi.org/10.1016/S1097-2765(02)00685-8) [Medline](#)
  52. I. M. Nilsson, G. von Heijne, Determination of the distance between the oligosaccharyltransferase active site and the endoplasmic reticulum membrane. *J. Biol. Chem.* **268**, 5798–5801 (1993). [Medline](#)
  53. A. Dell, A. Galadari, F. Sastre, P. Hitchen, Similarities and differences in the glycosylation mechanisms in prokaryotes and eukaryotes. *Int. J. Microbiol.* **2010**, 148178 (2010). [doi:10.1155/2010/148178](https://doi.org/10.1155/2010/148178) [Medline](#)
  54. A. Dumax-Vorzet, P. Roboti, S. High, OST4 is a subunit of the mammalian oligosaccharyltransferase required for efficient N-glycosylation. *J. Cell Sci.* **126**, 2595–2606 (2013). [doi:10.1242/jcs.115410](https://doi.org/10.1242/jcs.115410) [Medline](#)
  55. P. Burda, M. Aebi, The ALG10 locus of *Saccharomyces cerevisiae* encodes the alpha-1,2 glucosyltransferase of the endoplasmic reticulum: The terminal glucose of the lipid-linked oligosaccharide is required for efficient N-linked glycosylation. *Glycobiology* **8**, 455–462 (1998). [doi:10.1093/glycob/8.5.455](https://doi.org/10.1093/glycob/8.5.455) [Medline](#)
  56. C. Guthrie, R. G. Fink, Eds., *Guide to Yeast Genetics and Molecular Biology* (Elsevier, 1991). [Medline](#)
  57. M. Knop, K. Siegers, G. Pereira, W. Zachariae, B. Winsor, K. Nasmyth, E. Schiebel, Epitope tagging of yeast genes using a PCR-based strategy: More tags and improved practical routines. *Yeast* **15**, 963–972 (1999). [doi:10.1002/\(SICI\)1097-0061\(199907\)15:10B<963::AID-YEA399>3.0.CO;2-W](https://doi.org/10.1002/(SICI)1097-0061(199907)15:10B<963::AID-YEA399>3.0.CO;2-W) [Medline](#)
  58. U. Guedener, J. Heinisch, G. J. Koehler, D. Voss, J. H. Hegemann, A second set of loxP marker cassettes for Cre-mediated multiple gene knockouts in budding yeast. *Nucleic Acids Res.* **30**, e23 (2002). [doi:10.1093/nar/30.6.e23](https://doi.org/10.1093/nar/30.6.e23) [Medline](#)
  59. M. Schwarz, R. Knauer, L. Lehle, Yeast oligosaccharyltransferase consists of two functionally distinct sub-complexes, specified by either the Ost3p or Ost6p subunit. *FEBS Lett.* **579**, 6564–6568 (2005). [doi:10.1016/j.febslet.2005.10.063](https://doi.org/10.1016/j.febslet.2005.10.063) [Medline](#)
  60. R. Knauer, L. Lehle, The oligosaccharyltransferase complex from *Saccharomyces cerevisiae*. Isolation of the OST6 gene, its synthetic interaction with OST3, and analysis of the native complex. *J. Biol. Chem.* **274**, 17249–17256 (1999). [doi:10.1074/jbc.274.24.17249](https://doi.org/10.1074/jbc.274.24.17249) [Medline](#)
  61. R. S. Sikorski, P. Hieter, A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae*. *Genetics* **122**, 19–27 (1989). [Medline](#)
  62. M. K. Chee, S. B. Haase, New and redesigned pRS plasmid shuttle vectors for genetic manipulation of *Saccharomyces cerevisiae*. *G3* **2**, 515–526 (2012). [doi:10.1534/g3.111.001917](https://doi.org/10.1534/g3.111.001917) [Medline](#)
  63. J. D. Boeke, J. Trueheart, G. Natsoulis, G. R. Fink, 5-Fluoroorotic acid as a selective agent in yeast molecular genetics. *Methods Enzymol.* **154**, 164–175 (1987). [doi:10.1016/0076-6879\(87\)54076-9](https://doi.org/10.1016/0076-6879(87)54076-9) [Medline](#)
  64. T. K. Ritchie, Y. V. Grinkova, T. H. Bayburt, I. G. Denisov, J. K. Zolnerchik, W. M. Atkins, S. G. Sligar, Reconstitution of membrane proteins in phospholipid bilayer nanodiscs. *Methods Enzymol.* **464**, 211–231 (2009). [doi:10.1016/S0076-6879\(09\)64011-8](https://doi.org/10.1016/S0076-6879(09)64011-8) [Medline](#)
  65. H. Schagger, Tricine-SDS-PAGE. *Nat. Protoc.* **1**, 16–22 (2006). [doi:10.1038/nprot.2006.4](https://doi.org/10.1038/nprot.2006.4) [Medline](#)
  66. C. A. Schneider, W. S. Rasband, K. W. Eliceiri, NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **9**, 671–675 (2012). [doi:10.1038/nmeth.2089](https://doi.org/10.1038/nmeth.2089) [Medline](#)
  67. D. N. Mastronarde, Automated electron microscope tomography using robust prediction of specimen movements. *J. Struct. Biol.* **152**, 36–51 (2005). [doi:10.1016/j.jsb.2005.07.007](https://doi.org/10.1016/j.jsb.2005.07.007) [Medline](#)
  68. S. Q. Zheng, E. Palovcak, J.-P. Armache, K. A. Verba, Y. Cheng, D. A. Agard, MotionCor2: Anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017). [doi:10.1038/nmeth.4193](https://doi.org/10.1038/nmeth.4193) [Medline](#)
  69. S. H. W. Scheres, A Bayesian view on cryo-EM structure determination. *J. Mol. Biol.* **415**, 406–418 (2012). [doi:10.1016/j.jmb.2011.11.010](https://doi.org/10.1016/j.jmb.2011.11.010) [Medline](#)
  70. D. Kimanius, B. O. Forsberg, S. H. Scheres, E. Lindahl, Accelerated cryo-EM structure determination with parallelisation using GPUs in RELION-2. *eLife* **5**, e18722 (2016). [doi:10.7554/eLife.18722](https://doi.org/10.7554/eLife.18722) [Medline](#)

71. K. Zhang, Gctf: Real-time CTF determination and correction. *J. Struct. Biol.* **193**, 1–12 (2016). [doi:10.1016/j.jsb.2015.11.003](https://doi.org/10.1016/j.jsb.2015.11.003) [Medline](#)
72. L. Urnavicius, K. Zhang, A. G. Diamant, C. Motz, M. A. Schlager, M. Yu, N. A. Patel, C. V. Robinson, A. P. Carter, The structure of the dynactin complex and its interaction with dynein. *Science* **347**, 1441–1446 (2015). [doi:10.1126/science.aaa4080](https://doi.org/10.1126/science.aaa4080) [Medline](#)
73. X. C. Bai, E. Rajendra, G. Yang, Y. Shi, S. H. W. Scheres, Sampling the conformational space of the catalytic subunit of human  $\gamma$ -secretase. *eLife* **4**, e11182 (2015). [doi:10.7554/eLife.11182](https://doi.org/10.7554/eLife.11182) [Medline](#)
74. A. Punjani, J. L. Rubinstein, D. J. Fleet, M. A. Brubaker, cryoSPARC: Algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017). [doi:10.1038/nmeth.4169](https://doi.org/10.1038/nmeth.4169) [Medline](#)
75. A. Kucukelbir, F. J. Sigworth, H. D. Tagare, Quantifying the local resolution of cryo-EM density maps. *Nat. Methods* **11**, 63–65 (2014). [doi:10.1038/nmeth.2727](https://doi.org/10.1038/nmeth.2727) [Medline](#)
76. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of Coot. *Acta Crystallogr. D* **66**, 486–501 (2010). [doi:10.1107/S0907444910007493](https://doi.org/10.1107/S0907444910007493) [Medline](#)
77. P. D. Adams, P. V. Afonine, G. Bunkóczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L.-W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, P. H. Zwart, PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66**, 213–221 (2010). [doi:10.1107/S0907444909052925](https://doi.org/10.1107/S0907444909052925) [Medline](#)
78. A. Krogh, B. Larsson, G. von Heijne, E. L. Sonnhammer, Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. *J. Mol. Biol.* **305**, 567–580 (2001). [doi:10.1006/jmbi.2000.4315](https://doi.org/10.1006/jmbi.2000.4315) [Medline](#)
79. D. W. A. Buchan, F. Minneci, T. C. O. Nugent, K. Bryson, D. T. Jones, Scalable web services for the PSIPRED Protein Analysis Workbench. *Nucleic Acids Res.* **41**, W349–W357 (2013). [doi:10.1093/nar/gkt381](https://doi.org/10.1093/nar/gkt381) [Medline](#)
80. E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin, UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004). [doi:10.1002/jcc.20084](https://doi.org/10.1002/jcc.20084) [Medline](#)
81. C. Notredame, D. G. Higgins, J. Heringa, T-Coffee: A novel method for fast and accurate multiple sequence alignment. *J. Mol. Biol.* **302**, 205–217 (2000). [doi:10.1006/jmbi.2000.4042](https://doi.org/10.1006/jmbi.2000.4042) [Medline](#)
82. E. Krissinel, K. Henrick, Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774–797 (2007). [doi:10.1016/j.jmb.2007.05.022](https://doi.org/10.1016/j.jmb.2007.05.022) [Medline](#)
83. L. Holm, P. Rosenström, Dali server: Conservation mapping in 3D. *Nucleic Acids Res.* **38**, W545–W549 (2010). [doi:10.1093/nar/gkq366](https://doi.org/10.1093/nar/gkq366) [Medline](#)
84. X. Robert, P. Gouet, Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* **42**, W320–W324 (2014). [doi:10.1093/nar/gku316](https://doi.org/10.1093/nar/gku316) [Medline](#)

## ACKNOWLEDGMENTS

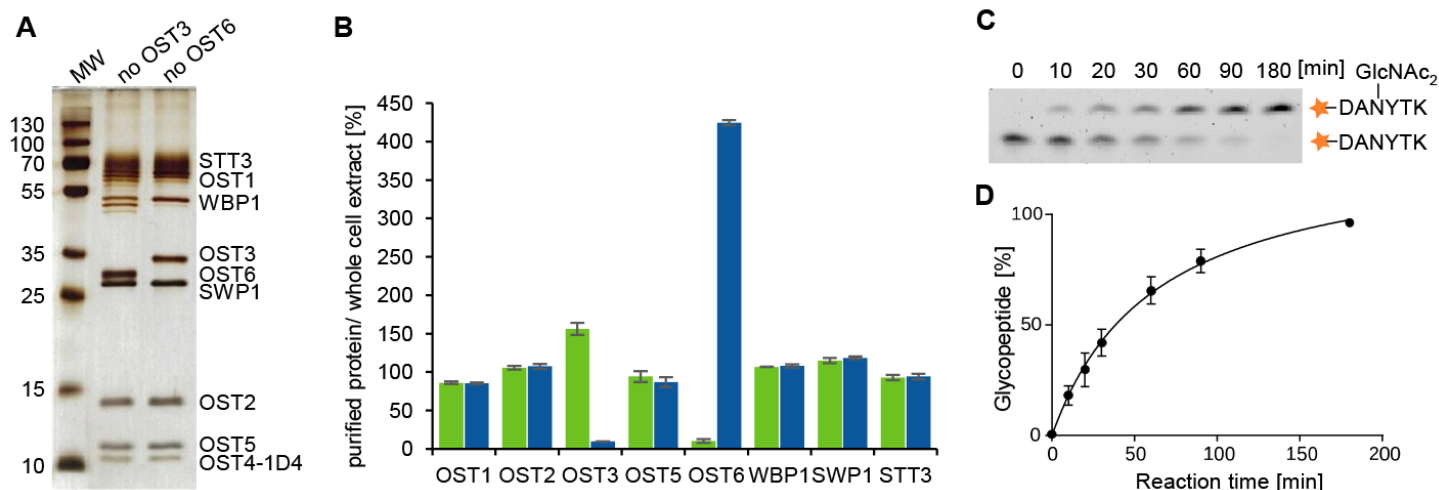
We thank J. Boilevin, T. Darbre and J.-L. Reymond for providing the LLO analog. This work was supported by the Swiss National Science Foundation Sinergia programs TransGlyco (CRSII3\_147632) and GlycoStart (CRSII5\_173709). R.W. acknowledges support from the ETH postdoctoral fellowship program. Cryo-EM data was collected at the electron microscopy facility of ETH Zurich (ScopeM). We thank P. Tittmann for technical support. Atomic coordinates of the de novo built yeast OST model have been deposited in the PDB under accession number 6EZN. The three-dimensional cryo-EM density maps have been deposited in the Electron Microscopy Data Bank (EMDB) under accession numbers EMD-4161. All data needed to evaluate the conclusions of this

paper are provided either in the paper or in the supplementary materials. The authors declare no conflicts of interest. Author contributions: K.P.L. and M.A. designed the project. J.E. generated the yeast strain, developed initial purification protocols, and purified OST for mass-spectrometry analyses. R.W. purified OST for structural studies, reconstituted OST in nanodiscs and performed in vitro glycosylation assays. E.M.N. carried out in vivo mutational analysis in yeast. J.K. and R.W. performed negative stain EM experiments and prepared cryo-EM grids. J.K. collected cryo-EM data and performed data analysis. R.W. built the OST model and performed model refinement. K.P.L. revised the model. R.W., J.K. and K.P.L. analyzed the structure. R.W. and K.P.L. wrote the manuscript with the help of J.K. and J.E.; all authors contributed to its revision.

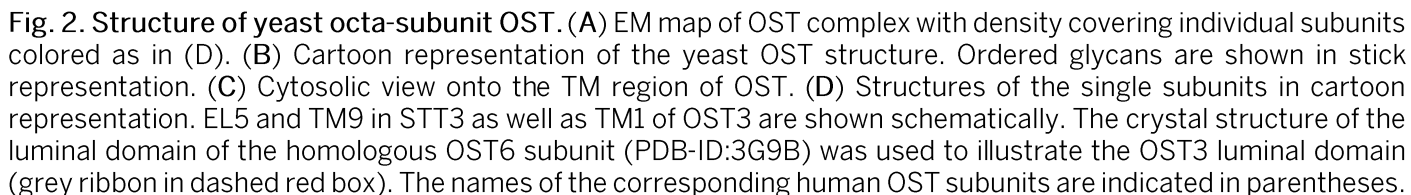
## SUPPLEMENTARY MATERIALS

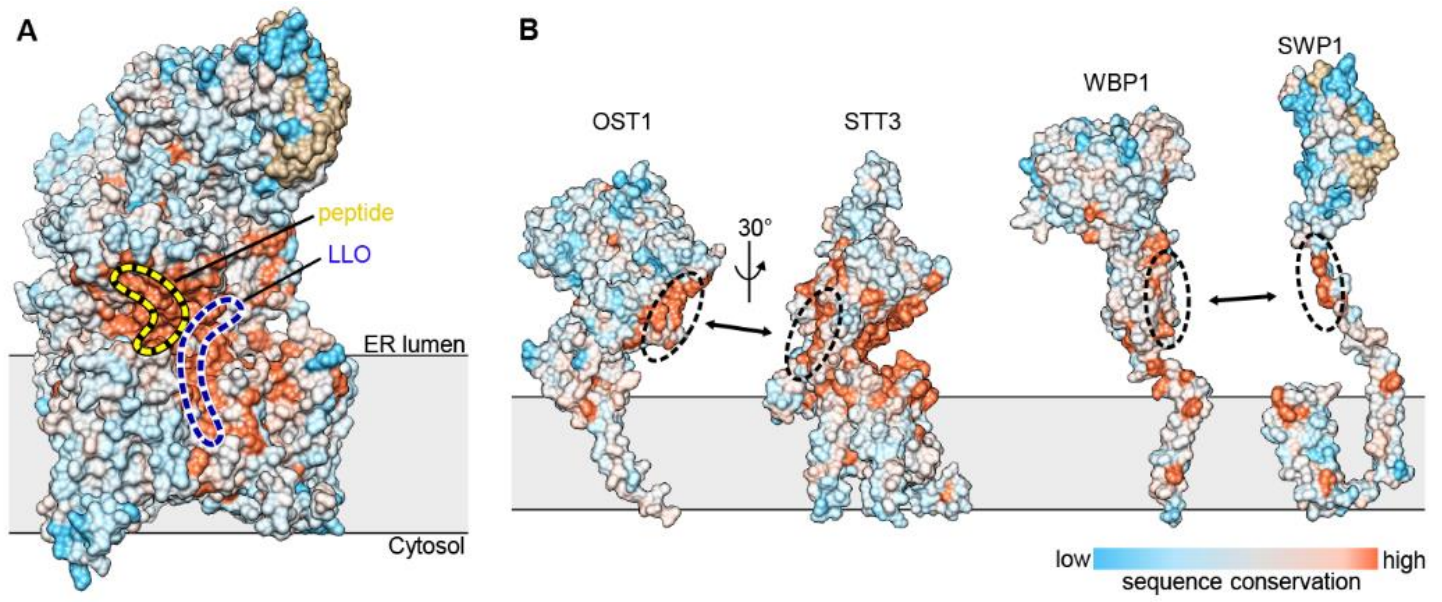
[www.sciencemag.org/cgi/content/full/science.aar5140/DC1](http://www.sciencemag.org/cgi/content/full/science.aar5140/DC1)  
Materials and Methods  
Table S1 and S2  
Figs. S1 to S8  
References (56–84)

16 November 2017; accepted 22 December 2017  
Published online 4 January 2018  
[10.1126/science.aar5140](https://doi.org/10.1126/science.aar5140)

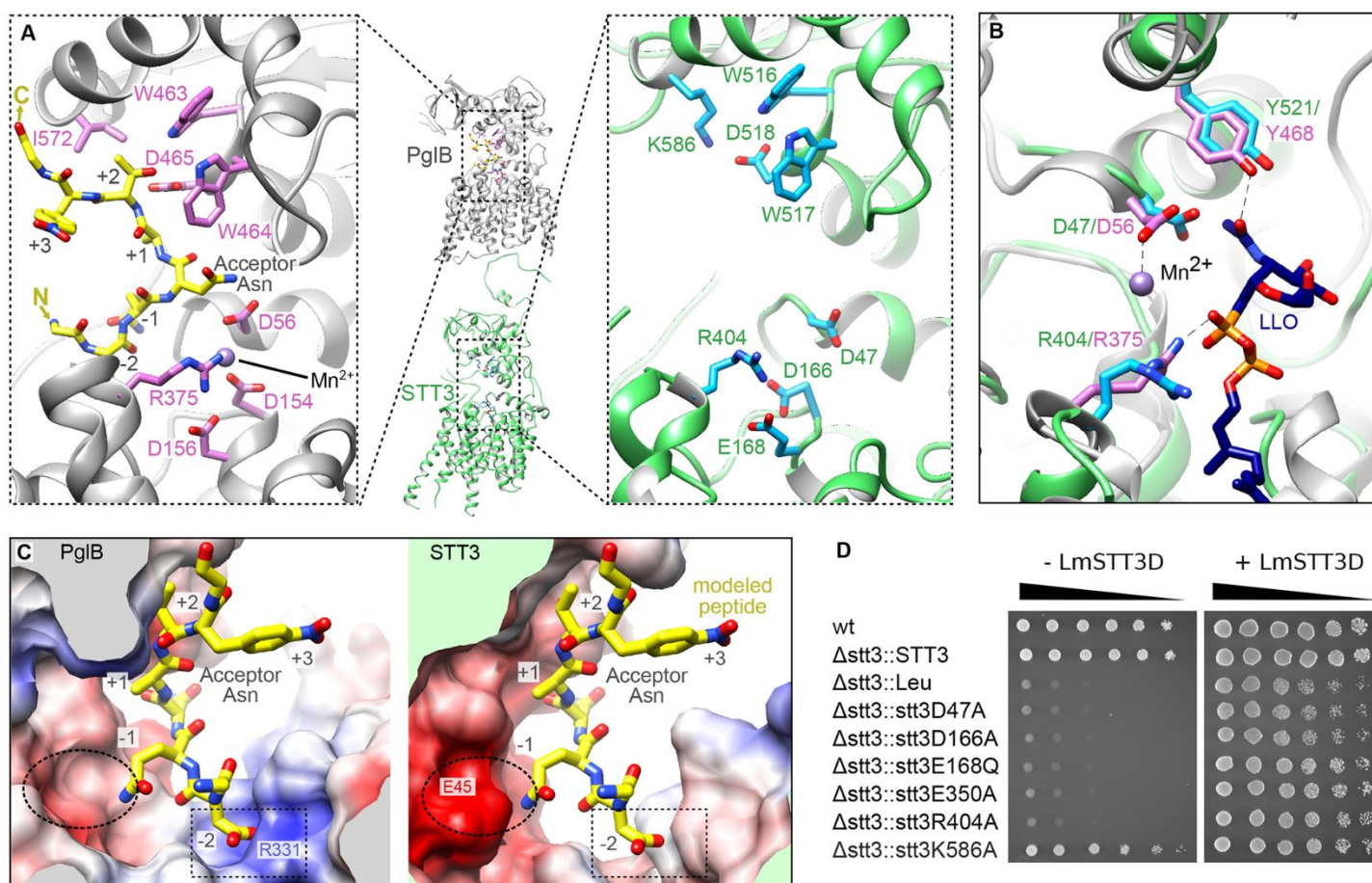


**Fig. 1. Purification, stoichiometry, and functional characterization of yeast OST.** (A) Silver-stained SDS-PAGE analysis of detergent-purified OST complexes containing either OST6 or OST3. A 1D4 affinity tag was fused to the C terminus of OST4. Multiple bands for OST1 and WBP1 indicate heterogeneous *N*-glycosylation. (B) Mass spectrometry-based quantification of subunit abundance in purified OST samples containing either OST3 (green bars) or OST6 (blue bars) relative to OST complexes from wild type cell extract ( $n = 3$ , error bars indicate SD). (C) In vitro glycosylation of a fluorescently labeled peptide (TAMRA-DANYTK) by nanodisc reconstituted, OST3-containing yeast OST complex, using a synthetic C<sub>20</sub>-LLO (NerylCitronellyl-PP-GlcNAc<sub>2</sub>) as a donor substrate. Glycosylated and non-glycosylated peptides were separated using a Tricine gel. (D) Following quantification of band intensities in (C), the ratio of glycosylated to unreacted peptide was plotted against the reaction time and fitted using a Michaelis-Menten saturation curve ( $n = 3$ , error bars indicate SD).

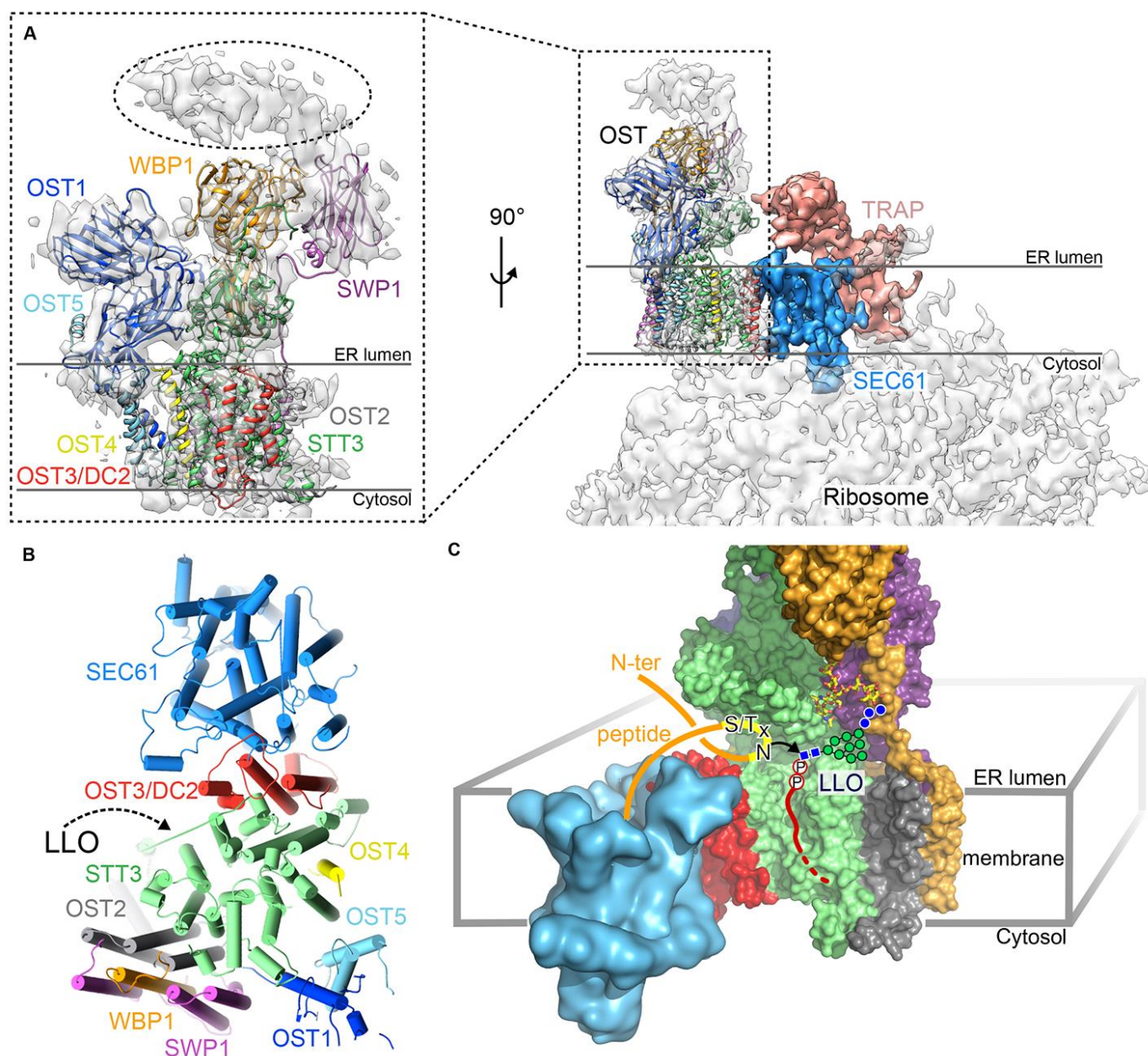




**Fig. 3. Sequence conservation in eukaryotic OST complexes.** (A) Surface representation of yeast OST complex colored according to sequence conservation. The predicted peptide and LLO binding sites are indicated. (B) Conserved subunit interfaces between OST1 and STT3 (left) and between WBP1 and SWP1 (right) are marked by dotted circles and arrows.



**Fig. 4. Active center of STT3.** (A) Ribbon representation of peptide-bound PglB (PDB-ID: 5OGL) and yeast STT3. Residues involved in substrate recognition or metal ion binding are shown as sticks. The numbering in the bound peptide (DQNATF sequence) is relative to the acceptor Asn residue. (B) Superposition of yeast STT3 (green ribbon) and LLO-bound PglB (grey ribbon). Functionally important residues (pink for PglB residues, blue for yeast STT3 residues) are shown as sticks, with interactions observed in PglB indicated with dashed lines. (C) Electrostatic surface representations of PglB and yeast STT3. The peptide bound to PglB was modeled into the yeast STT3 structure using the WWD motif and residues D47/D56 (PglB/STT3) as anchors. EL5 (aa294-aa322) of PglB was removed for clarity. Corresponding regions revealing the structural basis of distinct peptide specificity are indicated with dashed ovals or boxes. (D) In vivo activity assay of STT3 point mutants.



**Fig. 5. OST-translocon interactions.** (A) Yeast OST structure docked into electron cryo-tomography map of the mammalian ribosome-translocon-OST complex (EMDB-ID: 3068). A close-up of the dashed region is shown in the left panel. Additional EM density above SWP1 and WBP1 (dashed oval) probably corresponds to a ~300 aa N-terminal extension present in mammalian SWP1 homolog Ribophorin II. (B) View from the ER lumen onto the transmembrane regions of OST and of the Sec61 translocon (PDB-ID: 5A6U) (14) following the docking shown in (A). Although the OST3 subunits provides all contacts to Sec61 in this docking, it is replaced by the DC2 subunit in translocon-associated OST complexes. (C) Model of translocon-associated OST complex architecture and function. The orange, curved line depicts a nascent polypeptide entering the ER through the translocon and binding to the active site of STT3. An LLO molecule was manually placed in its likely binding pocket, with a red line representing dolichol, circled "P" phosphate moieties, and the Glc<sub>3</sub>Man<sub>9</sub>GlcNAc<sub>2</sub> moiety depicted by blue and green symbols according to standard glycan nomenclature. The black arrow depicts the proposed direction of the nucleophilic attack during glycan transfer.

Cite as: T. Bae *et al.*, *Science*  
10.1126/science.aan8690 (2017).

# Different mutational rates and mechanisms in human cells at pregastrulation and neurogenesis

Taejeong Bae,<sup>1</sup> Livia Tomasini,<sup>2</sup> Jessica Mariani,<sup>2</sup> Bo Zhou,<sup>3</sup> Tanmoy Roychowdhury,<sup>1</sup> Daniel Franjic,<sup>4</sup> Mihovil Pletikos,<sup>4</sup> Reenal Pattni,<sup>3</sup> Bo-Juen Chen,<sup>5</sup> Elisa Venturini,<sup>5</sup> Bridget Riley-Gillis,<sup>5</sup> Nenad Sestan,<sup>4,6</sup> Alexander E. Urban,<sup>3</sup> Alexej Abyzov,<sup>1\*</sup> Flora M. Vaccarino<sup>2,4,6\*</sup>

<sup>1</sup>Mayo Clinic, Department of Health Sciences Research, Rochester, MN 55905, USA. <sup>2</sup>Yale University, Child Study Center, Yale University, New Haven, CT 06520, USA. <sup>3</sup>Stanford University, Departments of Psychiatry and Genetics, Palo Alto, CA 94305, USA. <sup>4</sup>Yale University, Department of Neuroscience, New Haven, CT 06520, USA. <sup>5</sup>New York Genome Center, New York, NY 10013, USA. <sup>6</sup>Yale Kavli Institute for Neuroscience, New Haven, CT 06520, USA.

\*Corresponding author. Email: abyzov.alexej@mayo.edu (A.A.); flora.vaccarino@yale.edu (F.M.V.)

**Somatic mosaicism in the human brain may alter function of individual neurons. We analyzed genomes of single cells from the forebrains of three human fetuses (15 to 21 weeks post-conception) using clonal cell populations. We detected 200-400 single nucleotide variations (SNVs) per cell. SNV patterns resembled those found in cancer cell genomes, indicating a role of background mutagenesis in cancer. SNVs with a frequency of >2% in brain were shared with the spleen, revealing a pregastrulation origin. We reconstructed cell lineages for the first five post-zygotic cleavages and calculated a mutation rate of ~1.3 per division per cell. Later in development, during neurogenesis, the mutation spectrum shifted toward oxidative damage and the mutation rate increased. Both neurogenesis and early embryogenesis exhibit drastically more mutagenesis than adulthood.**

Somatic mutagenesis is one of the emerging areas of vertebrate genome biology. Several studies revealed extensive genomic mosaicism marked by hundreds of single nucleotide variants (SNVs) per cell in somatic tissues of the human body such as skin fibroblasts, intestine, liver, and colon (1–3). Mosaic copy number alterations are also common, and insertions of retrotransposable elements have been detected (4–10). Mosaicism is prominent in the central nervous system, with implications for brain evolution and the genomic underpinnings of human neuropsychiatric disorders (11, 12). Roughly 1500 SNVs might be present in mature neurons from the adult human cortex, only detectable in the analyzed cell and thought to be related to transcriptional activity (13). However, the temporal origin of these SNVs during development is unknown. Furthermore, the use of in vitro whole genome amplification (WGA) from single nuclei is prone to experimental artifacts mimicking SNVs (14, 15). Here we describe the discovery and analysis of mosaic SNVs in neuronal progenitor cells in three fetal human brains. Individual progenitor cells were allowed to proliferate into clonal cell populations, which provided insights into the genomes of the founder cells (fig. S1) and provide an estimation of the frequency and mutation spectrum of mosaic mutations in human development avoiding WGA-associated artifacts.

## Discovery and validation of mosaic SNVs

Brains were collected from phenotypically normal postmortem human fetuses ranging in age from 15 to 21 weeks post-

conception. Based on comparing counts of germline SNVs (3809591 for subject 316, 4316547 for subject 320, and 3746847 for subject 275) to those derived by the 1000 Genomes Project across different human sub-population, we concluded that subjects 316 and 275 were of non-African origin, whereas subject 320 (male, 17 pcw) was of African descent.

From a bulk culture of dissociated cells of the VZ/SVZ of the frontal cortex (FR), parietal cortex (PA) or basal ganglia (BG), we grew 31 single cell derived clonal cell populations each containing few thousand cells, using the limiting dilution approach. Few possible divisions prior to dilution are not likely to contribute significantly to the mutation landscape in each cell. DNA extracted from the individual clones, the source tissue of germinal zones, and the spleen was sequenced to a minimum of 30X genome coverage (fig. S1). For 3 clones we could not derive enough cells, hence DNA was amplified by multi-displacement amplification (MDA) before sequencing.

Mosaic SNVs present in the founder cell of the clones were discovered by comparing genomes of clones to each other and to genomes of the germinal zone tissue and spleen (table S1). We selected those calls with over 35% variant allele frequency (VAF) in clones as candidate mosaic SNVs. This limit was chosen to exclude mutations arising during culture, which should have a VAF of 25% or less. The distribution for the SNV discovery dataset is centered at around a VAF of 50% based on read depth, as expected for true mosaic

variants (figs. S1 and S2).

When comparing clones to each other, we filtered the resulting calls based on conformity of their recurrence to clones that are expected and are not expected to carry the same mosaic SNVs (fig. S3). Calls from such clone-to-clone comparisons were 98.9% concordant with calls from comparing clones to VZ/SVZ brain tissue or spleen (Fig. 1A). However, among 68 calls made exclusively from clone-to-clone comparison, 31 (46%) were missed from clone-to-brain tissue or clone-to-spleen comparisons because they corresponded to SNVs present in tissues at high frequency (Fig. 1C), demonstrating the advantage of the clone-to-clone comparison approach. Therefore, the clone-to-clone comparison represents an alternative design for the study of mosaicism to the use of familial trios (1).

Eight randomly selected SNVs were all confirmed in the clones using PCR and Sanger sequencing (table S2). As an additional validation strategy, we designed an oligonucleotide library complementary to the loci of all 6288 SNVs comprising the discovery dataset and performed capture and deep re-sequencing (to ~1000X coverage) in the DNA from 10 clones. This confirmed the 50%-centered VAF distribution for a majority of SNVs, with a minority (5.1%) having VAF lower than 35%, perhaps indicating that these could have arisen during cell culturing (figs. S1 and S4). Accordingly, we estimated our false positive rate at around 5%. From an in-silico comparison of our clones with the unrelated and well characterized cell line NA12878 we estimated that the sensitivity for discovering mosaic SNVs in the clones was ~83% (fig. S5).

### Mosaic SNV counts, mutation spectra, and distributions across brain regions

SNVs were found at rates of 108-572 per clone, with clones from older brains containing more variants (Fig. 1D), which averages to 200-400 SNVs after adjustment for discovery sensitivity and false positives. No differences in SNV counts for clones from frontal and parietal cortex and from frontal cortex and basal ganglia of the same brain were noticeable (Fig. 1D). Similarly, the relative contributions of substitution types to the mutation spectrum were the same for clones from different brains and from different brain regions (Fig. 1E). Overall, the transition to transversion ratio Ti/Tv was 0.6, with the most frequent substitution type being C:G>A:T transversion. This perhaps reflects DNA damage by oxidation, resulting in 8-Oxoguanine that is later fixated to threonine through incorrect base pairing with adenine (16). The second most common substitution was C:G>T:A transition, which is thought to be caused by deamination of cytosine and 5-methylcytosine (16). Linear approximation of the increase in SNV counts over time allowed estimating a mutation rate of 5.1 (95% CI 1.5 to 9) SNVs per day per progenitor during neurogenesis. This projects to a rate of roughly 8.6 (95% CI 1.6 to 20) SNVs

per division per progenitor, assuming that the length of cell cycle of cortical progenitors is between 27 to 54 hours based upon studies in primates (17). The large interval of this estimation is due to uncertainties in both the per day increase in SNVs and in the length of the cell cycle. Inter-individual variability may further widen the confidence interval of this estimate.

To genotype the presence of variants across brains we used the above referenced capture library to conduct deep sequencing (~1000X coverage) in the source tissue (dorsal and basal germinal layers), in the corresponding outer layers containing mature neurons (frontal, parietal cortex and basal ganglia, respectively), as well as in other brain regions (occipital cortex, cerebellum) and in a peripheral tissue, the spleen (table S3). A total of 144 SNVs was reliably genotyped, from 11 to 68 in each tissue, including spleen, with VAFs between 0.3% and 30% (Fig. 2A and fig. S6). High VAFs for 7 such SNVs were further cross-confirmed with an orthogonal technique, droplet digital PCR (table S2 and figs. S7 to S17). However, for hundreds of SNVs at much lower VAF (typically below 1%) the evidence for their presence in each tissue was not significant, likely because of their low VAFs in tissues.

Almost 60% of the genotyped SNVs (1.4% of total SNVs) and 92% of the SNVs with VAF above 2% in at least one brain region had nonzero VAF in the spleen (Fig. 2, B and C, and figs. S6 and S18). Since the brain is of neuroectodermal origin and the spleen of mesodermal origin, these shared SNVs likely occurred prior to gastrulation, when the mesoderm, ectoderm and endoderm differentiate from a single-layered epithelium. This suggestion is consistent with the range of VAFs of these shared SNVs, as there are about 12 cell divisions prior to gastrulation (18), corresponding to expected VAFs from 0.03% to 25% in somatic tissues, depending on how early in embryogenesis variants have occurred. Some SNVs were shared between the spleen and only some of the brain regions (Fig. 2, B and C, and figs. S6 and S18). This could indicate regional sub-lineages, i.e., that non-mixing sets of progenitors generate neurons in different brain regions, and that these distinct populations of progenitors may not have a common ancestor with cells in spleen. But conducting more sensitive assays is necessary to exclude the possibility of incomplete genotyping. Another observation pointing to an early origin of the mutations is that the overlap between SNVs in the basal ganglia VZ/SVZ and cortical VZ/SVZ with the respective differentiated regions was less than the overlap between the four regions together (fig. S18), suggesting that most of the genotyped mutations found in the brain arise at least before the splitting between basal and dorsal regions of the telencephalon, i.e., at around neural plate stage or earlier.

Genotyped mosaic SNVs clearly cluster by similar VAFs in each brain (Fig. 3 and fig. S19). Based on the average values of the frequencies for each cluster and sharing of such SNVs

between clones, we concluded that these clusters likely represent variants created during sequential post-zygotic cleavages. Assuming equal contribution of dividing cells to tissues, we reconstructed the cell progeny tree and determined the precise origin of 84 mutations during the first 5 post-zygotic cleavages (Fig. 3B and fig. S19). These mutations typically had VAF above 1%, while the remaining ones, typically with VAF below 1%, were assigned to later divisions. Only two SNVs had conflicting assignment between clusters and clones, which perhaps could be explained by mis-clustering or incorrect discovery/genotyping in clones. Alternatively, these conflicts along with the high VAFs for the very first SNV in the tree (Fig. 3A) may indicate an unequal lineage contribution to tissues due to asymmetric division, unequal proliferation, or positive/negative selection (19).

Using the trees we estimated the average mutation rates per division per daughter cell in the early human embryo as  $1.66 \pm 0.24$ ,  $1.18 \pm 0.33$ , and  $1.05 \pm 0.22$  for brains of the subjects 316, 320, and 275 respectively. The weighted average and variance of the 3 measurements is  $1.3 \pm 0.15$ , consistent with the rate of 1.2 estimated from analysis of de novo SNVs in familial trios (20), and lower than the lower bound estimate for mutability of neuronal progenitors, thereby suggesting that the mutation rate during neurogenesis is higher than that in the early embryo.

We then split the set of mosaic SNVs into early origin – those genotyped in tissues by capture experiments, and likely late origin – those not genotyped. The spectrum of early mosaic mutations (i.e., the frequency of nucleotide substitution in the context of tri-nucleotide motifs) bears little resemblance to the spectrum of SNVs occurring later in neuronal progenitors, revealing a shift in mutagenesis during development (Fig. 3C). Early mutations had the same 2.2  $Ti/Tv$  ratio as germline variants and had a larger fraction of C:G>T:A transitions overall (p-value of  $2.2 \times 10^{-16}$  by Fischer's exact test), and particularly in CpG motifs (p-value of  $4.3 \times 10^{-5}$  by Fisher's exact test), implicating the spontaneous deamination of 5-methylcytosine as a contributor to the mutagenic process (16). The signature of earlier mutations was also similar to the signature of de novo mutations in the human population (20). As some of early mutations can be passed to the next generation through the germline lineage, the convergence of their spectrum with de novo and germline variants is expected. Later mutations, on the other hand, had a larger contribution of C:G>A:T transversions (p-value of  $8.0 \times 10^{-12}$  by Fisher's exact test), implicating oxidative damage as a significant contributor to the mutagenic process (16). Furthermore, mutation spectrum for these transversions was most similar (correlation  $R=0.90$ ) to the spectrum observed in MUTYH-defective colorectal cancer, where repair of 8-oxoguanine, the product of oxidative damage, is compromised (21).

## Properties of mosaic SNVs

For the following analyses we used all mosaic SNVs in the discovery dataset from all clones, totaling to 6288, as mutation spectra for the three brains were extremely similar (fig. S20). The distribution between neighboring mosaic SNVs was consistent with uniform random placement across the genome (fig. S21). In line with this, no enrichment of exonic and intronic SNVs in any gene ontology category was observed either when assuming uniform background mutation rate or using as a background mosaic SNVs from liver, intestine, colon, or fibroblasts (1, 2). Roughly 3% of SNVs may have a functional consequence either by affecting protein coding sequence or gene regulation (fig. S21 and table S4). This projects to about 12 non-benign SNVs per progenitor cell at 20 post-conception weeks. A significant depletion of mosaic SNVs was observed in DNase hypersensitive sites relative to flanking regions (Fig. 4A). The depletion was more pronounced (10% vs 5%) when utilizing DNase hypersensitive sites for fetal brain rather than for a lymphoblastoid cell line, suggesting a relation between cells' epigenome and the genesis of mutations. Since no such depletion was observed in coding relative to intronic gene regions, the depletion is not the result of negative selection, rather likely reflects better repair efficiency in open DNA regions, as was observed for somatic mutations in cancers (22, 23).

Similar to somatic SNVs in cancers (24) and mosaic SNVs in skin fibroblasts (1), we found that the density of mosaic SNVs in the brain correlates negatively with most histone marks in fetal brain and embryonic stem cells (Fig. 4B). Comparison of our mosaic SNVs with mutational signatures found in cancer (25, 26) revealed that signatures 18 and 8, found, respectively, in neuroblastoma and medulloblastoma, as well as their combination, are the best descriptors for the spectrum of mosaic mutations in the developing brain (Fig. 4, C and D). Mosaic SNVs were equally well described by the combination of signatures 5 with 18 and 1B with 18. Therefore, signature 18 with suspected etiology of oxidative damage (27), consistently contributed to the mutation spectrum of mosaic SNVs in fetal brain progenitors. This signature was mostly similar to late SNVs, while signature 1B to the early ones (fig. S22).

## Implications for development and disease

Our study uncovered extensive mosaicism in human fetal brain, with 200-400 SNVs present per brain progenitor cell at 15-21 weeks of gestation. This amount of mosaicism is likely inherited by cortical postmitotic neurons, as neurogenesis ends around 20 weeks in humans (28). Indeed, our estimate is in good agreement with the estimate that postmitotic neurons have ~300-900 mosaic SNV within one year of birth {aao4426 Lodato}. There is an order of magnitude difference

between numbers of mosaic SNVs and de novo SNPs (20) implying a higher effect of mosaic SNVs on normal brain development and disease. Indeed we estimate that up to 12 non-benign mutations can be present in neuronal progenitors and consequently transmitted to a sizable fraction of daughter neurons. It is conceivable that in rare cases some of these mutations may have strong deleterious effect, for example, initiate overgrowth (29, 30) or neoplastic transformation by knocking out key genes. Indeed, the resemblance of mosaic SNVs in fetal brain to somatic mutations in brain cancers and particularly to medulloblastoma supports the theory that cancer driving mutations can happen by chance during background mutagenesis (31).

As dozens of discovered mutations happen prior to gastrulation, our study demonstrates that early post-zygotic mutations can be reconstructed from the analysis of a handful of clones and tissues, opening an avenue for charting individualized mosaicism maps. As mosaic variants can contribute to interindividual phenotypic differences and have been implicated in individual's disease risk, we suggest that knowing the individual "mosaicom" could be as important as knowing the individual germline genome, particularly given the much stronger selection acting on germline variants and the lower penetrance of mosaic variants that is likely to be translated in milder phenotypes.

We also discovered a shift in mutagenesis during development characterized by an increased mutation rate and a change in frequency of substitution types. We cannot rule out that the increased mutation rate can be partially explained by inter-individual variation, although we have no evidence for such variability. The shift occurs sometime between early cleavages and neurogenesis and may be the consequence of physiological, biochemical, and gene expression changes related to the generation of neurons from neural stem cells. Alternatively, the shift may reflect more general developmental processes common to all tissues during organogenesis, and, based on increased counts of mutations related to oxidative damage, could be coupled to higher availability of radical oxygen species after development of the cardio-vascular system of the embryo. If this is the case, we predict that mutation spectra and rates per division undergo a similar shift during development across all somatic lineages.

Our estimated average mutation rate of 5.1 SNVs per day per neuronal progenitor during neurogenesis implies that neurons generated at early and later stages of neurogenesis will carry different burdens of mosaic variants. This rate is three orders of magnitude higher than 0.4 to 2 mutations per year accumulated in the germline lineage of adults (20, 32, 33). It is also 50 times higher than the rate in postnatal stem cells of the small intestine, colon and liver, estimated to be 36 mutations per year (2). Therefore, our results show that the prenatal period is intrinsically highly mutagenic likely the

consequence of oxidative damage coupled with more frequent cell divisions.

We found no difference in SNV count between progenitors from cortex and from basal ganglia, implying that mosaicism accumulates at similar rates across the brain during neurogenesis. With the observed rate of 5.1 SNVs per day per neuronal progenitor one can project that cells in the forebrain SVZ and hippocampal subgranular zone, where neurogenesis and gliogenesis continues for more extended time periods (34–36), would accumulate about one thousand mosaic mutations by the time of birth. This estimate is consistent with the estimates of about 1000 mosaic SNVs present in skin fibroblast cells and in stem cells of colon and intestine in children (1, 2); indeed, mutations rates in all somatic proliferative cell lineages during prenatal development may be similar.

## REFERENCES AND NOTES

1. A. Abyzov, L. Tomasini, B. Zhou, N. Vasmatzis, G. Coppola, M. Amenduni, R. Pattini, M. Wilson, M. Gerstein, S. Weissman, A. E. Urban, F. M. Vaccarino, One thousand somatic SNVs per skin fibroblast cell set baseline of mosaic mutational load with patterns that suggest proliferative origin. *Genome Res.* **27**, 512–523 (2017). doi:10.1101/gr.215517.116 Medline
2. F. Blokzijl, J. de Ligt, M. Jager, V. Sasselli, S. Roerink, N. Sasaki, M. Huch, S. Boymans, E. Kuijk, P. Prins, I. J. Nijman, I. Martincorena, M. Mokry, C. L. Wiegerinck, S. Middendorp, T. Sato, G. Schwank, E. E. S. Nieuwenhuis, M. M. A. Verstegen, L. J. W. van der Laan, J. de Jonge, J. N. M. IJzermans, R. G. Vries, M. van de Wetering, M. R. Stratton, H. Clevers, E. Cuppen, R. van Boxtel, Tissue-specific mutation accumulation in human adult stem cells during life. *Nature* **538**, 260–264 (2016). doi:10.1038/nature19768 Medline
3. N. Saini, S. A. Roberts, L. J. Klimczak, K. Chan, S. A. Grimm, S. Dai, D. C. Fargo, J. C. Boyer, W. K. Kaufmann, J. A. Taylor, E. Lee, I. Cortes-Ciriano, P. J. Park, S. H. Schurman, E. P. Malt, P. A. Mieczkowski, D. A. Gordenin, The impact of environmental and endogenous damage on somatic mutation load in human skin fibroblasts. *PLOS Genet.* **12**, e1006385 (2016). doi:10.1371/journal.pgen.1006385 Medline
4. A. Abyzov, J. Mariani, D. Palejev, Y. Zhang, M. S. Haney, L. Tomasini, A. F. Ferrandino, L. A. Rosenberg Belmaker, A. Szekeley, M. Wilson, A. Kocabas, N. E. Calixto, E. L. Grigorenko, A. Huttner, K. Chawarska, S. Weissman, A. E. Urban, M. Gerstein, F. M. Vaccarino, Somatic copy number mosaicism in human skin revealed by induced pluripotent stem cells. *Nature* **492**, 438–442 (2012). doi:10.1038/nature11629 Medline
5. M. O'Huallachain, K. J. Karczewski, S. M. Weissman, A. E. Urban, M. P. Snyder, Extensive genetic variation in somatic human tissues. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 18018–18023 (2012). doi:10.1073/pnas.1213736109 Medline
6. M. J. McConnell, M. R. Lindberg, K. J. Brennand, J. C. Piper, T. Voet, C. Cowing-Zitron, S. Shumilina, R. S. Lasken, J. R. Vermeesch, I. M. Hall, F. H. Gage, Mosaic copy number variation in human neurons. *Science* **342**, 632–637 (2013). doi:10.1126/science.1243472 Medline
7. C. C. Laurie, C. A. Laurie, K. Rice, K. F. Doherty, L. R. Zelnick, C. P. McHugh, H. Ling, K. N. Hetrick, E. W. Pugh, C. Amos, Q. Wei, L. E. Wang, J. E. Lee, K. C. Barnes, N. N. Hansel, R. Mathias, D. Daley, T. H. Beaty, A. F. Scott, I. Ruczinski, R. B. Scharpf, L. J. Bierut, S. M. Hartz, M. T. Landi, N. D. Freedman, L. R. Goldin, D. Ginsburg, J. Li, K. C. Desch, S. S. Strom, W. J. Blot, L. B. Signorello, S. A. Ingles, S. J. Chanock, S. I. Berndt, L. Le Marchand, B. E. Henderson, K. R. Monroe, J. A. Heit, M. de Andrade, S. M. Armasu, C. Regnier, W. L. Lowe, M. G. Hayes, M. L. Marazita, E. Feingold, J. C. Murray, M. Melbye, B. Feenstra, J. H. Kang, J. L. Wiggs, G. P. Jarvik, A. N. McDavid, V. E. Seshan, D. B. Mirel, A. Crenshaw, N. Sharopova, A. Wise, J. Shen, D. R. Crosslin, D. M. Levine, X. Zheng, J. I. Udren, S. Bennett, S. C. Nelson, S. M. Gogarten, M. P. Conomos, P. Heagerty, T. Manolio, L. R. Pasquale, C. A. Haiman, N. Caporaso, B. S. Weir, Detectable clonal mosaicism from birth to old age and its

- relationship to cancer. *Nat. Genet.* **44**, 642–650 (2012). [doi:10.1038/ng.2271](https://doi.org/10.1038/ng.2271) [Medline](#)
8. K. B. Jacobs, M. Yeager, W. Zhou, S. Wacholder, Z. Wang, B. Rodriguez-Santiago, A. Hutchinson, X. Deng, C. Liu, M.-J. Horner, M. Cullen, C. G. Epstein, L. Burdett, M. C. Dean, N. Chatterjee, J. Sampson, C. C. Chung, J. Kovaks, S. M. Gapstur, V. L. Stevens, L. T. Teras, M. M. Gaudet, D. Albanes, S. J. Weinstein, J. Virtamo, P. R. Taylor, N. D. Freedman, C. C. Abnet, A. M. Goldstein, N. Hu, K. Yu, J.-M. Yuan, L. Liao, T. Ding, Y.-L. Qiao, Y.-T. Gao, W.-P. Koh, Y.-B. Xiang, Z.-Z. Tang, J.-H. Fan, M. C. Aldrich, C. Amos, W. J. Blot, C. H. Bock, E. M. Gillanders, C. C. Harris, C. A. Haiman, B. E. Henderson, L. N. Kolonel, L. Le Marchand, L. H. McNeill, B. A. Rybicki, A. G. Schwartz, L. B. Signorello, M. R. Spitz, J. K. Wiencke, M. Wrensch, X. Wu, K. A. Zanetti, R. G. Ziegler, J. D. Figueroa, M. Garcia-Closas, N. Malats, G. Marenne, L. Prokunina-Olsson, D. Baris, M. Schwenn, A. Johnson, M. T. Landi, L. Goldin, D. Consonni, P. A. Bertazzi, M. Rotunno, P. Rajaraman, U. Andersson, L. E. Beane Freeman, C. D. Berg, J. E. Buring, M. A. Butler, T. Carreon, M. Feychting, A. Ahlbom, J. M. Gaziano, G. G. Giles, G. Hallmans, S. E. Hankinson, P. Hartge, R. Henriksson, P. D. Inskip, C. Johansen, A. Landgren, R. McKean-Cowdin, D. S. Michaud, B. S. Melin, U. Peters, A. M. Ruder, H. D. Sesso, G. Severi, X.-O. Shu, K. Visvanathan, E. White, A. Wolk, A. Zeleniuch-Jacquotte, W. Zheng, D. T. Silverman, M. Kogevinas, J. R. Gonzalez, O. Villa, D. Li, E. J. Duell, H. A. Risch, S. H. Olson, C. Kooperberg, B. M. Wolpin, L. Jiao, M. Hassan, W. Wheeler, A. A. Arslan, H. B. Bueno-de-Mesquita, C. S. Fuchs, S. Gallinger, M. D. Gross, E. A. Holly, A. P. Klein, A. LaCroix, M. T. Mandelson, G. Petersen, M.-C. Boutron-Ruault, P. M. Bracci, F. Canzian, K. Chang, M. Cotterchio, E. L. Giovannucci, M. Goggins, J. A. Hoffman Bolton, M. Jenab, K.-T. Khaw, V. Krogh, R. C. Kurtz, R. R. McWilliams, J. B. Mendelsohn, K. G. Rabe, E. Riboli, A. Tjønneland, G. S. Tobias, D. Trichopoulos, J. W. Elena, H. Yu, L. Amundadottir, R. Z. Stolzenberg-Solomon, P. Kraft, F. Schumacher, D. Stram, S. A. Savage, L. Mirabello, I. L. Andriulis, J. S. Wunder, A. Patiño García, L. Sierrasesúmaga, D. A. Barkauskas, R. G. Gorlick, M. Purdue, W.-H. Chow, L. E. Moore, K. L. Schwartz, F. G. Davis, A. W. Hsing, S. I. Berndt, A. Black, N. Wentzensen, L. A. Brinton, J. Lissowska, B. Peplonska, K. A. McGlynn, M. B. Cook, B. I. Graubard, C. P. Kratz, M. H. Greene, R. L. Erickson, D. J. Hunter, G. Thomas, R. N. Hoover, F. X. Real, J. F. Fraumeni Jr., N. E. Caporaso, M. Tucker, N. Rothman, L. A. Pérez-Jurado, S. J. Chanock, Detectable clonal mosaicism and its relationship to aging and cancer. *Nat. Genet.* **44**, 651–658 (2012). [doi:10.1038/ng.2270](https://doi.org/10.1038/ng.2270) [Medline](#)
  9. G. D. Evrony, X. Cai, E. Lee, L. B. Hills, P. C. Elhosary, H. S. Lehmann, J. J. Parker, K. D. Atabay, E. C. Gilmore, A. Poduri, P. J. Park, C. A. Walsh, Single-neuron sequencing analysis of L1 retrotransposition and somatic mutation in the human brain. *Cell* **151**, 483–496 (2012). [doi:10.1016/j.cell.2012.09.035](https://doi.org/10.1016/j.cell.2012.09.035) [Medline](#)
  10. J. A. Erwin, A. C. Paquola, T. Singer, I. Gallina, M. Novotny, C. Quayle, T. A. Bedrosian, F. I. Alves, C. R. Butcher, J. R. Herdy, A. Sarkar, R. S. Lasken, A. R. Muotri, F. H. Gage, L1-associated genomic regions are deleted in somatic cells of the healthy human brain. *Nat. Neurosci.* **19**, 1583–1591 (2016). [doi:10.1038/nn.4388](https://doi.org/10.1038/nn.4388) [Medline](#)
  11. T. R. Insel, Brain somatic mutations: The dark matter of psychiatric genetics? *Mol. Psychiatry* **19**, 156–158 (2014). [doi:10.1038/mp.2013.168](https://doi.org/10.1038/mp.2013.168) [Medline](#)
  12. M. J. McConnell *et al.*, Brain Somatic Mosaicism Network, Intersection of diverse neuronal genomes and neuropsychiatric disease: The Brain Somatic Mosaicism Network. *Science* **356**, eaal1641 (2017). [doi:10.1126/science.aal1641](https://doi.org/10.1126/science.aal1641) [Medline](#)
  13. M. A. Lodato, M. B. Woodworth, S. Lee, G. D. Evrony, B. K. Mehta, A. Karger, S. Lee, T. W. Chittenden, A. M. D'Gama, X. Cai, L. J. Luquette, E. Lee, P. J. Park, C. A. Walsh, Somatic mutation in single human neurons tracks developmental and transcriptional history. *Science* **350**, 94–98 (2015). [doi:10.1126/science.aab1785](https://doi.org/10.1126/science.aab1785) [Medline](#)
  14. C. Chen, D. Xing, L. Tan, H. Li, G. Zhou, L. Huang, X. S. Xie, Single-cell whole-genome analyses by Linear Amplification via Transposon Insertion (LIANTI). *Science* **356**, 189–194 (2017). [doi:10.1126/science.aak9787](https://doi.org/10.1126/science.aak9787) [Medline](#)
  15. X. Dong, L. Zhang, B. Milholland, M. Lee, A. Y. Maslov, T. Wang, J. Vijg, Accurate identification of single-nucleotide variants in whole-genome-amplified single cells. *Nat. Methods* **14**, 491–493 (2017). [doi:10.1038/nmeth.4227](https://doi.org/10.1038/nmeth.4227) [Medline](#)
  16. A. Bacolla, D. N. Cooper, K. M. Vasquez, Mechanisms of base substitution mutagenesis in cancer genomes. *Genes (Basel)* **5**, 108–146 (2014). [doi:10.3390/genes5010108](https://doi.org/10.3390/genes5010108) [Medline](#)
  17. D. R. Kornack, P. Rakic, Changes in cell-cycle kinetics during the development and evolution of primate neocortex. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 1242–1246 (1998). [doi:10.1073/pnas.95.3.1242](https://doi.org/10.1073/pnas.95.3.1242) [Medline](#)
  18. K. L. Moore, T. V. N. Persaud, M. G. Torchia, *Before We Are Born* (Elsevier Health Sciences, 2015).
  19. Y. S. Ju, I. Martincorena, M. Gerstung, M. Petljak, L. B. Alexandrov, R. Rahbari, D. C. Wedge, H. R. Davies, M. Ramakrishna, A. Fullam, S. Martin, C. Alder, N. Patel, S. Gamble, S. O'Meara, D. D. Giri, T. Sauer, S. E. Pinder, C. A. Purdie, Å. Borg, H. Stunnenberg, M. van de Vijver, B. K. Tan, C. Caldas, A. Tutt, N. T. Ueno, L. J. van 't Veer, J. W. Martens, C. Sotiriou, S. Knappskog, P. N. Span, S. R. Lakhani, J. E. Eyfjörð, A. L. Børresen-Dale, A. Richardson, A. M. Thompson, A. Viari, M. E. Hurler, S. Nik-Zainal, P. J. Campbell, M. R. Stratton, Somatic mutations reveal asymmetric cellular dynamics in the early human embryo. *Nature* **543**, 714–718 (2017). [Medline](#)
  20. R. Rahbari, A. Wuster, S. J. Lindsay, R. J. Hardwick, L. B. Alexandrov, S. A. Turki, A. Dominiczak, A. Morris, D. Porteous, B. Smith, M. R. Stratton, M. E. Hurler, UK10K Consortium, Timing, rates and spectra of human germline mutation. *Nat. Genet.* **48**, 126–133 (2016). [doi:10.1038/ng.3469](https://doi.org/10.1038/ng.3469) [Medline](#)
  21. A. Viel, A. Bruselles, E. Meccia, M. Fornasari, M. Quaia, V. Canzonieri, E. Policicchio, E. D. Urso, M. Agostini, M. Genuardi, E. Lucci-Cordisco, T. Venesio, A. Martayan, M. G. Diodoro, L. Sanchez-Mete, V. Stigliano, F. Mazzei, F. Grasso, A. Giuliani, M. Baiocchi, R. Maestro, G. Giannini, M. Tartaglia, L. B. Alexandrov, M. Bignami, A specific mutational signature associated with DNA 8-oxoguanine persistence in MUTYH-defective colorectal cancer. *EBioMedicine* **20**, 39–49 (2017). [doi:10.1016/j.ebiom.2017.04.022](https://doi.org/10.1016/j.ebiom.2017.04.022) [Medline](#)
  22. D. Perera, R. C. Poulos, A. Shah, D. Beck, J. E. Pimanda, J. W. H. Wong, Differential DNA repair underlies mutation hotspots at active promoters in cancer genomes. *Nature* **532**, 259–263 (2016). [doi:10.1038/nature17437](https://doi.org/10.1038/nature17437) [Medline](#)
  23. R. Sabarinathan, L. Mularoni, J. Deu-Pons, A. Gonzalez-Perez, N. López-Bigas, Nucleotide excision repair is impaired by binding of transcription factors to DNA. *Nature* **532**, 264–267 (2016). [doi:10.1038/nature17661](https://doi.org/10.1038/nature17661) [Medline](#)
  24. P. Polak, R. Karlič, A. Koren, R. Thurman, R. Sandstrom, M. Lawrence, A. Reynolds, E. Rynes, K. Vlahović, J. A. Stamatoyannopoulos, S. R. Sunyaev, Cell-of-origin chromatin organization shapes the mutational landscape of cancer. *Nature* **518**, 360–364 (2015). [doi:10.1038/nature14221](https://doi.org/10.1038/nature14221) [Medline](#)
  25. L. B. Alexandrov *et al.*, Australian Pancreatic Cancer Genome Initiative, ICGC Breast Cancer Consortium, ICGC MMML-Seq Consortium, ICGC PedBrain, Signatures of mutational processes in human cancer. *Nature* **500**, 415–421 (2013). [doi:10.1038/nature12477](https://doi.org/10.1038/nature12477) [Medline](#)
  26. M. S. Lawrence, P. Stojanov, P. Polak, G. V. Kryukov, K. Cibulskis, A. Sivachenko, S. L. Carter, C. Stewart, C. H. Mermel, S. A. Roberts, A. Kiezun, P. S. Hammerman, A. McKenna, Y. Drier, L. Zou, A. H. Ramos, T. J. Pugh, N. Stransky, E. Helman, J. Kim, C. Sougnez, L. Ambrogio, E. Nickerson, E. Shefler, M. L. Cortés, D. Auclair, G. Saksena, D. Voet, M. Noble, D. DiCara, P. Lin, L. Lichtenstein, D. I. Heiman, T. Fennell, M. Imielinski, B. Hernandez, E. Hodis, S. Baca, A. M. Dulak, J. Lohr, D.-A. Landau, C. J. Wu, J. Melendez-Zajgla, A. Hidalgo-Miranda, A. Koren, S. A. McCarroll, J. Mora, R. S. Lee, B. Crompton, R. Onofrio, M. Parkin, W. Winckler, K. Ardlie, S. B. Gabriel, C. W. M. Roberts, J. A. Biegel, K. Stegmaier, A. J. Bass, L. A. Garraway, M. Meyerson, T. R. Golub, D. A. Gordenin, S. Sunyaev, E. S. Lander, G. Getz, Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature* **499**, 214–218 (2013). [doi:10.1038/nature12213](https://doi.org/10.1038/nature12213) [Medline](#)
  27. T. Helleday, S. Eshtad, S. Nik-Zainal, Mechanisms underlying mutational signatures in human cancers. *Nat. Rev. Genet.* **15**, 585–598 (2014). [doi:10.1038/nrg3729](https://doi.org/10.1038/nrg3729) [Medline](#)
  28. A. A. Pollen, T. J. Nowakowski, J. Chen, H. Retallack, C. Sandoval-Espinosa, C. R. Nicholas, J. Shuga, S. J. Liu, M. C. Oldham, A. Diaz, D. A. Lim, A. A. Leyrat, J. A. West, A. R. Kriegstein, Molecular identity of human outer radial glia during cortical development. *Cell* **163**, 55–67 (2015). [doi:10.1016/j.cell.2015.09.004](https://doi.org/10.1016/j.cell.2015.09.004) [Medline](#)
  29. A. Poduri, G. D. Evrony, X. Cai, P. C. Elhosary, R. Beroukhim, M. K. Lehtinen, L. B. Hills, E. L. Heinzen, A. Hill, R. S. Hill, B. J. Barry, B. F. D. Bourgeois, J. J. Riviello, A. J. Barkovich, P. M. Black, K. L. Ligon, C. A. Walsh, Somatic activation of AKT3 causes hemispheric developmental brain malformations. *Neuron* **74**, 41–48 (2012). [doi:10.1016/j.neuron.2012.03.010](https://doi.org/10.1016/j.neuron.2012.03.010) [Medline](#)
  30. J. H. Lee, M. Huynh, J. L. Silhavy, S. Kim, T. Dixon-Salazar, A. Heiberg, E. Scott, V. Bafna, K. J. Hill, A. Collazo, V. Funari, C. Russ, S. B. Gabriel, G. W. Mathern, J. G. Gleeson, De novo somatic mutations in components of the PI3K-AKT3-mTOR

- pathway cause hemimegalencephaly. *Nat. Genet.* **44**, 941–945 (2012).  
doi:10.1038/ng.2329 [Medline](#)
31. C. Tomasetti, B. Vogelstein, Variation in cancer risk among tissues can be explained by the number of stem cell divisions. *Science* **347**, 78–81 (2015).  
doi:10.1126/science.1260825 [Medline](#)
  32. J. J. Michaelson, Y. Shi, M. Gujral, H. Zheng, D. Malhotra, X. Jin, M. Jian, G. Liu, D. Greer, A. Bhandari, W. Wu, R. Corominas, A. Peoples, A. Koren, A. Gore, S. Kang, G. N. Lin, J. Estabillio, T. Gadowski, B. Singh, K. Zhang, N. Akshoomoff, C. Corsello, S. McCarroll, L. M. Iakoucheva, Y. Li, J. Wang, J. Sebat, Whole-genome sequencing in autism identifies hot spots for de novo germline mutation. *Cell* **151**, 1431–1442 (2012). doi:10.1016/j.cell.2012.11.019 [Medline](#)
  33. H. Jónsson, P. Sulem, B. Kehr, S. Kristmundsdóttir, F. Zink, E. Hjartarson, M. T. Hardarson, K. E. Hjorleifsson, H. P. Eggertsson, S. A. Gudjonsson, L. D. Ward, G. A. Arnadóttir, E. A. Helgason, H. Helgason, A. Gylfason, A. Jonasdóttir, A. Jonasdóttir, T. Rafnar, M. Frigge, S. N. Stacey, O. Th Magnusson, U. Thorsteinsdóttir, G. Masson, A. Kong, B. V. Halldorsson, A. Helgason, D. F. Gudbjartsson, K. Stefansson, Parental influence on human germline de novo mutations in 1,548 trios from Iceland. *Nature* **549**, 519–522 (2017).  
doi:10.1038/nature24018 [Medline](#)
  34. F. H. Gage, Neurogenesis in the adult brain. *J. Neurosci.* **22**, 612–613 (2002).  
[Medline](#)
  35. A. Ernst, K. Alkass, S. Bernard, M. Salehpour, S. Perl, J. Tisdale, G. Possnert, H. Druid, J. Frisén, Neurogenesis in the striatum of the adult human brain. *Cell* **156**, 1072–1083 (2014). doi:10.1016/j.cell.2014.01.044 [Medline](#)
  36. J. T. Gonçalves, S. T. Schafer, F. H. Gage, Adult neurogenesis in the hippocampus: from stem cells to behavior. *Cell* **167**, 897–914 (2016).  
doi:10.1016/j.cell.2016.10.021 [Medline](#)
  37. C. M. Hindson, J. R. Chevillet, H. A. Briggs, E. N. Gallichotte, I. K. Ruf, B. J. Hindson, R. L. Vessella, M. Tewari, Absolute quantification by droplet digital PCR versus analog real-time PCR. *Nat. Methods* **10**, 1003–1005 (2013).  
doi:10.1038/nmeth.2633 [Medline](#)
  38. H. Won, L. de la Torre-Ubieta, J. L. Stein, N. N. Parikhshak, J. Huang, C. K. Opland, M. J. Gandal, G. J. Sutton, F. Hormozdiari, D. Lu, C. Lee, E. Eskin, I. Voineagu, J. Ernst, D. H. Geschwind, Chromosome conformation elucidates regulatory relationships in developing human brain. *Nature* **538**, 523–527 (2016).  
doi:10.1038/nature19847 [Medline](#)
  39. L. H. Chadwick, The NIH Roadmap Epigenomics Program data resource. *Epigenomics* **4**, 317–324 (2012). doi:10.2217/epi.12.18 [Medline](#)
  40. L. Chen, P. Liu, T. C. Evans Jr., L. M. Ettwiller, DNA damage is a pervasive cause of sequencing errors, directly confounding variant identification. *Science* **355**, 752–756 (2017). doi:10.1126/science.aai8690 [Medline](#)

## ACKNOWLEDGMENTS

This work was supported by the HPC facilities operated by, and the staffs of, the Yale Center for Research Computing and Yale's W.M. Keck Biotechnology Laboratory, as well as NIH grants RR19895 and RR029676-01, which helped fund the cluster. The sequencing data from this study have been deposited to the NIH NIMH Data Archives (<https://data-archive.nimh.nih.gov>) under collection ID # 2330 and DOI:10.15154/1410419. This work was funded by the Mayo Clinic Center For Individualize Medicine and by NIH grants R01 MH100914 (FMV), U01 MH106876 (FMV, AA, AEU), U01 MH106874 (NS), P50 MH106934 (NS), and R03 CA191421 (AA). AA is also a Visiting Professor at Yale Child Study Center. Supplement contains additional data.

## SUPPLEMENTARY MATERIALS

[www.sciencemag.org/cgi/content/full/science.aan8690/DC1](http://www.sciencemag.org/cgi/content/full/science.aan8690/DC1)

Materials and Methods

Figs. S1 to S24

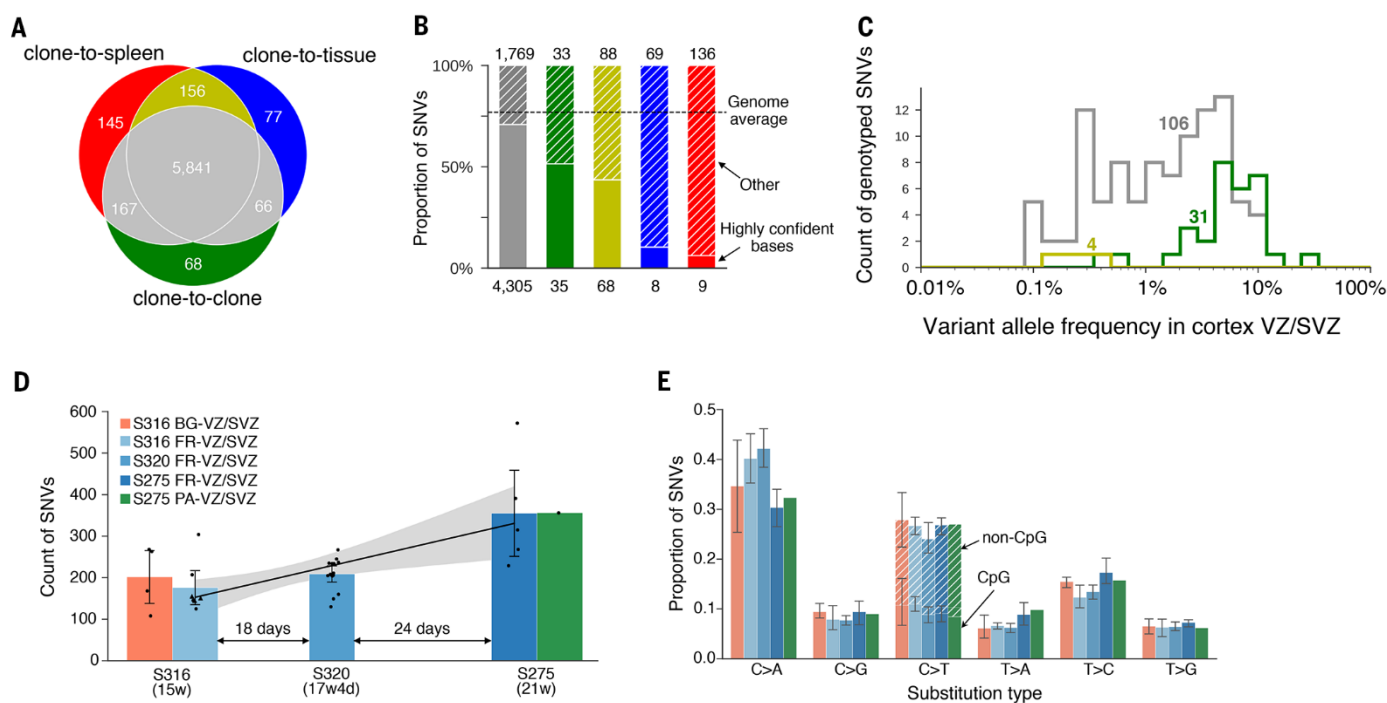
Tables S1 to S4

References (37–40)

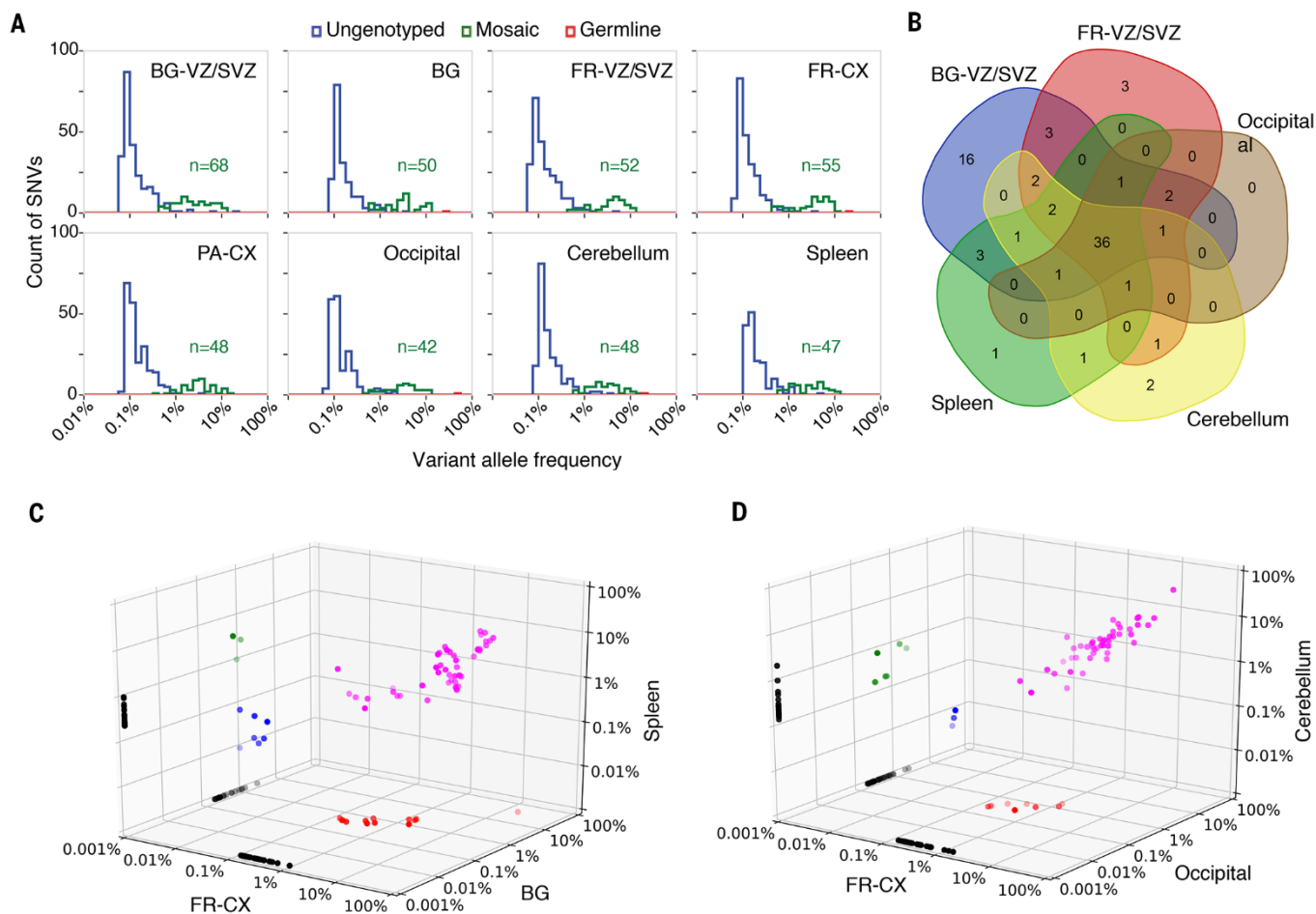
23 June 2017; accepted 28 November 2017

Published online 7 December 2017

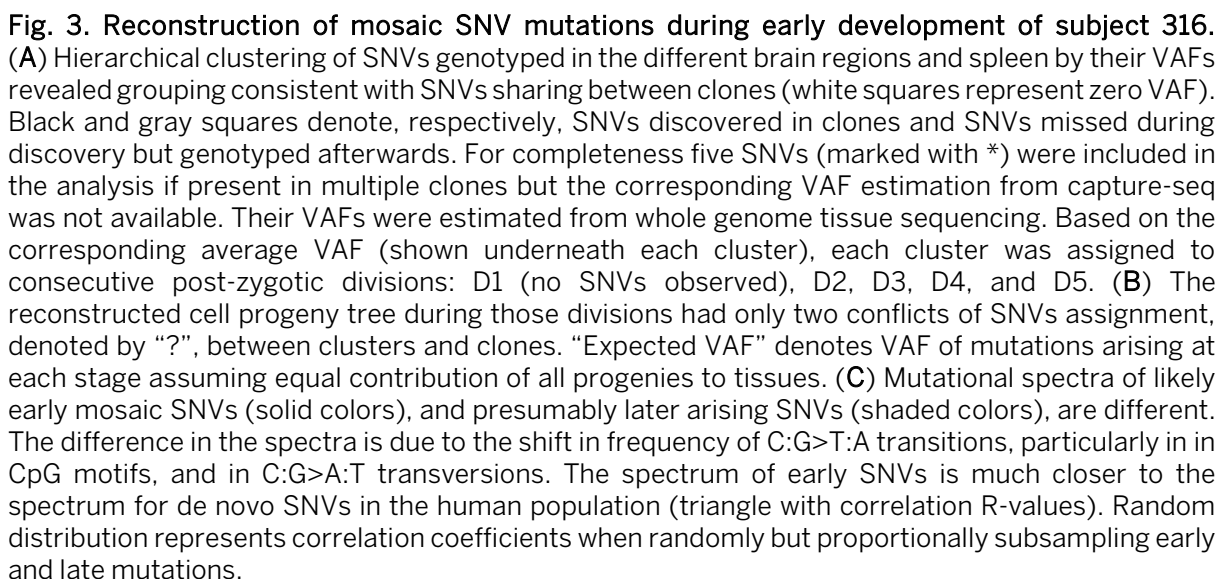
10.1126/science.aan8690

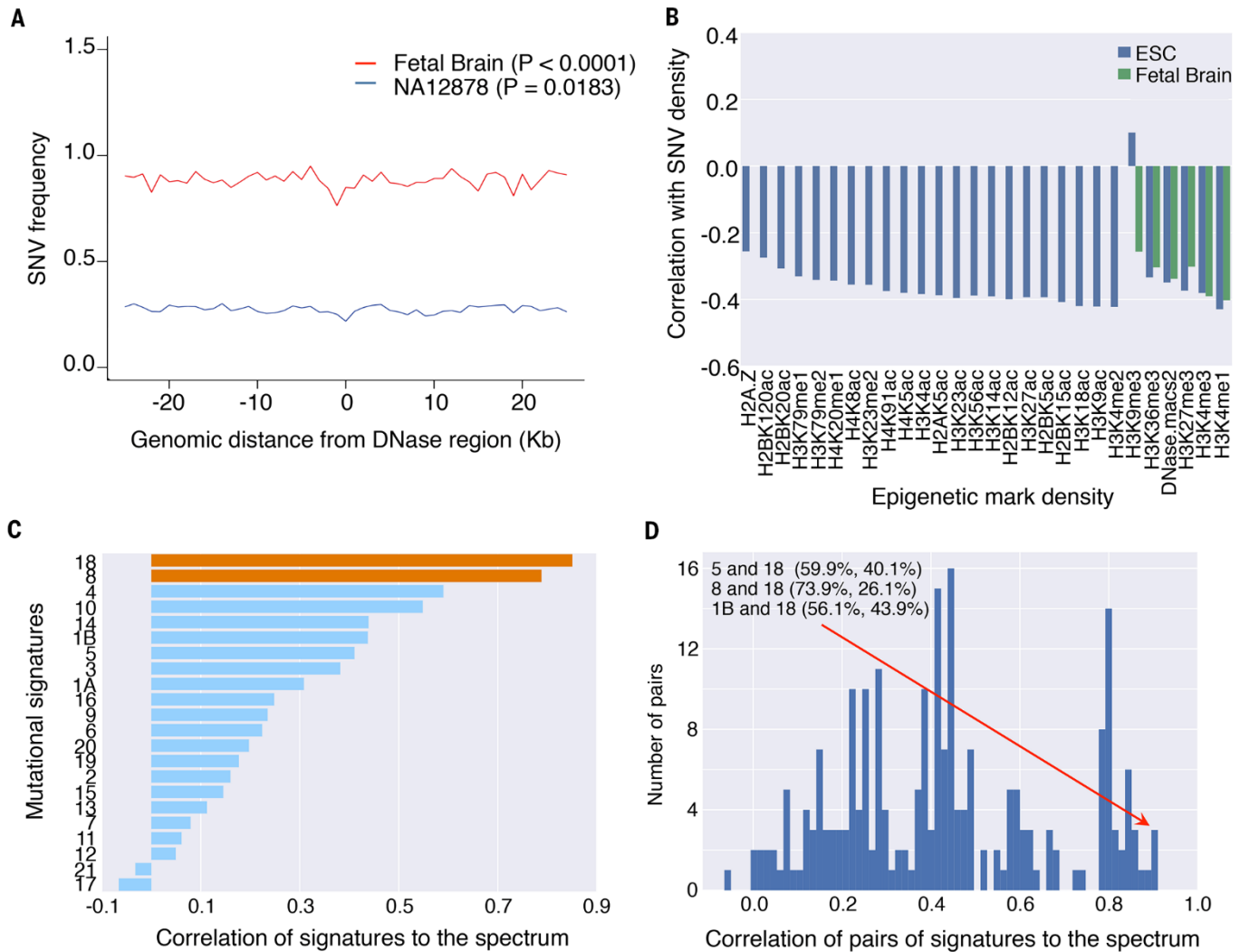


**Fig. 1. SNV discovery in brains.** (A) Three approaches of discovering mosaic SNVs were contrasted: comparing clones to the VZ/SVZ tissue of origin, comparing clones to the spleen, and comparing clones with each other (see fig. S3). The three approaches give largely concordant calls. The comparison is for calls from all three brains. (B) Calls unique to clone-to-original tissue (in blue) and clone-to-spleen (in red) discovery approaches are dramatically enriched for bases with less confident calling (as defined by the mask of the 1000 Genomes Project). These residual calls were not included in the final call set. (C) VAF of genotyped SNVs from deep re-sequencing in all three brains. The clone-to-clone discovery approach allows finding high frequency mosaic SNVs in brain tissue (green line) that are missed from clone-to-tissue/spleen comparisons. (D) Counts of mosaic SNVs per clone increase linearly with fetal age. (E) Contribution of each substitution type to the mutation spectrum is not different between different fetuses and brain regions.



**Fig. 2. Genotyping of SNV in original tissues.** (A) Several dozens of mosaic SNVs with VAF of 0.3% to 30% in tissues from various brain regions and from spleen are genotyped by the capture-resequencing approach (green line). For hundreds more SNVs, the evidence for presence in tissue is indistinguishable from background noise (blue line). (B) Venn diagram of genotyped mosaic SNVs across brain regions and spleen for subject 316. Almost 60% of mosaic SNVs could be genotyped in one or more brain regions and spleen and 44% could be genotyped in all brain regions and spleen. (C and D) Comparative VAFs for mosaic SNVs across different brain regions and spleen for the same subject. Many SNVs are shared by multiple brain regions and by brain and spleen with similar VAFs (shared across two tissues are in green, red and blue, while shared across three tissues are in magenta).





**Fig. 4. Properties of mosaic SNVs in brain.** (A) Depletion of mosaic SNVs in DNase hypersensitive sites, possibly indicating a better efficiency of DNA repair pathways in those regions (22, 23). (B) Density of mosaic SNVs correlates negatively with histone marks in embryonic stem cells and fetal brain, revealing similarity to somatic SNVs in cancers. (C) Mutational signatures 8 and 18 found in brain cancers have highest correlations with the mutation spectrum of mosaic SNVs. (D) Exhaustive combinations of pairs of signatures consistently shows that signatures 1B and 5 also contribute to description of the mutation spectrum in combination with signature 18. Thus, signature 18 is the best descriptor of mosaic SNVs in developing brain.

## REPORT

## NEURODEVELOPMENT

# Aging and neurodegeneration are associated with increased mutations in single human neurons

Michael A. Lodato,<sup>1,2,3\*</sup> Rachel E. Rodin,<sup>1,2,3,4\*</sup> Craig L. Bohrson,<sup>5\*</sup> Michael E. Coulter,<sup>1,2,3,4\*</sup> Alison R. Barton,<sup>5\*</sup> Minseok Kwon,<sup>5\*</sup> Maxwell A. Sherman,<sup>5</sup> Carl M. Vitzthum,<sup>5</sup> Lovelace J. Luquette,<sup>5</sup> Chandri N. Yandava,<sup>6</sup> Pengwei Yang,<sup>6</sup> Thomas W. Chittenden,<sup>6,7,8</sup> Nicole E. Hatem,<sup>1,2,3</sup> Steven C. Ryu,<sup>1,2,3</sup> Mollie B. Woodworth,<sup>1,2,3†</sup> Peter J. Park,<sup>5,9†</sup> Christopher A. Walsh<sup>1,2,3†</sup>

It has long been hypothesized that aging and neurodegeneration are associated with somatic mutation in neurons; however, methodological hurdles have prevented testing this hypothesis directly. We used single-cell whole-genome sequencing to perform genome-wide somatic single-nucleotide variant (sSNV) identification on DNA from 161 single neurons from the prefrontal cortex and hippocampus of 15 normal individuals (aged 4 months to 82 years), as well as 9 individuals affected by early-onset neurodegeneration due to genetic disorders of DNA repair (Cockayne syndrome and xeroderma pigmentosum). sSNVs increased approximately linearly with age in both areas (with a higher rate in hippocampus) and were more abundant in neurodegenerative disease. The accumulation of somatic mutations with age—which we term *genosenium*—shows age-related, region-related, and disease-related molecular signatures and may be important in other human age-associated conditions.

**A**ging in humans brings increased incidence of nearly all diseases, including neurodegenerative diseases (1). Markers of DNA damage increase in the brain with age (2), and genetic progeroid diseases such as Cockayne syndrome (CS) and xeroderma pigmentosum (XP), both caused by defects in DNA damage repair (DDR), are associated with neurodegeneration and premature aging (3). Mouse models of aging, CS, and XP have shown inconsistent relationships between these conditions and the accumulation of permanent somatic mutations in brain and nonbrain tissue (4–7). Although analysis of human bulk brain DNA, composed of multiple proliferative and nonproliferative cell types, revealed an accumulation of mutations during aging in the human brain (8), it is not known whether permanent somatic mutations accumulate with

age in mature neurons of the human brain. Here, we quantitatively examined whether aging or disorders of defective DDR result in more somatic mutations in single postmitotic human neurons.

Somatic mutations that occur in postmitotic neurons are specific to each cell and thus can only be comprehensively assayed by comparing the genomes of single cells (9). Therefore, we analyzed human neurons by single-cell whole-genome sequencing (WGS). Because alterations of the prefrontal cortex (PFC) have been linked to age-related cognitive decline and neurodegenerative disease (10), we analyzed 93 neurons from PFC of 15 neurologically normal individuals (Table 1 and tables S1 and S2) from ages 4 months to 82 years. We further examined 26 neurons from the hippocampal dentate gyrus (DG) of 6 of these individuals because the DG is a focal point for other age-related degenerative conditions such as Alzheimer's disease. Moreover, the DG is one of the few parts of the brain that appears to undergo neurogenesis after birth (11), which might create regional differences in number and type of somatic mutations. Finally, to determine whether defective DDR in early-onset neurodegenerative diseases is associated with increased somatic mutations, we analyzed 42 PFC neurons from 9 individuals diagnosed with CS or XP (Table 1 and table S3).

We isolated single neuronal nuclei using flow cytometry, lysed the nuclei on ice in alkaline conditions as previously performed (12, 13) (see the supplementary materials) to minimize lysis-induced artifacts, amplified their genomes using multiple displacement amplification (MDA), and subjected the amplified DNA to 45X WGS (Fig. 1A).

To identify somatic single-nucleotide variants (sSNVs) with high confidence, we developed a bioinformatic pipeline called linked-read analysis (LiRA) (14) to delineate true double-stranded sSNVs from single-stranded variants and artifacts. This method employs read-based linkage of candidate sSNVs with nearby germline SNPs and performs a model-based extrapolation of the genome-wide mutational frequency based on the ~20% of sSNVs that are sufficiently close to germline SNPs (Fig. 1B and supplementary materials). sSNVs determined by our algorithm (tables S4 and S5) showed alternate allele frequency distribution strikingly matching that of the germline SNVs (fig. S1). sSNV counts were not systematically influenced by technical metrics, such as post-mortem interval, time in storage, and coverage uniformity (fig. S2).

Across all normal neurons, genome-wide sSNV counts correlated with age (Fig. 2A and fig. S3) ( $P = 2 \times 10^{-5}$ , mixed effects model) (see the supplementary materials), despite some within-individual and within-age group heterogeneity. To explore potential variation in different brain regions, matched DG and PFC neurons were sequenced in six cases (Fig. 2A and fig. S3). Our analysis uncovered region-specific sSNV accumulation with age in both PFC ( $P = 4 \times 10^{-5}$ ) and DG single neurons ( $P = 2 \times 10^{-7}$ ), suggesting an almost twofold increase in the rate of accumulation in DG (~40 sSNVs/year) relative to PFC (~23 sSNVs/year) ( $P = 8 \times 10^{-4}$ ). Among the six cases, two had significant increases in DG, three had nominally increased counts in DG compared to PFC, and one had a nominally higher count in PFC (Fig. 2B).

Neurons from postmortem brains of CS individuals showed a ~2.3-fold excess of sSNVs relative to the expected age-adjusted normal PFC rate, and XP neurons showed a ~2.5-fold increase ( $P = 0.006$  for both) (Fig. 2C). Progeroid neurons showed a similar number of sSNVs as neurons from aged normal PFC, suggesting that defective nucleotide excision repair accelerates aging through sSNV accumulation.

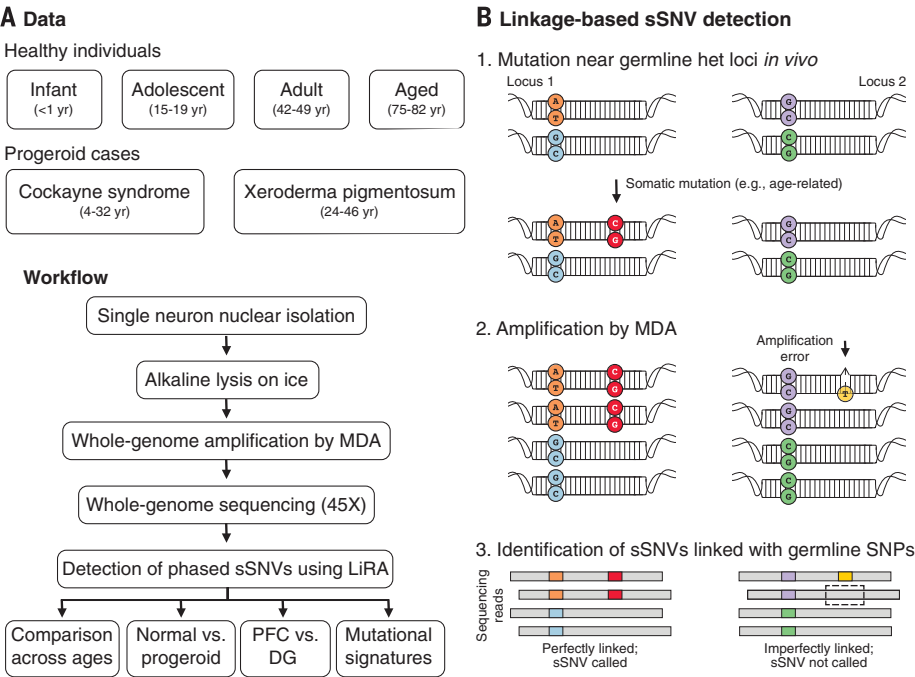
Molecular patterns of sSNVs also evolved with age. We previously reported that cytosine deamination influences patterns of human neuron sSNVs, resulting in abundant C>T mutations (13). C>T sSNVs accounted for most variants in the youngest PFC samples, but this fraction decreased with age (Fig. 2D and figs. S4 and S5). C>T mutations, although common in many biological contexts (15–18), are also a known artifact of MDA (19). Systematic differences in C>T burden during aging suggest that C>T variants are largely biological and not technical in nature. T>C variants increased in the PFC with age (Fig. 2E and figs. S4 and S5), possibly representing DNA damage linked to fatty-acid oxidation (20). As demonstrated previously (13), neuronal sSNVs in normal PFC were enriched in coding exons (fig. S6 and table S6) and displayed a transcriptional strand bias (fig. S7), and genes involved in neural function were enriched for neuronal sSNVs (fig. S8 and table S7). Coincident probability modeling suggested that the linear accumulation of sSNVs in our data set would result in an exponential

<sup>1</sup>Division of Genetics and Genomics, Manton Center for Orphan Disease, and Howard Hughes Medical Institute, Boston Children's Hospital, Boston, MA, USA. <sup>2</sup>Departments of Neurology and Pediatrics, Harvard Medical School, Boston, MA, USA. <sup>3</sup>Broad Institute of MIT and Harvard, Cambridge, MA, USA. <sup>4</sup>Program in Neuroscience and Harvard/MIT MD-PHD Program, Harvard Medical School, Boston, MA, USA. <sup>5</sup>Department of Biomedical Informatics, Harvard Medical School, Boston, MA, USA. <sup>6</sup>Computational Statistics and Bioinformatics Group, Advanced Artificial Intelligence Research Laboratory, WuXi NextCODE, Cambridge, MA, USA. <sup>7</sup>Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA, USA. <sup>8</sup>Division of Genetics and Genomics, Boston Children's Hospital, Harvard Medical School, Boston, MA, USA. <sup>9</sup>Division of Genetics, Brigham and Women's Hospital, Boston, MA, USA.

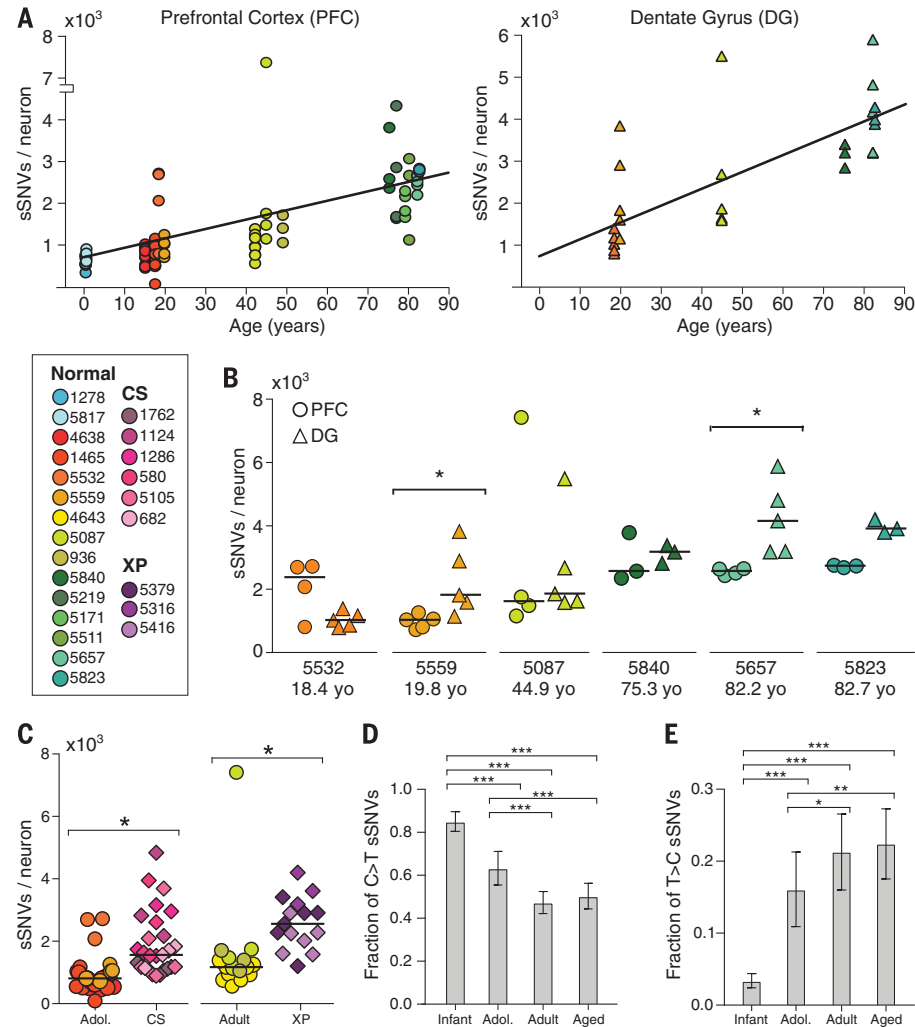
\*These authors contributed equally to this work. †Present address: Cañada College, Redwood City, CA, USA.

‡Corresponding author. Email: peter\_park@hms.harvard.edu (P.J.P.); christopher.walsh@childrens.harvard.edu (C.A.W.)

**Fig. 1. Detection of sSNVs across individuals and brain regions using single-neuron WGS and linkage-based analysis.** (A) Experimental outline of linkage-based mutation calling. A somatic mutation (red) may occur on one allele of a locus (Locus 1), potentially in close proximity to a germline heterozygous site (blue and orange), while other loci, such as Locus 2, remain unmutated. Later amplification errors could create a mismatch (yellow) on one strand of one allele of Locus 2 near a germline variant (purple). For Locus 1, any WGS read that covers both sites and contains the germline (orange) variant should also contain the somatic (red) variant; thus, these variants are perfectly linked. In contrast, at Locus 2, some but not all reads that cover the relevant germline variant (purple) will contain the somatic “variant” candidate (yellow), generating two classes of reads, some with the somatic variant on that allele, some without. Only perfectly linked sSNV candidates were considered in this study.



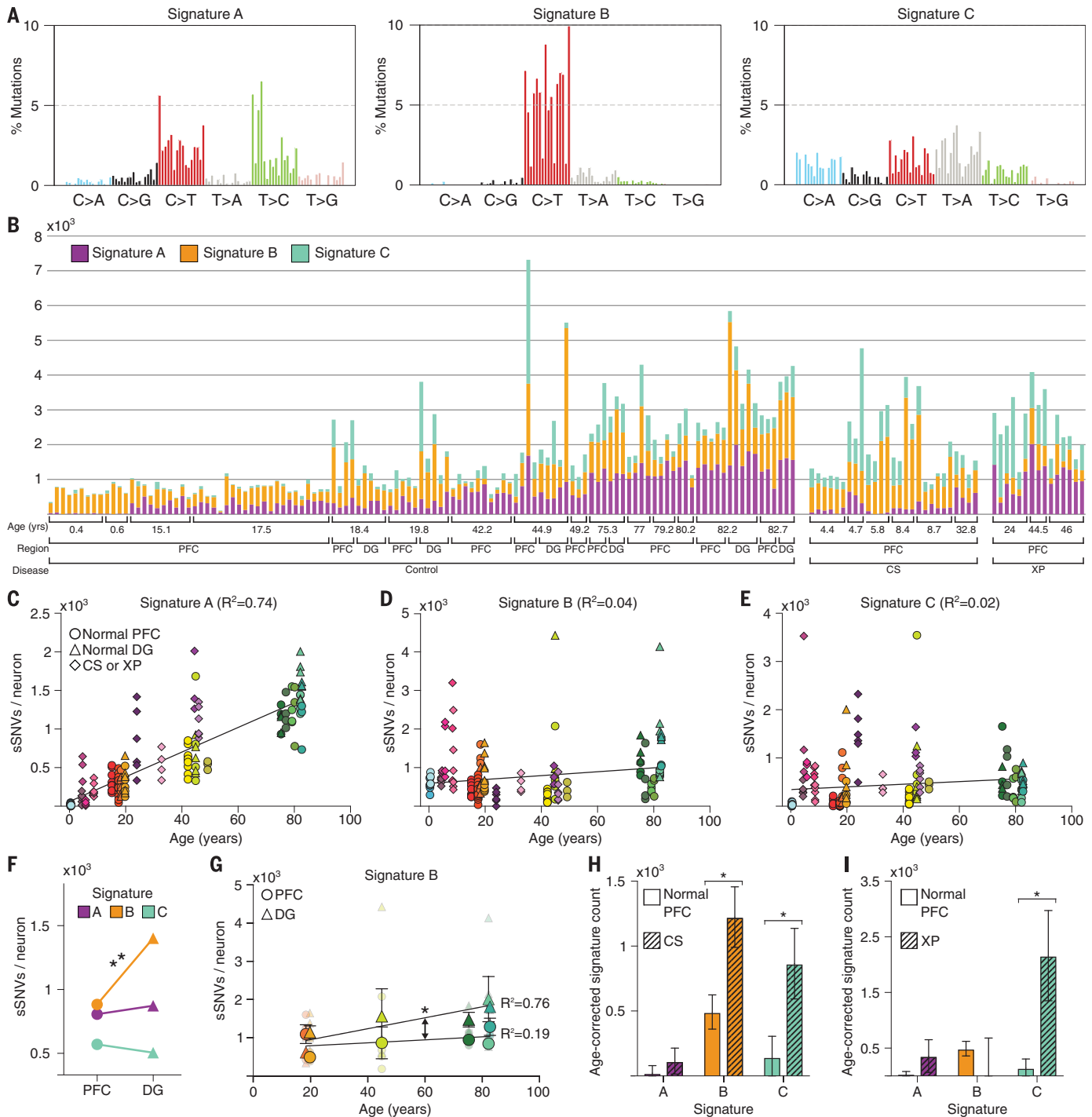
**Fig. 2. sSNVs increase with age in the PFC and DG and are elevated in CS and XP.** (A) sSNV counts plotted against age for neurons derived from PFC (circles) and DG (triangles), with linear regression lines. There is a strong correlation with age in both cases, with the rate of accumulation being nearly twofold higher in the DG than in the PFC. (B) Comparison of sSNV counts in matched PFC and DG within brains. (C) CS and XP patient neurons display elevated sSNV counts. \* denotes  $P < 0.05$  for (B) and (C). (D and E) Fraction of sSNVs composed of C>T (D) and T>C (E). \*, \*\*, and \*\*\* denote  $P \leq 0.05$ , 0.001, and 0.0001, respectively, using two-way analysis of variance with Sidak's correction.



accumulation of biallelic deleterious coding mutations, in agreement with classical hypotheses regarding the relationship between mutation and aging (fig. S9) (21), exacerbating differences in sSNV load in aging, across brain regions, and in disease.

Mutational signature analysis (22) revealed three signatures driving single-neuron mutational spectra (Fig. 3, A and B, and figs. S10 and S11). Signature A was composed mainly of C>T and T>C mutations and was the only signature to in-

crease with age ( $P = 9 \times 10^{-12}$ ), independent of brain region or disease status (Fig. 3, C to E). Signature A resembled a “clocklike” signature found in nearly all samples in a large-scale cancer genome analysis (Signature 5) (22) (fig. S10). Our data show



**Fig. 3. Signature analysis reveals mutational processes during aging, across brain regions, and in disease.** (A) Three mutational signatures identified by non-negative matrix factorization (each substitution is classified by its trinucleotide context). (B) Number of variants from Signatures A, B, and C in each of the 161 neurons in the data set. (C to E) Signature A strongly correlates with age, regardless of disease status or brain region, whereas Signatures B and C do not. (F) Signature B is

enriched in DG relative to PFC neurons. (G) Signature B increased with age in DG neurons, but not in matched PFC neurons, revealing a DG-specific aging signature. Solid shapes represent regional means, and transparent shapes represent individual neurons. (H and I) Comparison of age-corrected estimate of sSNVs per signature in CS and XP compared with PFC controls revealed enrichment in Signature C in both CS and XP. \* and \*\* denote  $P \leq 0.05$  and  $P \leq 0.001$ , respectively; mixed linear model.

that a similar clocklike signature is also active in postmitotic cells and hence independent of DNA replication.

Signature B consisted primarily of C>T mutations and did not correlate with age (Fig. 3D), suggesting a mutational mechanism active at very young ages, perhaps prenatally. Signature B may include technical artifacts, which are primarily C>T, but bona fide clonal sSNVs are also predominantly C>T (8, 12). This signature was enriched in DG compared with PFC ( $P = 2 \times 10^{-4}$ ) (Fig. 3F) and increased with age in DG, but not in PFC ( $P = 0.04$ , difference in slopes) (Fig. 3G). The observable difference in Signature B between these brain regions, and its correlation with age in DG alone, suggest that it is dominated by a biological mechanism, and these PFC-DG differences strikingly mirror differences in neurogenesis.

A third signature, Signature C, was distinguished from Signatures A and B by the presence

of C>A variants, the mutation class most closely associated with oxidative DNA damage (20). Indeed, CS and XP neurons, defective in DDR, were enriched for Signature C ( $P = 0.016$  and  $0.023$ , respectively) (Fig. 3, H and I), whereas Signature C also increased modestly with age in normal neurons ( $P = 0.03$ ). An outlier 5087 PFC neuron with the highest sSNV rate in our data set had a high proportion of Signature C mutations relative to other normal neurons, highlighting that even within a normal brain some neurons may be subject to catastrophic oxidative damage.

Our analysis revealed that sSNVs accumulated slowly but inexorably with age in the normal human brain, a phenomenon we term genosenium, and more rapidly still in progeroid neurodegeneration. Within 1 year of birth, postmitotic neurons already have ~300 to 900 sSNVs, strikingly dovetailing with the 200 to 400 sSNVs estimated to be already present in human progenitor cells at

20 weeks of gestation or earlier (23). Three signatures were associated with mutational processes in human neurons: a postmitotic, clocklike signature of aging, a possibly developmental signature that varied across brain regions, and a disease- and age-specific signature of oxidation and defective DNA damage repair. These associations were present even when data were reanalyzed after removal of all C>T mutations (fig. S11), demonstrating that any artifactual C>T mutations that may have escaped our filtering do not affect our main conclusions. The increase of oxidative mutations in aging and in disease presents a potential target for therapeutic intervention. Further, elucidating the mechanistic basis of the clocklike accumulation of mutations across brain regions and other tissues would increase our knowledge of age-related disease and cognitive decline. CS and XP cause neurodegeneration associated with higher rates of sSNVs, and it will be important to define how other, more common causes of neurodegeneration may influence genosenium as well.

Table 1. Case information and number of neurons analyzed in this study.					
Case ID	Sex	Age (years)	Diagnosis	PFC neurons	DG neurons
Infant					
1278	M	0.4	Normal	9	–
5817	M	0.6	Normal	4	–
				13	–
Adolescent					
4638	F	15.1	Normal	10	–
1465	M	17.5	Normal	22	–
5532	M	18.4	Normal	4	5
5559	F	19.8	Normal	5	5
				41	10
Adult					
4643	F	42.2	Normal	10	–
5087	M	44.9	Normal	4	5
936	F	49.2	Normal	3	–
				17	5
Aged					
5840	M	75.3	Normal	3	3
5219	F	77	Normal	4	–
5171	M	79.2	Normal	4	–
5511	F	80.2	Normal	3	–
5657	M	82	Normal	5	5
5823	F	82.7	Normal	3	3
				22	11
Cockayne syndrome					
1762	F	4.4	CS (CSB)	6	–
1124	F	4.7	CS (CSB)	3	–
1286	M	5.8	CS (CSB)	4	–
580	F	8.4	CS (CSB)	4	–
5105	M	8.7	CS (CSB)	6	–
682	M	32.8	CS (CSB)	4	–
				27	–
Xeroderma pigmentosum					
5379	F	24	XP (XPA)	6	–
5316	F	44.5	XP (XPA)	3	–
5416	F	46	XP (XPD)	6	–
				15	–
Total	24 cases			135 PFC neurons	26 DG neurons
				161 neurons	

REFERENCES AND NOTES

1. C. López-Otin, M. A. Blasco, L. Partridge, M. Serrano, G. Kroemer, *Cell* **153**, 1194–1217 (2013).  
2. T. Lu et al., *Nature* **429**, 883–891 (2004).  
3. J. A. Marteijn, H. Lans, W. Vermeulen, J. H. Hoeijmakers, *Nat. Rev. Mol. Cell Biol.* **15**, 465–481 (2014).  
4. H. Giese, M. E. Dollé, A. Hezel, H. van Steeg, J. Vijg, *Oncogene* **18**, 1257–1260 (1999).  
5. M. E. Dollé, W. K. Snyder, D. B. Dunson, J. Vijg, *Nucleic Acids Res.* **30**, 545–549 (2002).  
6. M. E. Dollé et al., *Mutat. Res.* **596**, 22–35 (2006).  
7. R. A. Busuttil et al., *PLoS ONE* **2**, e876 (2007).  
8. M. L. Hoang et al., *Proc. Natl. Acad. Sci. U.S.A.* **113**, 9846–9851 (2016).  
9. J. P. Dumanski, A. Piotrowski, *Methods Mol. Biol.* **838**, 249–272 (2012).  
10. S. J. van Veluw et al., *Brain Struct. Funct.* **217**, 797–808 (2012).  
11. P. S. Eriksson et al., *Nat. Med.* **4**, 1313–1317 (1998).  
12. G. D. Evrony et al., *Cell* **151**, 483–496 (2012).  
13. M. A. Lodato et al., *Science* **350**, 94–98 (2015).  
14. C. L. Bohrsen et al., *bioRxiv* 211169 [Preprint] 30 October 2017.  
15. G. Novese et al., *N. Engl. J. Med.* **371**, 2477–2487 (2014).  
16. J. L. Hazen et al., *Neuron* **89**, 1223–1236 (2016).  
17. I. Martincorena et al., *Science* **348**, 880–886 (2015).  
18. E. T. Lim et al., *Nat. Neurosci.* **20**, 1217–1224 (2017).  
19. Y. Hou et al., *Cell* **148**, 873–885 (2012).  
20. R. De Bont, N. van Larebeke, *Mutagenesis* **19**, 169–185 (2004).  
21. L. Szilard, *Proc. Natl. Acad. Sci. U.S.A.* **45**, 30–45 (1959).  
22. L. B. Alexandrov et al., *Nat. Genet.* **47**, 1402–1407 (2015).  
23. T. Bae et al., *Science* **359**, 550–555 (2018).

ACKNOWLEDGMENTS

We thank R. S. Hill, M. P. Anderson, J. Gulcher, Z. Chen, I. Cortez, the Dana Farber Cancer Institute Flow Cytometry Core, and the Research Computing group at Harvard Medical School for assistance. Human tissue was obtained from the NIH NeuroBioBank at the University of Maryland, and we thank the donors and their families for their invaluable donations for the advancement of science. This work was supported by K99 AG054749 01, F30 MH02909, 1S10RR028832-01, T32HG002295, U01MH106883, P50MH106933, R01 NS032457 U01 MH106883, the Harvard/MIT MD-PHD program, the Stuart H. Q. and Victoria Quan Fellowship in Neurobiology, and the Paul G. Allen Family Foundation. C.A.W. is an investigator of the Howard Hughes Medical Institute. See the supplementary materials for full acknowledgments. The supplementary materials also contain additional data.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/359/6375/555/suppl/DC1  
Materials and Methods  
Supplementary Text  
Figs. S1 to S11  
Tables S1 to S7  
References (24–52)

21 July 2017; accepted 22 November 2017  
Published online 7 December 2017  
10.1126/science.aao4426

## CARBON FIXATION

# A primordial and reversible TCA cycle in a facultatively chemolithoautotrophic thermophile

Takuro Nunoura,<sup>1\*</sup> Yoshito Chikaraishi,<sup>1,2,3</sup> Rikihisa Izaki,<sup>4</sup> Takashi Suwa,<sup>4</sup> Takaaki Sato,<sup>4</sup> Takeshi Harada,<sup>5</sup> Koji Mori,<sup>5</sup> Yumiko Kato,<sup>5</sup> Masayuki Miyazaki,<sup>1,6</sup> Shigeru Shimamura,<sup>6</sup> Katsunori Yanagawa,<sup>6†</sup> Aya Shuto,<sup>1</sup> Naohiko Ohkouchi,<sup>2</sup> Nobuyuki Fujita,<sup>5‡</sup> Yoshihiro Takaki,<sup>1,6</sup> Haruyuki Atomi,<sup>4,7</sup> Ken Takai<sup>6</sup>

Inorganic carbon fixation is essential to sustain life on Earth, and the reductive tricarboxylic acid (rTCA) cycle is one of the most ancient carbon fixation metabolisms. A combination of genomic, enzymatic, and metabolomic analyses of a deeply branching chemolithotrophic *Thermosulfidibacter takaii* ABI70S6<sup>T</sup> revealed a previously unknown reversible TCA cycle whose direction was controlled by the available carbon source(s). Under a chemolithoautotrophic condition, a rTCA cycle occurred with the reverse reaction of citrate synthase (CS) and not with the adenosine 5'-triphosphate-dependent citrate cleavage reactions that had been regarded as essential for the conventional rTCA cycle. Phylometabolic evaluation suggests that the TCA cycle with reversible CS may represent an ancestral mode of the rTCA cycle and raises the possibility of a facultatively chemolithomixotrophic origin of life.

**A**utotrophic carbon fixation is essential for sustaining life on Earth. The reductive tricarboxylic acid (rTCA) cycle and Wood-Ljungdahl pathway have been recognized as the most ancient carbon fixation pathways to ensure the biosynthesis of the five universal precursors of anabolism; acetyl-CoA (coenzyme A), pyruvate, phosphoenolpyruvate, oxaloacetate, and 2-oxoglutarate (1–5). However, essential components of the presently recognized rTCA cycle, the adenosine 5'-triphosphate (ATP)-dependent citrate cleavage enzymes ATP citrate lyase (ACL) or citryl-CoA synthetase (CCS)/citryl-CoA lyase (CCL), most likely emerged at a later stage from the domains of citrate synthase (CS) and succinyl-CoA synthetase (fig. S1) (6–9).

The Aquificae strain *Thermosulfidibacter takaii* ABI70S6<sup>T</sup> has been recognized as a chemolithoautotroph (10, 11), but this study revealed its abil-

ity for chemolithomixotrophic growth with organic acids including acetate and succinate in the presence of yeast extract (Fig. 1 and supplementary text). Hydrogen oxidation coupled with sulfur reduction was essential for growth. The genome sequence of *T. takaii* ABI70S6<sup>T</sup> (~1.8 million base pairs) (supplementary text) was not sufficient to identify the carbon fixation pathway, although the strain was expected to harbor a rTCA cycle because of its close phylogenetic relationship with other deeply branching chemolithoautotrophic Aquificae bacteria (9–11).

On the *T. takaii* genome, genes related to the (r)TCA cycle are encoded in one gene cluster, and only the isocitrate dehydrogenase (*ICDH*) gene is found in a different locus (table S1). Surprisingly, however, we could not find genes for ACL and CCS/CCL. Genes with essential motifs or signatures of these enzymes were also absent. Genes for the other known microbial carbon fixation pathways could not be identified. As for potential carbon dioxide fixation enzymes, genes for 2-oxoglutarate:ferredoxin oxidoreductase (OGOR), *ICDH*, pyruvate:ferredoxin oxidoreductase (POR), pyruvate carboxylase, malate dehydrogenase (malic enzyme), phosphoenolpyruvate carboxykinase (PEPCK), acetyl-CoA carboxylase, and propionyl-CoA carboxylase were present (fig. S2 and table S1).

The activities of these enzymes related to carbon fixation were examined with cell-free extract obtained from *T. takaii* cells grown chemolithomixotrophically. Among the potential carboxylases, activities of POR, OGOR, *ICDH*, PEPCK, acetyl-CoA carboxylase, and propionyl-CoA carboxylase were detected, whereas those of malic enzyme and pyruvate carboxylase were not. Most of the enzymatic activities in the rTCA cycle were detected (table S1). ATP-independent, CoA-dependent citrate lyase activity (reverse reaction of CS) was present, but ATP/GTP (gua-

nosine 5'-triphosphate)-citrate lyase and CoA-independent citrate lyase activities were absent (Table 1). Both citrate-generating and citrate-cleaving CS activities were detected in cells grown chemolithoautotrophically. The levels of CS activity in both directions and the malate dehydrogenase (MDH) activity were comparable in *T. takaii* cells grown chemolithoautotrophically and chemolithomixotrophically, suggesting that the two enzymes are expressed constitutively. Intriguingly, the ATP-independent turnover rate of citrate to oxaloacetate in *T. takaii* was comparable to the ATP-dependent conversion rate observed in other deeply branching chemolithoautotrophic Aquificae bacteria that utilize ACL or CCS/CCL in the classical rTCA cycle (11) (Table 1). To address the possibility that CS was involved in citrate cleavage in *T. takaii*, we examined the levels of CS activity in other deeply branching bacteria and compared them to that in *T. takaii*. As a result, the activity in *T. takaii* was at least two orders of magnitude higher than those observed in *Hydrogenobacter thermophilus* and *Persephonella marina*, which utilize the rTCA cycle for growth. All of the three citrate cleavage activities were not observed in *Thermodesulfator indicus* cells, consistent with the absence of genes for CS, ACL and CCS/CCL on its genome. These results suggest that there are no enzymes other than CS, ACL, and CCS/CCL that are involved in citrate cleavage in the deeply branching chemolithoautotrophic bacteria.

The CS activity in *T. takaii* was the highest among the enzyme activities tested. A high CS activity is unlikely to co-occur with carbon fixation pathways using acetyl-CoA as an intermediate such as the archaeal 3-hydroxypropionate/4-hydroxybutyrate (3HP/4HB) and dicarboxylate/4-hydroxybutyrate (DC/HB) cycles, as it would compete in consuming acetyl-CoA (12). In *Thermoproteus neutrophilus*, which utilizes the DC/HB cycle for CO<sub>2</sub> fixation and harbors a TCA cycle, the specific activity of CS was only 2% of that of MDH (13). Enzyme activity measurements in *T. takaii* also suggested the absence of these archaeal pathways (table S1).

RNA sequencing of *T. takaii* cells indicated that all of the genes involved in the TCA cycle were transcribed under both chemolithoautotrophic and chemolithomixotrophic (with acetate and/or succinate) growth conditions (table S2). Shotgun proteome analysis of the cell-free extract also confirmed the expression of most of these genes in cells grown chemolithoautotrophically and chemolithomixotrophically with acetate or succinate (table S2).

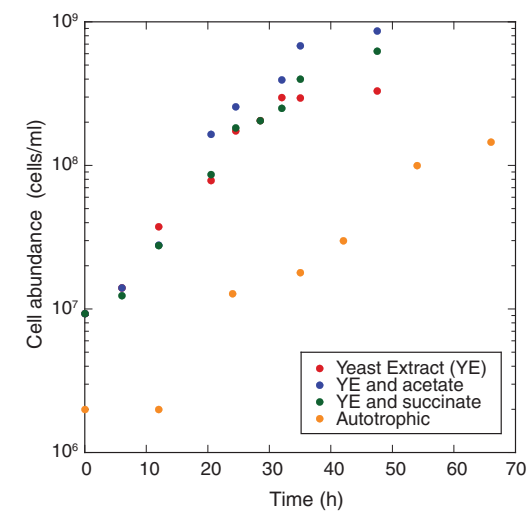
The carbon fixation pathway of *T. takaii* was further characterized by isotopologue analysis of protein-derived amino acids in cells grown with <sup>13</sup>C-labeled substrates (fig. S3). Genomic information suggests that alanine (Ala), aspartate (Asp), and glutamate (Glu) are synthesized from pyruvate, oxaloacetate, and 2-oxoglutarate by alanine aminotransferase, aspartate aminotransferase, and glutamate dehydrogenase, respectively (fig. S4). The isotopologues of amino acids from cells grown chemolithoautotrophically

<sup>1</sup>Research and Development Center for Marine Biosciences, Japan Agency for Marine-Earth Science and Technology (JAMSTEC), 2-15 Natsushima-cho, Yokosuka 237-0061, Japan. <sup>2</sup>Department of Biogeochemistry, Japan Agency for Marine-Earth Science and Technology (JAMSTEC), 2-15 Natsushima-cho, Yokosuka 237-0061, Japan. <sup>3</sup>Institute of Low Temperature Science, Hokkaido University, Kita-19, Nishi-8, Kita-ku, Sapporo 060-0819, Japan. <sup>4</sup>Department of Synthetic Chemistry and Biological Chemistry, Graduate School of Engineering, Kyoto University, Katsura, Nishikyo-ku, Kyoto 615-8510, Japan. <sup>5</sup>NITE Biological Resource Center (NBRC), National Institute of Technology and Evaluation (NITE), 2-49-10 Nishihara, Shibuya-ku, Tokyo 151-0066, Japan. <sup>6</sup>Department of Subsurface Geobiological Analysis and Research, Japan Agency for Marine-Earth Science and Technology (JAMSTEC), 2-15 Natsushima-cho, Yokosuka 237-0061, Japan. <sup>7</sup>Core Research for Evolutional Science and Technology (CREST), Japan Science and Technology Agency (JST), 7 Goban-cho, Chiyoda-ku, Tokyo 102-0076, Japan.

\*Corresponding author. Email: takuron@jamstec.go.jp †Present address: Faculty of Environmental Engineering, The University of Kitakyushu, Japan. ‡Present address: Faculty of Life Sciences, Tokyo University of Agriculture, Japan.

in the presence of [<sup>13</sup>C]bicarbonate for 3 hours can be explained by a citrate cleavage reaction and production of [1-<sup>13</sup>C] acetyl-CoA through the rTCA cycle (fig. S5 and tables S3 and S4). Intriguingly, although anaplerosis is necessary to sustain the level of cycle intermediates and maintain chemolithoautotrophy, the contribution of anaplerotic pathways was not clearly observed. Further studies will be necessary to clarify this issue. Shotgun proteome analysis of the cell-free extract of cells grown chemolithoautotrophically confirmed the presence of pyruvate carboxylase, PEPC, and malic enzyme (fig. S2), indicating the potential for anaplerosis with these enzymes.

The isotopologues from cells grown chemolithomixotrophically in the presence of succinate were examined with either [2,3-<sup>13</sup>C]succinate or [<sup>13</sup>C]bicarbonate. The isotopologue pattern can be explained by a bifurcated TCA cycle (fig. S5 and tables S3 and S4). The isotopologue pattern of Glu indicated the reductive generation of 2-oxoglutarate from succinate via succinyl-CoA. In contrast, that of Ala suggested that pyruvate was formed via decarboxylation of malate or oxaloacetate and not via conversion of acetyl-CoA to pyruvate. Moreover, that of Asp also indicated the dominance of oxidative generation



**Fig. 1. Growth curves of *T. takaii* grown under chemolithoautotrophic and chemolithomixotrophic conditions.** A gas mixture of H<sub>2</sub> and CO<sub>2</sub> (80:20) was used for all cultivation experiments in this study.

of oxaloacetate from succinate, and a portion was supplied by reduction from succinate via succinyl-CoA, 2-oxoglutarate, isocitrate, and citrate. Therefore, in this case, succinate proceeds in both the oxidative and reductive directions of the cycle, converging at acetyl-CoA via oxalo-

acetate and pyruvate in the former case and via 2-oxoglutarate, citrate, and citrate cleavage in the latter.

The isotopologue patterns in cells grown with acetate suggested a bifurcated TCA pathway from acetyl-CoA (fig. S5 and tables S3 and S4). Growth with acetate was examined in the presence of [1-<sup>13</sup>C] acetate, [1,2-<sup>13</sup>C]acetate, or [<sup>13</sup>C]bicarbonate. The isotopologue pattern of Ala suggested that pyruvate was generated from acetyl-CoA (carboxylation), whereas that of Glu suggested the generation of 2-oxoglutarate via oxidation (decarboxylation) of citrate, and not via reduction (carboxylation) of oxaloacetate. The presence of [5-<sup>13</sup>C] or [4,5-<sup>13</sup>C<sub>2</sub>]Glu in the cells grown with [1-<sup>13</sup>C] or [1,2-<sup>13</sup>C<sub>2</sub>]acetate, respectively, implied that a portion of cellular oxaloacetate was supplied from the yeast extract component in the medium. Examination of Asp and Ala indicated that pyruvate was generated from acetyl-CoA derived from acetate and oxaloacetate was formed via carboxylation of PEP or pyruvate. Therefore, in this case, acetate acts as the precursor of acetyl-CoA.

Intriguingly, in the presence of both succinate and acetate with [<sup>13</sup>C]bicarbonate, signatures of inorganic carbon fixation were not observed in the isotopologues of amino acids (fig. S5 and

Enzymatic activities (nmol product min <sup>-1</sup> mg <sup>-1</sup> protein)							
Species		<i>Thermosulfidibacter takaii</i>		<i>Thermovibrio ammonificans</i>	<i>Hydrogenobacter thermophilus</i>	<i>Persephonella marina</i>	<i>Thermodesulfator indicus</i>
Enzymes	Growth conditions	Chemolithomixotrophic	Chemolitho-autotrophic	Chemolitho-autotrophic	Chemolitho-autotrophic	Chemolitho-autotrophic	Chemolitho-mixotrophic
		acetate	succinate				
CS	Oxaloacetate production*	403	520	96	-	-	-
	CoA production	69,600	119,000	29,800	-	88 <sup>†</sup>	190
	CDS	TST_0783		-	-	PERMA_1528	-
ACL and CCS/CCL	Oxaloacetate production*	(372)	(620)	(71)	330	260	100
	ACL	-	-	Theam_1021, Theam_1022		-	PERMA_1714, PERMA_1715
	CCS	-	-	-	HTH_0201, HTH_1737	-	-
	CCL	-	-	-	HTH_0311	PERMA_1040	-
MDH	NADH consumption	16,400	11,500	10,300	5,100	5,500	17,300
	NADPH consumption	2,450	825	N.T.	2100	200	210
	CDS	TST_0784		Theam_0602	HTH_0756	PERMA_1130	Thein_1427

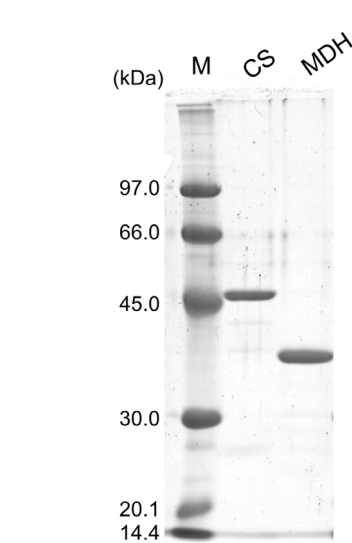
\*Oxaloacetate production was assessed by NADH consumption coupled with endogenous malate dehydrogenase.

†The CS activity was interpreted as a reverse reaction of CCL (8).

tables S3 and S4). This implied that the flux with both succinate and acetyl-CoA proceeded in the direction of decarboxylation. This is in contrast to the rTCA cycle with ACL from *Chlorobaculum tepidum*, which maintains its direction even in the presence of acetate (14). Inorganic carbon fixation examined with [ $^{13}\text{C}$ ]bicarbonate was also absent in cells grown chemolithomixotrophically with yeast extract as the sole organic carbon source. The result suggests that the directions of the TCA cycle in the presence of succinate and/or acetate were not affected by the presence of yeast extract.

A phylogenetic analysis of the CS domain indicated that the CS from *T. takaii* belonged to a cluster of CSs from deeply branching anaerobic bacteria, Deltaproteobacteria, Bacteroidetes, and eukaryotes (fig. S6). The CSs of *Desulfurella acetivorans* in Deltaproteobacteria, one of which functioned in the rTCA cycle (15), were also found in this cluster, whereas Deltaproteobacteria is distinct from Aquificae in the small subunit ribosomal RNA gene phylogeny. The phylogenetic topology suggests that both ACL and CCL likely evolved from a lineage of CSs. Among the characterized CSs, the CS from *T. takaii* was most related to the enzymes from eukaryotes, and there were no marked differences that suggest a bidirectional role for the *T. takaii* enzyme (fig. S7).

We produced and purified recombinant CS and MDH from *T. takaii* (Fig. 2). In the case of CS, we measured activity in both the direction of citrate generation and cleavage. With CS alone, only very small amounts of oxaloacetate were formed, and it was difficult to measure the activity using MDH. With an increase in enzyme and longer reaction times, we clearly observed generation of acetyl-CoA by high-performance liquid chromatography (HPLC) analysis (fig. S8A), indicating the occurrence of the citrate-cleaving activity. When we directly coupled the CS activity with the MDH activity, a significant amount of ATP-independent citrate lyase activity was observed, which was dependent on the presence of CS, citrate, and CoA (Table 2 and fig. S9). Production of acetyl-CoA under these conditions was also confirmed by HPLC (fig. S8B). The Michaelis constant ( $K_m$ ) values for citrate (8.6  $\mu\text{M}$ ) and CoA (17.4  $\mu\text{M}$ ) were extremely low, even when compared to those for previously characterized ATP-citrate lyase enzymes (Table 3). We next examined the citrate-generating activity. Measurements were carried out by quantifying the generation of CoA with varying concentrations of acetyl-CoA or oxaloacetate in the presence of 3000  $\mu\text{M}$  oxaloacetate or 1000  $\mu\text{M}$  acetyl-CoA, respectively (Table 2 and fig. S9). During the analysis at 70°C, we observed thermal degradation of acetyl-CoA to CoA, which was problematic particularly in measurements at low concentrations of oxaloacetate and 1000  $\mu\text{M}$  acetyl-CoA. We thus carried out measurements in the presence of 200  $\mu\text{M}$  acetyl-CoA when varying the concentration of oxaloacetate. The  $k_{\text{cat}}/K_m$  values for the citrate-generating reaction were



**Fig. 2. SDS-polyacrylamide gel electrophoresis of recombinant citrate synthase (CS) (TST\_0783) and malate dehydrogenase (MDH) (TST\_0784) proteins from *T. takaii* individually expressed in *E. coli*.** One microgram of protein was loaded onto each lane.

lower than expected, and this may be due to the thermal degradation of acetyl-CoA in our reactions. However, the high affinity and high  $k_{\text{cat}}/K_m$  value with citrate (and CoA) support our hypothesis that the CS from *T. takaii* functions in the direction of citrate cleavage in this organism. In the case of MDH, we could carry out kinetic analyses on the reaction from oxaloacetate to malate. The  $K_m$  value for oxaloacetate was  $20.1 \pm 6.2 \mu\text{M}$ . The  $V_{\text{max}}$  value was relatively high ( $1550 \pm 155 \mu\text{mol min}^{-1} \text{mg}^{-1}$ ), and the  $k_{\text{cat}}/K_m$  value was calculated as  $43,300 \text{ s}^{-1} \text{mM}^{-1}$  (Table 3 and fig. S10). We cannot judge whether these values are high or not, as there are no MDH proteins that have been characterized from organisms relying on the rTCA cycle for growth. We further examined the overall equilibrium of the CS and MDH reactions converting citrate, CoA, and reduced nicotinamide adenine dinucleotide (NADH) to malate, acetyl-CoA, and  $\text{NAD}^+$  at 50°C and obtained an equilibrium constant of  $K = 0.046 \pm 0.009$  (fig. S11). Analyses of the two enzymes from

*T. takaii* suggest that the enzymes can function in the direction of the rTCA cycle.

Accordingly, we conclude that *T. takaii* constitutively expresses and utilizes a reversible TCA cycle dependent on a bidirectional function of CS as its main carbon fixation pathway. Energetic and kinetic considerations theoretically support this possibility, given a sufficient energy supply via electron bifurcation. The proposed rTCA cycle in *T. takaii* includes three endergonic reactions: the ATP-independent citrate cleavage reaction (the free energy change under standard conditions  $\delta G^0 = +37.6 \text{ kJ/mol}$ ) (16) and the oxoglutarate synthase (OGOR with reduced ferredoxin) ( $\delta G^0 = +19 \text{ kJ/mol}$ ) and pyruvate synthase (POR with reduced ferredoxin) ( $\delta G^0 = +19 \text{ kJ/mol}$ ) reactions (17) (table S5). The overall reductive conversion of succinate to oxaloacetate via 2-oxoglutarate, isocitrate, and citrate is endergonic ( $\delta G^0 = +55 \text{ kJ}$ ). The reductive conversion from oxaloacetate to succinate is exergonic;  $\delta G^0 = -95 \text{ kJ}$  when fumarate is reduced by NADH (table S5). The conversion should become even more favorable when reduced ferredoxin is used for fumarate reduction. Several reducing reactions in anaerobic organisms are driven by electron donors with low redox potential, such as reduced ferredoxins. The supply of electron donors with low redox potential, critical in sustaining the driving force of the reaction in the reducing direction, is made possible by flavin-based electron bifurcation (17, 18). Therefore, sufficient depletion of both acetyl-CoA and oxaloacetate and an abundant supply of reducing equivalents via electron bifurcation should allow the cycle to overcome the energetically unfavorable citrate-cleavage reaction of CS. The obligate hydrogenotrophic chemolithotrophy of *T. takaii* may serve to ensure the sufficient supply of reduced ferredoxin for POR, OGOR, and possibly fumarate reductase. By contrast, when *T. takaii* is grown chemolithomixotrophically with acetate, isotopologue analysis indicated that 2-oxoglutarate is generated solely by oxidative conversion from acetyl-CoA and oxaloacetate, driven by the exergonic CS reaction (fig. S5). When *T. takaii* is grown chemolithomixotrophically with succinate, isotopologue analysis indicated that a major portion of the oxaloacetate is generated by oxidative conversion of succinate. A proton motive force

**Table 2. Kinetic parameters of citrate synthase (CS) from *T. takaii*.**

Substrate	$V_{\text{max}}$ ( $\mu\text{mol min}^{-1} \text{mg}^{-1}$ )	$k_{\text{cat}}$ ( $\text{s}^{-1}$ )	$K_m$ (mM)	$k_{\text{cat}}/K_m$ ( $\text{s}^{-1} \text{mM}^{-1}$ )
Acetyl-CoA	$298 \pm 13$	$250 \pm 11$	$0.0563 \pm 0.0074$	4440
Oxaloacetate	$298 \pm 13^*$	$250 \pm 11^*$	$0.0412 \pm 0.0065$	6060
CoA	$1.60 \pm 0.04$	$1.34 \pm 0.03$	$0.0174 \pm 0.0029$	771
Citrate	$1.11 \pm 0.02$	$0.931 \pm 0.017$	$0.00864 \pm 0.00116$	108

\*Taken from the data examining acetyl-CoA. The maximum initial velocity calculated from the data in the presence of 200  $\mu\text{M}$  acetyl-CoA was  $V_{\text{max}} = 239 \pm 12 (\mu\text{mol min}^{-1} \text{mg}^{-1})$ , corresponding to  $k_{\text{cat}} = 200 \pm 10 (\text{s}^{-1})$ .

Table 3. Kinetic parameters of citrate synthase (CS) for citrate and malate dehydrogenase (MDH) for oxaloacetate from *T. takaii*, and comparison with those of ATP-citrate lyases (ACLs).

Enzyme	Source organism	$V_{\max}$ citrate or specific activity ( $\mu\text{mol min}^{-1} \text{mg}^{-1}$ )	$k_{\text{cat}}$ ( $\text{s}^{-1}$ )	$K_{\text{m}}$ citrate (mM)	$k_{\text{cat}}/K_{\text{m}}$ ( $\text{s}^{-1} \text{mM}^{-1}$ )	Refs.
CS	<i>Thermosulfidibacter takaii</i>	1.11 $\pm$ 0.02	0.931 $\pm$ 0.017*	0.00864 $\pm$ 0.00116	108*	
ACL	<i>Hydrogenobacter thermophilus</i>	26.7	19.3*	6.25	3.09*	(23)
ACL	<i>Chlorobium limicola</i>	1.1	1.10*	0.23	4.78*	(24)
ACL	<i>Rattus norvegicus</i>	6.25	52.1	0.56	93.0	(25)
ACL	<i>Penicillium spiculisprum</i>	3	10.0	0.18	55.6	(26)
ACL	<i>Lipomyces starkeyi</i>	0.92	7.82	0.07	112	(27)
		$V_{\max}$ ( $\mu\text{mol min}^{-1} \text{mg}^{-1}$ )	$k_{\text{cat}}$ ( $\text{s}^{-1}$ )	$K_{\text{m}}$ oxaloacetate (mM)	$k_{\text{cat}}/K_{\text{m}}$ ( $\text{s}^{-1} \text{mM}^{-1}$ )	
MDH	<i>Thermosulfidibacter takaii</i>	1550 $\pm$ 155	871 $\pm$ 87*	0.0201 $\pm$ 0.0062	43,300*	

\*Per monomer

maintained by hydrogen oxidation coupled with S<sup>0</sup> reduction should be necessary for the endergonic oxidation of succinate in these cells (fig. S5), as in the case of *Desulfuromonas acetioxidans* (19).

The rTCA cycle is believed to be one of the most ancient carbon fixation pathways on Earth. Our results suggest that the ancient rTCA cycle in the early life on Earth did not necessarily require an ACL or CCS/CCL and originated from a bidirectional CS-dependent TCA cycle as identified in *T. takaii*. Identification of this previously unknown reversible TCA cycle and its functional regulation by the availability of carbon sources provides new insights that will help us understand the evolution of carbon fixation and central metabolisms in the earliest organisms on Earth. There is a long history of debate over heterotrophic versus autotrophic origin-of-life scenarios (20). Recently, abiotic synthesis of rTCA cycle components was observed in the presence of sulfate radicals (27). The directions of the individual conversions depended on the reaction, as there was no external energy input. When energy supplies are sufficient, the metabolic commitment toward autotrophy and heterotrophy driven by the primordial TCA cycle in *T. takaii* is flexible and would be controlled by environmental conditions such as availability of carbon sources. It has been pointed out that the core anabolism is universal at the ecosystem level, and the configuration of autotrophy or heterotrophy is an ecological distinction (22). The reversible TCA cycle, as seen in *T. takaii*, would meet these anabolic requirements and serve both autotrophic and heterotrophic life-styles. The versatile TCA cycle would allow the earliest possible chemolithotrophic forms of

life, with their limited metabolic network and catalytic enzyme (gene) resources, to make the most of what is available in their environment, such as in deep-sea hydrothermal vents. Thus, ancestral forms of life may have originated in the form of facultative autotrophs or mixotrophs. Further investigations into the possibilities of a facultative chemolithomixotrophic origin of life will provide key insights into proposing more likely scenarios for the origin of life on Earth.

REFERENCES AND NOTES

1. G. Wachtershäuser, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 200–204 (1990).  
2. G. D. Cody et al., *Geochim. Cosmochim. Acta* **65**, 3557–3576 (2001).  
3. E. Smith, H. J. Morowitz, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 13168–13173 (2004).  
4. R. Braakman, E. Smith, *PLOS Comput. Biol.* **8**, e1002455 (2012).  
5. V. Srinivasan, H. J. Morowitz, *Biol. Bull.* **216**, 126–130 (2009).  
6. L. B. Sánchez, M. Y. Galperin, M. Müller, *J. Biol. Chem.* **275**, 5794–5803 (2000).  
7. T. Kanao, T. Fukui, H. Atomi, T. Imanaka, *Eur. J. Biochem.* **268**, 1670–1678 (2001).  
8. M. Aoshima, M. Ishii, Y. Igarashi, *Mol. Microbiol.* **52**, 763–770 (2004).  
9. M. Hügl, H. Huber, S. J. Molyneux, C. Vetriani, S. M. Sievert, *Environ. Microbiol.* **9**, 81–92 (2007).  
10. T. Nunoura, H. Oida, M. Miyazaki, Y. Suzuki, *Int. J. Syst. Evol. Microbiol.* **58**, 659–665 (2008).  
11. T. Nunoura, in *The Prokaryotes, The Prokaryotes – Other Major Lineages of Bacteria and the Archaea*, E. Rosenberg, E. F. DeLong, S. Lory, E. Stackebrandt, F. Thompson, Eds. (Springer, 2014), 617–625.  
12. I. A. Berg et al., *Nat. Rev. Microbiol.* **8**, 447–460 (2010).  
13. W. H. Ramos-Vera, I. A. Berg, G. Fuchs, *J. Bacteriol.* **191**, 4286–4297 (2009).  
14. X. Feng, K. H. Tang, R. E. Blankenship, Y. J. Tang, *J. Biol. Chem.* **285**, 39544–39550 (2010).  
15. A. Mall et al., *Science* **359**, 563–567 (2018).  
16. R. W. Guynn, H. J. Gelberg, R. L. Veech, *J. Biol. Chem.* **248**, 6957–6965 (1973).  
17. G. Fuchs, *Annu. Rev. Microbiol.* **65**, 631–658 (2011).

18. W. Buckel, R. K. Thauer, *Biochim. Biophys. Acta* **1827**, 94–113 (2013).  
19. J. Paulsen, A. Kröger, R. K. Thauer, *Arch. Microbiol.* **144**, 78–83 (1986).  
20. A. Lazcano, *Cold Spring Harb. Perspect. Biol.* **2**, a002089 (2010).  
21. M. A. Keller, D. Kampjut, S. A. Harrison, M. Ralsler, *Nat. Ecol. Evol.* **1**, 0083 (2017).  
22. E. Smith, H. Morowitz, *Orig. Life Evol. Biosph.* **40**, 397–401 (2010).  
23. M. Ishii, Y. Igarashi, T. Kodama, *J. Bacteriol.* **171**, 1788–1792 (1989).  
24. G. Antranikian, C. Herzberg, G. Gottschalk, *J. Bacteriol.* **152**, 1284–1287 (1982).  
25. Y. Takeda, F. Suzuki, H. Inoue, *Methods Enzymol.* **13**, 153–160 (1969).  
26. A. Mählén, *Eur. J. Biochem.* **36**, 342–346 (1973).  
27. C. A. Boulton, C. Ratledge, *J. Gen. Microbiol.* **129**, 2863–2869 (1983).

ACKNOWLEDGMENTS

All methods and additional figures and text are available in the supplementary materials. T.N. designed the study; T.N. and K.M. contributed in cultivation analyses; T.N., Y.C., K.Y., M.M., and N.O. conducted metabolomics; T.H., Y.K., A.S., N.F., and Y.T. analyzed genome and/or transcriptome; S.S. conducted proteomics; T.N., T.Sa., and H.A. conducted enzymatic assays; and R.J., T.Su., T.Sa., and H.A. contributed in expression and characterization of recombinant enzymes. T.N., Y.C., T.Sa., H.A., and K.T. wrote the manuscript. All data and code required to understand and assess the conclusions of this research are available in the main text, supplementary materials, and via the following repositories. The genome sequence and transcriptome data set have been deposited in the DNA Data Bank of Japan, European Molecular Biology Laboratory, and GenBank with the accession numbers AP013035 and DRA004184, respectively. The proteome data set is available in jPOST with the accession numbers JPST000334 and PXD007930 (ProteomeXchange). The phylogenetic tree and alignment of CS domain are available in TreeBASE with the study ID S21638.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/359/6375/559/suppl/DC1  
Materials and Methods  
Supplementary Text  
Figs. S1 to S11  
Tables S1 to S5  
References (28–51)

11 July 2017; accepted 20 November 2017  
10.1126/science.aao3407

## CARBON FIXATION

# Reversibility of citrate synthase allows autotrophic growth of a thermophilic bacterium

Achim Mall,<sup>1,2</sup> Jessica Sobotta,<sup>3</sup> Claudia Huber,<sup>3</sup> Carolin Tschirner,<sup>1</sup> Stefanie Kowarschik,<sup>1</sup> Katarina Bačnik,<sup>1</sup> Mario Mergelsberg,<sup>1</sup> Matthias Boll,<sup>1</sup> Michael Hügler,<sup>4</sup> Wolfgang Eisenreich,<sup>3\*</sup> Ivan A. Berg<sup>1,2\*</sup>

Biological inorganic carbon fixation proceeds through a number of fundamentally different autotrophic pathways that are defined by specific key enzymatic reactions. Detection of the enzymatic genes in (meta)genomes is widely used to estimate the contribution of individual organisms or communities to primary production. Here we show that the sulfur-reducing anaerobic delta-proteobacterium *Desulfurella acetivorans* is capable of both acetate oxidation and autotrophic carbon fixation, with the tricarboxylic acid cycle operating either in the oxidative or reductive direction, respectively. Under autotrophic conditions, the enzyme citrate synthase cleaves citrate adenosine triphosphate independently into acetyl coenzyme A and oxaloacetate, a reaction that has been regarded as impossible under physiological conditions. Because this overlooked, energetically efficient carbon fixation pathway lacks key enzymes, it may function unnoticed in many organisms, making bioinformatical predictions difficult, if not impossible.

Most organic carbon on Earth derives from biological CO<sub>2</sub> fixation. Six different autotrophic pathways responsible for this process are known today (1–3). Each of these pathways is characterized by certain key enzymes, the genes for which can be confidently recognized in (meta)genomic databases, thus allowing an estimation of the autotrophic potential reaching from single species up to entire ecosystems. The reductive tricarboxylic acid cycle (rTCA cycle, or Arnon-Buchanan cycle) is among the most ancient metabolic processes (4, 5), probably emerged from a nonenzymatic precursor (6), and is present in bacteria belonging to various phylogenetic groups (1). It is a reversal of the oxidative tricarboxylic acid (oTCA) cycle, which provides redox equivalents and energy in organisms with a respiratory metabolism. The key step of the oTCA cycle is citrate synthesis from acetyl coenzyme A (CoA) and oxaloacetate, catalyzed by citrate synthase (CS). This reaction is regarded as one of the irreversible steps in the oTCA cycle (1, 2, 7). The current consensus is that in organisms using the rTCA cycle, CS is substituted either by a reversible adenosine triphosphate (ATP)-dependent citrate lyase (ACL) (8, 9) or by homologous enzymes catalyzing the same reaction in two steps,

citryl-CoA synthesis and cleavage (fig. S1) (10, 11). The presence of ACL in a bacterium is therefore regarded as a key indication for autotrophic CO<sub>2</sub> fixation via the rTCA cycle (12). Our study challenges this concept by demonstrating, through in vivo and in vitro experiments, that citrate cleavage in the autotrophic rTCA cycle can be catalyzed by CS, which was thought to function only in the oxidative direction.

*Desulfurella acetivorans* is a sulfur-reducing thermophilic (growth optimum at 52° to 57°C) delta-proteobacterium capable of growing either heterotrophically with acetate as an electron donor and carbon source (13) or autotrophically with molecular hydrogen (14). In the heterotrophic route, acetate is oxidized via the oTCA cycle using CS and 2-oxoglutarate:ferredoxin oxidoreductase (15). Because the genome of *D. acetivorans* encodes neither ACL nor full sets of key enzymes for other known autotrophic pathways (table S1), the pathway of CO<sub>2</sub> fixation could not be assigned yet. The genome contains pseudogenes for ACL with frame shifts in both the  $\alpha$  and  $\beta$  subunits (region 484760–487933), which make it nonfunctional. Amplification of the *D. acetivorans* DNA fragment containing these pseudogenes and its sequencing (16) confirmed that *acl* was interrupted (Fig. 1A). Furthermore, the genome contains a gene for 4-hydroxybutyryl-CoA dehydratase, the characteristic enzyme of the archaeal 3-hydroxypropionate/4-hydroxybutyrate and dicarboxylate/4-hydroxybutyrate cycles (17–19), but genes for other specific enzymes of these cycles are missing (table S1). Because the absence of key enzyme(s) of known CO<sub>2</sub> fixation pathways in an organism capable of autotrophic growth can be regarded as an indication of a

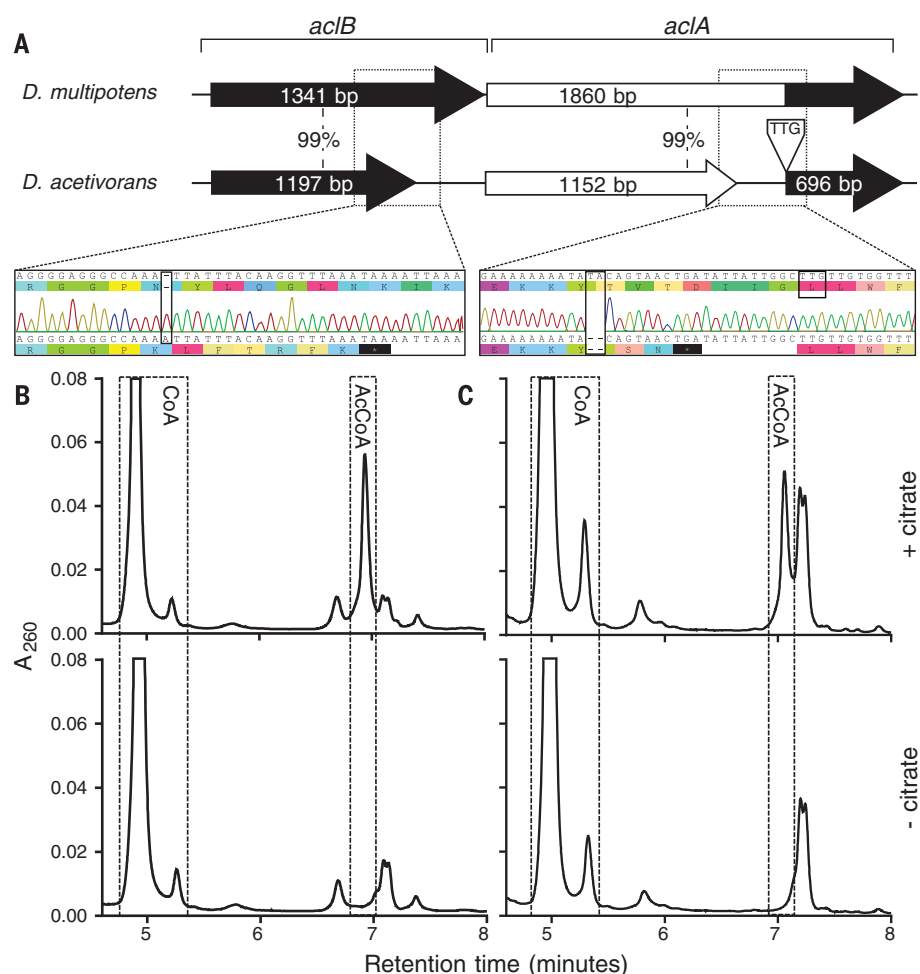
novel pathway, we decided to study CO<sub>2</sub> fixation in *D. acetivorans* in more detail.

In accord with the published data, *D. acetivorans* grew both heterotrophically with acetate/CO<sub>2</sub> and autotrophically on H<sub>2</sub>/CO<sub>2</sub> with minimal generation times of 5.7 and 4.9 hours to a density of  $1.5 \times 10^8$  and  $3.0 \times 10^8$  cells per ml, respectively. High activities of oTCA cycle enzymes were detected in extracts of cells grown under both conditions, with malate dehydrogenase (MDH) and CS activities being extremely high (Table 1). In contrast, no activities of key enzymes of known pathways of autotrophic CO<sub>2</sub> fixation were detected (table S1). By ultra performance liquid chromatography (UPLC) analysis, we determined a citrate (20 mM)-, CoA (1 mM)-, and NADH (reduced form of nicotinamide adenine dinucleotide) (5 mM)-dependent formation of acetyl-CoA (Fig. 1B and fig. S2). Furthermore, a CoA- and citrate-dependent oxidation of NADH was observed that could be attributed to oxaloacetate formation followed by NADH-dependent reduction to malate by endogenous MDH (Table 1). However, none of these reactions were dependent on the presence of ATP, as would be expected if ACL or citryl-CoA synthetase/citryl-CoA lyase were operating.

To further elucidate the fate of citrate, we incubated cell extracts of *D. acetivorans* under anaerobic conditions in buffer containing [U-<sup>13</sup>C<sub>6</sub>]citrate in the presence of NADH and CoA. After 10 min of incubation, the reaction was stopped and subjected to <sup>13</sup>C nuclear magnetic resonance (NMR) analysis. Though the specific <sup>13</sup>C NMR signals for [U-<sup>13</sup>C<sub>6</sub>]citrate could still be detected, we also observed six <sup>13</sup>C NMR multiplets as a result of product formation (Fig. 2 and table S2). On the basis of the chemical shifts, the <sup>13</sup>C-<sup>13</sup>C coupling constants, and titration experiments, these signals were unequivocally attributed to [U-<sup>13</sup>C<sub>4</sub>]malate and [U-<sup>13</sup>C<sub>2</sub>]acetyl-CoA. Together with the UPLC data, these results provide firm evidence for the cleavage of citrate to acetyl-CoA and oxaloacetate, with the latter being further reduced to malate by MDH, which was highly active in our cell extracts (Table 1). The only possible enzyme in *D. acetivorans* that can catalyze an ATP-independent citrate cleavage to oxaloacetate and acetyl-CoA is CS. However, with a free-energy difference  $\Delta G$  of  $-35.8 \text{ kJ mol}^{-1}$  in the canonical direction of citrate formation [at pH 7, ionic strength 0.25, and 38°C (20)], the cleavage reaction has been regarded as impossible under physiological conditions. Even though the next reaction of the rTCA cycle—oxaloacetate reduction to malate (catalyzed by MDH)—is highly exergonic [ $\Delta G$  of  $-27.1 \text{ kJ mol}^{-1}$  at pH 7, ionic strength 0.25, and 38°C (20)], the reversal of the CS reaction would require high substrate and low product concentrations. Using liquid chromatography-mass spectrometry (LC-MS), we found that the CoA/acetyl-CoA ratio in autotrophically grown *D. acetivorans* cells (93, table S3) was much higher than, for instance, in glucose-grown *Escherichia coli*

<sup>1</sup>Mikrobiologie, Fakultät für Biologie, Albert-Ludwigs-Universität Freiburg, 79104 Freiburg, Germany. <sup>2</sup>Institute for Molecular Microbiology and Biotechnology, University of Münster, 48149 Münster, Germany. <sup>3</sup>Lehrstuhl für Biochemie, Technische Universität München, 85748 München, Germany. <sup>4</sup>Department Microbiology and Molecular Biology, DVGW-Technologiezentrum Wasser (TZW), 76139 Karlsruhe, Germany.

\*Corresponding author. Email: ivan.berg@uni-muenster.de (I.A.B.); wolfgang.eisenreich@mytum.de (W.E.)



**Fig. 1. ACL-independent citrate cleavage in *D. acetivorans*.** (A) ACL pseudogene in *D. acetivorans* genome, as compared with the corresponding fragment of the genome of a closely related species, *D. multipotens*, which possesses an intact *acl* gene. Numbers portray the pairwise nucleotides identity. bp, base pairs. (B and C) UPLC analysis of the CoA and CoA-esters formed from citrate in the reaction catalyzed by *D. acetivorans* cell extracts (B) and by CS and MDH from porcine heart (Sigma) (C). A<sub>260</sub>, absorbance at 260 nm.

[2.3 (21)], whereas the citrate concentration in *D. acetivorans* was typical for bacterial cells [1.4 mM compared to 2 mM in *E. coli* (21)]. Using these concentrations of metabolites, we calculated an equilibrium oxaloacetate concentration of 0.13  $\mu$ M. Although it is low, this concentration is still in a physiological range, similar to, for example, the mitochondrial oxaloacetate concentration that is in the low micromolar range (22, 23). [U-<sup>13</sup>C<sub>4</sub>]malate and [U-<sup>13</sup>C<sub>2</sub>]acetyl-CoA could also be found when [U-<sup>13</sup>C<sub>6</sub>]citrate, CoA, and NADH were incubated with commercially available CS and MDH (from porcine heart, Sigma) (Fig. 1C and fig. S3), providing further evidence that the conversion of citrate in cell extracts is catalyzed by CS. The Michaelis constant ( $K_m$ ) values of CS and MDH measured in *D. acetivorans* cell extracts were close to those for the porcine enzymes (table S4). Apparently, the CS of *D. acetivorans* is not specifically adapted to citrate cleavage. This is also apparent from the high  $K_m$  value for citrate, thus

requiring the observed high specific activity of the enzyme.

In addition to citrate cleavage, we were able to detect activity of another characteristic enzyme of the rTCA cycle, fumarate reductase, in *D. acetivorans* cell extracts. This activity was measured as fumarate-dependent succinyl-CoA formation in an assay coupled with endogenous succinyl-CoA synthetase (Table 1). The identity of the product, succinyl-CoA, was confirmed by LC-MS analysis (fig. S4). The highest specific activity (9 nmol min<sup>-1</sup> mg<sup>-1</sup> protein) was observed with dithionite as the electron donor; lower activities could be detected when NADPH (reduced form of nicotinamide adenine dinucleotide phosphate) or NADH were used. No activity could be detected with reduced forms of benzyl viologen, methyl viologen, the menaquinone analogon 2,3-dimethyl-1,4-naphthoquinone, or flavin mononucleotide as electron donors. We assume that this reaction depends on a different, unidentified

electron donor with low redox potential in vivo, possibly coupling fumarate reduction by H<sub>2</sub> with the buildup of a proton-motive force, as was discussed for *Desulfurobacterium acetoxidans* (24).

To test the function of the rTCA cycle in *D. acetivorans* in vivo, we grew autotrophic cultures in the presence of 0.2 mM [1-<sup>13</sup>C]pyruvate, which was added in four 0.05 mM portions during the exponential growth phase. The cells were harvested and hydrolysed under acidic conditions. Four of the obtained amino acids—Ala, Asp, Glu, and Pro—were then converted into *tert*-butyl-dimethylsilyl derivatives, followed by gas chromatography–mass spectrometry analysis to determine <sup>13</sup>C enrichments and positions in the molecules (table S5). Alanine, reflecting its precursor pyruvate, displayed ~6% <sup>13</sup>C excess. Incorporation into aspartate, reflecting its precursor oxaloacetate, and into glutamate and proline, reflecting the precursor 2-oxoglutarate, was lower (1 to 2% <sup>13</sup>C excess) but still significant. A more detailed MS analysis of masses comprising all carbon atoms of the original amino acids (i.e., Ala, Asp, Glu, and Pro at a mass/charge ratio of 260, 418, 432, and 286, respectively, with fragments having lost one or two carbon atoms from the respective precursor amino acids) revealed some positional assignment of the <sup>13</sup>C label (fig. S5). Thus, comparison of <sup>13</sup>C excess in Ala<sup>260</sup> with that in Ala<sup>322</sup> (still carrying C-2 and C-3 of the original Ala molecule) immediately revealed that the <sup>13</sup>C label in Ala must have been entirely located at C-1, which is lost in the fragment, as expected from its origin from the [1-<sup>13</sup>C]pyruvate precursor. The apparently identical <sup>13</sup>C enrichments in Glu<sup>432</sup>/Glu<sup>104</sup> and Pro<sup>286</sup>/Pro<sup>184</sup> suggested that the <sup>13</sup>C label is located in positions 2 to 5 in glutamate and proline. The lower <sup>13</sup>C excess in Asp<sup>390</sup> (carrying C-2, C-3, and C-4 of Asp) in comparison with Asp<sup>418</sup> (comprising all carbon atoms of Asp) indicated that the label must be present in C-1 of Asp. As shown in fig. S6, these <sup>13</sup>C-distributions can be predicted by shuffling the [1-<sup>13</sup>C]pyruvate tracer via phosphoenolpyruvate (PEP) carboxylation into [1-<sup>13</sup>C]oxaloacetate, which enters the rTCA cycle and is converted into 2-oxoglutarate via the symmetric intermediates fumarate and succinate (fig. S6A). On the other hand, pyruvate assimilation via [1-<sup>13</sup>C]oxaloacetate followed by reactions of the oTCA cycle could not explain <sup>13</sup>C incorporation into Glu and Pro, as the <sup>13</sup>C label becomes lost during the conversion of isocitrate into 2-oxoglutarate (table S6 and figs. S5 and S6B). Together, these results clearly proved the existence of a functional rTCA cycle under in vivo conditions.

The operation of the rTCA cycle was also confirmed by inhibitor analysis of autotrophically growing *D. acetivorans* cultures. Two inhibitors of the cycle, fluoroacetate [which is converted through the CS reaction to the aconitase inhibitor fluorocitrate in the cell (25)] and glyoxylate [pyruvate:ferredoxin oxidoreductase inhibitor (26)], suppressed growth of *D. acetivorans* (fig. S7).

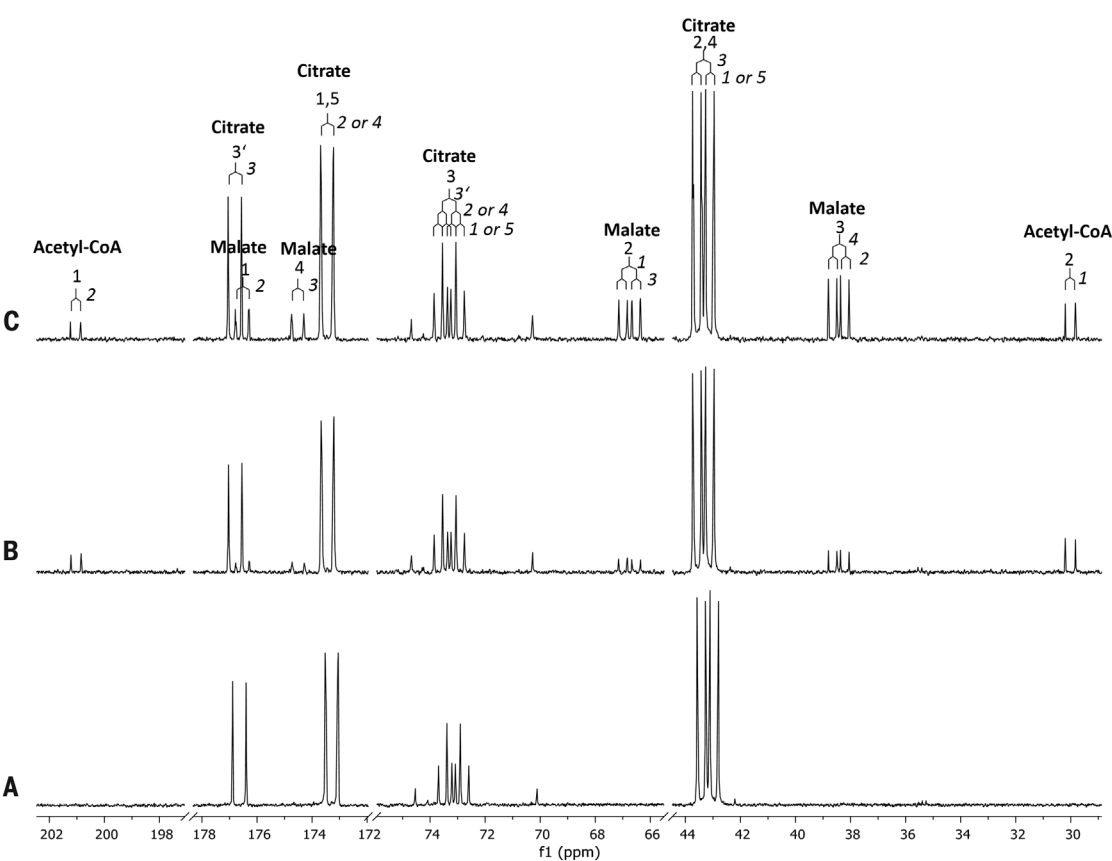
Table 1. Enzymes of central carbon metabolism in <i>D. acetivorans</i> . Activities were measured at 55°C. The number of biological repetitions ( <i>n</i> ) is indicated. ND, not determined; ADP, adenosine diphosphate; acc., accession number.			
Pathway/enzyme	Specific activity [ $\mu\text{mol min}^{-1} \text{mg}^{-1} \text{protein}$ ] $\pm$ SD		Candidate gene(s), GenBank acc.
	H <sub>2</sub> + CO <sub>2</sub>	Acetate + CO <sub>2</sub>	
oTCA cycle			
Citrate synthase	42.1 $\pm$ 0.5 ( <i>n</i> = 8)	55.3 $\pm$ 20.4 ( <i>n</i> = 4)	AHF97305, AHF97477, AHF97591
Aconitase	0.74 $\pm$ 0.11 ( <i>n</i> = 3)	0.51 $\pm$ 0.01 ( <i>n</i> = 2)	AHF96494, AHF96888
Isocitrate dehydrogenase	18.4 $\pm$ 0.6 ( <i>n</i> = 3)	41.5 $\pm$ 4.7 ( <i>n</i> = 3)	AHF97114
Succinyl-CoA synthetase	0.58 $\pm$ 0.10 ( <i>n</i> = 2)	0.024 $\pm$ 0.014 ( <i>n</i> = 2)	AHF96923, AHF96924, AHF96945, AHF96946, AHF97285, AHF97286
Acetate:succinyl-CoA CoA-transferase	0.28 $\pm$ 0.15 ( <i>n</i> = 3)	0.37 $\pm$ 0.17 ( <i>n</i> = 3)	AHF 97575, AHF96498, AHF96963
Succinate dehydrogenase	0.26 $\pm$ 0.002 ( <i>n</i> = 2)	0.32 $\pm$ 0.13 ( <i>n</i> = 3)	AHF96724, AHF96725, AHF96726, AHF96727
Fumarase	6.4 $\pm$ 5.3 ( <i>n</i> = 3)	7.0 $\pm$ 2.7 ( <i>n</i> = 3)	AHF96722, AHF96723
Malate dehydrogenase (NADH)	70.4 $\pm$ 7.6 ( <i>n</i> = 6)	138.1 $\pm$ 11.6 ( <i>n</i> = 3)	AHF96721, AHF97578
Malate dehydrogenase (NADPH)	2.76 $\pm$ 1.36 ( <i>n</i> = 4)	7.6 $\pm$ 2.4 ( <i>n</i> = 3)	
rTCA cycle			
Citrate synthase, reverse (acetyl-CoA formation)	0.20 $\pm$ 0.05 ( <i>n</i> = 3)	0.39 $\pm$ 0.02 ( <i>n</i> = 2)	AHF97305, AHF97477, AHF97591
Citrate synthase, reverse (NADH oxidation)	0.35 $\pm$ 0.09 ( <i>n</i> = 10)	1.86 $\pm$ 0.48 ( <i>n</i> = 3)	
Fumarate reductase	0.009 $\pm$ 0.002 ( <i>n</i> = 2)	<0.001 ( <i>n</i> = 2)	AHF96724, AHF96725, AHF96726, AHF96727
2-Oxoglutarate synthase	0.44 $\pm$ 0.12 ( <i>n</i> = 3)	0.21 $\pm$ 0.06 ( <i>n</i> = 3)	AHF96882, AHF96883, AHF97110, AHF97111, AHF97112, AHF97113
Central carbon metabolism			
Acetyl-CoA synthetase/acetate kinase + phosphate acetyltransferase	0.34 $\pm$ 0.05 ( <i>n</i> = 2)	0.06 $\pm$ 0.01 ( <i>n</i> = 2)	AHF97139, AHF97494, AHF97582, AHF97583
Pyruvate synthase	0.31 $\pm$ 0.08 ( <i>n</i> = 2)	0.017 $\pm$ 0.001 ( <i>n</i> = 2)	AHF96645, AHF96951, AHF97587, AHF96643, AHF96644
PEP carboxylase	<0.005 ( <i>n</i> = 2)	<0.005 ( <i>n</i> = 3)	ND
PEP carboxykinase (ADP)	0.17 $\pm$ 0.05 ( <i>n</i> = 2)	0.12 $\pm$ 0.05 ( <i>n</i> = 2)	AHF96904
PEP synthase	0.03 $\pm$ 0.01 ( <i>n</i> = 2)	0.016 $\pm$ 0.004 ( <i>n</i> = 2)	AHF97619
Pyruvate carboxylase	0.14 $\pm$ 0.03 ( <i>n</i> = 4)	0.014 $\pm$ 0.005 ( <i>n</i> = 3)	AHF96546

We designate this novel CS-dependent version of the rTCA cycle as “reversed oTCA cycle” (roTCA). The functioning of the roTCA cycle requires reduced ferredoxin, which is probably synthesized through electron bifurcation. Indeed, the *D. acetivorans* genome contains genes for a NAD-dependent ferredoxin:NADPH oxidoreductase (NfnAB) (27). Furthermore, the direct reduction of ferredoxin in a hydrogenase reaction is not implausible during growth at high H<sub>2</sub> concentration (80% in our experiments). The activities of the enzymes catalyzing the key reactions of the TCA cycle (Table 1 and supplementary text) were nearly the same for *D. acetivorans* grown organoheterotrophically in medium with acetate or lithoautotrophically with hydrogen. However, the same set of TCA cycle enzymes has to function in either the oxidative or reductive direction, depending on the growth conditions (Fig. 3, A and B). The question then arises as to how the regulation of carbon flux can occur. Two obvious possibilities are that either the presence of H<sub>2</sub> shifts the flux toward the roTCA cycle or that exogenous acetate drives the TCA cycle in the oxidative direction. Isotopologue profiling of *Desulfurella* metabolism during growth with or without H<sub>2</sub> in the presence

of CO<sub>2</sub> and [U-<sup>13</sup>C<sub>2</sub>]acetate revealed that acetate is mostly being oxidized under both conditions (tables S7 and S8 and figs. S8 and S9). These data disprove the regulatory effect of H<sub>2</sub> and suggest that the presence of acetate itself leads to increased intracellular acetyl-CoA concentration, triggering the direction of the cycle. The detected high activities of CS and MDH (Table 1) are not required for growth under heterotrophic conditions. However, the high activities allow fast switching from heterotrophic to autotrophic growth and reflect the adaptation of *D. acetivorans* to fluctuating acetate concentrations characteristic for the *D. acetivorans* natural environment, where both active acetate formation and oxidation take place (28). What makes this newly discovered variant of the rTCA cycle so special? First, the roTCA cycle requires one less ATP molecule per synthesized acetyl-CoA compared with the classical ACL-dependent cycle (1–3). In bioenergetic terms, the roTCA cycle seems to be the most efficient pathway of autotrophic CO<sub>2</sub> fixation known today (Fig. 3C). Still, the modified pathway is associated with an additional energy demand, as an organism using it must increase the amount of some catalysts to cope with their low substrate concentrations [note

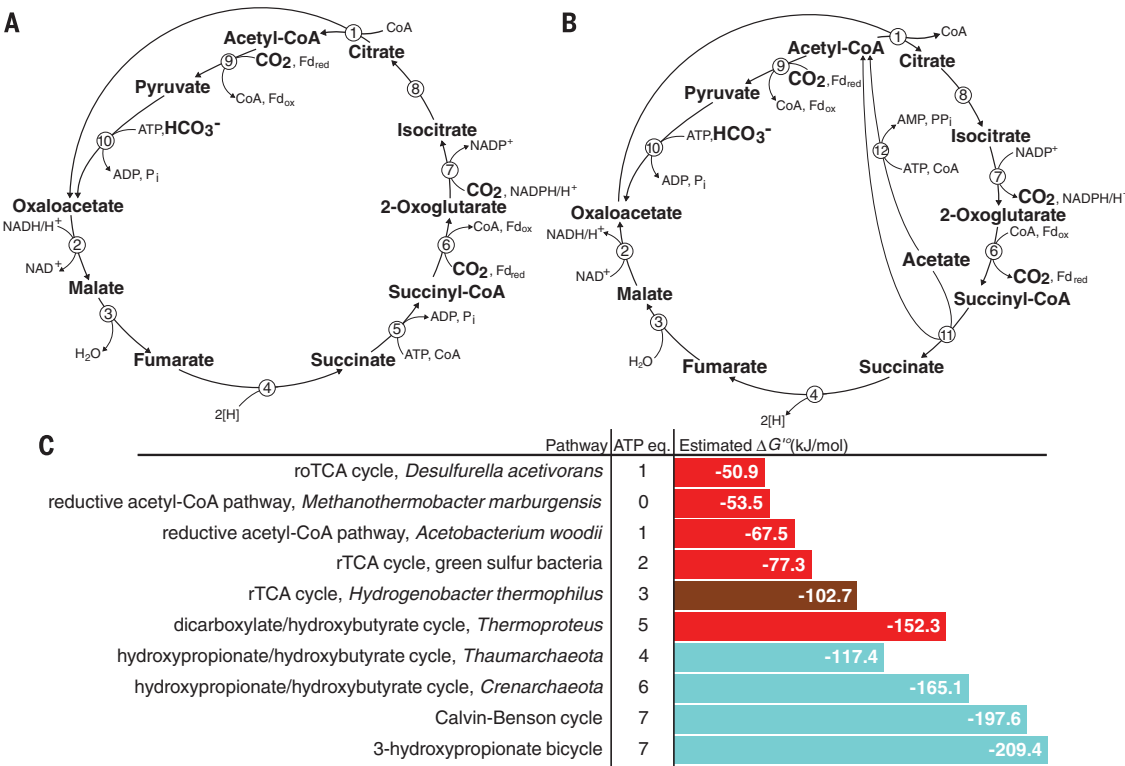
the very high activities of CS and MDH (Table 1)]. The high activities of the enzymes involved in thermodynamically unfavorable reactions ensure efficient coupling of the roTCA cycle to biosynthetic reactions. Anabolism is highly exergonic with a thermodynamic efficiency of ~40%, thus providing an additional driving force for the autotrophic pathway. Nevertheless, the required low product-to-substrate ratio and high enzymatic activities may render this strategy infeasible for some organisms or under certain growth conditions. Second, the roTCA cycle can hardly be recognized bioinformatically. It makes (meta)genome-based bioinformatic predictions of the autotrophic potential of an anaerobic organism (or microbial community) difficult, if not impossible. Third, many bacteria using the rTCA cycle to fix CO<sub>2</sub> also possess CS genes (29, 30), and conversion of the ACL version of the rTCA cycle into the roTCA cycle appears to be quite easy. Notably, another *Desulfurella* species, *D. multipotens*, possesses an intact *acl* gene (Fig. 1A) but apparently uses the roTCA cycle, as can be judged from the ACL and CS activity measurements and ATP independence of citrate cleavage (table S9). The usage of the roTCA cycle

**Fig. 2.  $^{13}\text{C}$  NMR analysis of  $[\text{U-}^{13}\text{C}_6]$  citrate cleavage catalyzed by cell extracts of autotrophically grown *D. acetivorans*.** The reaction mixture contained 2 mM NADH, 2 mM CoA, and 2 mM citrate. (A) Reaction after 0 min. (B) Reaction after 10 min of incubation. (C) Sample as in (B), plus 5  $\mu\text{g}$  of  $[\text{U-}^{13}\text{C}_4]$  malate reference. Signals and couplings of  $[\text{U-}^{13}\text{C}_6]$  citrate,  $[\text{U-}^{13}\text{C}_4]$  malate, and  $[\text{U-}^{13}\text{C}_2]$  acetyl-CoA are indicated. For numerical values of chemical shifts and couplings, see table S2. Note that citrate is a heat stable compound and did not degrade nonenzymatically. ppm, parts per million.



**Fig. 3. A reversible TCA cycle in *D. acetivorans* and its efficiency in comparison with other autotrophic  $\text{CO}_2$  fixation pathways.**

(A) roTCA cycle during growth on  $\text{CO}_2$ ,  $\text{H}_2$ , and elementary sulfur ( $\text{S}^0$ ). (B) oTCA cycle during growth on acetate and  $\text{S}^0$ . (C) ATP costs and estimated  $\Delta G^\circ$  values for the synthesis of acetyl-CoA via the known autotrophic  $\text{CO}_2$  fixation pathways (red, anaerobic; brown, microaerobic; blue, aerobic; see table S10 for details). Note that comparison of the pathways by their ATP costs does not take into account the costs for ferredoxin reduction and thus overestimates the energetic efficiency of anaerobic pathways. Enzymes: 1, citrate synthase; 2, malate dehydrogenase; 3, fumarase; 4, fumarate reductase (roTCA cycle) or succinate dehydrogenase (oTCA cycle); 5, succinyl-CoA synthetase; 6, 2-oxoglutarate synthase; 7, isocitrate dehydrogenase; 8, aconitase; 9, pyruvate synthase; 10, pyruvate carboxylase; 11, acetate:succinyl-CoA CoA-transferase; 12, acetyl-CoA synthetase/acetate kinase + phosphate acetyltransferase.



could be a widespread trait of anaerobic autotrophic organisms. Considering a wide distribution of its enzymes in anaerobes, the cycle could have evolved convergently in different microbial groups. Fourth, CS is structurally simpler than ACL, which consists of several domains, with one of them being homologous to CS (3). Therefore, the roTCA cycle may even predate the modern ACL-rTCA cycle, with an ancestral TCA cycle being fully reversible.

Our results show that unexpected discoveries are possible, even when studying well-known metabolic pathways. If the highly endergonic citrate synthase reaction can be reversed in vivo, which other apparently unidirectional metabolic reactions may be reversed under certain conditions? Despite the many advances of the omics era, these features cannot be identified bioinformatically but require classical biochemical studies for their discovery.

## REFERENCES AND NOTES

1. I. A. Berg, *Appl. Environ. Microbiol.* **77**, 1925–1936 (2011).
2. G. Fuchs, *Annu. Rev. Microbiol.* **65**, 631–658 (2011).
3. M. Hügler, S. M. Sievert, *Annu. Rev. Mar. Sci.* **3**, 261–289 (2011).
4. G. Wächtershäuser, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 200–204 (1990).
5. E. Smith, H. J. Morowitz, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 13168–13173 (2004).
6. M. A. Keller, D. Kampjut, S. A. Harrison, M. Ralsler, *Nat. Ecol. Evol.* **1**, 0083 (2017).
7. R. K. Thauer, D. Möller-Zinkhan, A. M. Spormann, *Annu. Rev. Microbiol.* **43**, 43–67 (1989).
8. R. N. Ivanovsky, N. V. Sintsov, E. N. Kondratieva, *Arch. Microbiol.* **128**, 239–241 (1980).
9. D. Möller, R. Schauder, G. Fuchs, R. K. Thauer, *Arch. Microbiol.* **148**, 202–207 (1987).
10. M. Aoshima, M. Ishii, Y. Igarashi, *Mol. Microbiol.* **52**, 751–761 (2004).
11. M. Aoshima, M. Ishii, Y. Igarashi, *Mol. Microbiol.* **52**, 763–770 (2004).
12. ACL has been found only in bacteria using the rTCA cycle for autotrophic CO<sub>2</sub> fixation and is thus regarded as its key enzyme. Two other characteristic enzymes of the cycle, 2-oxoglutarate:ferredoxin oxidoreductase and fumarate reductase, are present in many nonautotrophic anaerobes.
13. E. A. Bonch-Osmolovskaya, T. G. Sokolova, N. A. Kostrikin, G. A. Zavarzin, *Arch. Microbiol.* **153**, 151–155 (1990).
14. S. Pradella, H. Hippe, E. Stackebrandt, *FEMS Microbiol. Lett.* **159**, 137–144 (1998).
15. R. A. Schmitz, E. A. Bonch-Osmolovskaya, R. K. Thauer, *Arch. Microbiol.* **154**, 274–279 (1990).
16. Materials and methods are available as supplementary materials.
17. I. A. Berg, D. Kockelkorn, W. Buckel, G. Fuchs, *Science* **318**, 1782–1786 (2007).
18. H. Huber *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 7851–7856 (2008).
19. M. Könneke *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 8239–8244 (2014).
20. R. W. Guynn, H. J. Gelberg, R. L. Veech, *J. Biol. Chem.* **248**, 6957–6965 (1973).
21. B. D. Bennett *et al.*, *Nat. Chem. Biol.* **5**, 593–599 (2009).
22. E. A. Siess, D. G. Brocks, O. H. Wieland, *Hoppe Seylers Z. Physiol. Chem.* **359**, 785–798 (1978).
23. E. A. Siess, R. I. Kientsch-Engel, O. H. Wieland, *Biochem. J.* **218**, 171–176 (1984).
24. J. Paulsen, A. Kröger, R. K. Thauer, *Arch. Microbiol.* **144**, 78–83 (1986).
25. A. T. Proudfoot, S. M. Bradberry, J. A. Vale, *Toxicol. Rev.* **25**, 213–219 (2006).
26. R. K. Thauer, E. Rupprecht, K. Jungermann, *FEBS Lett.* **9**, 271–273 (1970).
27. A. P. Florentino, A. J. M. Stams, I. Sánchez-Andrea, *Front. Microbiol.* **8**, 222 (2017).
28. N. A. Chernykh *et al.*, *Extremophiles* **19**, 1157–1171 (2015).
29. T. P. Turova, O. L. Kovaleva, V. M. Gorlenko, R. N. Ivanovskii, *Microbiology (Engl. transl.)* **83**, 72–82 (2014).
30. K. H. Tang, R. E. Blankenship, *J. Biol. Chem.* **285**, 35848–35854 (2010).

## ACKNOWLEDGMENTS

We thank G. Fuchs, Albert-Ludwigs-Universität Freiburg, for his support, discussions, and suggestions during this work and for critical reading of the manuscript. This work was supported by the Deutsche Forschungsgemeinschaft (grant BE 4822/5-1 and Heisenberg Fellowship BE 4822/1-2 to I.A.B.). We also thank the Hans-Fischer Gesellschaft (Munich) for financial support. All data to understand and assess the conclusions of this research are available in the main text, supplementary materials, and the GenBank database ([www.ncbi.nlm.nih.gov/Genbank/](http://www.ncbi.nlm.nih.gov/Genbank/)). The genome sequences of *D. acetivorans* and *D. multipotens* are deposited under GenBank accession numbers CP007051.1 and PRJNA262210, respectively. Amino acid sequences of candidate enzymes are available under their respective GenBank accession numbers (see Table 1 and table S1).

## SUPPLEMENTARY MATERIALS

[www.sciencemag.org/content/359/6375/563/suppl/DC1](http://www.sciencemag.org/content/359/6375/563/suppl/DC1)  
Materials and Methods  
Supplementary Text  
Figs. S1 to S9  
Tables S1 to S11  
References (31–49)

10 July 2017; accepted 20 November 2017  
10.1126/science.aao2410

## ECOPHYSIOLOGY

# High-energy, high-fat lifestyle challenges an Arctic apex predator, the polar bear

A. M. Pagano,<sup>1,2\*</sup> G. M. Durner,<sup>1</sup> K. D. Rode,<sup>1</sup> T. C. Atwood,<sup>1</sup> S. N. Atkinson,<sup>3</sup> E. Peacock,<sup>1</sup> D. P. Costa,<sup>2</sup> M. A. Owen,<sup>4</sup> T. M. Williams<sup>2</sup>

Regional declines in polar bear (*Ursus maritimus*) populations have been attributed to changing sea ice conditions, but with limited information on the causative mechanisms. By simultaneously measuring field metabolic rates, daily activity patterns, body condition, and foraging success of polar bears moving on the spring sea ice, we found that high metabolic rates (1.6 times greater than previously assumed) coupled with low intake of fat-rich marine mammal prey resulted in an energy deficit for more than half of the bears examined. Activity and movement on the sea ice strongly influenced metabolic demands. Consequently, increases in mobility resulting from ongoing and forecasted declines in and fragmentation of sea ice are likely to increase energy demands and may be an important factor explaining observed declines in body condition and survival.

As the most carnivorous and only marine-living ursid, polar bears are lone among bears in their reliance on marine mammal prey. Evolution of this Arctic apex predator included behavioral and physiological adaptations that distinguish polar bears from terrestrial bears (1), which has made them dependent on the sea ice and may increase their vulnerability to climate change (2). As a consequence of living in this labile marine habitat, polar bears occupy expansive home ranges that are considerably larger than those occupied by other ursids (3) or predicted for similarly sized terrestrial carnivores (4). They also exhibit remarkable abilities to swim for extended distances (5). However, such long-distance movements, whether walking or swimming, necessitate substantial energetic resources to satisfy locomotor demands (6, 7). Historically, sufficient resources were afforded through the availability of fatty, energy-dense seal prey, which could be hunted efficiently from the sea ice (8). Presently, the sea ice minimum extent across the Arctic is decreasing at a rate of 14% per decade, spring break-up is occurring earlier, and fall freeze-up is occurring later (9). This decline in sea ice is likely reducing access to, and abundance of, seal prey (10).

Because metabolism determines the rate at which organisms require energy from their environment (11), measures of polar bear metabolic rates provide an important metric for linking declines in Arctic sea ice to polar bear survival. Changes in energy balance result-

ing from increased energy expenditure or reductions in foraging opportunities could lead to declines in body condition, survival, and reproductive success. Previous studies have reported that walking costs for polar bears are greater than predicted for other similarly sized mammals (6, 12). Yet, models predicting polar bear annual field metabolic rates (FMRs) assumed that FMRs would be relatively low because of the bear's sit-and-wait style of hunting and theorized ability to reduce metabolism while fasting (13, 14), similar to that of denning bears (13, 15). However, measures of mass loss and body temperature suggest that metabolic rates of fasting polar bears during the summer are in fact greater than those of denning bears (16–18). Despite this improved understanding of fasting metabolism, no study has provided quantitative estimates of the FMRs of active polar bears on the sea ice where they spend most—and in many areas of their range, all—of the year.

We measured the FMRs of female polar bears ( $n = 9$ ) during April 2014–2016 in the Beaufort Sea (Fig. 1A) over 8 to 11 days each year using doubly labeled water (DLW) (supplementary materials) (19). On the same animals, we deployed global positioning system (GPS)-equipped video camera collars and archival loggers with tri-axial accelerometers and conductivity sensors to determine activity and behavior in order to assess the causes of variation in FMRs (Fig. 1) (19). We determined individual energetic balance using video-derived observations of foraging and measures of blood biochemistry, body mass, and body composition. Additionally, we measured the resting metabolic rate (RMR) of a captive adult female polar bear using open-flow respirometry to assess baseline energetic costs of the species relative to free-ranging polar bears and other mammals.

Like other members of the order Carnivora whose diet is exclusively meat (20), polar bears exhibit greater RMRs as compared with predictions for other terrestrial mammals (derived from omnivores, carnivores, and folivores;  $\text{RMR} = 70 \times \text{mass}^{0.75}$ ) (Fig. 2) (21). RMRs measured from a single captive polar bear over six sessions averaged  $0.34 \text{ ml O}_2 \text{ g}^{-1} \text{ h}^{-1}$  ( $\text{SE} = 0.01$ ) with a low, post-absorptive (fasting) value of  $0.30 \text{ ml O}_2 \text{ g}^{-1} \text{ h}^{-1}$ , which converts to  $37.1 \text{ MJ day}^{-1}$ . On a mass-specific basis, this post-absorptive RMR was within 0.5 to 11.4% of post-absorptive RMRs previously reported for subadult polar bears (22, 23), which is 17% greater than measurements from polar bears while denning (24) and 2 to 21% greater than other ursids while resting or denning (Fig. 2) (19). Thus, carnivory and large body mass set a comparatively high maintenance cost for polar bears that must be satisfied to remain in energetic balance.

As a result, the FMR of polar bears is high relative to predictions for terrestrial mammals generally (25) and consistent with expected levels derived solely from other mammalian carnivores (26). Our measured FMRs averaged 1.6 times previously assumed values for polar bears ( $\text{SE} = 0.1$ , range = 1.0 to 2.6) (13, 14). Daily FMRs measured over 8 to 11 days, averaged  $0.45 \text{ ml CO}_2 \text{ g}^{-1} \text{ h}^{-1}$  ( $\text{SE} = 0.04$ ,  $n = 9$  bears), which converts to  $51.6 \text{ MJ day}^{-1}$  ( $\text{SE} = 6.2$ ,  $n = 9$  bears) or  $12,324.7 \text{ kcal day}^{-1}$ . These values ranged from 0.6 to 1.1 times predicted FMRs [mean ( $\bar{x}$ ) = 0.8,  $\text{SE} = 0.1$ ] for similarly sized marine and terrestrial mammalian carnivores (26) and 2.5 to 5.2 times predicted RMRs based on body mass ( $\bar{x} = 2.8$ ,  $\text{SE} = 0.3$ ) (Fig. 3A and table S1) (21). Daily FMR was 1.5 to 2.8 times the post-absorptive RMR ( $\bar{x} = 2.0$ ,  $\text{SE} = 0.2$ ) measured in this study. On average, daily FMR was 2.6 times (17) and 4 times (27) predicted values for male polar bears fasting on land.

To remain in energy balance with these elevated metabolic demands, polar bears have evolved hunting tactics to prey on high energy-content prey, such as ringed seals (*Pusa hispida*), and preferentially feed on their energy-dense blubber (13). Using our measures of daily FMR, we estimated that a solitary female bear on the spring sea ice would on average need to eat either one adult ringed seal, three subadult ringed seals, or 19 newborn ringed seal pups every 10 to 12 days to remain in energetic balance (Fig. 4A). Our estimates corroborate observations that bears in the early summer typically kill an adult or subadult ringed seal every 5 days (8). However, this rate of consumption would be necessary simply to energetically break even. Polar bears put on the majority of their body fat in the late spring and early summer (13) and can reach a relative fatness of 1 kg fat per kg lean body mass (28). To obtain this body condition, bears would either need to reduce their energy demands or increase their rate of food consumption.

Using video collar data, we documented bears' hunting behavior and foraging success. Bears used sit-and-wait tactics to hunt seals 90% of the time, and stalking comprised the remaining 10% of hunts (movies S1 to S4) (19). Bears

<sup>1</sup>U.S. Geological Survey, Alaska Science Center, Anchorage, AK 99508, USA. <sup>2</sup>Department of Ecology and Evolutionary Biology, University of California, Santa Cruz, Santa Cruz, CA 95060, USA. <sup>3</sup>Post Office Box 19, Group 7, RR2, Dugald, Manitoba ROE OK0, Canada. <sup>4</sup>Institute for Conservation Research, San Diego Zoo Global, San Diego, CA 92027, USA. \*Corresponding author. Email: apagano@usgs.gov

that successfully killed and ate adult or sub-adult ringed seals either gained or maintained body mass, whereas bears that only scavenged or showed no evidence of eating lost mass (Fig. 4, B, C, and D). Land-based remains of subsistence-harvested bowhead whale carcasses (*Balaena mysticetus*) provided an alternative food resource for two bears (Fig. 4, B and D), although 91% of their telemetry data were on the sea ice and away from this resource. Scavenging the muscle from previously killed seals was also common (Fig. 4D) and consistent with previous observations of polar bear foraging on the sea ice (8, 13).

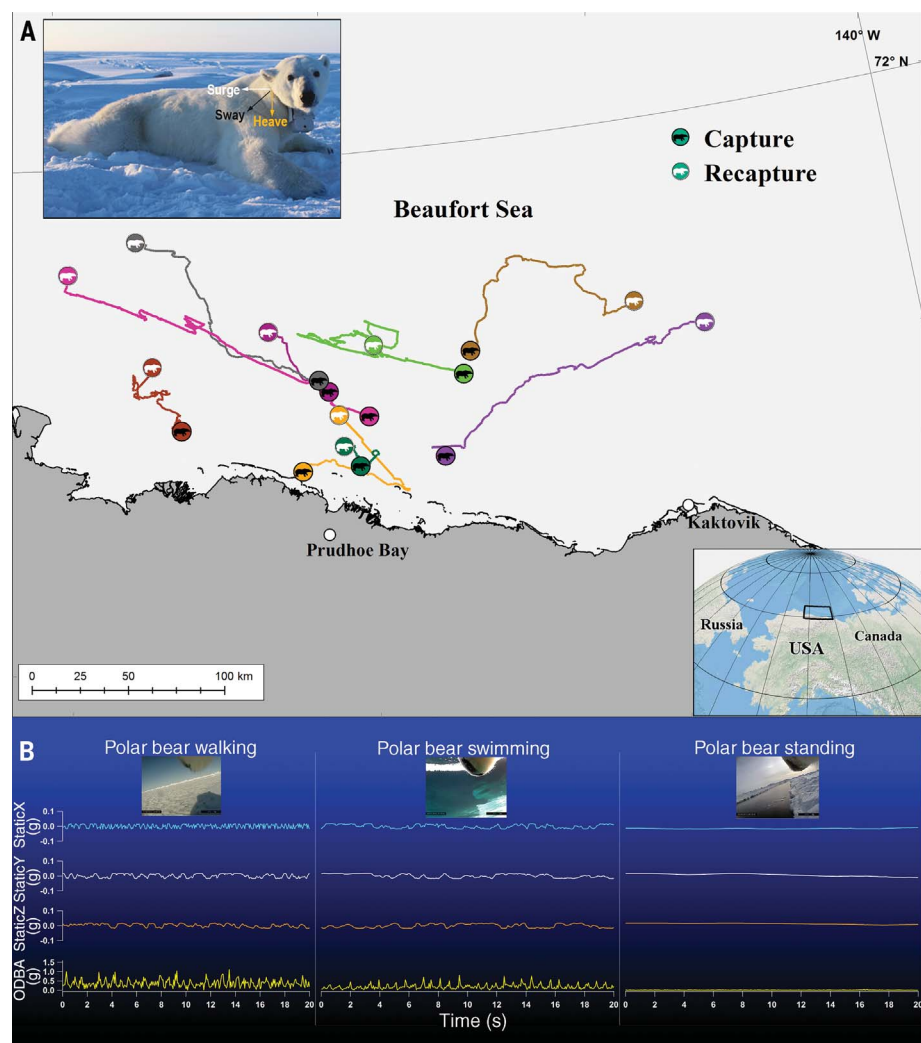
Four bears lost  $\geq 10\%$  of their body mass over the 8- to 11-day period (Fig. 4B and table S1), with an average loss of 1% per day ( $1.95 \text{ kg day}^{-1}$ ). This is 4 times the percent mass lost per day and 2.2 times the kilograms lost per day documented in fasting polar bears on land (16), which typically exhibit lower activity levels than those of the bears in this study (18). Nevertheless, this change in body mass is within the range of the percent mass loss observed in other carnivores (29, 30). Of these four bears, one lost a greater amount of lean body mass than fat mass (Fig. 4B and table S1), which suggests that she was in a prolonged fasting condition (16). This was confirmed with changes in serum urea/creatinine (U/C) ratios (an indicator of fasting for  $>7$  days) (table S1) (19, 31). Additionally, changes in U/C ratios over the 8- to 11-day period were strongly positively correlated with changes in body mass [coefficient of determination ( $R^2$ ) = 0.62,  $P = 0.01$ ,  $n = 9$  bears].

Despite the use of efficient sit-and-wait hunting tactics, polar bears in the spring exhibited activity patterns typical of other large terrestrial carnivores (19). Activity rates (the percentage of time engaged in nonresting behaviors) derived from accelerometers, which were recorded continuously throughout each day, ranged from 22 to 40% ( $\bar{x} = 34\%$ ,  $\text{SE} = 2.8$ ,  $n = 6$  bears) (fig. S2C). Bears spent 28% of the time walking ( $\text{SE} = 2.3$ ,  $n = 6$  bears) and only 0.3% of the time swimming ( $\text{SE} = 0.15$ ,  $n = 6$  bears) (fig. S2C). Activity rates derived from video collars (which recorded only during daylight hours) ranged from 13 to 60% of the day ( $\bar{x} = 32\%$ ,  $\text{SE} = 5.1$ ,  $n = 9$  bears) (fig. S2D). These activity rates are similar to those observed for adult and subadult polar bears on the summer sea ice (34.6% active) (18) and solitary female polar bears on the sea ice in April (16.9% active) in the Canadian Arctic (33), indicating marked variability in the ac-

tivity levels of this species depending on seasonal, geographical, and reproductive factors. The activity levels for polar bears were less than the activity levels documented in other bear species ( $\bar{x} = 54\%$  active) (34) but were similar to activity levels reported for other large terrestrial carnivores ( $\bar{x} = 39\%$  active) (supplementary materials) (19).

Ultimately, the ability of polar bears to achieve energy balance is dictated by the acquisition of metabolizable energy versus expenditure from basal metabolism, specific dynamic action, thermoregulation, reproduction, growth, and locomotion (35). We found that variation in daily FMR was primarily influenced by positive relationships with body mass (daily FMR =  $0.0002 \times \text{mass}^{2.41}$ ,  $R^2 = 0.91$ ,  $P < 0.001$ ,  $n = 9$  bears) (Fig. 3A),

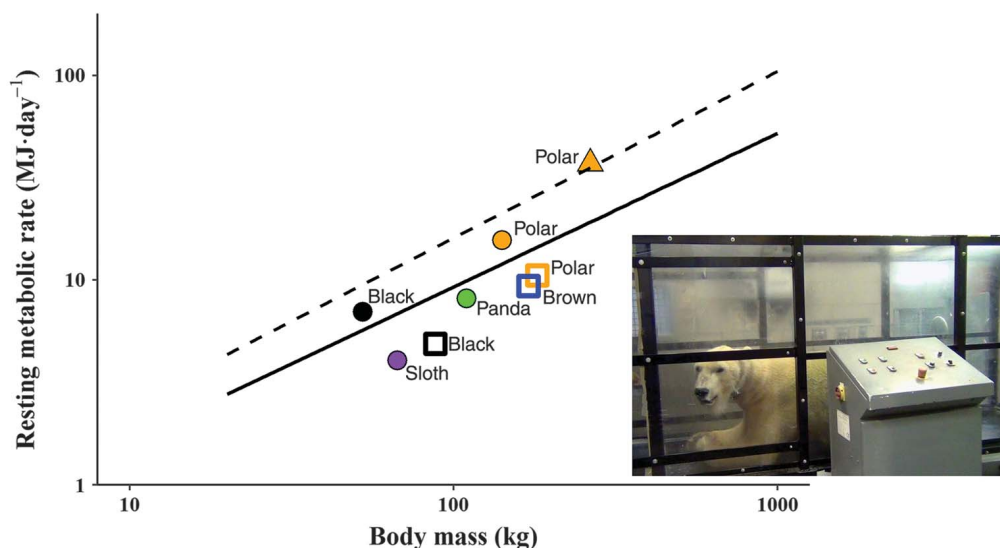
movement rate (daily FMR =  $167.3 \times \text{rate} + 153.0$ ,  $R^2 = 0.82$ ,  $P < 0.001$ ,  $n = 9$  bears) (Fig. 3B), and activity rate derived from video collars (daily FMR =  $336.73 \times \text{activity} + 180.5$ ,  $R^2 = 0.60$ ,  $P = 0.01$ ,  $n = 9$  bears) (Fig. 3C). We found a positive but nonsignificant relationship between activity rate derived from accelerometers and daily FMR ( $R^2 = 0.56$ ,  $P = 0.09$ ,  $n = 6$  bears). We further found a positive relationship between overall FMR and total distance traveled (overall FMR =  $0.006 \times \text{distance} + 1.5$ ,  $R^2 = 0.78$ ,  $P = 0.003$ ,  $n = 8$  bears) (Fig. 3D). These relationships suggest that basal metabolism, locomotion, and activity were the primary drivers of energy expenditure for the polar bears in our study. This reinforces that there is a substantial cost of locomotion in polar bears relative to other quadrupedal mammals, as



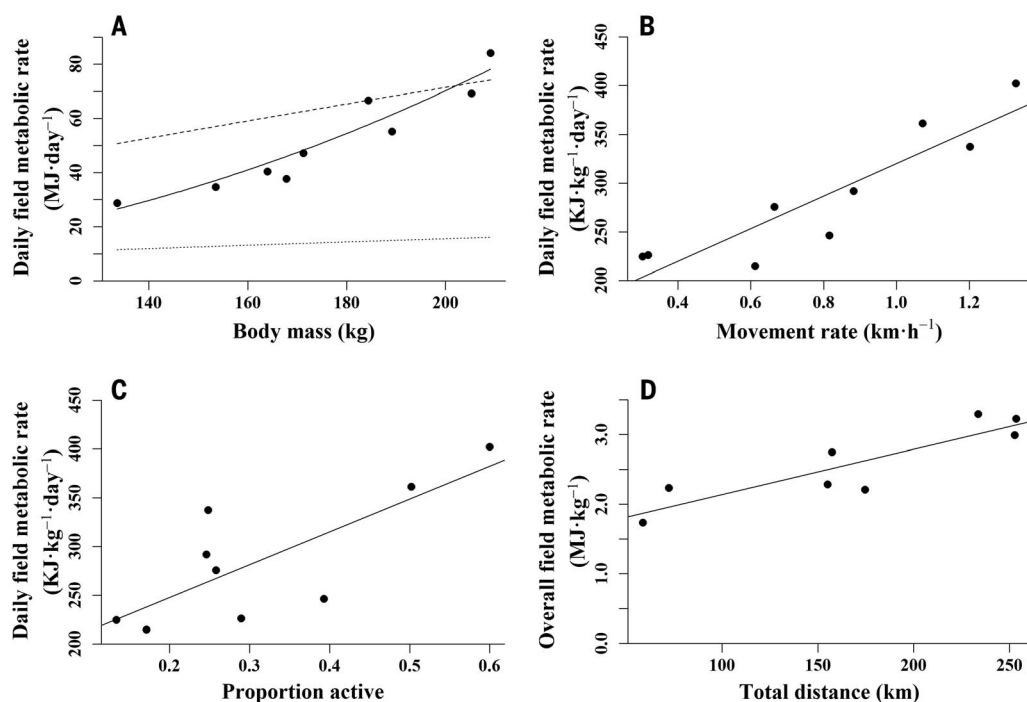
**Fig. 1. Field movements and accelerometer signatures of polar bears in April 2014 to 2016.**

(A) Capture and recapture locations and GPS movement paths of nine female polar bears dosed with DLW and equipped with GPS-equipped video camera collars and archival loggers with triaxial accelerometers and conductivity sensors. Gray area denotes land; white area shows sea ice cover. (Inset) Orientation of the accelerometer while attached to the video collar. (B) Accelerometer signatures of static acceleration in the surge (x), heave (y), and sway (z) directions and overall dynamic body acceleration (ODBA) while walking, swimming, and standing. Images show the corresponding behaviors derived from the animal-borne video camera.

**Fig. 2. Relationship between body mass and RMR in polar bears and other ursids.** RMR of an adult female polar bear (orange triangle, this study), compared with mean RMRs of subadult polar bears (orange circle), hibernating adult female polar bears (orange square), hibernating adult male brown bears (blue square), black bears (black circle), hibernating black bears (black square), adult and subadult panda bears (green circle), and female sloth bears (purple circle). Sources are available in the supplementary materials. The dashed line is the allometric regression for RMR in vertebrate-eating carnivores (20). The solid line is the allometric regression for RMR in eutherian mammals (21). (Inset) The adult female bear resting in the metabolic chamber from the present study.



**Fig. 3. Daily FMR and overall FMR in relation to body mass, movement, and activity rate of polar bears.** (A) Mean daily FMRs of female polar bears on the sea ice in relation to body mass. The allometric regression (solid line) is compared with predicted daily FMRs for marine and terrestrial mammalian carnivores (upper dashed line) (26), and predicted daily RMRs (lower dotted line) (21). (B) Least squares regression (solid line) of mean daily mass-specific FMR in comparison with mean movement rate. (C) Least squares regression (solid line) of mean daily mass-specific FMR in comparison with mean activity rate derived from video collars. (D) Least squares regression (solid line) of overall mass-specific FMR in comparison with total distance moved over 8 to 11 days. Regression statistics are provided in the main text. Each point represents a single value for one bear in (A) to (D).



has been documented in captive treadmill studies (6). For example, bears with movement rates averaging  $\geq 1.0 \text{ km h}^{-1}$  ( $n = 3$  bears) had 1.5 times greater FMRs than those of bears that moved  $< 1.0 \text{ km h}^{-1}$  ( $n = 6$  bears) (Fig. 3B).

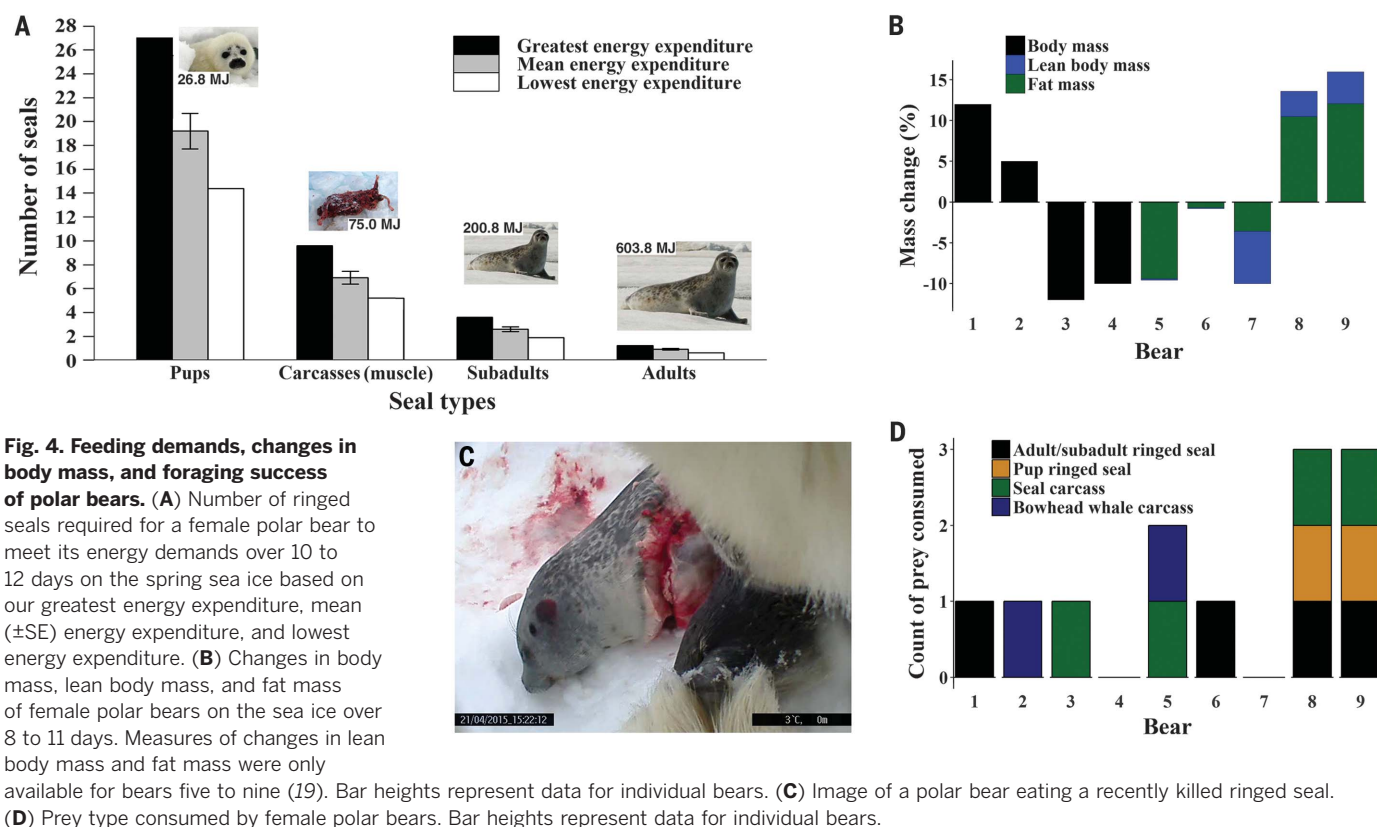
Admittedly, the activity levels and FMRs in this study may be biased low owing to the effects of recovery post-capture. On the basis of movement rate and activity sensor data, recovery post-capture for polar bears may last 2 to 3 days (36). Although one bear in this study successfully caught

and ate an adult or subadult ringed seal less than 24 hours after being collared (fig. S2B) (19), most bears exhibited lower activity rates during the first 24 hours after capture (fig. S3) (19). Hence, the FMRs and activity budgets presented in this study should be considered conservative levels for free-ranging polar bears.

We found that polar bears in the spring exhibit greater energetic demands than those of previous predictions (13, 14) both for maintenance functions and locomotion. Similar to other

marine mammals, polar bears likely transitioned to the marine ecosystem to take advantage of abundant prey resources (1) despite the increased energetic costs required for a marine and carnivorous existence (37). These demands necessitate access to high-energy-content prey in the form of ringed and bearded seals (*Erignathus barbatus*).

More than half of the bears in this study lost body mass, meaning that over the period of observation, their energy demand exceeded



**Fig. 4. Feeding demands, changes in body mass, and foraging success of polar bears.**

(A) Number of ringed seals required for a female polar bear to meet its energy demands over 10 to 12 days on the spring sea ice based on our greatest energy expenditure, mean ( $\pm$ SE) energy expenditure, and lowest energy expenditure. (B) Changes in body mass, lean body mass, and fat mass of female polar bears on the sea ice over 8 to 11 days. Measures of changes in lean body mass and fat mass were only available for bears five to nine (19). Bar heights represent data for individual bears. (C) Image of a polar bear eating a recently killed ringed seal. (D) Prey type consumed by female polar bears. Bar heights represent data for individual bears.

that gained by consuming prey. Although we cannot assess the effects of post-capture recovery on our observed foraging rates, previous researchers reported that 42% of adult female polar bears in the Beaufort Sea during the spring from 2000 to 2016 had not eaten for  $\geq 7$  days before capture (38). This rate of fasting was 12% greater than measurements from 1983 to 1999 (38), suggesting that spring ice conditions are affecting prey availability for polar bears even before the summer open water period. Additionally, access to optimal habitats (annual ice over the continental shelf) is expected to and in some areas has already declined as a result of climate change (39, 40). Survival rates of cubs, body condition of adult females, body size of young, litter mass, and yearling numbers have also exhibited declines in some regions of the Arctic (41, 42). Together with our data on the cost of activity and energy acquisition (Figs. 3 and 4), these studies suggest that an increasing proportion of bears are unable to meet their energy demands. Our results indicate that further increases in activity and movement resulting from declining and increasingly fragmented sea ice are likely to increase the demand side of the energy balance ratio (43). Inherently high energy demands create a physiological constraint that makes it difficult for polar bears to compensate for both increases in activity and declines in the availability of energy-dense prey as habitats become more fragmented (44). Hence,

increases in movement and activity rates mediated by the loss of sea ice habitat are likely to have negative cascading effects on polar bear reproductive success and, ultimately, their populations.

#### REFERENCES AND NOTES

1. A. Berta, *Return to the Sea: The Life and Evolutionary Times of Marine Mammals* (University of California Press, 2012).
2. K. L. Laidre et al., *Ecol. Appl.* **18**, S97–S125 (2008).
3. I. Stirling, A. E. Derocher, *Int. Conf. Bear Res. Manag.* **8**, 189–204 (1990).
4. S. H. Ferguson, M. K. Taylor, E. W. Born, A. Rosing-Asvid, F. Messier, *Ecol. Lett.* **2**, 311–318 (1999).
5. A. M. Pagano, G. M. Durner, S. C. Amstrup, K. S. Simac, G. S. York, *Can. J. Zool.* **90**, 663–676 (2012).
6. R. J. Hurst, N. A. Øritsland, P. D. Watts, *Acta Physiol. Scand.* **115**, 391–395 (1982).
7. B. D. Griffen, *Polar Biol.* 10.1007/s00300-017-2209-x (2017).
8. I. Stirling, *Can. J. Zool.* **52**, 1191–1198 (1974).
9. J. C. Stroeve, T. Markus, L. Boisvert, J. Miller, A. Barrett, *Geophys. Res. Lett.* **41**, 1216–1225 (2014).
10. I. Stirling, A. E. Derocher, *Arctic* **46**, 240–245 (1993).
11. J. H. Brown, J. F. Gillooly, A. P. Allen, V. M. Savage, G. B. West, *Ecology* **85**, 1771–1789 (2004).
12. R. J. Hurst, M. L. Leonard, P. D. Watts, P. Beckerton, N. A. Øritsland, *Can. J. Zool.* **60**, 40–44 (1982).
13. I. Stirling, N. A. Øritsland, *Can. J. Fish. Aquat. Sci.* **52**, 2594–2612 (1995).
14. M. C. S. Kingsley, in *Ringed Seals in the North Atlantic*, M.-P. Heide-Jørgensen, C. Lydersen, Eds. (The North Atlantic Marine Mammal Commission, 1998), vol. 1, pp. 181–196.
15. R. A. Nelson et al., *Int. Conf. Bear Res. Manag.* **5**, 284–290 (1983).
16. S. N. Atkinson, R. A. Nelson, M. A. Ramsay, *Physiol. Zool.* **69**, 304–316 (1996).
17. C. T. Robbins, C. Lopez-Alfaro, K. D. Rode, Ø. Tøien, O. L. Nelson, *J. Mammal.* **93**, 1493–1503 (2012).
18. J. P. Whiteman et al., *Science* **349**, 295–298 (2015).
19. Materials and methods are available as supplementary materials.
20. B. K. McNab, *Q. Rev. Biol.* **63**, 25–54 (1988).
21. M. Kleiber, *The Fire of Life: An Introduction to Animal Energetics* (John Wiley & Sons, 1975).
22. R. J. Hurst, thesis, University of Ottawa (1981).
23. P. D. Watts, K. L. Ferguson, B. A. Draper, *Comp. Biochem. Physiol. A. Comp. Physiol.* **98**, 191–193 (1991).
24. P. D. Watts, N. A. Øritsland, R. J. Hurst, *Physiol. Zool.* **60**, 687–691 (1987).
25. K. A. Nagy, *J. Exp. Biol.* **208**, 1621–1625 (2005).
26. K. A. Nagy, I. A. Girard, T. K. Brown, *Annu. Rev. Nutr.* **19**, 247–277 (1999).
27. P. K. Molnár, T. Klanjscek, A. E. Derocher, M. E. Obbard, M. A. Lewis, *J. Exp. Biol.* **212**, 2313–2323 (2009).
28. S. N. Atkinson, M. A. Ramsay, *Funct. Ecol.* **9**, 559–567 (1995).
29. E. Geffen, A. A. Degen, M. Kam, R. Hefner, K. A. Nagy, *J. Anim. Ecol.* **61**, 611–617 (1992).
30. J. B. Williams, M. D. Anderson, P. R. K. Richardson, *Ecology* **78**, 2588–2602 (1997).
31. A. E. Derocher, R. A. Nelson, I. Stirling, M. A. Ramsay, *Mar. Mamm. Sci.* **6**, 196–203 (1990).
32. I. Stirling, C. Spencer, D. Andriashek, *Mar. Mamm. Sci.* **32**, 13–37 (2016).
33. F. Messier, M. K. Taylor, M. A. Ramsay, *J. Zool. (Lond.)* **226**, 219–229 (1992).
34. S. Paisley, D. L. Garshelis, *J. Zool. (Lond.)* **268**, 25–34 (2006).
35. D. P. Costa, T. M. Williams, in *Biology of Marine Mammals*, J. E. Reynolds, S. A. Rommel, Eds. (Smithsonian Institution Press, 1999), pp. 176–217.
36. K. D. Rode et al., *Wildl. Res.* **41**, 311–322 (2014).
37. T. M. Williams, J. Haun, R. W. Davis, L. A. Fuiman, S. Kohin, *Comp. Biochem. Physiol. A Mol. Integr. Physiol.* **129**, 785–796 (2001).
38. K. D. Rode et al., *Glob. Chang. Biol.* **24**, 410–423 (2018).
39. G. M. Durner et al., *Ecol. Monogr.* **79**, 25–58 (2009).

40. J. V. Ware *et al.*, *Oecologia* **184**, 87–99 (2017).
41. I. Stirling, N. J. Lunn, J. Iacozza, *Arctic* **52**, 294–306 (1999).
42. J. F. Bromaghin *et al.*, *Ecol. Appl.* **25**, 634–651 (2015).
43. G. M. Durner *et al.*, *Glob. Chang. Biol.* **23**, 3460–3473 (2017).
44. V. Sahanatien, A. E. Derocher, *Anim. Conserv.* **15**, 397–406 (2012).

#### ACKNOWLEDGMENTS

This work was supported by the U.S. Geological Survey's Changing Arctic Ecosystems Initiative. Additional support was provided by Polar Bears International; the North Pacific Research Board; Washington State University; World Wildlife Fund (Canada); San Diego Zoo Global; SeaWorld and Busch Gardens Conservation

Fund; University of California, Santa Cruz; and the International Association for Bear Research and Management. Funding for the resting metabolic study was also provided by a NSF Instrument Development for Biological Research grant 1255913-015 (to T.M.W.). We thank M. Bakhtiari (Exeye) for developing the video collars used in this study. We thank T. Donnelly, K. Simac, and M. Spriggs for assistance in the field. We thank helicopter pilot F. Ross (Soloy Helicopters) for field support. We thank San Diego Zoo polar bear trainers T. Batson, N. Wagner, B. Wolf, and P. O'Neill. We thank members of the T.M.W. laboratory, D. Rizzolo, and B. Lyon for comments on previous drafts of the manuscript. This research used resources of the Core Science Analytics and Synthesis Applied Research Computing program at the U.S. Geological Survey. Data reported in this paper are archived at <https://doi.org/10.5066/F7XW4H0P>. The authors declare no

competing financial interests. Any use of trade, firm, or product names is for descriptive purposes only and does not reflect endorsement by the U.S. government.

#### SUPPLEMENTARY MATERIALS

[www.sciencemag.org/content/359/6375/568/suppl/DC1](http://www.sciencemag.org/content/359/6375/568/suppl/DC1)  
Materials and Methods  
Figs. S1 to S3  
Tables S1 to S2  
References (45–82)  
Movies S1 to S4

31 May 2017; accepted 19 December 2017  
10.1126/science.aan8677

## SURFACE SCIENCE

# Polarity compensation mechanisms on the perovskite surface $\text{KTaO}_3(001)$

Martin Setvin,<sup>1\*</sup> Michele Reticcioli,<sup>2</sup> Flora Poelzleitner,<sup>1</sup> Jan Hulva,<sup>1</sup> Michael Schmid,<sup>1</sup> Lynn A. Boatner,<sup>3</sup> Cesare Franchini,<sup>2</sup> Ulrike Diebold<sup>1</sup>

The stacking of alternating charged planes in ionic crystals creates a diverging electrostatic energy—a “polar catastrophe”—that must be compensated at the surface. We used scanning probe microscopies and density functional theory to study compensation mechanisms at the perovskite potassium tantalate ( $\text{KTaO}_3$ ) (001) surface as increasing degrees of freedom were enabled. The as-cleaved surface in vacuum is frozen in place but immediately responds with an insulator-to-metal transition and possibly ferroelectric lattice distortions. Annealing in vacuum allows the formation of isolated oxygen vacancies, followed by a complete rearrangement of the top layers into an ordered pattern of KO and  $\text{TaO}_2$  stripes. The optimal solution is found after exposure to water vapor through the formation of a hydroxylated overlayer with ideal geometry and charge.

All ionic crystals can be considered as stacks of alternately charged lattice planes along certain crystallographic directions. When a crystal is terminated perpendicular to such a “polar” direction, the electrostatic energy diverges (1, 2). The resulting instability can be compensated in various ways—for example, by charge transfer, structural reconstructions, changes in the surface morphology, or chemical doping (3–5). Such compensation mechanisms often create unusual physical and chemical materials properties (6, 7).

Surface and interface polarity has been discussed controversially, particularly for perovskite oxides. These materials have a chemical formula  $\text{ABO}_3$ , where the larger A cation is 12 coordinate and the smaller B cation is 6 coordinate. Along the [001] direction, they contain AO and  $\text{BO}_2$  planes that in many cases carry formal positive and negative charges. Tantalates and niobates (B is Ta or Nb) are prototypical examples of polar perovskites. Many exhibit (incipient) ferroelectricity (8), providing attractive options for electronics and sensors. The built-in field can also enhance electron-hole separation in light-harvesting schemes (9); indeed, these materials are highly efficient photocatalysts (10) with a record quantum efficiency for photochemical water splitting >50% (11).

Because surfaces play a major role in all these applications, it is important to understand which response the system “selects” to relieve the electrostatic instability. We studied the polar (001) surface of  $\text{KTaO}_3$  as a representative case using atomically resolved microscopy and spectroscopy, integral surface analysis techniques, and density functional theory (DFT). We show that  $\text{KTaO}_3(001)$

undergoes an entire series of routes to polarity compensation.

A scanning tunneling microscopy (STM) image of the surface of a  $\text{KTaO}_3(001)$  single crystal after being cleaved in ultrahigh vacuum (UHV) shows three steps running horizontally along the [100] direction (Fig. 1A). These steps have a height of  $\sim 0.4$  nm—i.e., the  $\text{KTaO}_3$  lattice constant. The entire area is covered by alternating KO and  $\text{TaO}_2$  terraces separated by half-unit-cell steps. Interestingly, these terraces run uninterrupted across the full-height steps. The KO planes are always on top, whereas  $\text{TaO}_2$  is in the valley region (see the sketch in Fig. 1D); apparently, the KO plane fractured more easily during cleavage. The KO and  $\text{TaO}_2$  terminations have a formal charge of  $-1e$  and  $+1e$  per unit cell, respectively. The characteristic terrace size is 3 to 8 nm, and their ratio is 1:1. Thus, the top layer is polarity-compensated on a long-range scale (2).

Although the electronic structure appeared delocalized with no measurable atomic corrugation in STM, noncontact atomic force microscopy (nc-AFM) readily showed at atomic resolution a perfect ( $1\times 1$ ) bulk-terminated structure with a low defect concentration. In a constant-height AFM image (Fig. 1B), the darker regions with atomic resolution correspond to the higher-lying KO terraces, and bright, uniform regions are lower-lying  $\text{TaO}_2$  terraces. All of the AFM images were taken in constant-height mode, and on  $\text{TaO}_2$  terraces the attractive forces were smaller because of the larger tip-surface distance. The AFM imaging mechanism on the KO terraces appears analogous to the prototypical  $\text{NaCl}(001)$  surface (12), which has an identical arrangement of surface ions. [Imaging the  $\text{TaO}_2$  on the lower terraces proved more difficult; see fig. S1 (13).]

Scanning tunneling spectroscopy (STS) (Fig. 1C) showed that the two terminations had distinctly different electronic structures. The KO terraces had a wide band gap spanning from  $-3.5$  to  $+1$  V, whereas  $\text{TaO}_2$  terraces appeared metallic (red/black curves in Fig. 1C), with a shallow state at

$\sim 0.2$  eV and several deeper-lying states. These in-gap states did not appear on the smaller  $\text{TaO}_2$  terraces (below  $\sim 8$  unit cells size), and were also absent in the border area ( $\sim 2$  unit cells wide) around the KO islands (see the green spectrum in Fig. 1C). Figure 1E shows a representative spatial distribution of the in-gap states (more details are in figs. S2 and S3). These states exhibited a characteristic wavelike pattern with  $\sim 1.6$  nm periodicity, indicative of their delocalized (bandlike) character (14). The derived wave vector is consistent with reported photoemission data (15).

The electric charge in the in-gap states counteracted the electrostatic potential arising from the  $\text{TaO}_2^+$  polarity, but these states formed only when the terrace width (and related electrostatic potential) exceeded a certain critical limit. The situation bears similarities to the formation of the two-dimensional electron gas (2DEG) at polar interfaces, where a critical film thickness is necessary (6, 16). These in-gap states already form on as-cleaved surfaces, which do not show any defects, and DFT calculations indicate a metallization of the surface (figs. S4 and S5).

The uncompensated electrostatic potential is a driving force for further polarity-compensation mechanisms, and we estimated its magnitude using Kelvin probe force microscopy (KPFM) (see Fig. 1F and fig. S6). On as-cleaved surfaces, the local contact potential varied by  $\sim 0.3$  V. Defect-free surfaces were only achieved for samples cleaved below room temperature. Annealing the surface in vacuum resulted in the development of more efficient polarity-compensation mechanisms. After raising the sample temperature to 312 K, defects formed on the KO terraces (Fig. 2, A and B). AFM images of these defects agreed well with Cl vacancies observed on the prototypical  $\text{NaCl}(001)$  surface (12) (see fig. S7). Accordingly, we attributed the defects here to O vacancies, the corresponding formal charge of  $+2e$  being suitable for compensating the negative charge of the KO regions. The vacancies appeared in central regions of larger terraces—i.e., at the locations with the highest electrostatic potential. The O vacancies were likely created by the migration of O atoms from the KO terraces toward the interface with  $\text{TaO}_2$ .

Annealing at higher temperature ( $T = 387$  K in Fig. 2C) resulted in small “holes” in the central regions of the KO terraces. The steps aligned more along  $\langle 110 \rangle$ , the nonpolar step direction of KO (see the structural model in Fig. 2G). Upon further annealing to 482 and 517 K, the holes interconnected (Fig. 2, D and E, respectively), and formed a labyrinth-like structure. At the atomic scale, the surface still had the original bulk ( $1\times 1$ ) termination, and KO and  $\text{TaO}_2$  terraces alternated in a quasi-periodic pattern separated by half-unit steps. The ratio of KO: $\text{TaO}_2$  remained 1:1, indicating that the process is only caused by KO rearrangement without desorption or diffusion from or to the bulk. KPFM measurements (fig. S6) showed that the O-vacancy formation reduced the original electrostatic potential by half, and after the surface rearranged into the labyrinth structure, the measured potential variations were negligible.

<sup>1</sup>Institute of Applied Physics, Technische Universität Wien, Vienna, Austria. <sup>2</sup>University of Vienna, Faculty of Physics and Center for Computational Materials Science, Vienna, Austria.

<sup>3</sup>Materials Science and Technology Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA.

\*Corresponding author. Email: setvin@iap.tuwien.ac.at

The labyrinth structures had a characteristic terrace width of  $\sim 1.5$  nm (4 to 5 atoms), which did not change upon further annealing. The pattern was remarkably stable up to 700 K. At even higher temperatures, the ordering was lost (fig. S8), likely because of K segregation from the bulk (17). The “magic” size of the stripes in Fig. 2, D to F, likely originated from a competition between two driving forces: Minimizing the terrace size suppressed the electrostatic potential arising from the polarity, whereas steps cost energy. This hypothesis was tested by calculating the surface energy and electrostatic potential as a function of terrace width. The structural model consisted of alternating KO and TaO<sub>2</sub> terraces of equal widths  $w$  (where  $w$  equals the number of K atoms in the 2D unit cell), separated by steps running in the  $\langle 110 \rangle$  direction (Fig. 2G). The energy minimum was achieved for a  $w$  of 4 to 5 atoms (red curve in Fig. 2H), in agreement with the experiment. The electrostatic potential changed smoothly in the range  $1 < w < 8$  and then started to deviate rapidly for higher  $w$ , indicating increasing polar instability.

Considering these calculations, the exact polarity-compensation mechanism of the as-cleaved KO terraces remains an open question. The low defect density and the absence of an electronic reconstruction should result in a relatively high electrostatic potential ( $>1$  V), yet the experimental KPFM data (Fig. 1F and fig. S6) showed considerably smaller potential corrugations. The incipient-ferroelectric character of the material could contribute to the polarity compensation, allowing for generation of electric fields inside the material at no energy cost. Our experimental indications are indirect, but the calculations showed ferroelectric-like lattice distortions (fig. S9). Further, the sample cleaving always resulted in a very characteristic layout of the KO/TaO<sub>2</sub> terraces that crossed full steps without interruption (Fig. 1A). Strain-induced (18) spinodal decomposition into ferroelectric domains could occur during the cleavage, and the resulting terrace layout would be a “printout” of these domains (fig. S9).

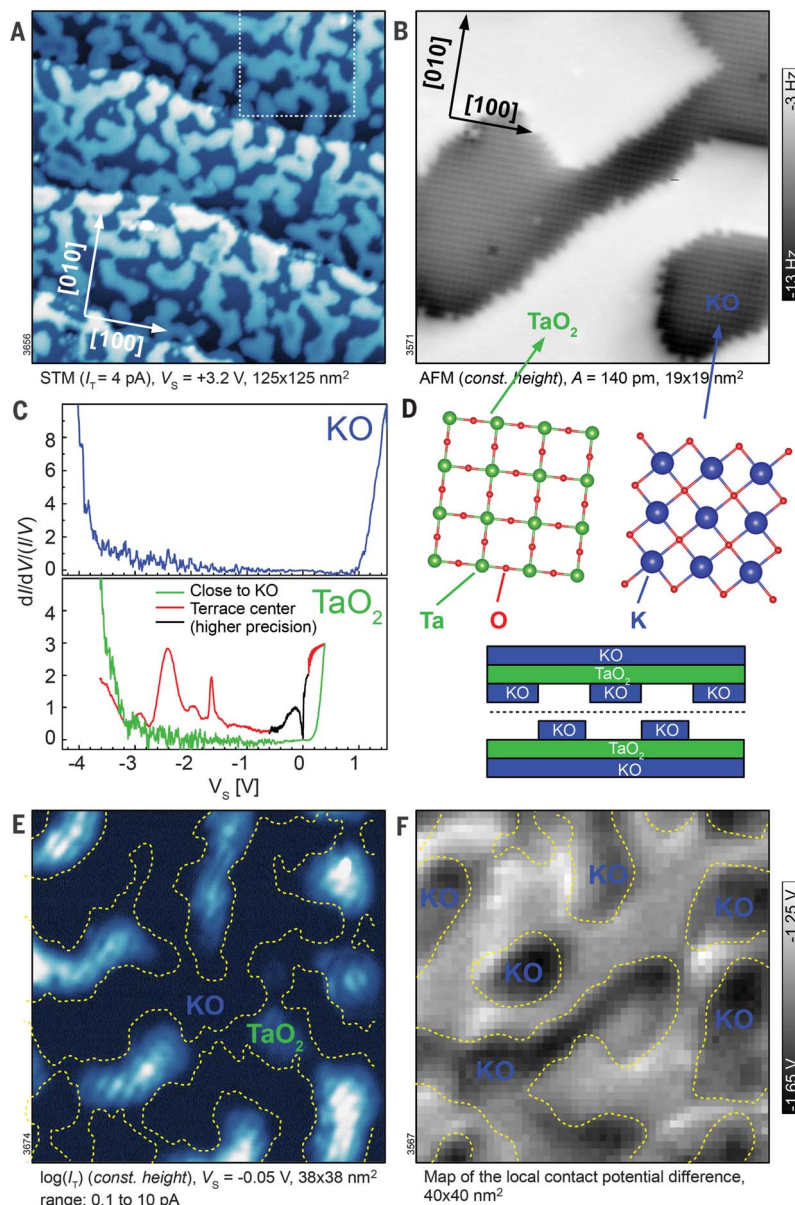
The most efficient polarity-compensation mechanism developed after exposing the surface to H<sub>2</sub>O vapor (Fig. 3). Water induced a complete restructuring of the surface layer, even in UHV and at room temperature. After dosing 300 Langmuir (L) of water vapor at 325 K, the surface was flat and had only full-unit-cell steps (Fig. 3A). More detailed AFM (Fig. 3B) showed a homogenous phase with a  $(2 \times 1)$  symmetry; low-energy electron diffraction (fig. S10) confirmed that the entire surface had been transformed.

This transition required high doses of water; Fig. 3 shows a surface exposed to 100 L H<sub>2</sub>O. Here, the occasional small island of the original bulk-terminated KO was found, interfaced with the  $(2 \times 1)$  overlayer. The  $(2 \times 1)$  termination could be switched back to the  $(1 \times 1)$ ; see Fig. 3D. Annealing above  $\sim 488$  K in UHV resulted in water desorption and formation of the same labyrinth-like structure of alternating KO/TaO<sub>2</sub> terraces as shown in Fig. 2, confirming reversibility between these two surface phases.

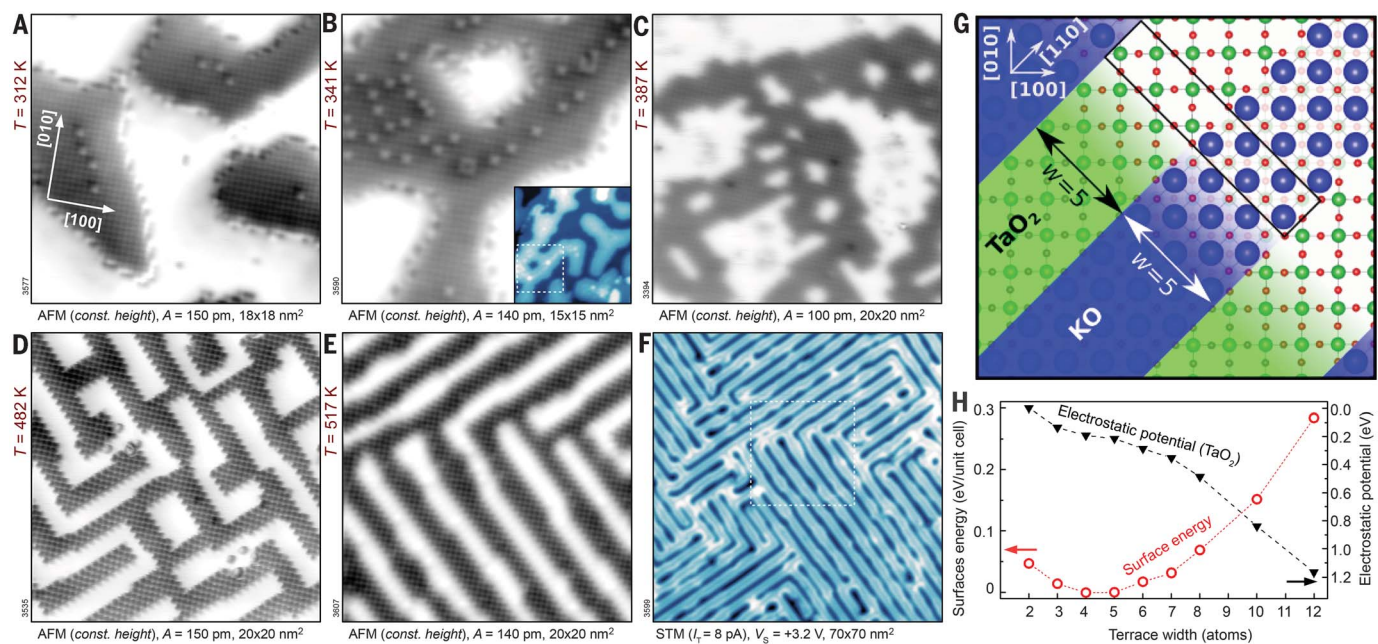
To determine the chemical composition of the  $(2 \times 1)$  superstructure, low-energy He<sup>+</sup> ion scattering is an ideal method because it is only sensitive to the very top layer of the surface. The red curve in Fig. 3E was measured on an as-cleaved surface. As expected, O, K, and Ta are present. After dosing with water, the Ta signal completely disappeared (blue curve). X-ray photoelectron spectroscopy (fig. S11) revealed the hydroxylated character of the  $(2 \times 1)$  termination and also excluded any possible contamination from the water exposure. It appears that water dissolved the KO islands and redistributed them across the TaO<sub>2</sub> planes. A proposed structural model is shown in

Fig. 3F. Each KO unit reacted with one H<sub>2</sub>O molecule, resulting in one K(OH)<sub>2</sub> unit per  $(2 \times 1)$  unit cell. This termination was also polarity-compensated, with a net charge of  $-0.5e$  per  $(1 \times 1)$  unit cell (2). A  $(2 \times 1)$  termination of KTaO<sub>3</sub>(001) has already been reported (19) and investigated theoretically (20), but the presence of water had not been considered.

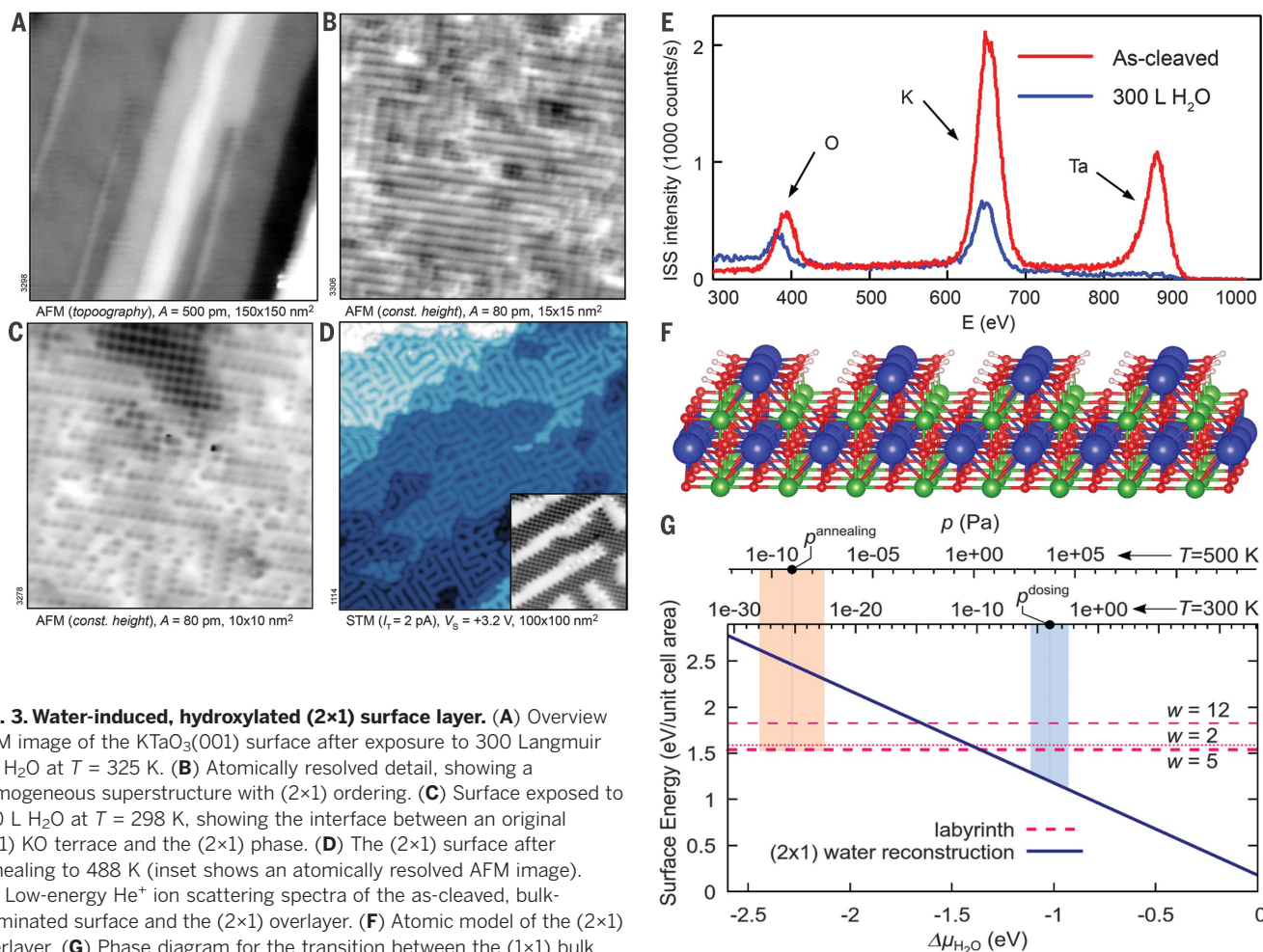
The KTaO<sub>3</sub>(001) surface exhibited a rich spectrum of mechanisms for compensating the surface polarity, and a proper control of the environment can be used to tailor which one is at work. The phenomena observed on the as-cleaved surfaces—i.e., electronic reconstruction and ferroelectric



**Fig. 1. Cleaved KTaO<sub>3</sub>(001).** (A) Overview STM image of a surface prepared by cleaving at  $T = 261$  K. (B) Atomically resolved AFM image. (C) STS spectra measured above the KO and TaO<sub>2</sub> terraces. (D) Atomic structure of the TaO<sub>2</sub> and KO terraces and a schematic drawing of the cleaving. (E) Spatial distribution of in-gap states, measured in the region marked by a dashed square in (A). Dashed lines mark steps. (F) KPFM image of a different region, showing the local electrostatic surface potential.



**Fig. 2. Progression of surface structure with increasing annealing temperature.** (A to E) AFM images of the surface after annealing in vacuum to 312, 341, 387, 482, and 517 K, respectively (measured at 4.8 K). The inset in (B) shows a larger-area STM image of the region. (F) Overview STM image of the surface shown in (E); the corresponding area is marked by a dashed square. (G) Top view of the model used for the DFT calculations. The unit cell is marked by a rectangle. (H) Calculated surface energy and electrostatic potential in the center of the TaO<sub>2</sub> terrace as functions of the terrace width.



**Fig. 3. Water-induced, hydroxylated (2x1) surface layer.** (A) Overview AFM image of the KTaO<sub>3</sub>(001) surface after exposure to 300 Langmuir (L) H<sub>2</sub>O at  $T = 325$  K. (B) Atomically resolved detail, showing a homogeneous superstructure with (2x1) ordering. (C) Surface exposed to 100 L H<sub>2</sub>O at  $T = 298$  K, showing the interface between an original (1x1) KO terrace and the (2x1) phase. (D) The (2x1) surface after annealing to 488 K (inset shows an atomically resolved AFM image). (E) Low-energy He<sup>+</sup> ion scattering spectra of the as-cleaved, bulk-terminated surface and the (2x1) overlayer. (F) Atomic model of the (2x1) overlayer. (G) Phase diagram for the transition between the (1x1) bulk termination and the (2x1) structure.

lattice distortions—are well known from perovskite heterostructures. We could directly probe the character of the 2DEG-like states and their spatial confinement to the highly polar regions. The labyrinth-like pattern formed upon annealing represents an interesting template for charge-carrier separation in light-harvesting devices, but protection against environmental influences would need to be facilitated for practical use. Although such a structure minimizes the surface energy, the corresponding energy gain is substantially lower compared with chemically induced polarity compensation by water. The phase diagram in Fig. 3G shows that the mechanisms encountered in UHV decrease the surface energy by tenths of electron volts, whereas hydroxylation brings almost one electron volt (at conditions corresponding to ambient water pressures). Perovskite surfaces in ambient conditions are typically hydroxylated and near the bulk termination, but little is known about the exact surface structure. The (2×1) reconstruction observed here could there-

fore serve as a working model for future studies into the surface chemistry of perovskite oxides.

#### REFERENCES AND NOTES

1. P. W. Tasker, *J. Phys. Chem.* **12**, 4977–4984 (1979).
2. C. Noguera, *J. Phys. Cond. Mat.* **12**, R367–R410 (2000).
3. A. F. Santander-Syro *et al.*, *Nature* **469**, 189–193 (2011).
4. B. Meyer, *Phys. Rev. B* **69**, 045416 (2004).
5. O. Dulub, U. Diebold, G. Kresse, *Phys. Rev. Lett.* **90**, 016102 (2003).
6. S. Thiel, G. Hammerl, A. Schmehl, C. W. Schneider, J. Mannhart, *Science* **313**, 1942–1945 (2006).
7. A. Bruix *et al.*, *Angew. Chem. Int. Ed.* **53**, 10525–10530 (2014).
8. D. J. Singh, *Phys. Rev. B* **53**, 176–180 (1996).
9. I. Grinberg *et al.*, *Nature* **503**, 509–512 (2013).
10. E. Grabowska, *Appl. Cat. B: Environmental* **186**, 97–126 (2016).
11. H. Kato, K. Asakura, A. Kudo, *J. Am. Chem. Soc.* **125**, 3082–3089 (2003).
12. L. Gross *et al.*, *Phys. Rev. B* **90**, 155455 (2014).
13. Materials and methods are available as supplementary materials.
14. M. Setvin *et al.*, *Phys. Rev. Lett.* **113**, 086402 (2014).
15. P. D. C. King *et al.*, *Phys. Rev. Lett.* **108**, 117602 (2012).
16. P. R. Willmott *et al.*, *Phys. Rev. Lett.* **99**, 155502 (2007).
17. K. Szot *et al.*, *J. Phys. Condens. Matter* **12**, 4687–4697 (2000).
18. J. H. Haeni *et al.*, *Nature* **430**, 758–761 (2004).
19. T. W. Trelenberg *et al.*, *J. Phys. Conds. Matter* **22**, 304009 (2010).
20. D. E. Deacon-Smith, D. O. Scanlon, C. R. A. Catlow, A. A. Sokol, S. M. Woodley, *Adv. Mater.* **26**, 7252–7256 (2014).

#### ACKNOWLEDGMENTS

This work was supported by European Research Council (ERC) advanced grant “Oxide Surfaces” (ERC-2011-ADG-20110209), Austrian Science Fund (FWF) Project Wittgenstein Prize (Z 250), Project F45 (FOXSI), and the FWF-SFB project VICOM (grant no. F41). Research at the Oak Ridge National Laboratory for one author (L.A.B.) was supported by the U.S. Department of Energy, Office of Science, Basic Energy Sciences, Materials Sciences and Engineering Division. All data are reported in the main text and supplement.

#### SUPPLEMENTARY MATERIALS

[www.sciencemag.org/content/359/6375/572/suppl/DC1](http://www.sciencemag.org/content/359/6375/572/suppl/DC1)  
Materials and Methods  
Figs. S1 to S11  
Coordinates Labyrinth-w5.cif  
References (21–27)

24 October 2017; accepted 19 December 2017  
10.1126/science.aar2287

## SUPERCONDUCTIVITY

# Probing optically silent superfluid stripes in cuprates

S. Rajasekaran,<sup>1\*</sup> J. Okamoto,<sup>2</sup> L. Mathey,<sup>2</sup> M. Fechner,<sup>1</sup> V. Thampy,<sup>3</sup>  
G. D. Gu,<sup>3</sup> A. Cavalleri<sup>1,4\*</sup>

Unconventional superconductivity in the cuprates coexists with other types of electronic order. However, some of these orders are invisible to most experimental probes because of their symmetry. For example, the possible existence of superfluid stripes is not easily validated with linear optics, because the stripe alignment causes interlayer superconducting tunneling to vanish on average. Here we show that this frustration is removed in the nonlinear optical response. A giant terahertz third harmonic, characteristic of nonlinear Josephson tunneling, is observed in  $\text{La}_{1.885}\text{Ba}_{0.115}\text{CuO}_4$  above the transition temperature  $T_c = 13$  kelvin and up to the charge-ordering temperature  $T_{co} = 55$  kelvin. We model these results by hypothesizing the presence of a pair density wave condensate, in which nonlinear mixing of optically silent tunneling modes drives large dipole-carrying supercurrents.

Single-layer cuprates of the type  $\text{La}_{2-x-y}(\text{Ba}, \text{Sr})_x(\text{Nd}, \text{Eu})_y\text{CuO}_4$  exhibit an anomalous suppression of the superconducting transition temperature  $T_c$  for doping levels near 12.5% (1). Several studies have shown that this suppression coincides with the formation of “stripes,” one-dimensional chains of charge rivers separated by regions of oppositely phased antiferromagnetism (2, 3). A schematic phase diagram of  $\text{La}_{2-x}\text{Ba}_x\text{CuO}_4$ , adapted from (1), is shown in Fig. 1, along with a sketch of the stripe geometry.

Recently, high-mobility in-plane transport was found in some of these striped phases at temperatures above the bulk superconducting  $T_c$  (4). The existence of a striped superfluid state with a spatially oscillating superconducting order parameter, a so-called pair density wave (PDW) state, was hypothesized (5) to explain the anomalously low in-plane resistivity.

Superfluid stripes are difficult to detect. Scanning tunneling microscopy (STM) experiments have reported spatial modulations in the superconducting condensate strength of  $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8+x}$  (6). However, STM is not sensitive to the phase of the order parameter, and these low-temperature measurements did not clarify the broader question of whether finite momentum condensation may be taking place above  $T_c$ . Also, according to the pair density wave model, superfluid transport perpendicular to the planes is frustrated owing to the stripe alignment (Fig. 1, inset). Hence, these stripes are invisible in linear  $c$ -axis optical measurements (7).

In this report, we show that superfluid stripes display characteristic signatures in the nonlinear terahertz frequency response. In  $\text{La}_{1.885}\text{Ba}_{0.115}\text{CuO}_4$ , we detect superfluid stripes above  $T_c = 13$  K and up to  $T_{co} = 55$  K.

Figure 2 displays the linear and nonlinear terahertz reflectivity spectra for two  $\text{La}_{2-x}\text{Ba}_x\text{CuO}_4$  crystals ( $x = 9.5$  and 15.5%), measured with single-cycle pulses (8) polarized along the  $c$  axis [see section S1 of (9)] and covering the spectral range between 150 GHz and 2.5 THz [see section S2 of (9)]. At these doping levels, the material exhibits homogeneous superconductivity and only weak stripe order, with transition temperatures of  $T_{co} \sim T_c = 34$  K ( $x = 9.5\%$ ) and  $T_{co} \sim 40$  K  $> T_c = 32$  K ( $x = 15.5\%$ ).

The low-temperature linear reflectivities ( $T = 5$  K  $< T_c$ ), show Josephson plasma edges at  $\omega_{JP0} = 500$  GHz and  $\omega_{JP0} = 1.4$  THz for  $x = 9.5$  and 15.5%, respectively (see dashed curves in Fig. 2, A and C). The reflectivity edges shifted to lower frequencies with increasing temperature, indicating a decrease in superfluid density [Fig. 2, B and D, and section S2 of (9)].

Nonlinear reflectivities, measured at field strengths between 20 and 80 kV/cm (see Fig. 2,

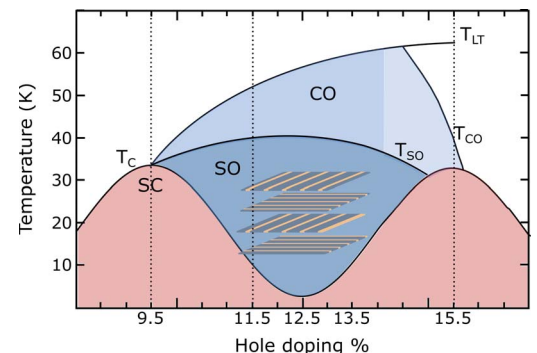
A and C) displayed two characteristic effects of nonlinear interlayer Josephson coupling: (i) a field-dependent red shift of the plasma edge (10–12) and (ii) a reflectivity peak at the third harmonic of the pump field  $\omega \approx 3\omega_{\text{pump}}$  (13). The third-harmonic field amplitude  $E(3\omega_{\text{pump}})$  scaled with the cube of the incident field strength  $E(\omega_{\text{pump}})$  (Fig. 2E), decreased in strength with temperature, and disappeared for  $T = T_c$ , tracking the superfluid density  $\omega\sigma_2(\omega \rightarrow 0)$  (see Fig. 2, C, D, and F).

These observations are well understood by a semiquantitative analysis of the interlayer phase dynamics of a homogeneous layered superconductor (10, 12, 14). For a  $c$ -axis electric field  $E(t) = E_0 \sin(\omega_{\text{pump}} t)$  the interlayer phase difference  $\theta(t)$  advances in time,  $t$ , according to the second Josephson equation  $\frac{\partial \theta(t)}{\partial t} = \frac{2eE(t)}{\hbar}$ , where  $d$  is the interlayer spacing ( $\sim 1$  nm),  $2e$  is the Cooper pair charge, and  $\hbar$  is Planck's constant  $h$  divided by  $2\pi$  (15). Because the  $c$ -axis superfluid density  $\rho_c$  scales with the order parameter phase difference  $\rho_c \propto \cos \theta$  and because  $\rho_c \propto \omega_J^2$ , the plasma frequency renormalizes as  $\omega_J^2 = \omega_{JP0}^2 \cos(\theta) = \omega_{JP0}^2 \cos[\theta_0 \cos(\omega_{\text{pump}} t)] \approx \omega_{JP0}^2 \left[ 1 - \frac{\theta_0^2}{4} - \frac{\theta_0^2 \cos(2\omega_{\text{pump}} t)}{4} \right]$ , where  $\theta_0 = 2edE_0/\hbar\omega_{\text{pump}}$ , an average red shift of the equilibrium plasma resonance  $\omega_{JP0}$  is hence estimated as  $\omega_J^2 = \omega_{JP0}^2 \left( 1 - \frac{\theta_0^2}{4} \right)$ . Secondly, a tunneling supercurrent,  $I$ , is excited at the first and third harmonic of the driving field,  $I(t) = I_c \sin[\theta_0 \cos(\omega_{\text{pump}} t)] \approx I_c \left[ \theta_0 \cos(\omega_{\text{pump}} t) - \frac{\theta_0^3}{6} \cos^3(\omega_{\text{pump}} t) \right]$ , giving rise to third harmonic radiation. Finally, because the third-harmonic signal is proportional to  $I_c$ , it is expected to follow the same temperature dependence as the superfluid density.

More comprehensive numerical simulations, based on space ( $x$ , not to be confused with the symbol for doping concentration)– and time ( $t$ )–dependent one-dimensional sine-Gordon equation for the Josephson phase  $\theta(x, t)$  [see section S3 of (9)] (12, 14, 16, 17), were used to obtain the electromagnetic field at the surface of the superconductor and to calculate the reflectivity for arbitrary field strengths. These simulations, reported in Fig. 3, reproduce the experimental data closely.

**Fig. 1. Phase diagram for  $\text{La}_{2-x}\text{Ba}_x\text{CuO}_4$ .**

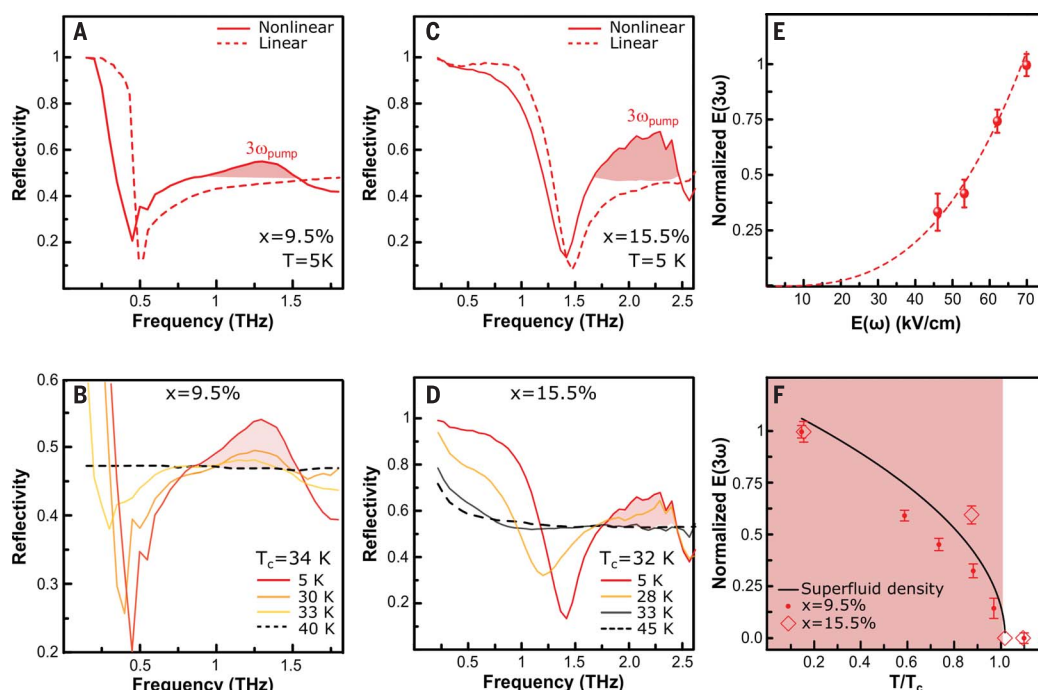
SC, SO, and CO denote the bulk superconducting, spin- and charge-ordered (striped), and charge-only-ordered phases, respectively.  $T_c$ ,  $T_{so}$ , and  $T_{co}$  are the corresponding ordering temperatures.  $T_{LT}$  denotes the orthorhombic-to-tetragonal structural transition temperature. The samples examined in this study are  $x = 9.5$ , 11.5, and 15.5% (dotted lines). Further, a schematic stripe-ordered state is shown wherein the tan stripes depict the charge rivers and the gray stripes depict the antiferromagnetic insulating region (inset). Figure adapted with permission from (1).



<sup>1</sup>Max Planck Institute for the Structure and Dynamics of Matter, Hamburg, Germany. <sup>2</sup>Centre for Optical Quantum Technologies and Institute for Laser Physics, University of Hamburg, Hamburg, Germany. <sup>3</sup>Condensed Matter Physics and Materials Science Department, Brookhaven National Laboratory, Upton, 11973 NY, USA. <sup>4</sup>Clarendon Laboratory, Department of Physics, University of Oxford, Oxford, UK. \*Corresponding author. Email: andrea.cavalleri@mpsd.mpg.de (A.C.); srivats.raijasekaran@mpsd.mpg.de (S.R.)

### Fig. 2. Third harmonic from homogeneous superconductor.

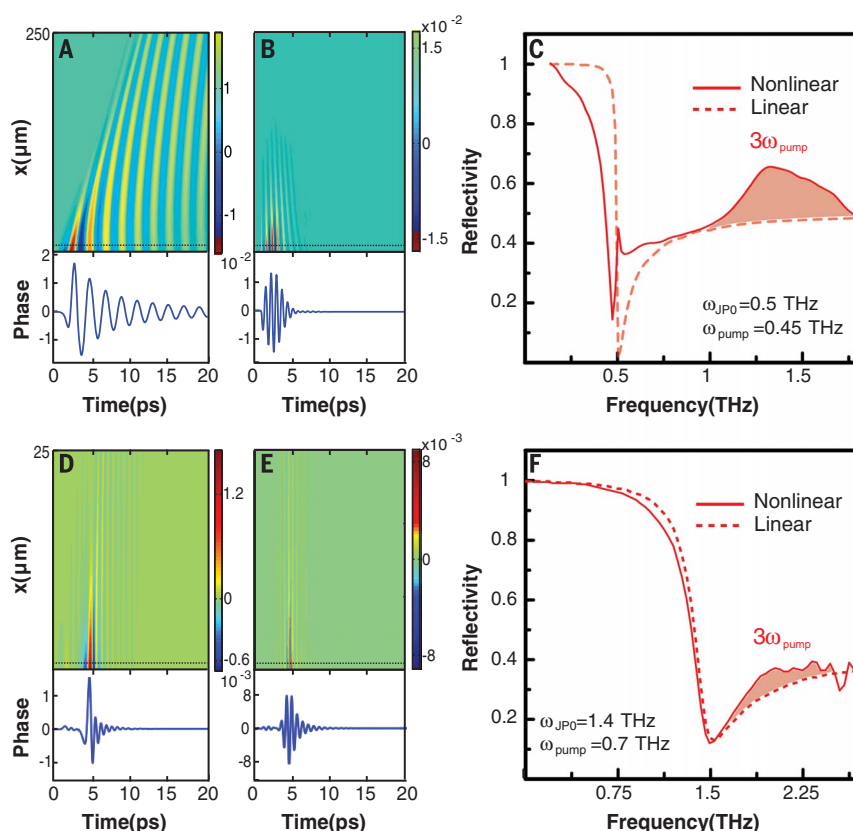
**(A)** Linear and nonlinear reflectivity of  $\text{La}_{2-x}\text{Ba}_x\text{CuO}_4$ , where  $x = 9.5\%$ , measured at  $T = 5$  K with  $\omega_{\text{pump}} = 450$  GHz. The linear reflectivity displays a Josephson plasma edge at  $\omega_{\text{JP0}} = 500$  GHz (0.5 THz), whereas the nonlinear reflectivity shows a red shift of the edge and third-harmonic generation (red shading). **(B)** Temperature dependence of nonlinear reflectivity for  $x = 9.5\%$ . The third-harmonic peak disappears above  $T_c = 34$  K. **(C)** Linear and nonlinear reflectivity for  $x = 15.5\%$ , measured with  $\omega_{\text{pump}} = 700$  GHz. **(D)** Temperature dependence of third-harmonic generation in  $x = 15.5\%$ . Third-harmonic generation (red shading) disappears above  $T_c = 32$  K. **(E)** Third-harmonic electric field strength (normalized to the highest signal) plotted as a function of the incident electric field strength [definition explained in section S1 of (9)] measured at  $T = 5$  K from the  $x = 9.5\%$  sample. The third-harmonic field displays a cubic dependence on the incident field strength. **(F)** Temperature dependence of the third-harmonic amplitude (normalized

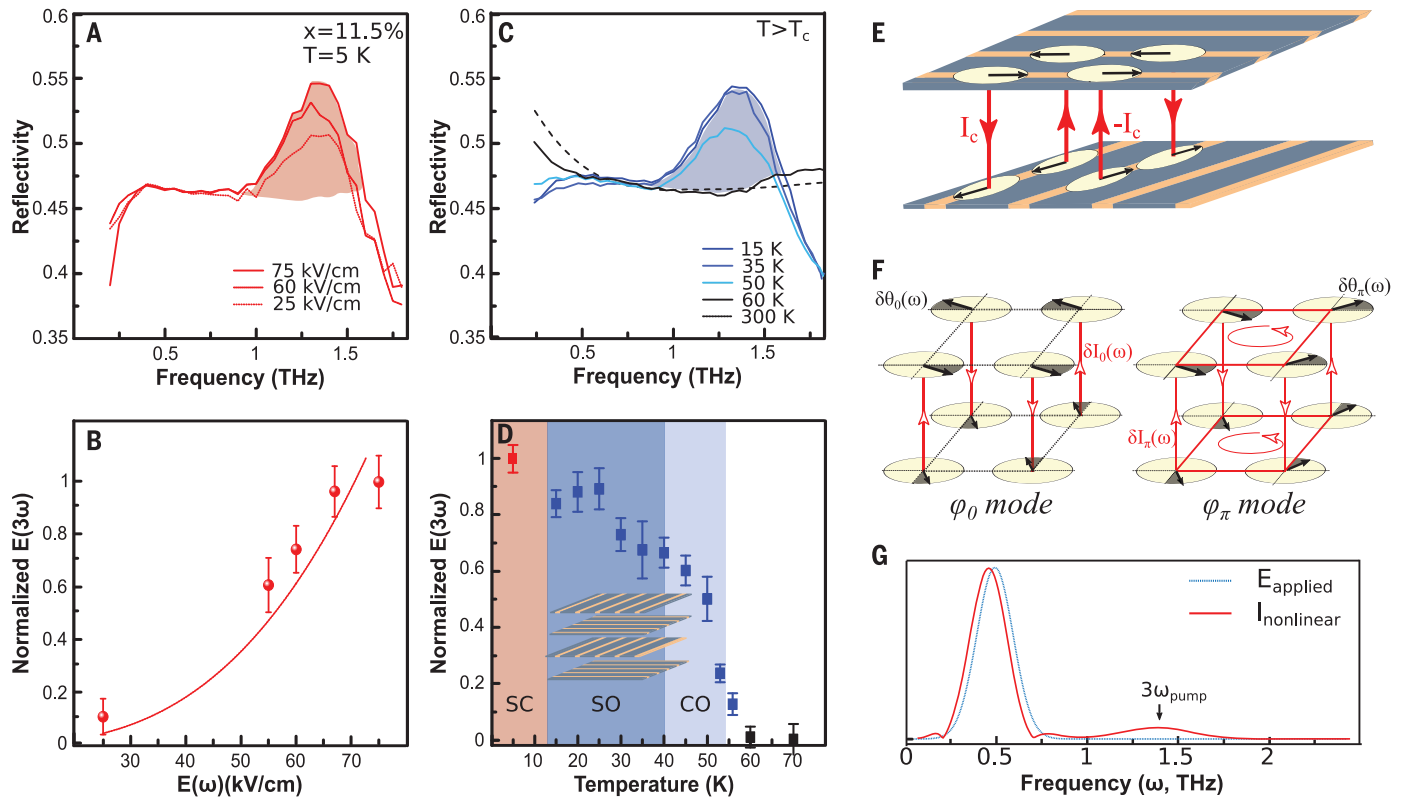


to the measurement at  $T = 5$  K) from the  $x = 9.5$  and  $15.5\%$  doping. The superfluid density  $[(\omega_{\text{JP0}}(\sigma \rightarrow 0))]$  (normalized to the measurement at  $T = 5$  K) extracted from the linear optical properties of the  $x = 9.5\%$  sample is also shown. All of the quantities vanish above  $T_c$ . Red shading indicates the bulk superconducting phase.

### Fig. 3. Simulated nonlinear reflectivity for homogeneous superconductivity.

Simulations at  $x = 9.5$  and  $15.5\%$  doping. **(A)** Simulated space ( $x$ , not to be confused with the symbol for doping concentration)– and time ( $t$ )–dependent order parameter phase  $[\theta(x, t)]$  obtained by numerically solving the sine-Gordon equation on  $x = 9.5\%$  samples [see section S3 of (9)]. The equation makes use of equilibrium superfluid density extracted from the linear optical properties and assumes excitation with terahertz pulses of shape and strength used in the experiment. The horizontal dotted lines indicate the spatial coordinate  $x$  at which the line cuts are displayed (lower panel). **(B)** Simulated order parameter phase (A) after frequency filtering centered at  $3\omega_{\text{pump}}$  with its corresponding line cut (lower panel). **(C)** Simulated reflectivity in the linear ( $E = 0.1$  kV/cm) and the nonlinear ( $E = 80$  kV/cm) regime. The third-harmonic generation component is highlighted (red shading). **(D)** to **(F)** Same as in (A) to (C) but for  $x = 15.5\%$ .





**Fig. 4. Third-harmonic generation in the striped phase.** (A) Nonlinear frequency-dependent reflectivity measured in the striped  $x = 11.5\%$  samples recorded for three different field strengths at  $T = 5$  K ( $T < T_c = 13$  K). (B) Electric field dependence of the third-harmonic amplitude for  $T = 5$  K. (C) Temperature dependence of the nonlinear reflectivity for  $T > T_c = 13$  K. (D) Temperature dependence of the third-harmonic signal (normalized to the highest field measurements at  $T = 5$  K). (E) Schematic of the order parameter phase in a pair density wave condensate. The black arrows represent the superconducting order parameter phase at each lattice point. Interlayer tunneling from perpendicularly aligned superfluid stripes is

equivalent to a checkerboard lattice of alternating  $\pi/2$  and  $-\pi/2$  phase Josephson junctions. Such a lattice has tunneling currents of  $I_c$  and  $-I_c$  flowing at the neighboring junctions at equilibrium (thick red lines). (F) Excitation modes of the PDW indicating the  $\varphi_0$  and  $\varphi_\pi$  modes (see text). The shaded region under the black arrow represents the phase excursion from the equilibrium geometry ( $\delta\theta_0$  and  $\delta\theta_\pi$ ). Corresponding current fluctuations  $\delta I_0$  and  $\delta I_\pi$  produced by such excitations are also depicted (thin red lines). (G) Calculated nonlinear current response for the unit cell of (E) after application of a single-cycle optical pulse centered at 500 GHz frequency [see section S5 of (9) for details].

We next turn to the key results of this report, which provide evidence for superfluidity in the normal-state striped phase in  $\text{La}_{1.885}\text{Ba}_{0.115}\text{CuO}_4$  ( $x = 11.5\%$ ). In this compound, striped charge order coexists with superconductivity below  $T_c = 13$  K and extends into the normal state up to  $T_{co} = 55$  K (1).

Figure 4A displays the results for  $T < T_c = 13$  K. Note that in this material, the equilibrium Josephson plasma resonance was at lower frequency than in the other two compounds with higher  $T_c$  (at  $\sim 150$  GHz) and could not be observed [see section S2 of (9)]. However, a third harmonic at  $\omega = 3\omega_{\text{pump}} \approx 1.4$  THz was clearly observed. Notably, the third-harmonic signal remained finite also for  $T > T_c = 13$  K and up to  $T \sim T_{co} = 55$  K (Fig. 4, C and D). A giant third harmonic, amounting to several percent of the driving field, can be understood only in the presence of superconducting tunneling above  $T_c$  and up to  $T_{co}$ .

In the following paragraphs, we show that a pair density wave, which does not show features of superfluidity in the linear optical properties, retains the large nonlinear optical signal of the

homogeneous condensate. As shown in Fig. 4E, the order parameter phase of the PDW, encoded by the vector angle (black arrow), changes between neighboring stripes and rotates by  $90^\circ$  from one plane to the next (18), resulting in a checkerboard lattice of  $\pi/2$  and  $-\pi/2$  Josephson junctions. Hence, the equilibrium PDW supports a lattice of staggered tunneling supercurrents, which average out to zero (thick red arrows in Fig. 4E).

The interlayer phase fluctuations probed in the optical response are described by two normal modes, termed here  $\varphi_0$  and  $\varphi_\pi$ . The  $\varphi_0$  mode is optically active. For an optical field  $E(t) = E_0 \sin(\omega_{\text{pump}} t)$ , one has identical phase excursions  $\delta\theta_0(t) = \frac{2e d E_0}{\hbar \omega_{\text{pump}}} \cos(\omega_{\text{pump}} t)$  at each site (shading under black arrows in Fig. 4F). However, current fluctuations  $\delta I_0(t)$  of equal magnitude, but opposite sign, at neighboring  $\pi/2$  and  $-\pi/2$  junctions, makes this mode silent (red arrows in Fig. 4F). The second  $\varphi_\pi$  mode consists of phase excursions  $\delta\theta_\pi(t)$  occurring in opposite directions at neighboring sites and is optically inactive (see Fig. 4F).

In the nonlinear regime, the optical response of the PDW is no longer zero. Because  $\varphi_0$  is odd and  $\varphi_\pi$  is even, a nonlinear expansion of the Josephson energy ( $U$ ) can be written as  $U(\varphi_0, \varphi_\pi) \propto J_0 \varphi_\pi^2 - J'' \varphi_\pi + C \varphi_0^2 \varphi_\pi$ , where  $J_0$  and  $J''$  are the in-plane and out-of-plane Josephson energies and  $C = J''/2$ , with  $J_0 \gg J''$ , where  $C$  is the coupling constant between the two normal modes [see section S5 of (9)].

A large third-harmonic signal is readily predicted from the  $\varphi_0^2 \varphi_\pi$  coupling term. For an optical field of the type  $E(t) = E_0 \sin(\omega_{\text{pump}} t)$ , which acts only on the mode  $\varphi_0$ , the equations of motion of the phase are  $\varphi_0 \approx \omega_{\text{pump}} E_0 \cos(\omega_{\text{pump}} t)$  and  $\varphi_\pi \approx -2C \varphi_0^2$ . These coupled equations imply not only phase oscillations in  $\varphi_0$  mode at  $\omega_{\text{pump}}$ , i.e.,  $\delta\theta_0(\omega_{\text{pump}})$ , but also indirect excitation of the optically inactive  $\varphi_\pi$  mode. Because in the equation of motion for  $\varphi_\pi$  the driving force is proportional to  $\varphi_0^2$ , the phase of this mode  $\delta\theta_\pi$  is driven at  $2\omega_{\text{pump}}$ . Because the total nonlinear current  $I_{\text{tot}} = \sum I_{\text{junctions}}$  contains terms of the type  $I_{\text{tot}}^* \sim 2I_c \varphi_0 \varphi_\pi$ , the two phase coordinates  $\varphi_0(\omega_{\text{pump}})$  and  $\varphi_\pi(2\omega_{\text{pump}})$  are mixed

and produce current oscillations at the difference and sum frequencies  $\omega_{\text{pump}}$  and  $3\omega_{\text{pump}}$ . A numerical solution of the two equations of motion with realistic parameters for the Josephson coupling energies  $J_0$  and  $J''$  displays oscillatory currents at the fundamental and third harmonic [Fig. 4G, also see section S5 of (9) for details on the simulations].

Finally, from the model above, the third-harmonic current is predicted to scale linearly with the out-of-plane critical current and, hence, the superfluid density of the stripes. From the plot in Fig. 4D, the local superfluid density at spin- and charge-ordering temperature  $T_{\text{so}}$  is found to be 60% of that measured below  $T_c$  and subsequently decreases continuously as the temperature is increased further, before vanishing at  $T_{\text{co}}$ .

Although the model discussed above provides a plausible description of the experimental observations, other hypotheses for the origin of the third-harmonic signal for  $T > T_c$  should be considered. The measured and simulated third harmonic is far larger than, and hence easily distinguished from, the effect of noncondensed quasiparticles. The nonlinear susceptibility detected in the present experiments  $\chi^{(3)} \sim 10^{-15} \text{ m}^2/\text{V}^2$  (obtained from the electric field strengths at the fundamental and third harmonic) is several orders of magnitude bigger than typical cubic nonlinearities ( $\chi^{(3)} \sim 10^{-18} - 10^{-20} \text{ m}^2/\text{V}^2$ ) (19–21). Furthermore, first-principle calculations show that, for this compound, the value of  $\chi^{(3)}$  from quasiparticle transport within anharmonic bands is at least three orders of magnitude smaller than what is measured here [see section S6 of (9)].

A second alternative may involve the sliding of a charge density wave along the  $c$  axis, as discussed in (22) for blue bronze. However, the efficiency of the sliding of a charge density wave, reported in (22) for kilohertz frequency excita-

tion, is expected to reduce strongly at higher excitation frequencies and can be ruled out for the terahertz irradiation [see section S7 of (9)]. The results for the 15.5% sample, where the third-harmonic signal disappears at  $T_c = 32 \text{ K}$  ( $< T_{\text{co}} = 40 \text{ K}$ ), further indicate that the third harmonic results from superconducting tunneling rather than charge ordering.

The observation of a colossal third-harmonic signal in the stripe-ordered state of  $\text{La}_{1.885}\text{Sr}_{0.115}\text{CuO}_4$  provides compelling experimental evidence for finite momentum condensation in the normal state of cuprates and underscores the power of nonlinear terahertz optics as a sensitive probe of frustrated excitations in quantum solids. A natural direction for this line of research involves the study of other forms of charge order that compete or coexist with superconductivity, such as those found in  $\text{YBa}_2\text{Cu}_3\text{O}_{6+x}$  (23, 24). One may also find application of these techniques in other regimes of the cuprate pseudogap, with finite superfluid density, with vanishing range phase correlations (25, 26), or where other forms of density waves (27, 28) have been discussed.

## REFERENCES AND NOTES

1. M. Hückler *et al.*, *Phys. Rev. B* **83**, 104506 (2011).
2. P. Abbamonte *et al.*, *Nat. Phys.* **1**, 155–158 (2005).
3. J. M. Tranquada, B. J. Sternlieb, J. D. Axe, Y. Nakamura, S. Uchida, *Nature* **375**, 561–563 (1995).
4. Q. Li, M. Hückler, G. D. Gu, A. M. Tselik, J. M. Tranquada, *Phys. Rev. Lett.* **99**, 067001 (2007).
5. E. Berg *et al.*, *Phys. Rev. Lett.* **99**, 127003 (2007).
6. M. H. Hamidian *et al.*, *Nature* **532**, 343–347 (2016).
7. C. C. Homes *et al.*, *Phys. Rev. B* **85**, 134510 (2012).
8. J. Hebling, K.-L. Yeh, M. C. Hoffmann, B. Bartal, K. A. Nelson, *J. Opt. Soc. Am. B* **25**, B6 (2008).
9. Materials and methods are available as supplementary materials.
10. Y. Laplace, A. Cavalleri, *Adv. Phys. X* **1**, 387–411 (2016).
11. A. Dienst *et al.*, *Nat. Photonics* **5**, 485–488 (2011).
12. A. Dienst *et al.*, *Nat. Mater.* **12**, 535–541 (2013).
13. (1) In addition to the two features, a reflectivity spike at  $\omega_{\text{JPO}}$  is observed for  $x = 9.5\%$  doping. This feature corresponds to a parametric amplification of the Josephson plasma waves [see (14)]. (2) The central frequency of the terahertz pulses for experiments in  $x = 9.5$  and  $11.5\%$  was  $\omega_{\text{pump}} = 0.45 \text{ THz}$ .
14. S. Rajasekaran *et al.*, *Nat. Phys.* **12**, 1012–1016 (2016).
15. B. D. Josephson, *Rev. Mod. Phys.* **36**, 216–220 (1964).
16. S. Savel'ev, V. A. Yampol'skii, A. L. Rakhmanov, F. Nori, *Rep. Prog. Phys.* **73**, 026501 (2010).
17. X. Hu, S.-Z. Lin, *Supercond. Sci. Technol.* **23**, 053001 (2010).
18. E. Berg, E. Fradkin, S. A. Kivelson, J. M. Tranquada, *New J. Phys.* **11**, 115004 (2009).
19. R. W. Boyd, Z. Shi, I. De Leon, *Opt. Commun.* **326**, 74–79 (2014).
20. B. Buchalter, G. R. Meredith, *Appl. Opt.* **21**, 3221–3224 (1982).
21. O. Schubert *et al.*, *Nat. Photonics* **8**, 119–123 (2014).
22. T. Chen, L. Mihály, G. Grüner, *Phys. Rev. B Condens. Matter* **36**, 2931–2934 (1987).
23. G. Ghiringhelli *et al.*, *Science* **337**, 821–825 (2012).
24. J. Chang *et al.*, *Nat. Phys.* **8**, 871–876 (2012).
25. J. Corson, R. Mallozzi, J. Orenstein, J. N. Eckstein, I. Bozovic, *Nature* **398**, 221–223 (1999).
26. L. S. Bilbro *et al.*, *Nat. Phys.* **7**, 298–302 (2011).
27. K. Fujita *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **111**, E3026–E3032 (2014).
28. P. A. Lee, *Phys. Rev. X* **4**, 031017 (2014).

## ACKNOWLEDGMENTS

We thank S. A. Kivelson and J. M. Tranquada for helpful feedback on the manuscript. We thank D. Nicoletti for providing us with the linear reflectivity data of the  $x = 11.5\%$  sample. The research leading to these results received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007–2013) and ERC grant agreement no. 319286 [Frontiers in Quantum Materials' Control (Q-MAC)]. Work performed at Brookhaven was supported by the U.S. Department of Energy, Division of Materials Science under contract no. DE-AC02-98CH10886. The data from the terahertz measurements and the simulations are kept at the Cavalleri Laboratory at Max Planck Institute of Structure and Dynamics, Hamburg, and are available from the authors upon request.

## SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/359/6375/575/suppl/DC1  
Materials and Methods  
Figs. S1 to S8  
References (29–34)

30 March 2017; accepted 22 December 2017  
10.1126/science.aan3438

## TOPOLOGICAL MATTER

# Observation of chiral phonons

Hanyu Zhu,<sup>1,2</sup> Jun Yi,<sup>1</sup> Ming-Yang Li,<sup>3</sup> Jun Xiao,<sup>1</sup> Lifa Zhang,<sup>4</sup> Chih-Wen Yang,<sup>3</sup> Robert A. Kaindl,<sup>2</sup> Lain-Jong Li,<sup>3</sup> Yuan Wang,<sup>1,2\*</sup> Xiang Zhang<sup>1,2,\*</sup>

Chirality reveals symmetry breaking of the fundamental interaction of elementary particles. In condensed matter, for example, the chirality of electrons governs many unconventional transport phenomena such as the quantum Hall effect. Here we show that phonons can exhibit intrinsic chirality in monolayer tungsten diselenide. The broken inversion symmetry of the lattice lifts the degeneracy of clockwise and counterclockwise phonon modes at the corners of the Brillouin zone. We identified the phonons by the intervalley transfer of holes through hole-phonon interactions during the indirect infrared absorption, and we confirmed their chirality by the infrared circular dichroism arising from pseudoangular momentum conservation. The chiral phonons are important for electron-phonon coupling in solids, phonon-driven topological states, and energy-efficient information processing.

Chirality is a fundamental property of an object not identical to its mirror image. For elementary and quasiparticles, it is an important quantum concept at the heart of modern physics. For example, the discovered handedness of neutrinos in electroweak interaction revolutionized our understanding of the universal parity conservation law. Chiral fermions with spin-momentum locking can emerge in solid-state systems such as the recently observed Weyl semimetals with inversion symmetry-breaking lattices (1). Electron chirality in graphene defined through pseudospin was found to cause distinctive transport properties such as unconventional Landau quantization and Klein tunneling (2, 3). Meanwhile, electrons in monolayer transition metal dichalcogenides with broken inversion symmetry display optical helicity, which opens the field of valleytronics (4). The outstanding question is whether bosonic collective excitations such as phonons can attain chirality. Recently, the intrinsic chirality of phonons without applied external magnetic fields in an atomic lattice was predicted theoretically at the corners of the Brillouin zone of an asymmetric two-dimensional (2D) hexagonal lattice (5). Because of the threefold rotational symmetry of the crystal and the momentum vectors, the vibrational plane wave is composed of unidirectional atomic circular rotation. Because the phonons in the same valley have distinct energy levels, the rotation is not the superposition of linear modes, in contrast to the circular motion in nonchiral cases (6). The hypothetical chiral phonons are potentially important for the control of intervalley scattering (7, 8), lattice modulation-driven electronic phase transitions (9, 10) and topological states (11), as well as information carriers that can be robust

against decoherence for solid-state quantum information applications (12).

Here we report the observation of intrinsic chiral phonons via transient infrared (IR) spectroscopy in the atomic lattice of monolayer WSe<sub>2</sub>. A particle's angular momentum is calculated by the phase change under a rotational transformation. This definition can be extended for quasiparticles in solid lattices with discrete spatial symmetry, as long as the rotation is in the crystalline symmetry group (13). For an inversion symmetry-breaking 2D hexagonal lattice such as monolayer WSe<sub>2</sub> (Fig. 1A), phonons with well-defined pseudoangular momentum (PAM) are located at the  $\Gamma$ , K, and K' points in the reciprocal space (5). Because the phonons inherit the threefold rotational ( $C_3$ ) symmetry of the lattice, their function of motion must be  $C_3$  invariant except for a phase difference (i.e., circular) (14). The K and K' phonons are particularly notable because there is no reflection line that preserves both the lattice and their momentum (fig. S5). Therefore, the atomic rotation cannot be reversed without changing its momentum or energy. For example, if we attempt to reverse the rotational direction of the Se atoms in a longitudinal optical phonon LO(K) while keeping the relative phase determined by the  $\mathbf{q} = \mathbf{K}$  momentum, the mode changes to a longitudinal acoustic phonon LA(K) that oscillates at a different frequency. Thus, unlike the  $\Gamma$  phonon, such a rotation cannot be decomposed into orthogonal linear vibrations with a  $\pm\pi/2$  phase shift. On the other hand, to maintain the same mode, the relative phase must be reversed, resulting in a phonon at the K' point.

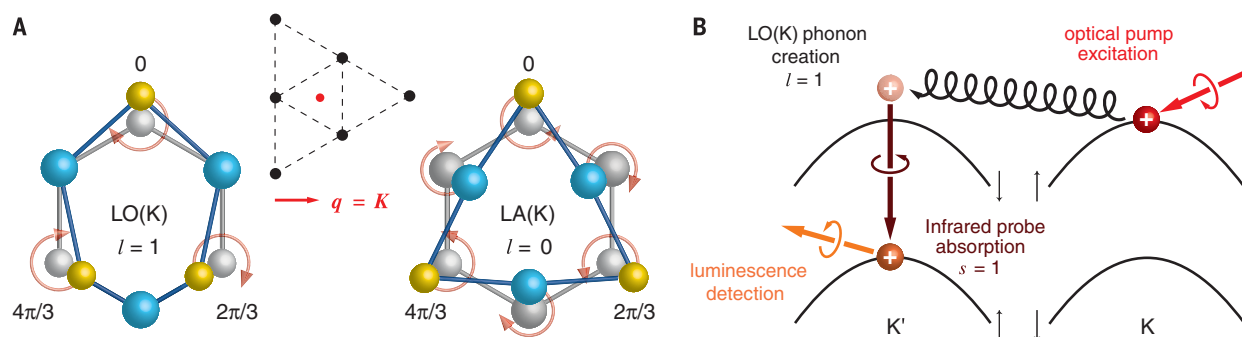
We deduce the phonon PAM  $l$  for each mode from the phase change after counterclockwise 120° rotation:  $\hat{C}_3(u_k) = e^{-i(2\pi/3)l} u_k$ , where  $u_k$  is the function of atomic motion (fig. S6). Because LA(K) is always  $C_3$  symmetric, it has  $l = 0$ ; in contrast, LO(K) gains negative phase after rotation (becomes ahead of time) and thus  $l = 1$ . The K phonons with different PAM are completely nondegenerate because of the mass difference between W and Se. The lifting of degeneracy

guarantees that each mode has a distinctive selection rule for the electron-phonon and optical scattering (table S1). These features differentiate the intrinsically chiral phonons from any previously investigated phonons where circular atomic motion results from superposing linear polarized eigenmodes (6, 15, 16). Chirality in the absence of an external magnetic flux is also distinguished from the magnetic field-induced split of degenerate superposition modes (17). Finally, both the physical origin and properties of the atomic chiral phonons are fundamentally different from previous phononic crystal-based chirality from geometric engineering or topological edge states (18–20).

We use an optical pump-probe technique to identify the chiral phonon by its scattering with holes (Fig. 1B and fig. S2). The linear momentum, PAM, and energy of the respective phonon modes are determined by characterizing all of the other particles involved in the intervalence band transition (IVBT) process (14). First, we inject holes at the K valley by a left-circularly polarized (LCP) optical pump pulse. The K valley-polarized hole relaxes to the valence band edge with initial linear momentum  $\mathbf{p}_i = -\mathbf{K}$ . It can transit to the K' point ( $\mathbf{p}_f = \mathbf{K}$ ) by emitting a K phonon ( $\mathbf{q} = \mathbf{p}_i - \mathbf{p}_f = -2\mathbf{K} = \mathbf{K}$ ), but the intermediate state is virtual because the spin-conserving state has much higher energy due to large spin-orbit coupling. We then send an IR pulse to satisfy the energy conservation and place the hole in the spin-split band at K'. Because the hole states have zero PAM, the PAM of the phonon must be equal to the spin angular momentum of the absorbed IR photon. Therefore, the LCP pulse controls the creation of only LO(K) phonons. Eventually, we recognize the final spin-split state in the opposite valley by the energy and right circular polarization of its radiative decay. The selection rule of phonon creation will be identical if the hole first absorbs an IR photon, transits to an intermediate state within the valley, and then emits a phonon, as long as the final state is the same. The phonon annihilation is negligible because the population is very low for all phonon modes at the K or K' points at a base temperature of 82 K. The selection is also not affected by a small momentum distribution of the initial and final holes (14) due to finite thermal energy or the strong many-body interaction (21). This is because, like the optical valley-selectivity, the contrast of electron-phonon coupling strength remains large in the vicinity of K and K', even if the PAM is only defined on these points.

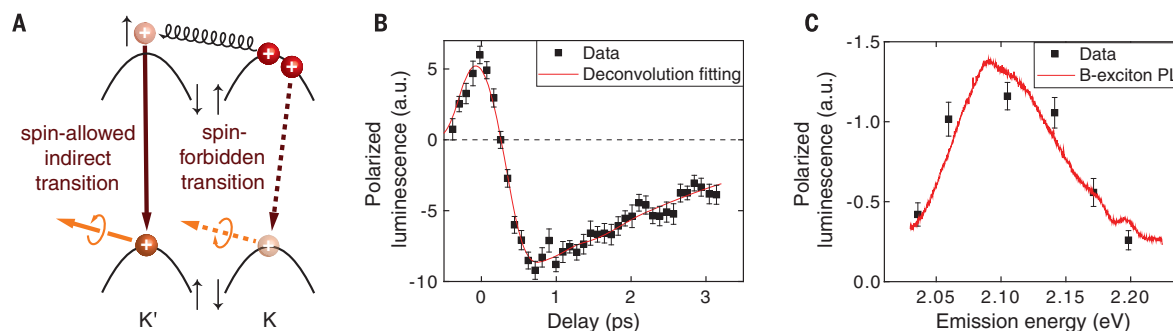
To determine the linear momentum of the involved phonon, we measure the linear momentum of the initial and final holes. Under circularly polarized pump excitation, the majority of holes maintain their linear momentum near the original valley, as shown by the positive helicity of the photoluminescence. The final linear momentum of the spin-split states is read out by the polarization of their luminescence at higher energy than the original pump photon (Fig. 2A). We measure the difference of luminescence with the same polarization as the pump light ( $I_{\text{same}}$ )

<sup>1</sup>Nanoscale Science and Engineering Center, University of California, Berkeley, CA 94720, USA. <sup>2</sup>Materials Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA. <sup>3</sup>Physical Sciences and Engineering Division, King Abdullah University of Science and Technology, Thuwal 23955-6900, Kingdom of Saudi Arabia. <sup>4</sup>Department of Physics, Nanjing Normal University, Nanjing, Jiangsu 210023, China. \*Corresponding author. Email: xzhang@me.berkeley.edu (X.Z.); yuanwang@berkeley.edu (Y.W.)



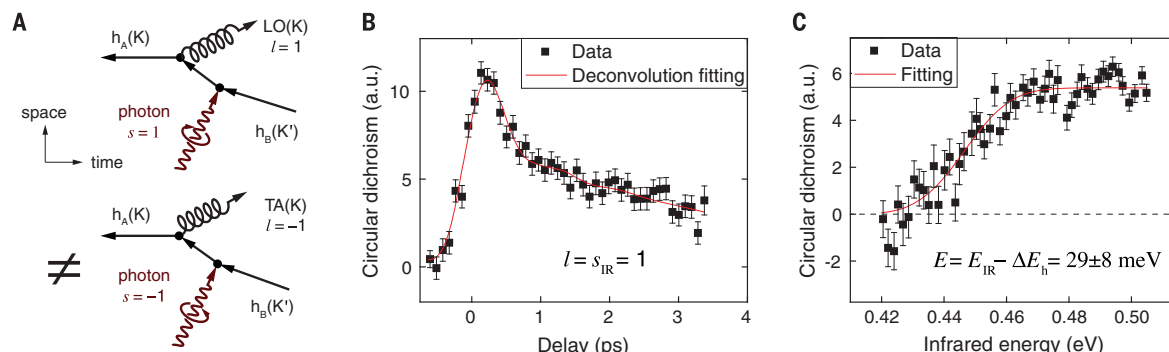
**Fig. 1. Nondegenerate chiral phonons in monolayer WSe<sub>2</sub> and the selection rule of hole-phonon interactions.** (A) The atomic motion of W and Se atoms (blue and yellow spheres, respectively) in the real space is intrinsically circular for the chiral phonons residing at the K point (red dot) in the reciprocal lattice (array of black dots), due to the threefold rotational symmetry. Because the momentum vector determines the relative phase of the Se motion, opposite rotations correspond to completely different modes. The two modes

have distinct PAM with respect to the center of the hexagon. (B) In the intervalley optical transition of holes, the PAM of the K phonon is equal to the spin of the IR photon due to angular momentum conservation. A hole injected by the LCP optical pump moves from the K to the K' valley through virtual scattering with a LO(K) phonon. Owing to the spin and energy mismatch, a real phonon is created only if the hole simultaneously absorbs an IR probe photon and transits to the spin-split state, which is signaled by the RCP luminescence.



**Fig. 2. Chiral phonon creation from the intervalley transition of holes measured by polarized luminescence.** (A) The spin-allowed intervalley transition of holes from the K point creates a K phonon, yet the spin-forbidden transition within the same valley does not. The two processes are distinguished by the polarized luminescence signal. (B) The observed initial positive polarization indicates a spin-flipping transition of nonequilibrium holes, with a lifetime of ~0.2 ps. After that, the negative polarization shows a prevailing

phonon-creating intervalley transition for valley-polarized holes, with a lifetime of ~2 ps at 82 K. The pump and probe energies are tuned to 1.97 and 0.51 eV, respectively, whereas the signal is integrated from 2.05 to 2.19 eV. (C) The spectrum of the polarized luminescence taken at  $\tau = 0.8$  ps agrees with the B-exciton photoluminescence (PL) and confirms its spin-split hole origin. The error bars in (B) and (C) indicate SE calculated from multiple measurements. a.u., arbitrary units.



**Fig. 3. Chirality of phonons measured by transient IR CD.** (A) The phonons participating in the indirect optical transition are different for the LCP and RCP IR photons because of the conservation of linear and angular momentum. Their distinct electron-phonon scattering strengths dictate that the two processes have different amplitudes, leading to polarized IR absorption.  $h_A(K)$ , hole at the valence band edge of the K valley;  $h_B(K')$ , hole in the spin-split band of the K' valley. (B) The measured CD at 82 K (black squares and red fitting curve) is positive for both the initial intravalley spin-flipping transition and the later intervalley phonon-creating

transition. It not only proves the chirality of phonons but also indicates that the LO phonon process is dominant. (C) The spectrum of the transient CD acquired at  $\tau = 0.8$  ps shows a photonic energy threshold of  $0.448 \pm 0.002$  eV. Its step shape is expected from the 2D excitonic density of states. Subtracting the energy difference between the initial and final hole states, a phonon energy of  $29 \pm 8$  meV is deduced, in agreement with that of the LO mode. The error bars denote SE calculated from multiple measurements, and the threshold uncertainty indicates the SE of the fitting parameter.

and the opposite polarization ( $I_{\text{oppo}}$ ) as a function of the probe delay time  $\tau$  (Fig. 2B). We find that the polarized luminescence ( $I_{\text{same}} - I_{\text{oppo}}$ ) is positive for  $\tau < 0.2$  ps, indicating that the spin-flipping intervalley IVBT has a higher probability when the IR probe arrives immediately after the excitation of holes. This is because spin is not strictly conserved for optical transitions away from the high-symmetry points. However, the process is forbidden at the valley center with incident light perpendicular to the lattice plane, so its magnitude decays with a lifetime of 0.2 ps as the nonequilibrium carriers thermalize to the band edge (22, 23). After that, the luminescence switches polarization as a result of the dominant spin-conserving intervalley IVBT. It shows that the majority of final holes have a linear momentum opposite that of the initial holes. The momentum must have been transferred to phonons instead of defects, as seen from the comparison of the second-order Raman scattering through the defect-assisted one-phonon process with that of the two-phonon process, which is an order of magnitude stronger (fig. S7). Directional intervalley transfer through electron-electron interactions (24) is excluded because the pump is not in resonance with the A exciton, and the signal is linearly proportional to the hole density (fig. S8). The exchange interaction reduces the polarization of the holes but should not create an opposite polarization (25). Finally, the depolarization plus the decay of the hole population yields a joint lifetime of  $\sim 2$  ps for the polarized luminescence, in agreement with previous exciton studies (26). To further verify the nature of the final holes, we measure the polarized luminescence at different collection energies with  $\tau = 0.8$  ps. The emission spectra obtained with different IR photon energies are all close to the B-exciton photoluminescence (Fig. 2C and fig. S12), which proves that the emission is the radiative decay of a real state near the edge of the spin-split band, rather than a virtual state. Therefore, we use the luminescence from 2.05 to 2.19 eV to quantify the indirect IR absorption and the phonon creation.

The PAM of valley phonons is determined by measuring the polarization selection of the absorbed IR photon, resulting from the chiral electron-phonon interaction. With LCP excitation, the intervalley transition of holes from the K valley can either absorb an LCP IR photon to produce LO phonons or absorb a right-circularly polarized (RCP) IR photon to produce  $\text{TA}/\text{A}_1$  phonons (Fig. 3A). We observe different IR absorption when the probe has the same polarization as the pump light ( $\alpha_{\text{same}}$ ) versus when it has the opposite polarization ( $\alpha_{\text{oppo}}$ ). This IR circular dichroism ( $\text{CD} = \alpha_{\text{same}} - \alpha_{\text{oppo}}$ ) demonstrates that scattering cross sections of the two processes are not equal, owing to the difference of electron-phonon coupling strength. As shown in Fig. 3B, the CD at 82 K is always positive for valley-polarized holes. Because the holes have zero PAM regardless of valley index or spin-split bands, the angular momentum of

the IR photon is transferred to either the spin or the phonon. The initial intervalley IVBT produces positive CD because the electronic spin flips from  $-1/2$  in the spin-split band to  $+1/2$  at the valence band edge at the K point. However, this contribution decays fast as a function of IR delay in accordance with Fig. 2B. Thus, the intervalley phonon-creating transition is the dominant source of positive CD for  $\tau \geq 0.8$  ps. It indicates that the LO branch contributes most to the indirect IR absorption. For nonchiral systems, although photons with opposite circular polarization may interact with different particles, the probabilities are always equal. Therefore, the CD is the signature of intrinsic phonon chirality.

The energy of this phonon mode is measured according to energy conservation: The energy sum of the incoming particles, the IR photon, and the initial holes must be equal to that of the outgoing particles, the chiral phonon, and the final holes (14). We observe that the CD spectrum at  $\tau = 0.8$  ps (Fig. 3C) is a step function with a clear threshold near  $0.448 \pm 0.002$  eV, broadened by the spectral width of the IR pulse. The shape corresponds to a transition from a hole at the valence edge to a band with parabolic dispersion in two dimensions (fig. S9). This value is distinguished from either the intra-excitonic transition (27) or the exciton dissociation energy (27). Next, we find that the ground-state configuration of the initial holes is the dark A trion at  $E_i = 1.671 \pm 0.006$  eV (fig. S11) (28) based on the dominant bright A-trion photoluminescence (fig. S1). We verify that the dark trions are near thermal equilibrium because the IR spectrum is very different from that of nonequilibrium carriers with excessive energy at  $\tau = 0$  ps (fig. S10). The final state is the bright B trion at  $E_f = 2.090 \pm 0.005$  eV, as measured through its emission (fig. S12). Summarizing these values, we deduce the phonon energy  $E_{\text{phonon}} = E_{\text{photon}} + E_i - E_f = 29 \pm 8$  meV, in agreement with the first-principles calculation and the Raman spectrum for the LO(K) mode (figs. S7 and S9). The uncertainty may be reduced in the future by improving the uniformity of the sample and the spectral resolution of the B-exciton emission. Better spectroscopy will also potentially reveal details about the chiral phonon-exciton coupling—such as the contribution from LA phonons and the various pathways of the indirect transition—that are not resolvable with the current signal-to-noise ratio.

Our findings of chiral phonons are fundamentally important for potential experimental tests of quantum theories with chiral bosons (29) in the solid state. Our work also provides a possible route for controlling valley and spin through electron-phonon scattering and strong spin-phonon interactions (16). Furthermore, the lifting of degeneracy by chirality offers robust PAM information against decoherence and long-range perturbation and offers a new degree of freedom to the design and implementation of phononic circuitry (30) at the atomic scale without magnetic fields.

## REFERENCES AND NOTES

- M. Z. Hasan, S.-Y. Xu, I. Belopolski, S.-M. Huang, *Annu. Rev. Condens. Matter Phys.* **8**, 289–309 (2017).
- F. D. M. Haldane, *Phys. Rev. Lett.* **61**, 2015–2018 (1988).
- S. Das Sarma, S. Adam, E. H. Hwang, E. Rossi, *Rev. Mod. Phys.* **83**, 407–470 (2011).
- J. R. Schaibley et al., *Nat. Rev. Mater.* **1**, 16055 (2016).
- L. Zhang, Q. Niu, *Phys. Rev. Lett.* **115**, 115502 (2015).
- H. Katsuki et al., *Nat. Commun.* **4**, 3801 (2013).
- H. Zeng, J. Dai, W. Yao, D. Xiao, X. Cui, *Nat. Nanotechnol.* **7**, 490–493 (2012).
- B. R. Carvalho et al., *Nat. Commun.* **8**, 14670 (2017).
- M. Rini et al., *Nature* **449**, 72–74 (2007).
- M. Först, R. Mankowsky, A. Cavalleri, *Acc. Chem. Res.* **48**, 380–387 (2015).
- G. Jotzu et al., *Nature* **515**, 237–240 (2014).
- K. C. Lee et al., *Science* **334**, 1253–1256 (2011).
- R. C. Johnson, *Phys. Lett. B* **114**, 147–151 (1982).
- Materials, methods, and additional information are available as supplementary materials.
- S.-Y. Chen, C. Zheng, M. S. Fuhrer, J. Yan, *Nano Lett.* **15**, 2526–2532 (2015).
- T. F. Nova et al., *Nat. Phys.* **13**, 132–136 (2017).
- G. Schaack, *J. Phys. C* **9**, L297–L301 (1976).
- A. Spadoni, M. Ruzzene, S. Gonella, F. Scarpa, *Wave Motion* **46**, 435–450 (2009).
- R. Süsstrunk, S. D. Huber, *Science* **349**, 47–50 (2015).
- C. Brendel, V. Peano, O. J. Painter, F. Marquardt, *Proc. Natl. Acad. Sci. U.S.A.* **114**, E3390–E3395 (2017).
- T. Low et al., *Nat. Mater.* **16**, 182–194 (2017).
- F. Ceballos, Q. Cui, M. Z. Bellus, H. Zhao, *Nanoscale* **8**, 11681–11688 (2016).
- P. Steinleitner et al., *Nano Lett.* **17**, 1455–1460 (2017).
- M. Manca et al., *Nat. Commun.* **8**, 14927 (2017).
- T. Yu, M. W. Wu, *Phys. Rev. B* **89**, 205303 (2014).
- T. Yan, J. Ye, X. Qiao, P. Tan, X. Zhang, *Phys. Chem. Chem. Phys.* **19**, 3176–3181 (2017).
- C. Poellmann et al., *Nat. Mater.* **14**, 889–893 (2015).
- X.-X. Zhang et al., *Nat. Nanotechnol.* **12**, 883–888 (2017).
- M. M. H. Barreira, C. Wotzasek, *Phys. Rev. D Part. Fields* **45**, 1410–1415 (1992).
- N. Li et al., *Rev. Mod. Phys.* **84**, 1045–1066 (2012).

## ACKNOWLEDGMENTS

We thank F. Wang and Q. Niu for helpful discussions. This work was primarily supported by the U.S. Department of Energy, Office of Science, Basic Energy Sciences, Materials Sciences and Engineering Division under contract no. DE-AC02-05-CH11231 within the van der Waals Heterostructures program (KCWF16) for sample preparation and theory and data analysis, and within the Subwavelength Metamaterials Program (KC12XZ) for optical design and measurement. R.A.K. was supported under the same contract within the Ultrafast Materials Science program (KC2203) for mid-IR frequency conversion. J.Y. acknowledges a scholarship from the China Scholarship Council (CSC) under grant no. 201606310094. L.Z. thanks M. Gao for helpful calculation and discussion and acknowledges support from the National Natural Science Foundation of China (grant no. 11574154). L.-J.L. acknowledges support from the King Abdullah University of Science and Technology through a Competitive Research Grant (CRG5) for monolayer  $\text{WSe}_2$  synthesis. All data needed to evaluate the conclusions in the paper are present in the paper and the supplementary materials.

## SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/359/6375/579/suppl/DC1  
Materials and Methods  
Figs. S1 to S12  
Table S1  
References (31–46)  
Movies S1 and S2

20 October 2017; accepted 19 December 2017  
10.1126/science.aar2711

Cite as: D. Chowell *et al.*, *Science*  
10.1126/science.aao4572 (2017).

# Patient HLA class I genotype influences cancer response to checkpoint blockade immunotherapy

Diego Chowell,<sup>1,2</sup> Luc G.T. Morris,<sup>2,3\*</sup> Claud M. Grigg,<sup>4\*</sup> Jeffrey K. Weber,<sup>5</sup> Robert M. Samstein,<sup>1,2</sup> Vladimir Makarov,<sup>1,2</sup> Fengshen Kuo,<sup>1,2</sup> Sviatoslav M. Kendall,<sup>1,2</sup> David Requena,<sup>6</sup> Nadeem Riaz,<sup>1,2,7</sup> Benjamin Greenbaum,<sup>8</sup> James Carroll,<sup>9</sup> Edward Garon,<sup>9</sup> David M. Hyman,<sup>10,14</sup> Ahmet Zehir,<sup>11</sup> David Solit,<sup>1,10,12</sup> Michael Berger,<sup>1,11,12</sup> Ruhong Zhou,<sup>5,13</sup> Naiyer A. Rizvi,<sup>4†</sup> Timothy A. Chan<sup>1,2,7,14†</sup>

<sup>1</sup>Human Oncology and Pathogenesis Program, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA. <sup>2</sup>Immunogenomics and Precision Oncology Platform, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA. <sup>3</sup>Department of Surgery, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA.

<sup>4</sup>NewYork-Presbyterian/Columbia University Medical Center, 177 Fort Washington Ave., New York, NY 10032, USA. <sup>5</sup>IBM Thomas J. Watson Research Center, Yorktown Heights, NY 10598, USA. <sup>6</sup>The Rockefeller University, New York, NY 10065, USA. <sup>7</sup>Department of Radiation Oncology, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA. <sup>8</sup>Tisch Cancer Institute, Departments of Medicine, Oncological Sciences, and Pathology, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA.

<sup>9</sup>David Geffen School of Medicine at UCLA, 2825 Santa Monica Boulevard, Suite 200, Santa Monica, CA 90404, USA. <sup>10</sup>Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA. <sup>11</sup>Department of Pathology, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA. <sup>12</sup>Marie-Josée and Henry R.

Kravis Center for Molecular Oncology, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA. <sup>13</sup>Department of Chemistry, Columbia University, New York, NY 10027, USA. <sup>14</sup>Weill Cornell School of Medicine, New York, NY 10065, USA.

†Corresponding author. E-mail: chant@mskcc.org (T.A.C.); nar2144@cumc.columbia.edu (N.A.R.)

\*These authors contributed equally to this work.

**CD8<sup>+</sup> T cell-dependent killing of cancer cells requires efficient presentation of tumor antigens by human leukocyte antigen class I (HLA-I) molecules. However, the extent to which patient-specific HLA-I genotype influences response to anti-PD-1 or anti-CTLA-4 is currently unknown. We determined the HLA-I genotype of 1,535 advanced cancer patients treated with immune checkpoint blockade (ICB). Maximal heterozygosity at HLA-I loci (*A*, *B*, and *C*) improved overall survival after ICB compared to patients who were homozygous for at least one HLA locus. In two independent melanoma cohorts, patients with the HLA-B44 supertype had extended survival, whereas the HLA-B62 supertype (including HLA-B\*15:01) or somatic LOH at HLA-I, was associated with poor outcome. Molecular dynamics simulations of HLA-B\*15:01 revealed unique elements that may impair CD8<sup>+</sup> T cell recognition of neoantigens. Our results have important implications for predicting response to ICB and for the design of neoantigen-based therapeutic vaccines.**

Immune checkpoint inhibitors that target cytotoxic T-lymphocyte-associated protein 4 (CTLA-4) or programmed cell death protein 1 (PD-1) or its ligand (PDL-1) have markedly improved the treatment of patients with metastatic cancer (1–3). However, tumor responses to these drugs are variable, and treatment resistance is common (4–6). To date, most research to predict clinical efficacy to immune checkpoint blockade (ICB) therapies has focused on tumor immune phenotype, somatic genomic features, or the gut microbiome (7–21), but how host germline genetics impacts response is unclear.

HLA class I (HLA-I) genotype has been linked with differential immune responses to infection, inflammatory conditions, and autoimmune diseases (22–30). Each HLA-I molecule binds specific peptides derived from intracellular proteins for presentation on the cell surface to CD8<sup>+</sup> T cells (31–33). The anti-tumor activity of ICB has been shown to depend on CD8<sup>+</sup> T cell, HLA class I-dependent immune activity (34–36). We performed survival and genetic association analyses to address two hypotheses: (i) zygosity at HLA-I genes

influences survival of cancer patients to ICB, and (ii) individual HLA-I germline alleles influences survival to ICB differently.

We examined two sets of cancer patients (henceforth called cohort 1 and cohort 2) treated with ICB. Cohort 1 ( $n = 369$  patients) was treated with anti-CTLA-4 or anti-PD-1 therapy, and exome sequencing and clinical data were obtained. Within cohort 1, 269 patients had advanced melanoma [previously reported (7, 11, 12, 17)], and 100 patients had advanced non-small cell lung cancer (NSCLC) (10) (table S1). Patients with NSCLC were treated mainly with anti-PD-1 monotherapy. Cohort 2 ( $n = 1,166$  patients) comprised different cancer types including melanoma and NSCLC (table S1), and tumors were subjected to targeted next-generation sequencing (MSK-IMPACT) (37). These patients were treated with drugs targeting CTLA-4, PD-1/PD-L1, or a combination of both, at the Memorial Sloan Kettering Cancer Center (37). For all patients in both cohorts, we performed high-resolution HLA class I genotyping from normal DNA using DNA sequencing data or a clinically validated HLA typing assay (LabCorp).

HLA class I molecules are highly polymorphic with variation located in the peptide-binding region, as each variant binds a select repertoire of peptide ligands. As such, an individual homozygous in at least one HLA-I locus would be predicted to present a smaller, less diverse repertoire of tumor-derived neoantigens to cytotoxic T lymphocytes (CTLs) compared to a person who is heterozygous at each class I locus (32). We therefore asked whether greater diversity (heterozygosity) in the repertoire of antigen-presenting HLA-I molecules could be associated with better survival following ICB therapy. We examined HLA-I variation at each of the genes (HLA-A, -B, and -C) in cohort 1 and cohort 2, by employing a Cox proportional hazard regression model to examine overall survival probability. HLA-I homozygosity in at least one locus was associated with reduced survival in cohort 1 ( $n = 369$ ;  $P = 0.036$ , HR = 1.40, 95% CI 1.02 – 1.9) (Fig. 1A), and was validated in the independent cohort of 1,166 patients (cohort 2;  $P = 0.028$ , HR = 1.31, 95% CI 1.03 – 1.70) (Fig. 1B). The number of somatic mutations in tumors was not statistically different between homozygous and heterozygous patients (fig. S1, A and B). Furthermore, the association of HLA-I homozygosity with reduced survival remained significant in multivariable Cox regression modeling when analyzed for mutation load, tumor stage, age, and drug class in cohort 1 ( $P = 0.02$ , HR = 1.50, 95% CI 1.07 – 2.10) (table S2) and in cohort 2 ( $P = 0.028$ , HR = 1.31, 95% CI 1.03 – 1.67) (table S3).

We next examined all 1,535 patients from cohort 1 and 2 together, to determine whether the effect of homozygosity may be due to a single HLA-I locus, or a combination of different loci. This analysis revealed that homozygosity at one HLA-I locus (A, or B, or C) was associated with significant reduction of overall survival ( $P = 0.003$ , HR = 1.38, 95% CI 1.11 – 1.70) (Fig. 1C). Interestingly, the effect of homozygosity on survival due to specific HLA-I locus seemed mostly associated with HLA-B ( $P = 0.052$ , HR = 1.66, 95% CI 0.93 – 2.94) (Fig. 1C) and HLA-C ( $P = 0.004$ , HR = 1.60, 95% CI 1.16 – 2.21) (Fig. 1C). Of note, the number of patients available likely limited the interpretability of analyses involving combinations of loci (e.g., HLA-A and -B). Our findings may be explained because HLA-B is generally expressed at higher levels on the cell surface than HLA-A and HLA-C, and because HLA-B alleles bind to a greater diversity of peptides (38, 39). Amino acids that bind to the B pocket of HLA-A alleles are broadly hydrophobic. In contrast, the B pocket of HLA-B alleles can accommodate a greater variety of residues (29, 39). Interestingly, antigen-presenting cells express higher levels of HLA-C on the cell surface than other cell types (40), suggesting that heterozygous HLA-C may facilitate continuous CTL priming (41).

Previous reports have shown that the total number of somatic coding mutations in a cancer genome correlates with

response to ICB (7, 8, 10–12, 17). An explanation for this observation is that the number of tumor mutations presented on the cell surface increases the probability of neoantigen recognition by cytotoxic T cells (42). We found that HLA-I homozygosity and low mutation burden were strongly associated with decreased survival compared to patients who were heterozygous at each class I locus and whose tumors had high mutation burden, in cohort 1 ( $P = 0.003$ , HR = 2.03, 95% CI 1.27 – 3.30) (Fig. 1D) and in cohort 2 ( $P < 0.0001$ , HR = 2.98, 95% CI 1.84 – 4.82) (Fig. 1E). Notably, the combined effect of HLA class I heterozygosity and mutation load on improved survival was greater compared with mutation load alone (Fig. 1, F and G).

Previous work has reported loss of heterozygosity (LOH) of HLA class I genes in cancer (43, 44). We thus analyzed all tumor exomes from cohort 1 and identified 32 patients who were heterozygous at all HLA-I loci, but had LOH in at least one HLA-I locus in their tumors (table S1). Patients with LOH of HLA-I were associated with reduced survival ( $P = 0.05$ , HR = 1.60, 95% CI 1.03 – 2.43) (Fig. 1H). Furthermore, the effect of LOH of HLA-I on survival was greater in patients whose tumors contained low mutation load ( $P = 0.0006$ , HR = 3.68, 95% CI 1.64 – 8.23) (Fig. 1I). Given that only a small fraction of presented tumor mutations are immunogenic in cancer patients (45, 46), our findings suggest that relatively small differences in the number of available HLA-I molecules in a given individual can present major challenges to effective anti-tumor T cell responses and efficacy to ICB. Furthermore, the demonstration of a significant survival advantage to HLA-I heterozygosity in patients treated with ICB both at the germline and somatic level, highlights its importance in the dynamic anti-tumor immune response and immune evasion.

As an exploratory analysis, we also found that HLA class II homozygosity at HLA-DP was associated with reduced survival ( $P = 0.018$ , HR = 1.45, 95% CI 1.06 – 2.00) (fig. S2A). Additionally, homozygosity at HLA-DPB locus was associated with decreased survival ( $P = 0.04$ , HR = 1.37, 95% CI 1.07 – 1.87). This effect was independent of the associations of HLA-I homozygosity and mutation burden (tables S4 and S5). It is noteworthy that mismatched HLA-DP has been shown to be associated with graft-versus-host disease (47).

Additionally, we employed next-generation deep sequencing of T-cell receptor CDR3 regions (TCR-seq) (48, 49) from a subset of tumor samples collected on-therapy (4 weeks post-Nivolumab initiation) (17). We found significantly higher on-therapy clonality of TCR CDR3s in HLA heterozygous patients compared to HLA homozygotes (in at least one class I locus or at HLA-DP) (Wilcoxon rank-sum test  $P = 0.0093$ ) (Fig. 3F and table S1). To refine the interpretation of this result with respect to the antigen-binding properties of the TCR repertoire, the clonality of CDR3s encoded by a single VJ cassette combination was analyzed individually (17, 50). Notably,

higher on-therapy clonality of TCR CDR3s per VJ was observed in HLA heterozygotes (Wilcoxon rank-sum test  $P = 0.023$ ) (Fig. 3G). Altogether, these results indicate that the diversity of HLA molecules in a given patient influences the selection and the resulting T-cell clonal expansion reactive against neoantigens following ICB (51).

To investigate the clinical relevance of individual HLA-I alleles after ICB therapy, we examined the effects of HLA-I supertypes on overall survival. Individual HLA-I alleles are classified into twelve discrete supertypes (52, 53), based upon similar peptide-anchor-binding specificities (26, 52, 53). These supertypes together cover most HLA-A and HLA-B alleles found in distinct populations (52, 53).

To assess the effect of HLA supertype on survival, we focused on melanoma patients, as there were a sufficient number of patients in the two patient sets for meaningful analysis. Based on the biological definition of supertypes, we classified the 27 HLA-A alleles present in the patients with melanoma into six A supertypes, and the 50 HLA-B alleles into six B supertypes (Fig. 2A and table S1). We found two B supertypes, associated with survival outcome in patients with advanced melanoma treated with anti-CTLA-4. Patients with B44 superfamily alleles had significantly better survival ( $P = 0.01$ , HR = 0.61, 95% CI 0.42 – 0.89) (Table 1) and patients with B62 alleles had significantly reduced survival ( $P = 0.0007$ , HR = 2.29, 95% CI 1.40 – 3.74) (Table 1). The B44 supertype was present at a prevalence of 45% and the B62 supertype at 15% (Fig. 2A). We did not find any supertype significantly associated with overall survival in patients with NSCLC, likely due to the limited sample size.

We then examined whether these supertype associations were influenced by the presence of specific component HLA-I alleles. The B44 association was influenced by HLA-B\*18:01, HLA-B\*44:02, HLA-B\*44:03, HLA-B\*44:05, and HLA-B\*50:01 ( $P = 0.001$ , HR = 0.49, 95% CI 0.32 – 0.76) (Fig. 2C) (Table 1). And, the B62 association was significantly driven by HLA-B\*15:01 ( $P = 0.002$ , HR = 2.21, 95% CI 1.33 – 3.70) (Fig. 3A) (Table 1). Both of these B44 and B62 allele associations remained statistically significant ( $P = 0.01$  and  $P = 0.02$ , respectively) after a Bonferroni correction. The variability in the effect on survival across these B44 alleles might be explained due to allele frequency in the cohort or due to particular differences in the peptide motifs inside or outside primary anchor pockets (54, 55).

In the independent cohort 2, melanoma patients treated with anti-PD-1 or anti-CTLA-4 who had these B44 supertype alleles had significantly better overall survival on univariate ( $P = 0.054$ , HR = 0.32, 95% CI 0.09 – 1.1) (Fig. 2, B and D), and multivariable analysis (tables S6 and S7). Furthermore, the effect of B44s on extended survival was greater when somatic mutational load was also considered in cohort 1 ( $P < 0.0001$ , HR = 0.23, 95% CI 0.13 – 0.41) (Fig. 2E) and in cohort 2 ( $P =$

0.023, HR = 0.13, 95% CI 0.02 – 1.03) (Fig. 2F). The combined effect of the B44 alleles and mutation load was greater than simply considering mutation burden alone (Fig. 2, G and H). We note that, in general, outcomes of melanoma patients in cohort 2 tended to be better than in cohort 1 because patients who received ICB and were accrued to our protocol for MSK-IMPACT testing tended to have longer survival. Yet, despite this trend, we still observed a significant effect from the B44 alleles. Notably, the B44 alleles did not associate with survival in patients with melanoma from The Cancer Genome Atlas (TCGA), suggesting that the presence of B44 was predictive of response to ICB and was not prognostic (Fig. 2I).

Most members of the B44 supertype share a preference for peptides with Glu (E) at anchor position P2, and polar and hydrophobic residues at the C terminus (54, 56) (Fig. 2J). Interestingly, we found that one out of the six enriched amino acid mutations across these tumors was G > E (fig. S3). This observation suggests that there might be an enrichment of presentation of B44-restricted neoantigens. Additionally, a number of previously identified immunogenic antigens expressed by melanomas are HLA-B44 restricted (Fig. 2J and table S8), including the testis antigen MAGEA3, restricted to HLA-B\*44:03 and HLA-B\*18:01 (both members of B44), and a clonal immunogenic neoantigen (FAM3C: TESPFEQHI) identified in a melanoma patient with long-term response to CTLA-4 blockade from cohort 1 (table S8) (11, 13).

In contrast, the B62 association with poor survival driven by the HLA-B\*15:01 allele was intriguing (Fig. 3A) (Table 1). In an exploratory analysis, we sought to determine whether any molecular features in HLA-B\*15:01 are associated with its effect on survival. Out of all the HLA-B alleles that were available for three-dimensional structural analysis ( $n = 119$ ) (table S9), we identified three alleles at their highest resolutions, HLA-B\*15:01, HLA-B\*07:02, and HLA-B\*53:01, as possessing a structural bridge in the peptide-binding groove (Arg62, Ile66, and Leu163) (Fig. 3, B and C).

We postulated that this specific structural feature may modulate the effective T cell recognition of neopeptides presented on HLA-B\*15:01. To evaluate the validity of this hypothesis, we conducted molecular dynamics (MD) simulations following similar protocols used in previous studies (57–59).

In the case of HLA-B\*07:02 and HLA-B\*53:07, MD simulations demonstrated that the bound peptide expands the respective HLA binding cleft, effectively breaking the bridge (fig. S4, A to D). Conversely, in the HLA-B\*15:01 molecule, the bridge was largely maintained with the peptide present, and the bridging residues were also made much less flexible (Fig. 3, D and E). While the mean bridge separation remained nearly constant ( $\sim 6\text{\AA}$ ) in both systems of HLA-B\*15:01 (Fig. 3E), the fluctuations in this distance were less dramatic in the peptide-bound complex. Altogether, these unique structural

and dynamical elements of HLA-B\*15:01 may impair the total strength of the interaction with T-cell receptor for effective neoantigen recognition. However, further experimental work will be necessary to test this hypothesis. It is noteworthy that we found several mutations in genes that have been recently reported to contribute to HLA and cytolytic activity (4–6) (fig. S5). However, we did not find any particular gene mutation associated with decreased overall survival.

Our findings reveal that HLA-I genes influence patient survival to ICB. Both patient-specific HLA-I genotype as well as somatic alterations in tumors impacted clinical outcome to ICB, suggesting these factors could be considered in the design of future clinical trials. The observation that the B44 is associated with extended overall survival may provide an opportunity for the development of therapeutic vaccines that potentially target immunodominant HLA-B44-restricted neoantigens expressed by melanomas. Our findings indicate that HLA-I homozygosity and LOH represent a genetic barrier to effective immunotherapy, and alternative ways to harness the immune system may be necessary to maximize clinical benefit.

## REFERENCES AND NOTES

1. D. T. Le *et al.*, Mismatch-repair deficiency predicts response of solid tumors to PD-1 blockade. *Science* **356**, eaan6733 (2017).
2. F. S. Hodi, S. J. O'Day, D. F. McDermott, R. W. Weber, J. A. Sosman, J. B. Haanen, R. Gonzalez, C. Robert, D. Schadendorf, J. C. Hassel, W. Akerley, A. J. M. van den Eertwegh, J. Lutzky, P. Lorigan, J. M. Vaubel, G. P. Linette, D. Hogg, C. H. Ottensmeier, C. Lebbé, C. Peschel, I. Quidt, J. I. Clark, J. D. Wolchok, J. S. Weber, J. Tian, M. J. Yellin, G. M. Nichol, A. Hoos, W. J. Urba, Improved survival with ipilimumab in patients with metastatic melanoma. *N. Engl. J. Med.* **363**, 711–723 (2010). [doi:10.1056/NEJMoa1003466](https://doi.org/10.1056/NEJMoa1003466) [Medline](#)
3. C. Robert, G. V. Long, B. Brady, C. Dutriaux, M. Maio, L. Mortier, J. C. Hassel, P. Rutkowski, C. McNeil, E. Kalinka-Warzocho, K. J. Savage, M. M. Hernberg, C. Lebbé, J. Charles, C. Mihalciou, V. Chiarion-Sileni, C. Mauch, F. Cognetti, A. Arance, H. Schmidt, D. Schadendorf, H. Gogas, L. Lundgren-Eriksson, C. Horak, B. Sharkey, I. M. Waxman, V. Atkinson, P. A. Ascierto, Nivolumab in previously untreated melanoma without BRAF mutation. *N. Engl. J. Med.* **372**, 320–330 (2015). [doi:10.1056/NEJMoa1412082](https://doi.org/10.1056/NEJMoa1412082) [Medline](#)
4. J. M. Zaretsky, A. Garcia-Diaz, D. S. Shin, H. Escuin-Ordinas, W. Hugo, S. Hu-Lieskovan, D. Y. Torrejon, G. Abril-Rodriguez, S. Sandoval, L. Barthly, J. Saco, B. Homet Moreno, R. Mezzadra, B. Chmielowski, K. Ruchalski, I. P. Shintaku, P. J. Sanchez, C. Puig-Saus, G. Cherry, E. Seja, X. Kong, J. Pang, B. Berent-Maoz, B. Comin-Anduix, T. G. Graeber, P. C. Tumeh, T. N. M. Schumacher, R. S. Lo, A. Ribas, Mutations Associated with Acquired Resistance to PD-1 Blockade in Melanoma. *N. Engl. J. Med.* **375**, 819–829 (2016). [doi:10.1056/NEJMoa1604958](https://doi.org/10.1056/NEJMoa1604958) [Medline](#)
5. J. Gao, L. Z. Shi, H. Zhao, J. Chen, L. Xiong, Q. He, T. Chen, J. Roszik, C. Bernatchez, S. E. Woodman, P.-L. Chen, P. Hwu, J. P. Allison, A. Futreal, J. A. Wargo, P. Sharma, Loss of IFN- $\gamma$  Pathway Genes in Tumor Cells as a Mechanism of Resistance to Anti-CTLA-4 Therapy. *Cell* **167**, 397–404.e9 (2016). [doi:10.1016/j.cell.2016.08.069](https://doi.org/10.1016/j.cell.2016.08.069) [Medline](#)
6. F. Zhao, A. Sucker, S. Horn, C. Heeke, N. Bielefeld, B. Schrörs, A. Bicker, M. Lindemann, A. Roesch, G. Gaudernack, M. Stiller, J. C. Becker, V. Lennerz, T. Wölfel, D. Schadendorf, K. Griewank, A. Paschen, Melanoma Lesions Independently Acquire T-cell Resistance during Metastatic Latency. *Cancer Res.* **76**, 4347–4358 (2016). [doi:10.1158/0008-5472.CAN-16-0008](https://doi.org/10.1158/0008-5472.CAN-16-0008) [Medline](#)
7. W. Hugo, J. M. Zaretsky, L. Sun, C. Song, B. H. Moreno, S. Hu-Lieskovan, B. Berent-Maoz, J. Pang, B. Chmielowski, G. Cherry, E. Seja, S. Lomeli, X. Kong, M. C. Kelley, J. A. Sosman, D. B. Johnson, A. Ribas, R. S. Lo, Genomic and Transcriptomic Features of Response to Anti-PD-1 Therapy in Metastatic Melanoma. *Cell* **165**, 35–44 (2016). [doi:10.1016/j.cell.2016.02.065](https://doi.org/10.1016/j.cell.2016.02.065) [Medline](#)
8. D. T. Le, J. N. Uram, H. Wang, B. R. Bartlett, H. Kemberling, A. D. Eyring, A. D. Skora, B. S. Luber, N. S. Azad, D. Laheru, B. Biedrzycki, R. C. Donehower, A. Zaheer, G. A. Fisher, T. S. Crocenzi, J. J. Lee, S. M. Duffy, R. M. Goldberg, A. de la Chapelle, M. Koshiji, F. Bhajee, T. Huebner, R. H. Hruban, L. D. Wood, N. Cuka, D. M. Pardoll, N. Papadopoulos, K. W. Kinzler, S. Zhou, T. C. Cornish, J. M. Taube, R. A. Anders, J. R. Eshleman, B. Vogelstein, L. A. Diaz Jr., PD-1 Blockade in Tumors with Mismatch-Repair Deficiency. *N. Engl. J. Med.* **372**, 2509–2520 (2015). [doi:10.1056/NEJMoa1500596](https://doi.org/10.1056/NEJMoa1500596) [Medline](#)
9. N. Riaz, J. J. Havel, S. M. Kendall, V. Makarov, L. A. Walsh, A. Desrichard, N. Weinhold, T. A. Chan, Recurrent SERPINB3 and SERPINB4 mutations in patients who respond to anti-CTLA4 immunotherapy. *Nat. Genet.* **48**, 1327–1329 (2016). [doi:10.1038/ng.3677](https://doi.org/10.1038/ng.3677) [Medline](#)
10. N. A. Rizvi, M. D. Hellmann, A. Snyder, P. Kvistborg, V. Makarov, J. J. Havel, W. Lee, J. Yuan, P. Wong, T. S. Ho, M. L. Miller, N. Rekhtman, A. L. Moreira, F. Ibrahim, C. Bruggeman, B. Gasmir, R. Zappasodi, Y. Maeda, C. Sander, E. B. Garon, T. Merghoub, J. D. Wolchok, T. N. Schumacher, T. A. Chan, Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. *Science* **348**, 124–128 (2015). [doi:10.1126/science.1257451](https://doi.org/10.1126/science.1257451) [Medline](#)
11. A. Snyder, V. Makarov, T. Merghoub, J. Yuan, J. M. Zaretsky, A. Desrichard, L. A. Walsh, M. A. Postow, P. Wong, T. S. Ho, T. J. Hollmann, C. Bruggeman, K. Kannan, Y. Li, C. Elipenahli, C. Liu, C. T. Harbison, L. Wang, A. Ribas, J. D. Wolchok, T. A. Chan, Genetic basis for clinical response to CTLA-4 blockade in melanoma. *N. Engl. J. Med.* **371**, 2189–2199 (2014). [doi:10.1056/NEJMoa1406498](https://doi.org/10.1056/NEJMoa1406498) [Medline](#)
12. E. M. Van Allen, D. Miao, B. Schilling, S. A. Shukla, C. Blank, L. Zimmer, A. Sucker, U. Hillen, M. H. G. Foppen, S. M. Goldinger, J. Utikal, J. C. Hassel, B. Weide, K. C. Kaehler, C. Loguaj, P. Mohr, R. Gutzmer, R. Dummer, S. Gabriel, C. J. Wu, D. Schadendorf, L. A. Garraway, Genomic correlates of response to CTLA-4 blockade in metastatic melanoma. *Science* **350**, 207–211 (2015). [doi:10.1126/science.1257451](https://doi.org/10.1126/science.1257451) [Medline](#)
13. N. McGranahan, A. J. S. Furness, R. Rosenthal, S. Ramskov, R. Lyngaa, S. K. Saini, M. Jamal-Hanjani, G. A. Wilson, N. J. Birkbak, C. T. Hiley, T. B. K. Watkins, S. Shafi, N. Murugaesu, R. Mitter, A. U. Akarca, J. Linares, T. Marafioti, J. Y. Henry, E. M. Van Allen, D. Miao, B. Schilling, D. Schadendorf, L. A. Garraway, V. Makarov, N. A. Rizvi, A. Snyder, M. D. Hellmann, T. Merghoub, J. D. Wolchok, S. A. Shukla, C. J. Wu, K. S. Peggs, T. A. Chan, S. R. Hadrup, S. A. Quezada, C. Swanton, Clonal neoantigens elicit T cell immunoreactivity and sensitivity to immune checkpoint blockade. *Science* **351**, 1463–1469 (2016). [doi:10.1126/science.1257451](https://doi.org/10.1126/science.1257451) [Medline](#)
14. R. S. Herbst, J.-C. Soria, M. Kowanetz, G. D. Fine, O. Hamid, M. S. Gordon, J. A. Sosman, D. F. McDermott, J. D. Powderly, S. N. Gettinger, H. E. K. Kohrt, L. Horn, D. P. Lawrence, S. Rost, M. Leabman, Y. Xiao, A. Moktrin, H. Koepfen, P. S. Hegde, I. Mellman, D. S. Chen, F. S. Hodi, Predictive correlates of response to the anti-PD-L1 antibody MPDL3280A in cancer patients. *Nature* **515**, 563–567 (2014). [doi:10.1038/nature14011](https://doi.org/10.1038/nature14011) [Medline](#)
15. P. C. Tumeh, C. L. Harview, J. H. Yearley, I. P. Shintaku, E. J. M. Taylor, L. Robert, B. Chmielowski, M. Spasic, G. Henry, V. Ciobanu, A. N. West, M. Carmona, C. Kivork, E. Seja, G. Cherry, A. J. Gutierrez, T. R. Grogan, C. Mateus, G. Tamasic, J. A. Glaspy, R. O. Emerson, H. Robins, R. H. Pierce, D. A. Elashoff, C. Robert, A. Ribas, PD-1 blockade induces responses by inhibiting adaptive immune resistance. *Nature* **515**, 568–571 (2014). [doi:10.1038/nature13954](https://doi.org/10.1038/nature13954) [Medline](#)
16. S. L. Topalian, J. M. Taube, R. A. Anders, D. M. Pardoll, Mechanism-driven biomarkers to guide immune checkpoint blockade in cancer therapy. *Nat. Rev. Cancer* **16**, 275–287 (2016). [doi:10.1038/nrc.2016.36](https://doi.org/10.1038/nrc.2016.36) [Medline](#)
17. N. Riaz, J. J. Havel, V. Makarov, A. Desrichard, W. J. Urba, J. S. Sims, F. S. Hodi, S. Martín-Algarra, R. Mandal, W. H. Sharfman, S. Bhatia, W.-J. Hwu, T. F. Gajewski, C. L. Slingluff Jr., D. Chowell, S. M. Kendall, H. Chang, R. Shah, F. Kuo, L. G. T. Morris, J.-W. Sidhom, J. P. Schneck, C. E. Horak, N. Weinhold, T. A. Chan, Tumor and Microenvironment Evolution during Immunotherapy with Nivolumab. *Cell* **171**, 934–949.e15 (2017). [doi:10.1016/j.cell.2017.09.028](https://doi.org/10.1016/j.cell.2017.09.028) [Medline](#)
18. T. Davoli, H. Uno, E. C. Wooten, S. J. Elledge, Tumor aneuploidy correlates with markers of immune evasion and with reduced response to immunotherapy. *Science* **355**, eaaf8399 (2017). [doi:10.1126/science.1257451](https://doi.org/10.1126/science.1257451) [Medline](#)
19. W. Roh, P.-L. Chen, A. Reuben, C. N. Spencer, P. A. Prieto, J. P. Miller, V. Gopalakrishnan, F. Wang, Z. A. Cooper, S. M. Reddy, C. Gumbs, L. Little, Q. Chang,

- W.-S. Chen, K. Wani, M. P. De Macedo, E. Chen, J. L. Austin-Breneman, H. Jiang, J. Roszik, M. T. Tetzlaff, M. A. Davies, J. E. Gershenwald, H. Tawbi, A. J. Lazar, P. Hwu, W.-J. Hwu, A. Diab, I. C. Glitza, S. P. Patel, S. E. Woodman, R. N. Amaria, V. G. Prieto, J. Hu, P. Sharma, J. P. Allison, L. Chin, J. Zhang, J. A. Wargo, P. A. Futreal, Integrated molecular analysis of tumor biopsies on sequential CTLA-4 and PD-1 blockade reveals markers of response and resistance. *Sci. Transl. Med.* **9**, eaah3560 (2017). [doi:10.1126/scitranslmed.aah3560](https://doi.org/10.1126/scitranslmed.aah3560) [Medline](#)
20. V. Gopalakrishnan, C. N. Spencer, L. Nezi, A. Reuben, M. C. Andrews, T. V. Karpinets, P. A. Prieto, D. Vicente, K. Hoffman, S. C. Wei, A. P. Cogdill, L. Zhao, C. W. Hudgens, D. S. Hutchinson, T. Manzo, M. Petaccia de Macedo, T. Cotechini, T. Kumar, W. S. Chen, S. M. Reddy, R. S. Sloane, J. Galloway-Pena, H. Jiang, P. L. Chen, E. J. Shpall, K. Rezvani, A. M. Alousi, R. F. Chemaly, S. Shelburne, L. M. Vence, P. C. Okhuysen, V. B. Jensen, A. G. Swennes, F. McAllister, E. M. R. Sanchez, Y. Zhang, E. Le Chatelier, L. Zitvogel, N. Pons, J. L. Austin-Breneman, L. E. Haydu, E. M. Burton, J. M. Gardner, E. Sirmans, J. Hu, A. J. Lazar, T. Tsujikawa, A. Diab, H. Tawbi, I. C. Glitza, W. J. Hwu, S. P. Patel, S. E. Woodman, R. N. Amaria, M. A. Davies, J. E. Gershenwald, P. Hwu, J. E. Lee, J. Zhang, L. M. Coussens, Z. A. Cooper, P. A. Futreal, C. R. Daniel, N. J. Ajami, J. F. Petrosino, M. T. Tetzlaff, P. Sharma, J. P. Allison, R. R. Jenq, J. A. Wargo, Gut microbiome modulates response to anti-PD-1 immunotherapy in melanoma patients. *Science* eaan4236 (2017). [doi:10.1126/science.aan4236](https://doi.org/10.1126/science.aan4236) [Medline](#)
  21. B. Routy, E. Le Chatelier, L. Derosa, C. P. M. Duong, M. T. Alou, R. Daillère, A. Fluckiger, M. Messaoudene, C. Rauber, M. P. Roberti, M. Fidelle, C. Flament, V. Poirier-Colame, P. Opolon, C. Klein, K. Iribarren, L. Mondragón, N. Jacquilot, B. Qu, G. Ferrere, C. Clémenson, L. Mezquita, J. R. Masip, C. Naltet, S. Brosseau, C. Kaderbhai, C. Richard, H. Rizvi, F. Levenez, N. Galleron, B. Quinquis, N. Pons, B. Ryffel, V. Minard-Colin, P. Gonin, J.-C. Soria, E. Deutsch, Y. Loriot, F. Ghiringhelli, G. Zalcman, F. Goldwasser, B. Escudier, M. D. Hellmann, A. Eggermont, D. Raoult, L. Albiges, G. Kroemer, L. Zitvogel, Gut microbiome influences efficacy of PD-1-based immunotherapy against epithelial tumors. *Science* eaan3706 (2017). [doi:10.1126/science.aan3706](https://doi.org/10.1126/science.aan3706) [Medline](#)
  22. F. Pereyra, X. Jia, P. J. McLaren, A. Telenti, P. I. de Bakker, B. D. Walker, S. Ripke, C. J. Brumme, S. L. Pulit, M. Carrington, C. M. Kadie, J. M. Carlson, D. Heckerman, R. R. Graham, R. M. Plenge, S. G. Deeks, L. Gianniny, G. Crawford, J. Sullivan, E. Gonzalez, L. Davies, A. Camargo, J. M. Moore, N. Beattie, S. Gupta, A. Crenshaw, N. P. Burt, C. Guiducci, N. Gupta, X. Gao, Y. Qi, Y. Yuki, A. Piechocka-Trocha, E. Cutrell, R. Rosenberg, K. L. Moss, P. Lemay, J. O'Leary, T. Schaefer, P. Verma, I. Toth, B. Block, B. Baker, A. Rothchild, J. Lian, J. Proudfoot, D. M. Alvino, S. Vine, M. M. Addo, T. M. Allen, M. Altfeld, M. R. Henn, S. Le Gall, H. Streeck, D. W. Haas, D. R. Kuritzkes, G. K. Robbins, R. W. Shafer, R. M. Gulick, C. M. Shikuma, R. Haubrich, S. Riddler, P. E. Sax, E. S. Daar, H. J. Ribaud, B. Agan, S. Agarwal, R. L. Ahern, B. L. Allen, S. Altidor, E. L. Altschuler, S. Ambardar, K. Anastos, B. Anderson, V. Anderson, U. Andrad, D. Antoniskis, D. Bangsberg, D. Barbaro, W. Barrie, J. Bartczak, S. Barton, P. Basden, N. Basgoz, S. Bazner, N. C. Bellos, A. M. Benson, J. Berger, N. F. Bernard, A. M. Bernard, C. Birch, S. J. Bodner, R. K. Bolan, E. T. Boudreaux, M. Bradley, J. F. Braun, J. E. Brndjar, S. J. Brown, R. Brown, S. T. Brown, J. Burack, L. M. Bush, V. Cafaro, O. Campbell, J. Campbell, R. H. Carlson, J. K. Carmichael, K. K. Casey, C. Cavacuiti, G. Celestin, S. T. Chambers, N. Chez, L. M. Chirch, P. J. Cimoch, D. Cohen, L. E. Cohn, B. Conway, D. A. Cooper, B. Cornelson, D. T. Cox, M. V. Cristofano, G. Cuchural Jr., J. L. Czartoski, J. M. Dahman, J. S. Daly, B. T. Davis, K. Davis, S. M. Davod, E. DeJesus, C. A. Dietz, E. Dunham, M. E. Dunn, T. B. Ellerlin, J. J. Eron, J. J. Fangman, C. E. Farel, H. Ferlazzo, S. Fidler, A. Fleener-Ford, R. Frankel, K. A. Freedberg, N. K. French, J. D. Fuchs, J. D. Fuller, J. Gaberman, J. E. Gallant, R. T. Gandhi, E. Garcia, D. Garmon, J. C. Gathe Jr., C. R. Gaultier, W. Gebre, F. D. Gilman, I. Gilson, P. A. Goepfert, M. S. Gottlieb, C. Goulston, R. K. Groger, T. D. Gurley, S. Haber, R. Hardwicke, W. D. Hardy, P. R. Harrigan, T. N. Hawkins, S. Heath, F. M. Hecht, W. K. Henry, M. Hladek, R. P. Hoffman, J. M. Horton, R. K. Hsu, G. D. Huhn, P. Hunt, M. J. Hupert, M. L. Illeman, H. Jaeger, R. M. Jellinger, M. John, J. A. Johnson, K. L. Johnson, H. Johnson, K. Johnson, J. Joly, W. C. Jordan, C. A. Kauffman, H. Khanlou, R. K. Killian, A. Y. Kim, D. D. Kim, C. A. Kinder, J. T. Kirchner, L. Kogelman, E. M. Kojic, P. T. Korthuis, W. Kurisu, D. S. Kwon, M. LaMar, H. Lampiris, M. Lanzafame, M. M. Lederman, D. M. Lee, J. M. Lee, M. J. Lee, E. T. Lee, J. Lemoine, J. A. Levy, J. M. Llibre, M. A. Liguori, S. J. Little, A. Y. Liu, A. J. Lopez, M. R. Loutfy, D. Loy, D. Y. Mohammed, A. Man, M. K. Mansour, V. C. Marconi, M. Markowitz, R. Marques, J. N. Martin, H. L. Martin Jr., K. H. Mayer, M. J. McElrath, T. A. McGhee, B. H. McGovern, K. McGowan, D. McIntyre, G. X. Mcleod, P. Menezes, G. Mesa, C. E. Metroka, D. Meyer-Olson, A. O. Miller, K. Montgomery, K. C. Mounzer, E. H. Nagami, I. Nagin, R. G. Nahass, M. O. Nelson, C. Nielsen, D. L. Norene, D. H. O'Connor, B. O. Ojikutu, J. Okulicz, O. O. Oladehin, E. C. Oldfield 3rd, S. A. Olender, M. Ostrowski, W. F. Owen Jr., E. Pae, J. Parsonnet, A. M. Pavlatos, A. M. Perlmutter, M. N. Pierce, J. M. Pincus, L. Pisani, L. J. Price, L. Proia, R. C. Prokesch, H. C. Pujet, M. Ramgopal, A. Rathod, M. Rausch, J. Ravishankar, F. S. Rhome, C. S. Richards, D. D. Richman, B. Rodes, M. Rodriguez, R. C. Rose 3rd, E. S. Rosenberg, D. Rosenthal, P. E. Ross, D. S. Rubin, E. Rumbaugh, L. Saenz, M. R. Salvaggio, W. C. Sanchez, V. M. Sanjana, S. Santiago, W. Schmidt, H. Schuitemaker, P. M. Sestak, P. Shalit, W. Shay, V. N. Shirvani, V. I. Silebi, J. M. Sizemore Jr., P. R. Skolnik, M. Sokol-Anderson, J. M. Sosman, P. Stabile, J. T. Stapleton, S. Starrett, F. Stein, H. J. Stellbrink, F. L. Sterman, V. E. Stone, D. R. Stone, G. Tambussi, R. A. Taplitz, E. M. Tedaldi, A. Telenti, W. Theisen, R. Torres, L. Tosiello, C. Tremblay, M. A. Tribble, P. D. Trinh, A. Tsao, P. Ueda, A. Vaccaro, E. Valadas, T. J. Vanig, I. Vecino, V. M. Vega, W. Veikley, B. H. Wade, C. Walworth, C. Wanidworanun, D. J. Ward, D. A. Warner, R. D. Weber, D. Webster, S. Weis, D. A. Wheeler, D. J. White, E. Wilkins, A. Winston, C. G. Wlodaver, A. van't Wout, D. P. Wright, O. O. Yang, D. L. Yurdin, B. W. Zabukovic, K. C. Zachary, B. Zeeman, M. Zhao; International HIV Controllers Study, The major genetic determinants of HIV-1 control affect HLA class I peptide presentation. *Science* **330**, 1551–1557 (2010). [doi:10.1126/science.1195271](https://doi.org/10.1126/science.1195271) [Medline](#)
  23. M. Carrington, G. W. Nelson, M. P. Martin, T. Kissner, D. Vlahov, J. J. Goedert, R. Kaslow, S. Buchbinder, K. Hoots, S. J. O'Brien, HLA and HIV-1: Heterozygote advantage and B\*35-Cw\*04 disadvantage. *Science* **283**, 1748–1752 (1999). [doi:10.1126/science.283.5408.1748](https://doi.org/10.1126/science.283.5408.1748) [Medline](#)
  24. J. Fellay, K. V. Shianna, D. Ge, S. Colombo, B. Ledergerber, M. Weale, K. Zhang, C. Gumbs, A. Castagna, A. Cossarizza, A. Cozzi-Lepri, A. De Luca, P. Easterbrook, P. Francioli, S. Mallal, J. Martinez-Picado, J. M. Miro, N. Obel, J. P. Smith, J. Wyniger, P. Descombes, S. E. Antonarakis, N. L. Letvin, A. J. McMichael, B. F. Haynes, A. Telenti, D. B. Goldstein, A whole-genome association study of major determinants for host control of HIV-1. *Science* **317**, 944–947 (2007). [doi:10.1126/science.1143767](https://doi.org/10.1126/science.1143767) [Medline](#)
  25. X. Gao, G. W. Nelson, P. Karacki, M. P. Martin, J. Phair, R. Kaslow, J. J. Goedert, S. Buchbinder, K. Hoots, D. Vlahov, S. J. O'Brien, M. Carrington, Effect of a single amino acid change in MHC class I molecules on the rate of progression to AIDS. *N. Engl. J. Med.* **344**, 1668–1675 (2001). [doi:10.1056/NEJM200105313442203](https://doi.org/10.1056/NEJM200105313442203) [Medline](#)
  26. E. Trachtenberg, B. Korber, C. Sollars, T. B. Kepler, P. T. Hraber, E. Hayes, R. Funkhouser, M. Fugate, J. Theiler, Y. S. Hsu, K. Kunstman, S. Wu, J. Phair, H. Erlich, S. Wolinsky, Advantage of rare HLA supertype in HIV disease progression. *Nat. Med.* **9**, 928–935 (2003). [doi:10.1038/nm893](https://doi.org/10.1038/nm893) [Medline](#)
  27. P. J. R. Goulder, B. D. Walker, HIV and HLA class I: An evolving relationship. *Immunity* **37**, 426–440 (2012). [doi:10.1016/j.immuni.2012.09.005](https://doi.org/10.1016/j.immuni.2012.09.005) [Medline](#)
  28. A. V. S. Hill, C. E. M. Allsopp, D. Kwiatkowski, N. M. Anstey, P. Twumasi, P. A. Rowe, S. Bennett, D. Brewster, A. J. McMichael, B. M. Greenwood, Common west African HLA antigens are associated with protection from severe malaria. *Nature* **352**, 595–600 (1991). [doi:10.1038/352595a0](https://doi.org/10.1038/352595a0) [Medline](#)
  29. P. Kiepiela, A. J. Leslie, I. Honeyborne, D. Ramduth, C. Thobakgale, S. Chetty, P. Rathnavalu, C. Moore, K. J. Pfaffert, L. Hilton, P. Zimbwa, S. Moore, T. Allen, C. Brander, M. M. Addo, M. Altfeld, I. James, S. Mallal, M. Bunce, L. D. Barber, J. Slinger, C. Day, P. Klenerman, J. Mullins, B. Korber, H. M. Coovadia, B. D. Walker, P. J. R. Goulder, Dominant influence of HLA-B in mediating the potential co-evolution of HIV and HLA. *Nature* **432**, 769–775 (2004). [doi:10.1038/nature03113](https://doi.org/10.1038/nature03113) [Medline](#)
  30. T. J. T. Pi, *HLA and disease associations*. (Springer Science & Business Media, 2012).
  31. D. Chowell, S. Krishna, P. D. Becker, C. Cocita, J. Shu, X. Tan, P. D. Greenberg, L. S. Klavinskis, J. N. Blattman, K. S. Anderson, TCR contact residue hydrophobicity is a hallmark of immunogenic CD8+ T cell epitopes. *Proc. Natl. Acad. Sci. U.S.A.* **112**, E1754–E1762 (2015). [doi:10.1073/pnas.1500973112](https://doi.org/10.1073/pnas.1500973112) [Medline](#)
  32. P. C. Doherty, R. M. Zinkernagel, A biological role for the major histocompatibility antigens. *Lancet* **1**, 1406–1409 (1975). [doi:10.1016/S0140-6736\(75\)92610-0](https://doi.org/10.1016/S0140-6736(75)92610-0) [Medline](#)

33. P. Parham, T. Ohta, Population biology of antigen presentation by MHC class I molecules. *Science* **272**, 67–74 (1996). [doi:10.1126/science.272.5258.67](https://doi.org/10.1126/science.272.5258.67) Medline
34. M. M. Gubin, X. Zhang, H. Schuster, E. Caron, J. P. Ward, T. Noguchi, Y. Ivanova, J. Hundal, C. D. Arthur, W.-J. Kriebler, G. E. Mulder, M. Toebes, M. D. Vesely, S. S. K. Lam, A. J. Korman, J. P. Allison, G. J. Freeman, A. H. Sharpe, E. L. Pearce, T. N. Schumacher, R. Aebbersold, H.-G. Rammensee, C. J. M. Melief, E. R. Mardis, W. E. Gillanders, M. N. Artyomov, R. D. Schreiber, Checkpoint blockade cancer immunotherapy targets tumour-specific mutant antigens. *Nature* **515**, 577–581 (2014). [doi:10.1038/nature13988](https://doi.org/10.1038/nature13988) Medline
35. E. Tran, M. Ahmadzadeh, Y.-C. Lu, A. Gros, S. Turcotte, P. F. Robbins, J. J. Gartner, Z. Zheng, Y. F. Li, S. Ray, J. R. Wunderlich, R. P. Somerville, S. A. Rosenberg, Immunogenicity of somatic mutations in human gastrointestinal cancers. *Science* **350**, 1387–1390 (2015). [doi:10.1126/science.1253](https://doi.org/10.1126/science.1253) Medline
36. E. Tran, P. F. Robbins, Y.-C. Lu, T. D. Prickett, J. J. Gartner, L. Jia, A. Pasetto, Z. Zheng, S. Ray, E. M. Groh, I. R. Kriley, S. A. Rosenberg, T-Cell Transfer Therapy Targeting Mutant KRAS in Cancer. *N. Engl. J. Med.* **375**, 2255–2262 (2016). [doi:10.1056/NEJMoa1609279](https://doi.org/10.1056/NEJMoa1609279) Medline
37. A. Zehir, R. Benayed, R. H. Shah, A. Syed, S. Middha, H. R. Kim, P. Srinivasan, J. Gao, D. Chakravarty, S. M. Devlin, M. D. Hellmann, D. A. Barron, A. M. Schram, M. Hameed, S. Dogan, D. S. Ross, J. F. Hechtman, D. F. Delair, J. Yao, D. L. Mandelker, D. T. Cheng, R. Chandramohan, A. S. Mohanty, R. N. Ptashkin, G. Jayakumar, M. Prasad, M. H. Syed, A. B. Rema, Z. Y. Liu, K. Nafa, L. Borsu, J. Sadowska, J. Casanova, R. Bacares, I. J. Kiecka, A. Razumova, J. B. Son, L. Stewart, T. Baldi, K. A. Mullaney, H. Al-Ahmadie, E. Vakiani, A. A. Abeshouse, A. V. Penson, P. Jonsson, N. Camacho, M. T. Chang, H. H. Won, B. E. Gross, R. Kundra, Z. J. Heins, H.-W. Chen, S. Phillips, H. Zhang, J. Wang, A. Ochoa, J. Wills, M. Eubank, S. B. Thomas, S. M. Gardos, D. N. Reales, J. Galle, R. Durany, R. Cambria, W. Abida, A. Cercek, D. R. Feldman, M. M. Gounder, A. A. Hakimi, J. J. Harding, G. Iyer, Y. Y. Janjigian, E. J. Jordan, C. M. Kelly, M. A. Lowery, L. G. T. Morris, A. M. Omuro, N. Raj, P. Razavi, A. N. Shoushtari, N. Shukla, T. E. Soumerai, A. M. Varghese, R. Yaeger, J. Coleman, B. Bochner, G. J. Riely, L. B. Saltz, H. I. Scher, P. J. Sabbatini, M. E. Robson, D. S. Klimstra, B. S. Taylor, J. Baselga, N. Schultz, D. M. Hyman, M. E. Arcila, D. B. Solit, M. Ladanyi, M. F. Berger, Mutational landscape of metastatic cancer revealed from prospective clinical sequencing of 10,000 patients. *Nat. Med.* **23**, 703–713 (2017). [doi:10.1038/nm.4333](https://doi.org/10.1038/nm.4333) Medline
38. D. Snary, C. J. Barnstable, W. F. Bodmer, M. J. Crumpton, Molecular structure of human histocompatibility antigens: The HLA-C series. *Eur. J. Immunol.* **7**, 580–585 (1977). [doi:10.1002/eji.1830070816](https://doi.org/10.1002/eji.1830070816) Medline
39. S. G. Marsh, P. Parham, L. D. Barber, *The HLA factsbook*. (Academic Press, 1999).
40. M. R. Schaefer, M. Williams, D. A. Kulpa, P. K. Blakely, A. Q. Yaffee, K. L. Collins, A novel trafficking signal within the HLA-C cytoplasmic tail allows regulated expression upon differentiation of macrophages. *J. Immunol.* **180**, 7804–7817 (2008). [doi:10.4049/jimmunol.180.12.7804](https://doi.org/10.4049/jimmunol.180.12.7804) Medline
41. D. S. Chen, I. Mellman, Elements of cancer immunity and the cancer-immune set point. *Nature* **541**, 321–330 (2017). [doi:10.1038/nature21349](https://doi.org/10.1038/nature21349) Medline
42. T. N. Schumacher, R. D. Schreiber, Neoantigens in cancer immunotherapy. *Science* **348**, 69–74 (2015). [doi:10.1126/science.1253](https://doi.org/10.1126/science.1253) Medline
43. N. Aptsiauri, T. Cabrera, A. Garcia-Lora, M. A. Lopez-Nevot, F. Ruiz-Cabello, F. Garrido, MHC class I antigens and immune surveillance in transformed cells. *Int. Rev. Cytol.* **256**, 139–189 (2007). [doi:10.1016/S0074-7696\(07\)56005-5](https://doi.org/10.1016/S0074-7696(07)56005-5) Medline
44. N. McGranahan, R. Rosenthal, C. T. Hiley, A. J. Rowan, T. B. K. Watkins, G. A. Wilson, N. J. Birkbak, S. Veeriah, P. Van Loo, J. Herrero, C. Swanton, TRACERx Consortium, Allele-Specific HLA Loss and Immune Escape in Lung Cancer Evolution. *Cell* **171**, 1259–1271.e11 (2017). [doi:10.1016/j.cell.2017.10.001](https://doi.org/10.1016/j.cell.2017.10.001) Medline
45. N. van Rooij, M. M. van Buuren, D. Philips, A. Velds, M. Toebes, B. Heemskerk, L. J. A. van Dijk, S. Behjati, H. Hilkmann, D. El Atmioui, M. Nieuwland, M. R. Stratton, R. M. Kerkhoven, C. Kegmir, J. B. Haanen, P. Kvistborg, T. N. Schumacher, Tumor exome analysis reveals neoantigen-specific T-cell reactivity in an ipilimumab-responsive melanoma. *J. Clin. Oncol.* **31**, e439–e442 (2013). [doi:10.1200/JCO.2012.47.7521](https://doi.org/10.1200/JCO.2012.47.7521) Medline
46. M. Yadav, S. Jhunjunwala, Q. T. Phung, P. Lupardus, J. Tanguay, S. Bumbaca, C. Franci, T. K. Cheung, J. Fritsch, T. Weinschenck, Z. Modrusan, I. Mellman, J. R. Lill, L. Delamarre, Predicting immunogenic tumour mutations by combining mass spectrometry and exome sequencing. *Nature* **515**, 572–576 (2014). [doi:10.1038/nature14001](https://doi.org/10.1038/nature14001) Medline
47. E. W. Petersdorf, M. Malkki, C. O'hUigin, M. Carrington, T. Gooley, M. D. Haegenson, M. M. Horowitz, S. R. Spellman, T. Wang, P. Stevenson, High HLA-DP Expression and Graft-versus-Host Disease. *N. Engl. J. Med.* **373**, 599–609 (2015). [doi:10.1056/NEJMoa1500140](https://doi.org/10.1056/NEJMoa1500140) Medline
48. C. S. Carlson, R. O. Emerson, A. M. Sherwood, C. Desmarais, M.-W. Chung, J. M. Parsons, M. S. Steen, M. A. LaMadrid-Herrmannsfeldt, D. W. Williamson, R. J. Livingston, D. Wu, B. L. Wood, M. J. Rieder, H. Robins, Using synthetic templates to design an unbiased multiplex PCR assay. *Nat. Commun.* **4**, 2680 (2013). [doi:10.1038/ncomms3680](https://doi.org/10.1038/ncomms3680) Medline
49. H. S. Robins, P. V. Campregher, S. K. Srivastava, A. Wachter, C. J. Turtle, O. Khsai, S. R. Riddell, E. H. Warren, C. S. Carlson, Comprehensive assessment of T-cell receptor  $\beta$ -chain diversity in alphabeta T cells. *Blood* **114**, 4099–4107 (2009). [doi:10.1182/blood-2009-04-217604](https://doi.org/10.1182/blood-2009-04-217604) Medline
50. J. S. Sims, B. Grinshpun, Y. Feng, T. H. Ung, J. A. Neira, J. L. Samanamud, P. Canoll, Y. Shen, P. A. Sims, J. N. Bruce, Diversity and divergence of the glioma-infiltrating T-cell receptor repertoire. *Proc. Natl. Acad. Sci. U.S.A.* **113**, E3529–E3537 (2016). [doi:10.1073/pnas.1601012113](https://doi.org/10.1073/pnas.1601012113) Medline
51. B. Li, T. Li, J.-C. Pignon, B. Wang, J. Wang, S. A. Shukla, R. Dou, Q. Chen, F. S. Hodi, T. K. Choueiri, C. Wu, N. Hacohen, S. Signoretti, J. S. Liu, X. S. Liu, Landscape of tumor-infiltrating T cell repertoire of human cancers. *Nat. Genet.* **48**, 725–732 (2016). [doi:10.1038/ng.3581](https://doi.org/10.1038/ng.3581) Medline
52. A. Sette, J. Sidney, Nine major HLA class I supertypes account for the vast preponderance of HLA-A and -B polymorphism. *Immunogenetics* **50**, 201–212 (1999). [doi:10.1007/s002510050594](https://doi.org/10.1007/s002510050594) Medline
53. J. Sidney, B. Peters, N. Frahm, C. Brander, A. Sette, HLA class I supertypes: A revised and updated classification. *BMC Immunol.* **9**, 1 (2008). [doi:10.1186/1471-2172-9-1](https://doi.org/10.1186/1471-2172-9-1) Medline
54. N. Hillen, G. Mester, C. Lemmel, A. O. Weinzierl, M. Müller, D. Wernet, J. Hennenlotter, A. Stenzl, H.-G. Rammensee, S. Stevanović, Essential differences in ligand presentation and T cell epitope recognition among HLA molecules of the HLA-B44 supertype. *Eur. J. Immunol.* **38**, 2993–3003 (2008). [doi:10.1002/eji.200838632](https://doi.org/10.1002/eji.200838632) Medline
55. H.-G. Rammensee, T. Friede, S. Stevanović, MHC ligands and peptide motifs: First listing. *Immunogenetics* **41**, 178–228 (1995). [doi:10.1007/BF00172063](https://doi.org/10.1007/BF00172063) Medline
56. M. DiBriano, K. C. Parker, D. H. Margulies, J. Shiloach, R. V. Turner, W. E. Biddison, J. E. Coligan, Identification of the peptide binding motif for HLA-B44, one of the most common HLA-B alleles in the Caucasian population. *Biochemistry* **34**, 10130–10138 (1995). [doi:10.1021/bi00032a005](https://doi.org/10.1021/bi00032a005) Medline
57. P. Liu, X. Huang, R. Zhou, B. J. Berne, Observation of a dewetting transition in the collapse of the melittin tetramer. *Nature* **437**, 159–162 (2005). [doi:10.1038/nature03926](https://doi.org/10.1038/nature03926) Medline
58. Y. Tu, M. Lv, P. Xiu, T. Huynh, M. Zhang, M. Castelli, Z. Liu, Q. Huang, C. Fan, H. Fang, R. Zhou, Destructive extraction of phospholipids from Escherichia coli membranes by graphene nanosheets. *Nat. Nanotechnol.* **8**, 594–601 (2013). [doi:10.1038/nnano.2013.125](https://doi.org/10.1038/nnano.2013.125) Medline
59. R. Zhou, X. Huang, C. J. Margulis, B. J. Berne, Hydrophobic collapse in multidomain protein folding. *Science* **305**, 1605–1609 (2004). [doi:10.1126/science.1101176](https://doi.org/10.1126/science.1101176) Medline
60. S. J. Patel, N. E. Sanjana, R. J. Kishton, A. Eidzadeh, S. K. Vodnala, M. Cam, J. J. Gartner, L. Jia, S. M. Steinberg, T. N. Yamamoto, A. S. Merchant, G. U. Mehta, A. Chichura, O. Shalem, E. Tran, R. Eil, M. Sukumar, E. P. Gujjarro, C.-P. Day, P. Robbins, S. Feldman, G. Merlino, F. Zhang, N. P. Restifo, Identification of essential genes for cancer immunotherapy. *Nature* **548**, 537–542 (2017). [doi:10.1038/nature23477](https://doi.org/10.1038/nature23477) Medline
61. S. A. Shukla, M. S. Rooney, M. Rajasagi, G. Tiao, P. M. Dixon, M. S. Lawrence, J. Stevens, W. J. Lane, J. L. Dellagatta, S. Steelman, C. Sougne, K. Cibulskis, A. Kiezun, N. Hacohen, V. Brusica, C. J. Wu, G. Getz, Comprehensive analysis of cancer-associated somatic mutations in class I HLA genes. *Nat. Biotechnol.* **33**, 1152–1158 (2015). [doi:10.1038/nbt.3344](https://doi.org/10.1038/nbt.3344) Medline
62. K. Kiyotani, T. H. Mai, Y. Nakamura, Comparison of exome-based HLA class I genotyping tools: Identification of platform-specific genotyping errors. *J. Hum. Genet.* **62**, 397–405 (2017). Medline
63. H. Cao, J. Wu, Y. Wang, H. Jiang, T. Zhang, X. Liu, Y. Xu, D. Liang, P. Gao, Y. Sun, B. Gifford, M. D'Ascenzo, X. Liu, L. C. A. M. Tellier, F. Yang, X. Tong, D. Chen, J. Zheng, W. Li, T. Richmond, X. Xu, J. Wang, Y. Li, An integrated tool to study MHC region:

- Accurate SNV detection and HLA genes typing in human MHC region using targeted high-throughput sequencing. *PLOS ONE* **8**, e69388 (2013). [doi:10.1371/journal.pone.0069388](https://doi.org/10.1371/journal.pone.0069388) [Medline](#)
64. A. Szolek, B. Schubert, C. Mohr, M. Sturm, M. Feldhahn, O. Kohlbacher, OptiType: Precision HLA typing from next-generation sequencing data. *Bioinformatics* **30**, 3310–3316 (2014). [doi:10.1093/bioinformatics/btu548](https://doi.org/10.1093/bioinformatics/btu548) [Medline](#)
  65. D. T. Cheng, T. N. Mitchell, A. Zehir, R. H. Shah, R. Benayed, A. Syed, R. Chandramohan, Z. Y. Liu, H. H. Won, S. N. Scott, A. R. Brannon, C. O'Reilly, J. Sadowska, J. Casanova, A. Yannes, J. F. Hechtman, J. Yao, W. Song, D. S. Ross, A. Oultache, S. Dogan, L. Borsu, M. Hameed, K. Nafa, M. E. Arcila, M. Ladanyi, M. F. Berger, Memorial Sloan Kettering-Integrated Mutation Profiling of Actionable Cancer Targets (MSK-IMPACT): A Hybridization Capture-Based Next-Generation Sequencing Clinical Assay for Solid Tumor Molecular Oncology. *J. Mol. Diagn.* **17**, 251–264 (2015). [doi:10.1016/j.jmoldx.2014.12.006](https://doi.org/10.1016/j.jmoldx.2014.12.006) [Medline](#)
  66. D. M. Hyman, D. B. Solit, M. E. Arcila, D. T. Cheng, P. Sabbatini, J. Baselga, M. F. Berger, M. Ladanyi, Precision medicine at Memorial Sloan Kettering Cancer Center: Clinical next-generation sequencing enabling next-generation targeted therapy trials. *Drug Discov. Today* **20**, 1422–1428 (2015). [doi:10.1016/j.drudis.2015.08.005](https://doi.org/10.1016/j.drudis.2015.08.005) [Medline](#)
  67. M. A. DePristo, E. Banks, R. Poplin, K. V. Garimella, J. R. Maguire, C. Hartl, A. A. Philippakis, G. del Angel, M. A. Rivas, M. Hanna, A. McKenna, T. J. Fennell, A. M. Kernysky, A. Y. Sivachenko, K. Cibulskis, S. B. Gabriel, D. Altshuler, M. J. Daly, A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* **43**, 491–498 (2011). [doi:10.1038/ng.806](https://doi.org/10.1038/ng.806) [Medline](#)
  68. H. Li, R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009). [doi:10.1093/bioinformatics/btp324](https://doi.org/10.1093/bioinformatics/btp324) [Medline](#)
  69. A. McKenna, M. Hanna, E. Banks, A. Sivachenko, K. Cibulskis, A. Kernysky, K. Garimella, D. Altshuler, S. Gabriel, M. Daly, M. A. DePristo, The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010). [doi:10.1101/gr.107524.110](https://doi.org/10.1101/gr.107524.110) [Medline](#)
  70. K. Cibulskis, M. S. Lawrence, S. L. Carter, A. Sivachenko, D. Jaffe, C. Sougnez, S. Gabriel, M. Meyerson, E. S. Lander, G. Getz, Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat. Biotechnol.* **31**, 213–219 (2013). [doi:10.1038/nbt.2514](https://doi.org/10.1038/nbt.2514) [Medline](#)
  71. D. C. Koboldt, Q. Zhang, D. E. Larson, D. Shen, M. D. McLellan, L. Lin, C. A. Miller, E. R. Mardis, L. Ding, R. K. Wilson, VarScan 2: Somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res.* **22**, 568–576 (2012). [doi:10.1101/gr.129684.111](https://doi.org/10.1101/gr.129684.111) [Medline](#)
  72. D. E. Larson, C. C. Harris, K. Chen, D. C. Koboldt, T. E. Abbott, D. J. Dooling, T. J. Ley, E. R. Mardis, R. K. Wilson, L. Ding, SomaticSniper: Identification of somatic point mutations in whole genome sequencing data. *Bioinformatics* **28**, 311–317 (2012). [doi:10.1093/bioinformatics/btr665](https://doi.org/10.1093/bioinformatics/btr665) [Medline](#)
  73. C. T. Saunders, W. S. W. Wong, S. Swamy, J. Becq, L. J. Murray, R. K. Cheetham, Strelka: Accurate somatic small-variant calling from sequenced tumor-normal sample pairs. *Bioinformatics* **28**, 1811–1817 (2012). [doi:10.1093/bioinformatics/bts271](https://doi.org/10.1093/bioinformatics/bts271) [Medline](#)
  74. Z. R. Chalmers, C. F. Connelly, D. Fabrizio, L. Gay, S. M. Ali, R. Ennis, A. Schrock, B. Campbell, A. Shlien, J. Chmielecki, F. Huang, Y. He, J. Sun, U. Tabori, M. Kennedy, D. S. Lieber, S. Roels, J. White, G. A. Otto, J. S. Ross, L. Garraway, V. A. Miller, P. J. Stephens, G. M. Frampton, Analysis of 100,000 human cancer genomes reveals the landscape of tumor mutational burden. *Genome Med.* **9**, 34 (2017). [doi:10.1186/s13073-017-0424-2](https://doi.org/10.1186/s13073-017-0424-2) [Medline](#)
  75. R. Shen, V. E. Seshan, FACETS: Allele-specific copy number and clonal heterogeneity analysis tool for high-throughput DNA sequencing. *Nucleic Acids Res.* **44**, e131 (2016). [doi:10.1093/nar/gkw520](https://doi.org/10.1093/nar/gkw520) [Medline](#)
  76. S. Arnaud-Haond, C. M. Duarte, F. Alberto, E. A. Serrão, Standardizing methods to address clonality in population studies. *Mol. Ecol.* **16**, 5115–5139 (2007). [doi:10.1111/j.1365-294X.2007.03535.x](https://doi.org/10.1111/j.1365-294X.2007.03535.x) [Medline](#)
  77. W. Humphrey, A. Dalke, K. Schulten, VMD: Visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38, 27–28 (1996). [doi:10.1016/0263-7855\(96\)00018-5](https://doi.org/10.1016/0263-7855(96)00018-5) [Medline](#)
  78. J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé, K. Schulten, Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26**, 1781–1802 (2005). [doi:10.1002/jcc.20289](https://doi.org/10.1002/jcc.20289) [Medline](#)
  79. M. J. Abraham, Performance enhancements for GROMACS nonbonded interactions on BlueGene. *J. Comput. Chem.* **32**, 2041–2046 (2011). [doi:10.1002/jcc.21766](https://doi.org/10.1002/jcc.21766) [Medline](#)
  80. K. Fleischhauer, D. Avila, F. Vilbois, C. Traversari, C. Bordignon, H.-J. Wallny, Characterization of natural peptide ligands for HLA-B\*4402 and -B\*4403: Implications for peptide involvement in allorecognition of a single amino acid change in the HLA-B44 heavy chain. *Tissue Antigens* **44**, 311–317 (1994). [doi:10.1111/j.1399-0039.1994.tb02401.x](https://doi.org/10.1111/j.1399-0039.1994.tb02401.x) [Medline](#)
  81. J. Herman, P. van der Bruggen, I. F. Luescher, S. Mandruzzato, P. Romero, J. Thonnard, K. Fleischhauer, T. Boon, P. G. Coulie, A peptide encoded by the human MAGE3 gene and presented by HLA-B44 induces cytolytic T lymphocytes that recognize tumor cells expressing MAGE3. *Immunogenetics* **43**, 377–383 (1996). [doi:10.1007/BF02199806](https://doi.org/10.1007/BF02199806) [Medline](#)
  82. V. G. Brichard, J. Herman, A. Van Pel, C. Wildmann, B. Gaugler, T. Wölfel, T. Boon, B. Lethé, A tyrosinase nonapeptide presented by HLA-B44 is recognized on a human melanoma by autologous cytolytic T lymphocytes. *Eur. J. Immunol.* **26**, 224–230 (1996). [doi:10.1002/eji.1830260135](https://doi.org/10.1002/eji.1830260135) [Medline](#)
  83. P. G. Coulie, F. Lehmann, B. Lethé, J. Herman, C. Lurquin, M. Andrawiss, T. Boon, A mutated intron sequence codes for an antigenic peptide recognized by cytolytic T lymphocytes on a human melanoma. *Proc. Natl. Acad. Sci. U.S.A.* **92**, 7976–7980 (1995). [doi:10.1073/pnas.92.17.7976](https://doi.org/10.1073/pnas.92.17.7976) [Medline](#)
  84. R. Chiari, F. Foury, E. De Plaen, J. F. Baurain, J. Thonnard, P. G. Coulie, Two antigens recognized by autologous cytolytic T lymphocytes on a melanoma result from a single point mutation in an essential housekeeping gene. *Cancer Res.* **59**, 5785–5792 (1999). [Medline](#)
  85. N. Vigneron, A. Ooms, S. Morel, G. Degiovanni, B. J. Van Den Eynde, Identification of a new peptide recognized by autologous cytolytic T lymphocytes on a human melanoma. *Cancer Immun.* **2**, 9 (2002). [Medline](#)
  86. M. Nielsen, C. Lundegaard, P. Wornig, S. L. Lauemøller, K. Lamberth, S. Buus, S. Brunak, O. Lund, Reliable prediction of T-cell epitopes using neural networks with novel sequence representations. *Protein Sci.* **12**, 1007–1017 (2003). [doi:10.1110/ps.0239403](https://doi.org/10.1110/ps.0239403) [Medline](#)

## ACKNOWLEDGMENTS

Data reported in this study are tabulated in the main text and supplementary materials. This work was supported by NIH/NCI Cancer Center Support Grant P30 CA008748, NIH K08 DE024774 (L.G.T.M.), Damon Runyon Cancer Research Foundation (L.G.T.M.), Pershing Square Sohn Cancer Research Alliance (T.A.C.), STARR Cancer Consortium (T.A.C.), Stand Up 2 Cancer (T.A.C.), and NIH 1R01CA205426 (N.A.R. and T.A.C.). T.A.C. is a co-founder of Gritstone Oncology. We thank the Chan lab and members of the BMS biomarker team for helpful and stimulating discussions. We thank all the patients for their participation in the clinical trials, which is critical to improving care. We thank members of the melanoma service at MSK, including Drs. J. Wolchok, A. Shoushtari, M. Postow, and M. Callahan for their care of clinical trial patients. We thank members of the Columbia oncology team for their excellent care of trial patients. We thank J. Sims, J. Havel, Y. Shen, and R. Srivastava for very stimulating discussions and helpful suggestions. We thank the Marie-Josée and Henry R. Kravis Center for Molecular Oncology, the Cycle for Survival, and members of the Molecular Pathology & Diagnostics for their help with MSK-IMPACT sequencing. The data are available at the following accession numbers: dbGAP, phs001041.v1.p1; dbGAP, phs000452.v2.p1; SRA, SRP067938 and SRP090294; dbGAP, phs000980.v1.p1; SRA, PRJNA419415, PRJNA419422, and PRJNA419530; cBioPortal for Cancer Genomics, <http://cbioportal.org/msk-impact> (37). The results published here are in part based on data generated by TCGA pilot project established by the National Cancer Institute and National Human Genome Research Institute. Information about TCGA and the investigators and institutions that constitute the TCGA research network can be found at <http://cancergenome.nih.gov/>. T.A.C. receives research funding from Bristol Myers Squibb. N.A.R. is a consultant/advisory board member for AstraZeneca, BMS, Roche, Merck, Novartis, Lilly, and Pfizer. T.A.C. and N.A.R. are cofounders of Gritstone Oncology.

## **SUPPLEMENTARY MATERIALS**

[www.sciencemag.org/cgi/content/full/science.aao4572/DC1](http://www.sciencemag.org/cgi/content/full/science.aao4572/DC1)

Materials and Methods

Supplementary Text

Figs. S1 to S5

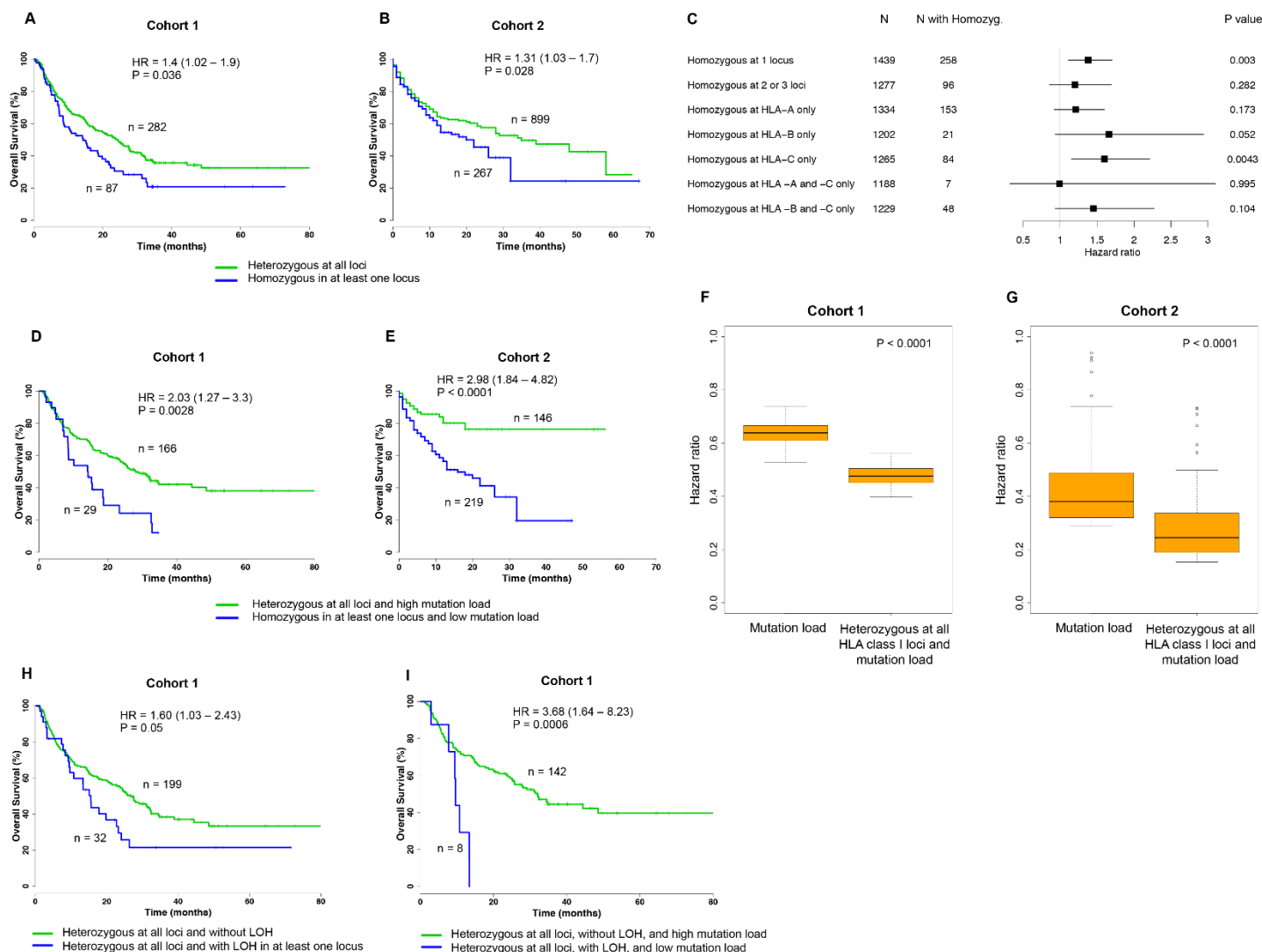
Tables S1 to S9

References (61–86)

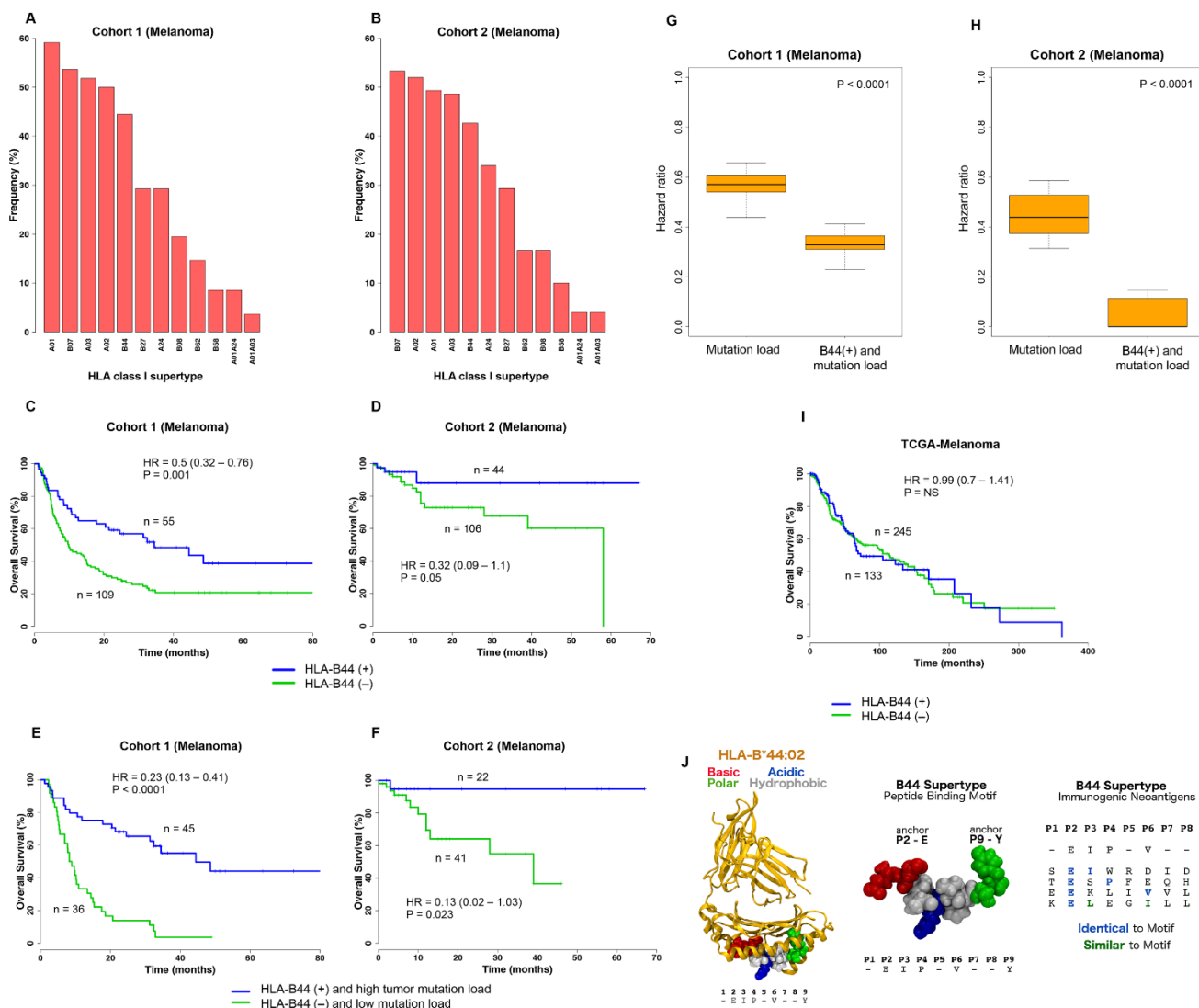
24 July 2017; accepted 29 November 2017

Published online 7 December 2017

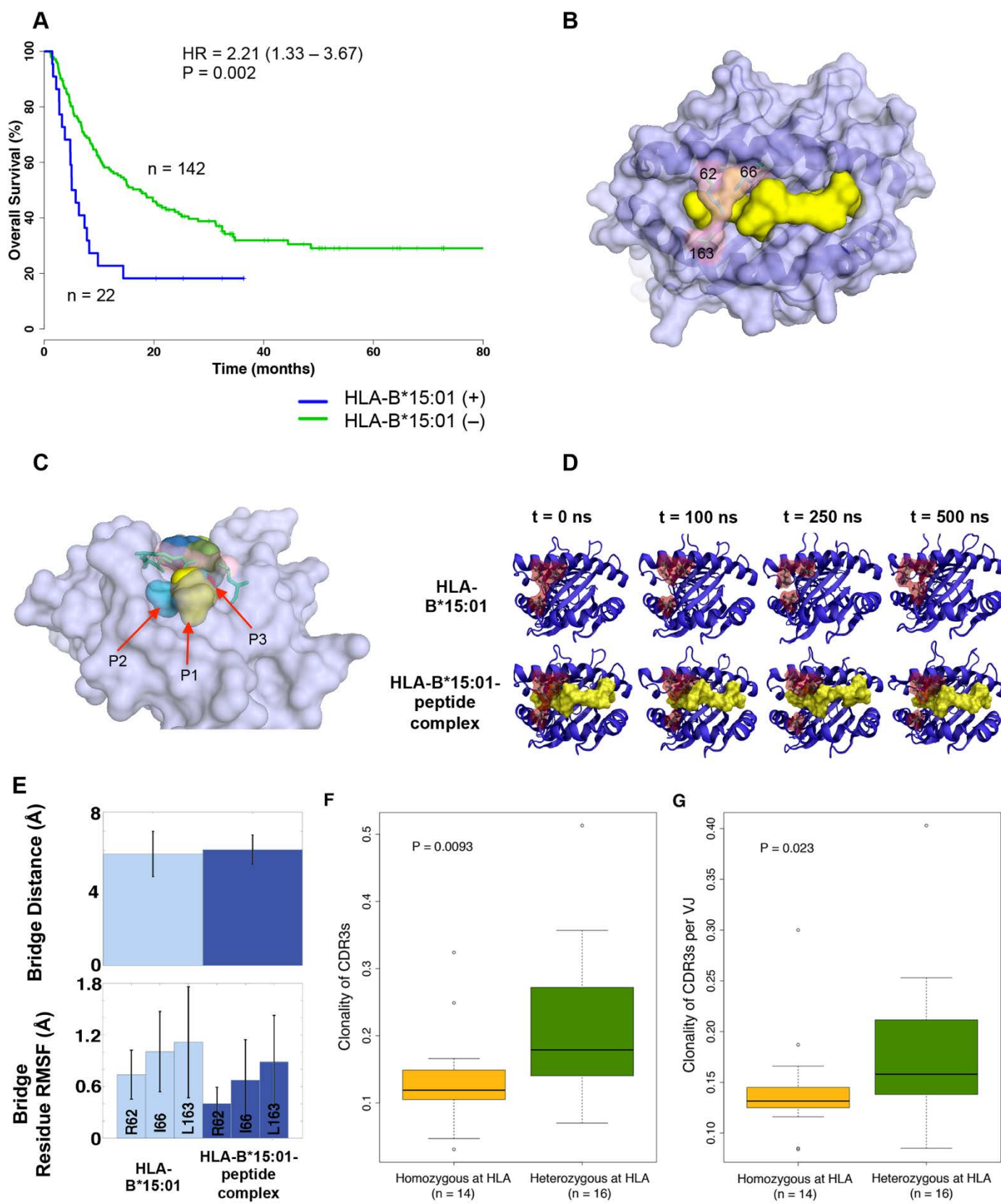
10.1126/science.aao4572



**Fig. 1. Effect of HLA class I homozygosity on survival in patients treated with immune checkpoint inhibitors.** (A) Association between homozygosity in at least one HLA-I locus and reduced overall survival in cohort 1. (B) Association between homozygosity in at least one HLA-I locus and reduced survival in cohort 2. (C) Association between HLA-I homozygosity and decreased survival from all 1,535 patients. Data show one or more HLA-I loci or individual loci (HLA-A, HLA-B, and HLA-C). Indicated are the number of patients and hazard ratio (HR). Horizontal lines represent the 95% confidence interval. *P* value was calculated using the Log-rank test. (D) Patients in cohort 1 with heterozygosity at all HLA-I loci and a high mutation load (defined as >113 mutations) compared to patients that are homozygous for at least one HLA-I locus and have a low mutation load. (E) Patients in cohort 2 with heterozygosity at all HLA-I loci and a high tumor mutational load (defined as >16.72 mutations) compared to patients that are homozygous in at least one HLA-I locus and have a low mutation load. (F) Distribution of hazard ratios to stratify cohort 1 patients based on tumor mutational load. The combined effect of HLA-I heterozygosity at all loci and mutation load on improved survival was greater compared to mutation load alone. (G) Distribution of hazard ratios to stratify cohort 2 patients based on mutation load. A range of cutoffs across the quartiles of mutation load were used. *P* values were calculated using the Wilcoxon-rank sum test. (H) Survival analysis showing that LOH of heterozygous germline HLA-I is associated with decreased overall survival in patients treated with ICB. (I) Survival analysis showing that the effect of LOH of heterozygous germline HLA-I is greater in tumors with low mutation burden compared to tumors with high mutation load and without LOH. High mutation load is defined as >113 mutations.



**Fig. 2. Influence of the HLA-B44 supertype on survival of melanoma patients treated with immune checkpoint inhibitors.** (A) Prevalence of the different HLA supertypes in patients with melanoma from cohort 1. (B) Prevalence of the different HLA supertypes in the patients with melanoma from cohort 2. (C and D) Survival analysis of patients possessing the B44 alleles [B44 (+)] compared with patients without the B44 alleles [B44 (-)] from cohort 1 (C) and cohort 2 (D). (E and F) Survival analysis of patients with the B44 alleles and with high mutation burden versus patients without B44 and with low mutation load, from cohort 1 (E) and cohort 2 (F). (G) Distribution of hazard ratios to stratify cohort 1 patients based on mutation load. The combined effect of B44 and mutation load on increased survival was greater compared to simply considering mutation load alone. (H) Distribution of hazard ratios to stratify cohort 2 patients based on mutation load. A range of cutoffs across the quartiles of mutation load was used. *P* values were calculated using the Wilcoxon-rank sum test. (I) Survival analysis of melanoma patients with and without the B44 alleles from the TCGA cohort. (J) Left: Example of peptide motif common among B44 alleles, docked in complex with HLA-B\*44:02 based on an available crystal structure (PDB: 1M60). The five common residues (E2, I3, P4, V6, and Y9) of the motif were reported in (56). Peptide residues are colored according to their properties as basic (red), acidic (blue), polar (green), or hydrophobic (grey). Center: Close up view of an example peptide conforming to the B44 motif (54, 56). Residues at positions 2 and 9 are important for anchoring the peptide in the HLA binding groove (54). Right: Alignment between B44 peptide motif and known immunogenic neoantigens (table S8) restricted to B44 expressed by melanomas. All neoepitopes feature Glu (E) at position 2; neoantigens are also either identical or similar to the motif at one or two additional positions. The neoantigen, FAM3C: TESPFEQHI, was identified in a melanoma patient with long-term response to anti-CTLA-4 from cohort 1. Sequence similarity was determined using standard residue classes (GAVLI, FYW, CM, ST, KRH, DENQ, and P).



**Fig. 3. Effect of the HLA-B\*15:01 allele on overall survival of melanoma patients treated with immune checkpoint inhibitors.** (A) Survival analysis showing reduced survival in ICB-treated melanoma patients from cohort 1 with and without the HLA-B\*15:01 allele. (B) Overview of the three-dimensional structure of the peptide-binding groove of HLA-B\*15:01, light purple; bound peptide, yellow; bridging residues, light pink. (C) Side view of the bridge-sequestration effect over bound-peptide residue positions P2 (light blue) and P3 (red). (D) MD simulation snapshots of both the isolated HLA B\*15:01 molecule and its complex with a 9-mer UBCH6 peptide; each trajectory was run over the course of 500 ns of simulation time. (E) Observables from the MD simulations described in D. The mean bridge distances in the HLA-B\*15:01 molecule and in the HLA-B\*15:01-peptide complex are comparable. The residue-position root mean square fluctuations (RMSFs) indicate that each of the bridging residues becomes more rigid in the presence of the peptide. (F) On-therapy clonality of TCR CDR3s between HLA heterozygous patients and patients who are HLA homozygous (in at least one class I locus or at HLA-DP). (G) On-therapy clonality of TCR CDR3s per VJ combination between HLA heterozygous patients and patients with HLA homozygosity (in at least one class I locus or at HLA-DP).

**Table 1. Association of HLA supertype with overall survival of melanoma patients treated with immune checkpoint blockade.** Data show the influence of specific HLA class I alleles on patient survival.

HLA class I supertype	Frequency	HR	P value
A24	0.29	0.67 (0.44–1.03)	0.07
A01	0.59	0.87 (0.60–1.27)	0.47
A03	0.52	1.39 (0.96–2.03)	0.08
A02	0.5	1.13 (0.76–1.63)	0.53
B58	0.09	0.98 (0.51–1.88)	0.96
B62	0.15	2.29 (1.40–3.74)	0.0007
B27	0.29	1.09 (0.73–1.63)	0.67
B44	0.45	0.61 (0.42–0.89)	0.009
B07	0.54	1.35 (0.92–1.97)	0.12
B08	0.2	0.85 (0.52–1.39)	0.51
A01A03	0.04	1.20 (0.49–2.94)	0.69
A01A24	0.09	0.89 (0.43–1.83)	0.76
Alleles influencing the significant associations			
B44s, B*18:01, B*44:02, B*44:03, B*44:05, B*50:01	0.34	0.49 (0.32–0.76)	0.001
B62s, B*15:01	0.13	2.21 (1.33–3.70)	0.002

## MOLECULAR BIOLOGY

# The piRNA targeting rules and the resistance to piRNA silencing in endogenous genes

Donglei Zhang,<sup>1,2\*</sup> Shikui Tu,<sup>3,4\*</sup> Michael Stubna,<sup>1</sup> Wei-Sheng Wu,<sup>5</sup> Wei-Che Huang,<sup>5</sup> Zhiping Weng,<sup>3</sup> Heng-Chi Lee<sup>1†</sup>

Piwi-interacting RNAs (piRNAs) silence transposons to safeguard genome integrity in animals. However, the functions of the many piRNAs that do not map to transposons remain unknown. Here, we show that piRNA targeting in *Caenorhabditis elegans* can tolerate a few mismatches but prefer perfect pairing at the seed region. The broad targeting capacity of piRNAs underlies the germline silencing of transgenes in *C. elegans*. Transgenes engineered to avoid piRNA recognition are stably expressed. Many endogenous germline-expressed genes also contain predicted piRNA targeting sites, and periodic An/Tn clusters (PATCs) are an intrinsic signal that provides resistance to piRNA silencing. Together, our study revealed the piRNA targeting rules and highlights a distinct strategy that *C. elegans* uses to distinguish endogenous from foreign nucleic acids.

Element induced wimpy testis (Piwi) proteins and their associated Piwi-interacting RNAs (piRNAs) function as a guardian of animal genomes through transposon silencing in various animals (1–5). However, many animals produce piRNAs that do not match transposon sequences. For example, the vast majority of the 15,000 piRNAs encoded by the *Caenorhabditis elegans* genome do not exhibit extensive complementarity to transposons (3, 4, 6). In mice, tens of

thousands of distinct piRNAs produced at the pachytene stage during spermatogenesis do not map to transposons (7). These observations suggest additional targets and functions of piRNAs.

Identification of piRNA targets and the piRNA targeting rules has proven to be rather difficult. Cross-linking immunoprecipitation (CLIP) analyses of Piwi proteins suggest that they associate with diverse mRNAs (8–10). However, because diverse piRNAs engage with many mRNAs, it is difficult

to infer the target of a given piRNA from these CLIP analyses. Therefore, additional approaches are required to identify piRNA sites in vivo. In some cases, targets of piRNAs can be inferred if the mRNA target is cleaved by Piwi (11, 12). However, these cleaved mRNAs likely present only a fraction of piRNA targets in vivo because the slicer activity of Piwi is dispensable for silencing in some animals, including *C. elegans* (13–16). Because only few piRNA targets other than transposons have been identified, the piRNA targeting rules remain undefined, and both sequence-specific and sequence-nonspecific functions of the Piwi/piRNA complex have been proposed (8–10, 12, 17, 18).

To gain insight into the piRNA targeting mechanism, we identified the targets of a single piRNA and examined how the piRNA recognizes its targets. In *C. elegans*, piRNA targeting leads to the recruitment of RNA-dependent RNA polymerases (RdRPs) that produce secondary small RNAs named 22G-RNAs (fig. S1A) (3, 13, 15). These 22G-RNAs are loaded onto worm-specific Argonautes (WAGOs) to induce gene silencing (19–21). Because these 22G-RNAs are produced around the targeting

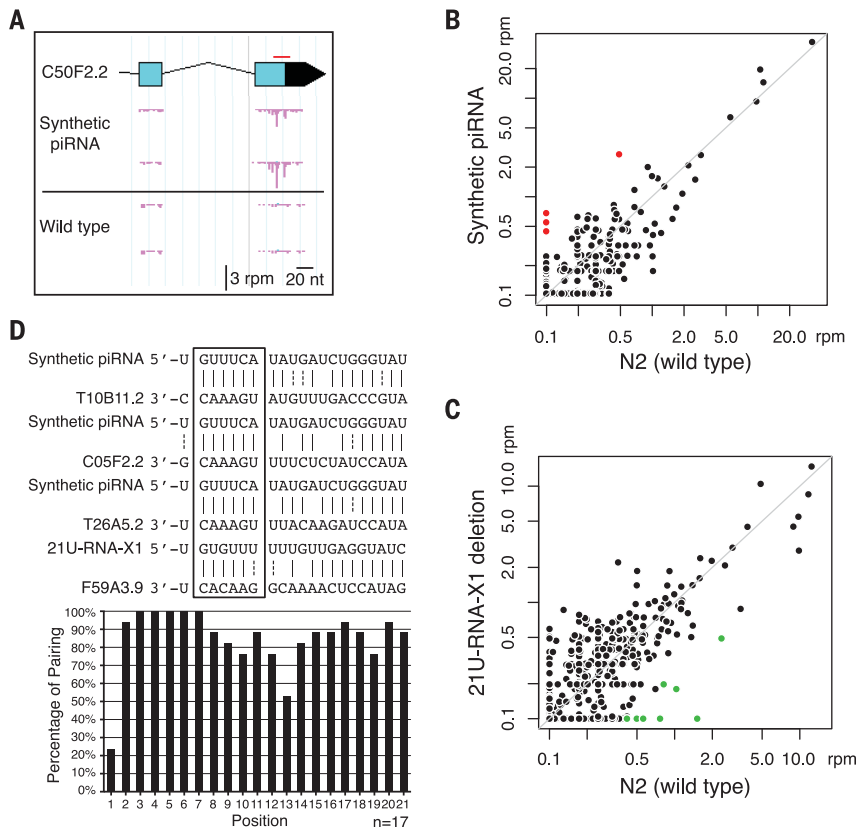
<sup>1</sup>Department of Molecular Genetics and Cell Biology, University of Chicago, Chicago, IL 60637, USA. <sup>2</sup>Department of Biochemistry and Molecular Biology, School of Basic Medicine, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, Hubei, China. <sup>3</sup>Program in Bioinformatics and Integrative Biology, University of Massachusetts Medical School, Worcester, MA 01605, USA. <sup>4</sup>Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai, China. <sup>5</sup>Department of Electrical Engineering, National Cheng Kung University, Tainan, Taiwan.

\*These authors contributed equally to this work.

†Corresponding author. Email: hengchilee@uchicago.edu

**Fig. 1. 22G-RNA loci as a proxy to identify the targets of specific piRNAs.** (A) An example of 22G-RNA distributions at one of the RNA targets of the synthetic piRNA (GFP-targeting piRNA #1) in the indicated strains with biological replicates. Each pink bar indicates the first nucleotide position and abundance of 22G-RNAs. The red bar indicates the position targeted by the synthetic piRNA. rpm, reads per million.

(B) A scatter plot showing the abundance of 22G-RNAs around each potential targeting site [100-nucleotide (nt) window centered with each target site] of the synthetic piRNA (GFP-targeting #1) in the control strain and in the strain expressing the synthetic piRNA. The potential targeting sites are sites of RNA transcripts that pair to the specific piRNA with six or fewer mismatches. Marked in red are sites at which 22G-RNA levels increased more than fourfold in the strain expressing the synthetic piRNA relative to the control strain. (C) A scatter plot showing the abundance of 22G-RNAs at each potential targeting site of 21U-RNA-X1 in the N2 (wild-type) strain and in the strain containing a deletion of the 21U-RNA-X1 coding loci. Marked in green are sites at which 22G-RNA levels decreased more than 75% in the strain loss of 21U-RNA-X1 relative to the N2 wild type. (D) The pairing between piRNAs and identified targets. (Top) Examples of pairings between the piRNAs and their targets. (Bottom) A bar graph showing the percentage of base pairing at each position within the piRNAs with all 17 identified targets. GU wobble pairing is considered as paired here in order to highlight the near-perfect pairing at the seed region when all GU pairs are allowed.



site, the 22G-RNAs can serve as a “signature” for piRNA targeting sites in vivo (13, 15). Therefore, we identified the targets of a piRNA by examining the 22G-RNA species gained in animals expressing a synthetic piRNA or the 22G-RNA species lost in animals carrying a deletion of a specific piRNA (fig. S1B). We obtained animals expressing a synthetic piRNA or losing an endogenous piRNA through a CRISPR/Cas9-based genome-editing strategy that modified the locus of an endogenous piRNA (fig. S1B). Small-RNA sequencing confirmed the expression or loss of specific piRNAs in these animals (fig. S1, C to F) and was used to identify changes in 22G-RNA levels. Together, we identified six RNA targets in the animals producing the synthetic piRNAs and 11 RNA targets in animals lacking the endogenous piRNAs (Fig. 1, A to C, and table S1). We noticed that a region of the piRNAs, from the second to seventh nucleotide, pairs well to the identified targeting sites (Fig. 1D). This implies a critical role for the pairing of this region in piRNA targeting, which we define as the piRNA seed. The piRNA seed is reminiscent of the microRNA (miRNA) seed, which is essential for miRNA target recognition (22). In addition, we observed apparent pairing outside of the piRNA seed region (Fig. 1D and fig. S1G). These observations suggest that base pairing outside of the seed region also contributes to piRNA target recognition, but a few mismatches can be tolerated. Furthermore, we noticed that GU wobble pairs are over represented, relative to other non-Watson-Crick pairs, in these targeting events (table S1). Last, the first nucleotide does not appear to contribute to piRNA targeting (Fig. 1D).

In light of these findings, we developed a piRNA reporter assay so as to gain further insights into the piRNA targeting rules. In this assay, we examined whether synthetic *gfp*-targeting piRNAs with various mismatches to the green fluorescent pro-

tein (GFP) sequence can trigger the silencing of an expressed GFP<sup>dpRNA::CDK-1</sup> transgene (dpRNA stands for depletion of piRNA targeting sites, in which the GFP sequence has been recoded to avoid silencing by endogenous piRNAs) (Fig. 2A, left, and fig. S2A). Because we noticed that the synthetic piRNAs can be produced from animals carrying extrachromosomal arrays with synthetic piRNA loci, we chose this method to systemically produce various *gfp*-targeting piRNAs (fig. S2A). We observed that GFP<sup>dpRNA::CDK-1</sup> was silenced in the animals expressing synthetic piRNAs that are perfectly complementary to GFP mRNAs or contain two mismatches outside of the piRNA seed region (Fig. 2, A to C, and fig. S2B). On the contrary, we failed to observe the silencing of GFP<sup>dpRNA::CDK-1</sup> when one or two mismatches were located at the piRNA seed region (Fig. 2, B and C, and fig. S2B). In addition, our reporter assay suggests that piRNAs tolerate up to three nonseed mismatches but not RNA bulges (fig. S2, C to E). We also observed that one GU wobble pair is tolerated in the seed region, and GU pairs are moderately more tolerated than mismatches in the nonseed region (fig. S2, C and D). Last, we obtained consistent results in our reporter analyses using gene-edited worms that express synthetic *gfp*-targeting piRNAs from an endogenous piRNA locus (fig. S2F). Overall, our reporter assay revealed a similar but more stringent piRNA-targeting logic than that from our analyses of synthetic piRNA targets. Together, our analyses suggest that piRNA targeting in *C. elegans* prefers near-perfect pairing at the piRNA seed region. In addition, supplementary pairing outside of the seed region also contributes to piRNA targeting, but few mismatches are tolerated.

It has been known for decades that transgenes carrying various foreign sequences, such as GFP or mCherry, are frequently silenced in the germline of *C. elegans* (23, 24). A previous study has

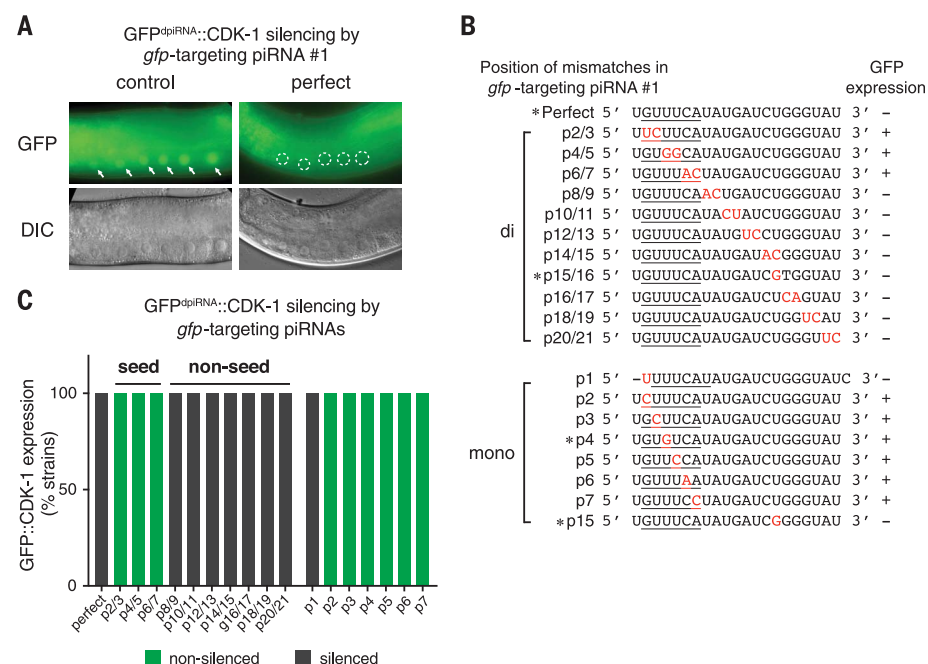
shown that Piwi protein PRG-1 is required for the silencing of the transgene GFP::CDK-1 (19). If piRNAs recognize GFP sequences, then removal of piRNA targeting sites from the GFP sequences should allow transgene GFP::CDK-1 expression in the germline. To predict more piRNA targeting sites on transgenes, we used the relaxed piRNA targeting criteria similar to those derived from our analyses of synthetic piRNA targets (Data File S1, algorithms of target prediction). These criteria predicted 17 piRNA targeting sites on GFP mRNA (Fig. 3A and table S2). We introduced silent mutations in the GFP sequences so that we no longer identified piRNA targeting sites, yielding the recoded-GFP<sup>dpRNA</sup> sequences. Remarkably, although the GFP::CDK-1 transgene is always silenced in the germline of wild-type animals, we observed strong GFP expression from all five independent transgenic strains that we obtained with recoded GFP<sup>dpRNA::CDK-1</sup> inserted at the same locus (Fig. 3B).

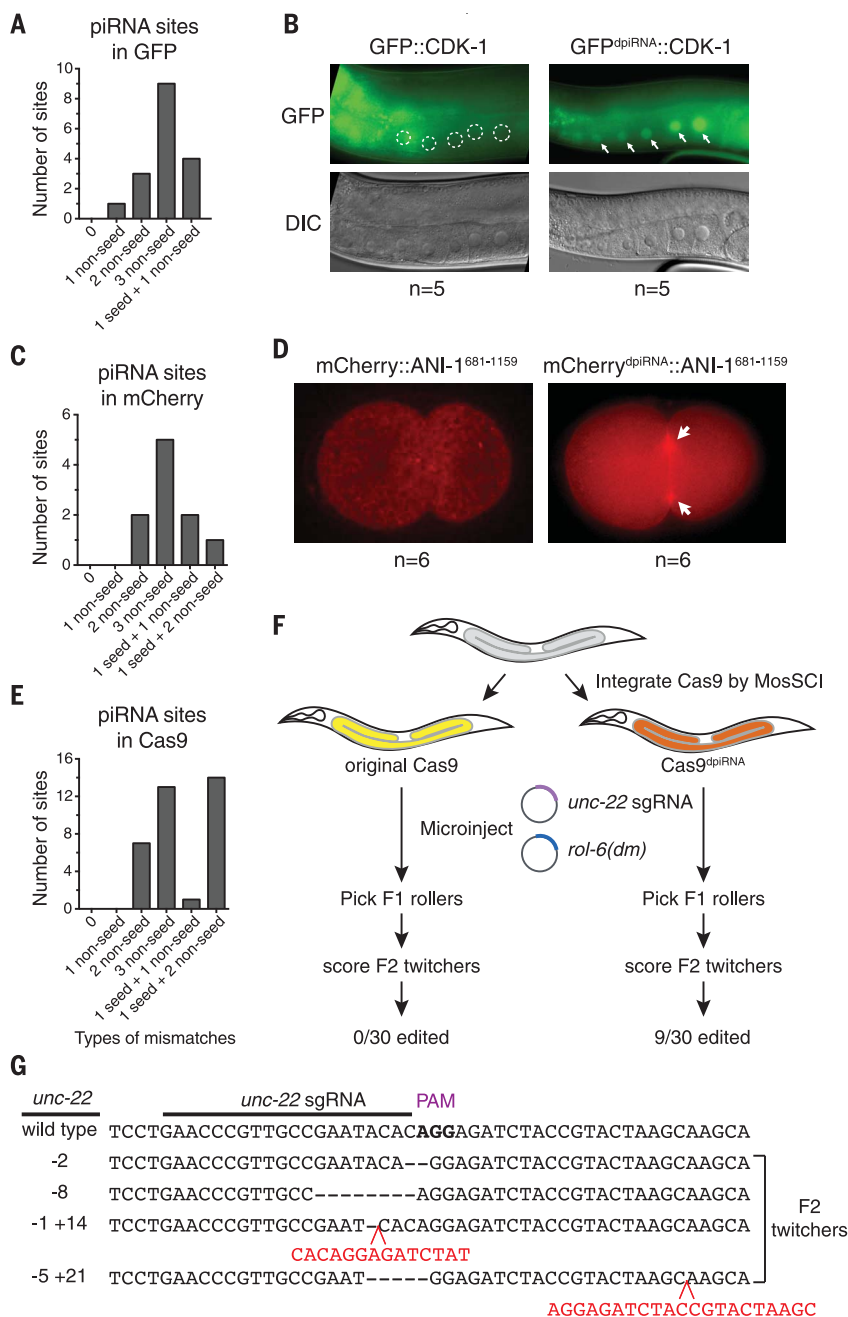
To test whether our approach can be generally applied to other transgenes in order to avoid gene silencing, we chose to modify the mCherry-tagged C-terminal region of Anillin (mCherry::ANI-1<sup>681-1159</sup>), another transgene that is always silenced in the germline (25). We predicted 10 piRNA targeting sites in mCherry mRNA and introduced silent mutations to disrupt predicted piRNA targeting sites (Fig. 3C and table S2). Whereas the original mCherry::ANI-1<sup>681-1159</sup> transgene was silenced in all six transgenic lines, the modified mCherry<sup>dpRNA::ANI-1</sup> was robustly expressed at the cleavage furrow of the one-cell embryo in all six transgenic lines we obtained (Fig. 3D).

As a last test, we applied this approach to modify Cas9 sequences. Transgenic *C. elegans* strains stably expressing Cas9 have not been successfully obtained (26). Again, we introduced silent mutations in order to remove all predicted piRNA targeting

**Fig. 2. A piRNA reporter assay to investigate the piRNA targeting rules. (A)** Fluorescence micrographs showing the expression of transgene GFP<sup>dpRNA::CDK-1</sup> in worms carrying an extrachromosomal array to express the *gfp*-targeting piRNA with perfect pairing or in the control strain that does not express the synthetic piRNA.

Arrows indicate the germline nuclei with expressed transgene. Circles indicate the germline nuclei with silenced transgene. The unmarked green fluorescent signals are autofluorescent signals generated from worm intestinal granules. **(B)** The sequences of the *gfp*-targeting piRNAs, the positions of the mismatches (red), and their effects on the expression of GFP<sup>dpRNA::CDK-1</sup>. Asterisk indicates *gfp*-targeting piRNAs produced by gene-edited animals modified at an endogenous piRNA locus (21U-5499). **(C)** Percentage of transgenic animals that exhibit the silencing of GFP<sup>dpRNA::CDK-1</sup> in animals expressing specific *gfp*-targeting piRNAs. At least eight independent strains carrying extrachromosomal arrays (roller) are examined for each piRNA.



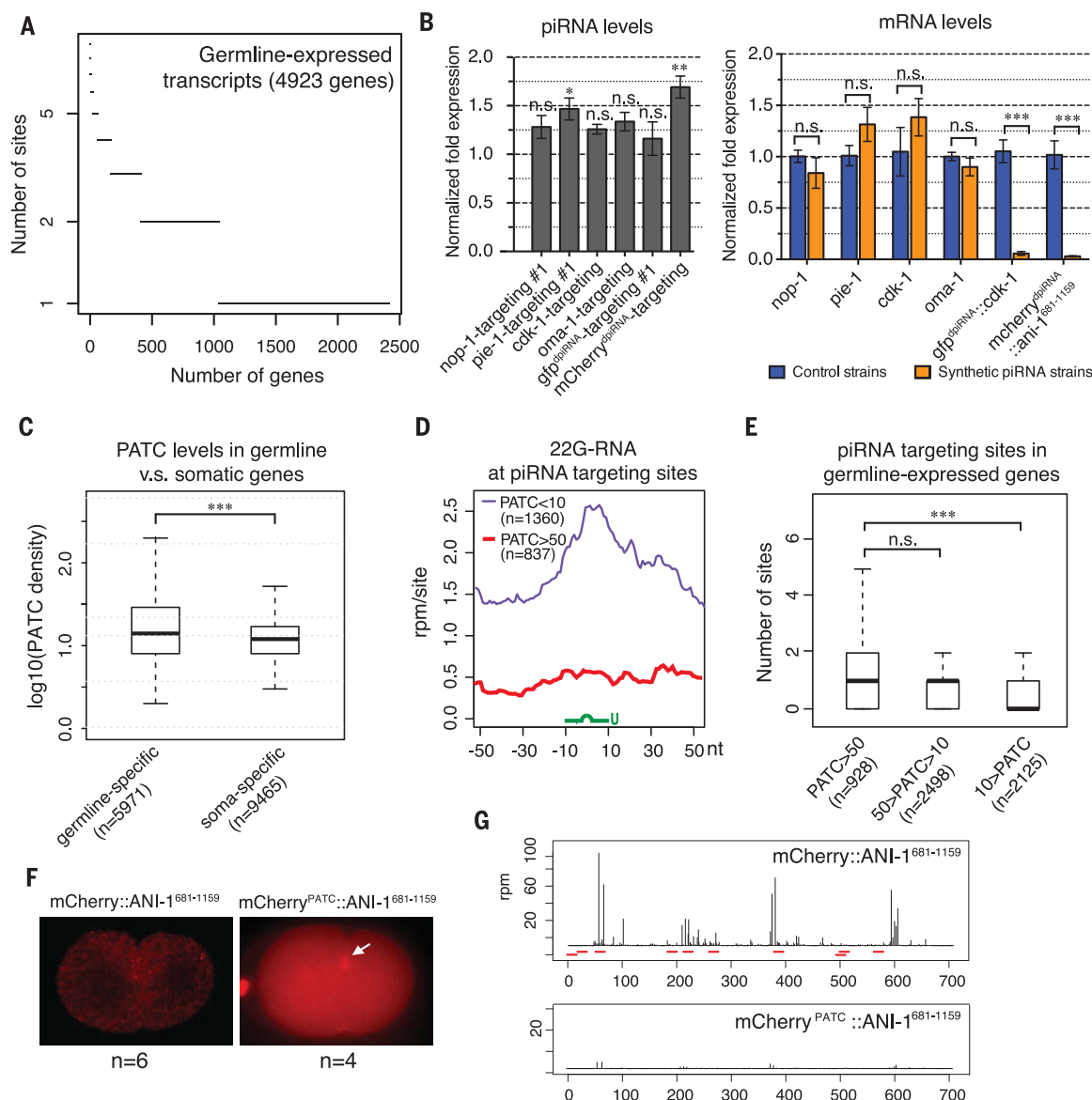


sites (Fig. 3E and table S2) and obtained transgenic animals carrying the original or the modified Cas9 transgene. To test whether Cas9 is stably expressed, we injected the transgenic animals with an *unc-22* single-guide RNA (sgRNA)-expressing plasmid and a *rol-6(su1006)* plasmid as a dominant transformation marker (Fig. 3F). The animals carrying *unc-22* mutations exhibit a visible twitcher phenotype and can be easily identified (27). Out of 30 F1 transformed progeny (roller), nine animals produced F2 twitcher progeny from animals carrying the modified Cas9<sup>dpiRNA</sup> transgene, whereas no F2 twitcher progeny were observed from animals carrying the original Cas9 transgene. DNA sequencing of these F2 twitcher animals confirmed that they carry various *unc-22*-edited alleles (Fig.

3F). These observations functionally demonstrate that the modified Cas9<sup>dpiRNA</sup> transgene is stably expressed and thus can create edited alleles. Taken together, these experiments verify that our predictions of piRNA targeting sites encompass the critical sites that trigger gene silencing.

We next wondered whether endogenous germline genes have evolved to avoid piRNA recognition. Previous studies have shown that most *C. elegans* germline transcripts are targeted by either WAGO Argonaute-associated 22G-RNA, which correlates with expression of the transcript (28, 29). Using stringent piRNA targeting criteria corresponding to the ones derived from our reporter analyses, we predicted that

**Fig. 3. Silencing-prone transgenes can be expressed in the germline by avoiding piRNA targeting.** (A) Predicted piRNA sites in GFP mRNA sequence. The numbers of piRNA sites that contain different types of mismatches are shown. The relaxed criteria are used to predict piRNA sites on transgenes: All GU wobble pairing is allowed (considered as paired), and up to three non-GU mismatches are allowed when sites have perfect seed pairing, or up to one non-GU mismatches are allowed when sites have one non-GU mismatches in the seed region. The mismatch at the first nucleotide of a piRNA is not counted or considered. (B) The expression of original GFP::CDK-1 that contains the predicted piRNA targeting sites, or the modified GFP<sup>dpiRNA</sup>::CDK-1 where all predicted piRNA sites have been removed by introducing silent mutations (right). Arrows indicate the germline nuclei with expressed GFP<sup>dpiRNA</sup>::CDK-1. Circles indicate the germline nuclei with silenced GFP::CDK-1. (C) Predicted piRNA sites in mCherry mRNA sequence (left). (D) The expression of original mCherry::ANI-1<sup>681-1159</sup> that contains the predicted piRNA targeting sites, or the modified mCherry<sup>dpiRNA</sup>::ANI-1<sup>681-1159</sup> in which the predicted piRNA sites have been removed by introducing silent mutations (right). Arrows indicate the expression of mCherry<sup>dpiRNA</sup>::ANI-1<sup>681-1159</sup> at cleavage furrows of the one-cell embryo. (E) Predicted piRNA sites in Cas9 mRNA sequence. (F) A schematic showing the procedure followed to examine if genome editing occurs in transgenic animals that carry the original or modified Cas9 transgenes. Plasmids containing *unc-22* sgRNA and *rol-6(su1006)* dominant transformation marker plasmid are coinjected into transgenic animals that have been carrying the Cas9 transgene for more than five generations. F1 transformed roller animals are picked, and their F2 progeny are scored for *unc-22* gene editing through twitcher phenotype. (G) Sequences of various *unc-22* edited alleles obtained in the animals carrying the modified Cas9 transgene injected with plasmid encoding *unc-22* sgRNA. Indels are highlighted in red.



**Fig. 4. Germline-expressed genes exhibit resistance to piRNA silencing through their intrinsic signals, such as PATCs.** (A) Numbers of predicted piRNA sites on germline-expressed RNA transcripts. To predict more confident targeting sites, the stringent piRNA targeting criteria are used here, in which up to one GU wobble pair was allowed in the seed region, and overall only up to two mismatches plus an additional GU mismatch were allowed. In addition, the mismatch at the first nucleotide of the piRNA is not counted or considered. The RNA targets of CSR-1 Argonaute (CSR-1 targets) are used to define the germline-expressed genes. (B) Quantitative reverse transcription polymerase chain reaction measurements of (left) the abundance of the synthetic piRNAs in comparison with the level of endogenous 21U-5499 (value = 1) in the control strain and (right) the expression levels of corresponding mRNA targets in the indicated strains. *nop-1*-, *cdk-1*-, and *oma-1*-targeting piRNAs were produced by gene-edited animals, whereas *pie-1*-targeting piRNAs were produced by animals carrying extrachromosomal arrays. Error bars represent standard error of the mean from biological duplicated samples. The statistics for synthetic piRNA expression were calculated by comparing the levels of specific piRNAs and 21U-5499. n.s., not significant; \* $P < 0.05$ ,

\*\* $P < 0.01$ , \*\*\* $P < 0.001$ , Student's  $t$  test. (C) A box and whisker plot showing the density of PATC in the germline-specific and somatic-specific genes. \*\*\* $P < 0.001$ , Wilcoxon rank sum test. (D) The density of 22G-RNAs within a 100-nt window around predicted piRNA target sites of germline-specific transcripts with high PATC density (PATC > 50) or low PATC density (PATC < 10). The plots are centered at the sequence targeted by piRNAs (green). The stringent piRNA targeting criteria were used here to predict piRNA target sites.  $n$  = number of predicted piRNA sites. (E) The box-and-whisker plots showing the number of predicted piRNA-targeted sites on germline-expressed genes that contain the indicated range of PATC density. The stringent piRNA targeting criteria were used here to predict piRNA target sites. n.s., not significant; \*\*\* $P < 0.001$ , Wilcoxon rank sum test. (F) Fluorescence micrographs showing the expression of the original mCherry::ANI-1<sup>681-1159</sup> harboring synthetic introns (no PATC) and mCherry<sup>PATC</sup>::ANI-1<sup>681-1159</sup> harboring PATC-containing introns. (G) 22G-RNA distribution at mCherry coding sequence of the indicated transgenes. Each bar indicates the first nucleotide position and abundance of 22G-RNAs. The red bars mark the location of piRNA targeting sites predicted by using the relaxed piRNA targeting criteria.

that are perfectly complementary to several of these genes, including *pie-1*, *nop-1*, *cdk-1*, and *oma-1*. We engineered these piRNAs using the same locus (21U-5499) we used for producing *gfp*-targeting piRNAs, and these synthetic piRNAs were expressed at similar levels (Fig. 4B). In the animals expressing synthetic piRNAs targeting endogenous genes, we did not observe a reduction of mRNA levels or the phenotypes associated with silencing of these genes (Fig. 4B and fig. S4, A to D). In addition, no phenotype associated with silencing was observed in animals expressing any of six additional synthetic piRNAs that target various regions of the *pie-1* or *nop-1* genes (fig. S4, A and B). This is in stark contrast to the animals expressing seven distinct *gfp* or mCherry-targeting piRNAs, which all trigger potent silencing of GFP<sup>ΔpiRNA</sup>::CDK-1 or mCherry<sup>ΔpiRNA</sup>::ANI-1, respectively (Fig. 4B and figs. S4, E and F). Together, our results suggested that at least some endogenous germline genes exhibit resistance to piRNA-mediated gene silencing in *C. elegans*.

Previous studies have proposed that CSR-1 Argonaute-associated 22G-RNAs may form an epigenetic memory of “self” to promote gene expression in the germline (30, 31). We therefore examined whether the *nop-1*-targeting piRNA can trigger *nop-1* silencing in *csr-1* mutants. In either *csr-1* F2 one-cell embryos with or without the treatment of *csr-1* RNAi, we did not observe that the synthetic piRNA conferred the phenotype associated with *nop-1* silencing (fig. S5A). To further test whether the piRNA resistance of endogenous genes is mediated by epigenetic signals, we used a Cas9-based gene-editing strategy to delete the *nop-1* gene and its untranslated regions (UTRs) from the genome, which would remove all chromatin-based signals as well as the DNA/RNA templates that are required to produce CSR-1 22G-RNAs targeting *nop-1* (fig. S5B). We then reinserted the *nop-1* gene back to its original locus or to the locus where our silenced transgenes are inserted. The reinserted *nop-1* gene remained resistant to silencing by endogenous or synthetic *nop-1*-targeting piRNA (fig. S6, C and D). Although these data represented negative results, our analyses provided no evidence for epigenetic mechanisms in licensing germline gene expression.

We therefore investigated whether intrinsic sequences of germline genes provide resistance to piRNA silencing. One such candidate is 10-base periodic An/Tn clusters (PATCs), an intrinsic DNA sequence element found in the introns or promoters of some germline genes in *C. elegans* (32). A recent study has reported that PATCs can promote the expression of transgenes inserted at heterochromatin in *C. elegans*, but whether PATCs can provide resistance to piRNA silencing has not been explored (33). We found that PATCs are enriched in germline genes and particularly enriched in the germline-expressed genes (Fig. 4C and fig. S6, A and B). To examine the global effect of PATCs on piRNA silencing, we compared the local 22G-RNA distribution at predicted piRNA sites in germline-specific transcripts with high or low PATC density. We observed that local 22G-RNAs accumulated around the targets only for genes with low, but not

high, PATC density (Fig. 4D and fig. S6, C and D). These observations imply that PATCs negatively affect the ability of the piRNA pathway to induce and/or maintain 22G-RNAs at the piRNA targeting sites. Furthermore, we observed that germline-specific genes of higher PATC density contain more predicted piRNA targeting sites than genes with lower PATC density (Fig. 4E and fig. S7, A and B), implying that PATCs allow the expression of germline genes despite harboring piRNA targeting sites. Last, if PATCs can provide resistance to piRNA silencing, insertion of PATC introns to a silencing-prone transgene should license its expression. Indeed, replacing the mCherry introns of the mCherry::ANI-1 transgene with PATC-containing introns from the *smu-1* gene led to the stable expression of the transgene (Fig. 4F). Small-RNA sequencing further showed that dramatically fewer mCherry antisense 22G-RNAs are produced in the worms carrying mCherry<sup>PATC</sup>::ANI-1<sup>681-1159</sup> than in those carrying the original mCherry::ANI-1<sup>681-1159</sup> (Fig. 4G). Together, our findings suggest that PATCs act as a licensing signal that provides resistance to piRNA silencing.

Overall, our study revealed the piRNA targeting logic in *C. elegans*. In addition, our research suggested that diverse piRNAs can recognize and silence various foreign nucleic acids because of their broad targeting capacity. Because several different modes of miRNA targeting have been described in animals and plants (22, 34–38), additional modes of piRNA-targeting are likely to exist as well. Nonetheless, our study demonstrated that piRNA-mediated gene silencing underlies the transgene silencing phenomenon in the germline of *C. elegans* and provided a simple solution to achieve transgene expression by avoiding piRNA recognition.

Our study showed that many endogenous genes also contain piRNA targeting sites but exhibit resistance to piRNA silencing. Our analyses suggested PATCs to be a licensing signal protecting endogenous genes from piRNA silencing. How PATCs counter against piRNA silencing remains unknown. A recent study showed that PATCs are enriched in germline genes within repressive chromatin domain, suggesting that PATCs may prevent piRNAs from establishing heterochromatin at their target (33). Our data suggest that PATCs function not simply by promoting euchromatin formation but also by inhibiting the production of 22G-RNA at piRNA targeting sites (Fig. 4, D and G). If so, it will suggest that the formation of heterochromatin may feedback to promote the production of 22G-RNAs. Such a relationship between chromatin and small-RNA production is reminiscent of RNA-induced transcriptional silencing described in *Schizosaccharomyces pombe*, in which small-RNA-guided heterochromatin recruits RdRPs to produce more small RNAs (39). In addition, because our data suggested that some endogenous genes—such as *nop-1*, *cdk-1*, or *oma-1*—exhibit resistance to piRNA silencing despite low PATC density (Fig. 4B) (33), other mechanisms may exist to license self genes for expression. Taken together, our studies revealed a strategy by which *C. elegans* defends its genome against foreign nucleic acids, whereby diverse piRNAs

silence foreign genes that are not licensed for expression.

## REFERENCES AND NOTES

1. J. Brennecke et al., *Cell* **128**, 1089–1103 (2007).
2. K. Saito et al., *Genes Dev.* **20**, 2214–2222 (2006).
3. P. J. Batista et al., *Mol. Cell* **31**, 67–78 (2008).
4. P. P. Das et al., *Mol. Cell* **31**, 79–90 (2008).
5. M. A. Carmell et al., *Dev. Cell* **12**, 503–514 (2007).
6. J. G. Ruby et al., *Cell* **127**, 1193–1207 (2006).
7. A. A. Aravin, G. J. Hannon, J. Brennecke, *Science* **318**, 761–764 (2007).
8. A. Vourekas et al., *Nat. Struct. Mol. Biol.* **19**, 773–781 (2012).
9. L.-T. Gou et al., *Cell Res.* **24**, 680–700 (2014).
10. J. A. Toombs et al., *RNA* **23**, 504–520 (2017).
11. M. Reuter et al., *Nature* **480**, 264–267 (2011).
12. W. S. S. Goh et al., *Genes Dev.* **29**, 1032–1044 (2015).
13. M. P. Bagijn et al., *Science* **337**, 574–578 (2012).
14. N. Darricarrère, N. Liu, T. Watanabe, H. Lin, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 1297–1302 (2013).
15. H. C. Lee et al., *Cell* **150**, 78–87 (2012).
16. K. Saito et al., *Genes Dev.* **24**, 2493–2498 (2010).
17. P. Rajasethupathy et al., *Cell* **149**, 693–707 (2012).
18. A. Vourekas, P. Alexiou, N. Vrettos, M. Maragkakias, Z. Mourelatos, *Nature* **531**, 390–394 (2016).
19. M. Shirayama et al., *Cell* **150**, 65–77 (2012).
20. B. A. Buckley et al., *Nature* **489**, 447–451 (2012).
21. A. Ashe et al., *Cell* **150**, 88–99 (2012).
22. D. P. Bartel, *Cell* **136**, 215–233 (2009).
23. W. G. Kelly, A. Fire, *Development* **125**, 2451–2456 (1998).
24. C. Merritt, C. M. Gallo, D. Rasoloson, G. Seydoux, *Transgenic solutions for the germline*, in *WormBook*, O. Hobert, S. Mango, Eds. (The *C. elegans* Research Community, 2010).
25. Y. C. Tse, A. Piekny, M. Glotzer, *Mol. Biol. Cell* **22**, 3165–3175 (2011).
26. S. Waaijers et al., *Genetics* **195**, 1187–1191 (2013).
27. H. Kim et al., *Genetics* **197**, 1069–1080 (2014).
28. W. Gu et al., *Mol. Cell* **36**, 231–244 (2009).
29. J. M. Claycomb et al., *Cell* **139**, 123–134 (2009).
30. M. Seth et al., *Dev. Cell* **27**, 656–663 (2013).
31. C. J. Wedeles, M. Z. Wu, J. M. Claycomb, *Dev. Cell* **27**, 664–671 (2013).
32. A. Fire, R. Alcazar, F. Tan, *Genetics* **173**, 1259–1273 (2006).
33. C. Frøkjær-Jensen et al., *Cell* **166**, 343–357 (2016).
34. C. Shin et al., *Mol. Cell* **38**, 789–802 (2010).
35. I. Ha, B. Wightman, G. Ruvkun, *Genes Dev.* **10**, 3041–3050 (1996).
36. A. Lal et al., *Mol. Cell* **35**, 610–625 (2009).
37. E. Allen, Z. Xie, A. M. Gustafson, J. C. Carrington, *Cell* **121**, 207–221 (2005).
38. M. W. Jones-Rhoades, D. P. Bartel, *Mol. Cell* **14**, 787–799 (2004).
39. K. Noma et al., *Nat. Genet.* **36**, 1174–1180 (2004).

## ACKNOWLEDGMENTS

We thank C. Mello for his guidance of H.-C.L. during his postdoctoral training and for resources to initiate the project; J. Staley, E. Ferguson, A. Ruthenburg, and J. Brown for critical comments on the manuscripts; E. Xiao for assistance on designing and cloning transgenes; M. Glotzer and K. Longhini for sharing reagents and unpublished results; and members of the H.-C.L. laboratory and J. Staley laboratory for helpful discussions. The deep sequencing data described in this manuscript are available at the Sequence Read Archive (SRA) of the National Center for Biotechnology Information (accession no. SRP108932). This work is supported by Ministry of Science of Technology of Taiwan grants (MOST-105-2918-I-006-002, MOST-105-2221-E-006-203-MY2, and MOST-106-2628-E-006-006-MY2) to W.-S.W., an NIH P01 grant (HD078253) to Z.W., and an NIH R00 grant (GM108866) to H.-C.L.

## SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/359/6375/587/suppl/DC1  
Materials and Methods  
Supplementary Text  
Figs. S1 to S9  
Tables S1 to S4  
References (40–45)  
Data File S1

10 July 2017; resubmitted 8 December 2017  
Accepted 2 January 2018  
10.1126/science.aao2840

## COLON CANCER

# Patients with familial adenomatous polyposis harbor colonic biofilms containing tumorigenic bacteria

Christine M. Dejea,<sup>1,2\*</sup> Payam Fathi,<sup>1,2,3†</sup> John M. Craig,<sup>4</sup> Annemarie Boleij,<sup>1,5</sup> Rahwa Taddese,<sup>5</sup> Abby L. Geis,<sup>1,2‡</sup> Xinqun Wu,<sup>1,3</sup> Christina E. DeStefano Shields,<sup>1,2</sup> Elizabeth M. Hechenbleikner,<sup>6§</sup> David L. Huso,<sup>7||</sup> Robert A. Anders,<sup>8</sup> Francis M. Giardiello,<sup>2,3</sup> Elizabeth C. Wick,<sup>6¶</sup> Hao Wang,<sup>1,2</sup> Shaoguang Wu,<sup>1,3</sup> Drew M. Pardoll,<sup>1,2</sup> Franck Housseau,<sup>1,2</sup> Cynthia L. Sears<sup>1,2,3#</sup>

Individuals with sporadic colorectal cancer (CRC) frequently harbor abnormalities in the composition of the gut microbiome; however, the microbiota associated with precancerous lesions in hereditary CRC remains largely unknown. We studied colonic mucosa of patients with familial adenomatous polyposis (FAP), who develop benign precursor lesions (polyps) early in life. We identified patchy bacterial biofilms composed predominately of *Escherichia coli* and *Bacteroides fragilis*. Genes for colibactin (*clbB*) and *Bacteroides fragilis* toxin (*bft*), encoding secreted oncotoxins, were highly enriched in FAP patients' colonic mucosa compared to healthy individuals. Tumor-prone mice cocolonized with *E. coli* (expressing colibactin), and enterotoxigenic *B. fragilis* showed increased interleukin-17 in the colon and DNA damage in colonic epithelium with faster tumor onset and greater mortality, compared to mice with either bacterial strain alone. These data suggest an unexpected link between early neoplasia of the colon and tumorigenic bacteria.

Colorectal cancer (CRC) is very common globally and develops through accumulation of colonic epithelial cell (CEC) mutations that promote transition of normal mucosa to adenocarcinoma. Around 5% of CRC occurs in individuals with an inherited mutation (1). One hereditary condition, familial adenomatous polyposis (FAP), is caused by germline mutation in the *APC* tumor suppressor gene. Individuals with FAP are born with their first mutation in the transition to CRC, and as somatic mutations accumulate, develop hundreds to thousands of colorectal polyps. The onset and frequency of polyp formation within families bearing the same *APC* gene mutation varies substantially (2), suggesting that additional factors contribute to disease

onset, including components of the microbiome (3).

The colon contains trillions of bacteria that are separated from the colonic epithelium by a dense mucus layer. This mucus layer promotes tolerance to foreign antigens by limiting bacterial-epithelial cell contact and, thus, mucosal inflammatory responses. In contrast, bacterial breaches into the colonic mucus layer with, in some, biofilm formation fosters chronic mucosal inflammation (4–6).

We previously reported that biofilms on normal mucosa of sporadic CRC patients were associated with a pro-oncogenic state (6, 7), suggesting that biofilm formation is an important epithelial event influencing CRC. To test the hypothesis that biofilm formation may be an early event in the progression of hereditary colon cancer, we examined the mucosa of FAP patients at clinically indicated colectomy.

We initially screened surgically resected tissue preserved in Carnoy's fixative from five patients with FAP and one with juvenile polyposis syndrome (table S1). Colon biopsies from individuals undergoing screening colonoscopy or surgical resections served as controls ( $n = 20$ , table S2). Polyps and macroscopically normal tissue were labeled with a panbacterial 16S ribosomal RNA (rRNA) fluorescence in situ hybridization (FISH) probe. Each FAP patient exhibited bacterial invasion through the mucus layer scattered along the colonic axis (Fig. 1A, table S3, and fig. S1). Unlike the continuous mucosal biofilms in sporadic CRC patients (6), FAP tissue displayed patchy bacterial mucus invasion (average length, 150  $\mu\text{m}$ ) on ~70% of the surgically resected colon specimens collected from four of six hereditary tumor patients. Biofilms were not

restricted to polyps, nor did they display right colon geographic preference as observed in sporadic CRC (table S3 and figs. S1 and S2). Biofilms were not detected in the mucus layer of the FAP patient who received oral antibiotics 24 hours before surgery (table S1 and fig. S2).

Specimens with bacterial biofilms were further screened by additional 16S rRNA probes to recognize the major phyla detected in biofilms of sporadic CRC; namely, *Bacteroides/Prevotella*, *Proteobacteria*, *Lachnospiraceae*, and *Fusobacteria* (table S4). Notably, FAP biofilms were composed predominantly of mucus-invasive *Proteobacteria* (~60 to 70%) and *Bacteroides* (10 to 32%) (table S3). *Fusobacteria* were not detected, and *Lachnospiraceae* were rare (<3%) by quantitative FISH analysis (table S3).

Additional probe sets (table S4) identified the predominant biofilm members as *E. coli* and *B. fragilis* (Fig. 1A, bottom panels; table S3). Invasion of the epithelial cell layer by biofilm community members was detected in all patients harboring biofilms (Fig. 1B), a finding similar to that in sporadic CRC patients. Further, FISH of mucosal biopsies from ileal pouches or anorectal remnants of additional, longitudinally followed, postcolectomy FAP patients revealed biofilms in 36% and mucosal-associated *E. coli* or *B. fragilis* in 50% (table S5). Thus, *E. coli* and *B. fragilis* are frequent, persistent mucosal colonizers of the FAP gastrointestinal tract. Of note, semiquantitative colon mucosa bacterial cultures of *Apc*<sup>Min716/+</sup> mice (truncation at the 716 codon of *Apc*), a murine correlate of FAP, displayed similar enrichment of *Bacteroides* and *Enterobacteriaceae* compared to wild-type (WT) littermates, consistent with data reported for *Apc*<sup>Min850/+</sup> mice (fig. S3) (8). These results suggest that *Apc* mutations enhance mucosal bacterial adherence, altering the bacterial-host epithelial interaction.

Strong experimental evidence exists supporting the carcinogenic potential of molecular subtypes of both *E. coli* and *B. fragilis* (9, 10); the two dominant biofilm members identified in direct contact with host colonic epithelial cells in our FAP patients. *E. coli* containing the polyketide synthase (*pks*) genotoxin island (encodes the genes responsible for synthesis of the colibactin genotoxin) induces DNA damage in vitro and in vivo along with colon tumorigenesis in azoxymethane (AOM)-treated interleukin-10 (IL-10)-deficient mice (10), whereas, enterotoxigenic *Bacteroides fragilis* (ETBF) induces colon tumorigenesis in *Apc*<sup>Min/+</sup> mice (9). Human epidemiological studies have associated ETBF and *pks+* *E. coli* with inflammatory bowel disease and sporadic CRC (10–13). Thus, we cultured banked frozen mucosal tissues from 25 FAP patients (two polyps and two normal tissues per patient when available, table S1) and 23 healthy individuals (mucosal sample from surgical resection or one ascending and one descending colon biopsy per colonoscopy subject, table S2) for the presence of *pks+* *E. coli* and ETBF. The mucosa of FAP patients was significantly associated with *pks+* *E. coli* (68%) and ETBF (60%) compared

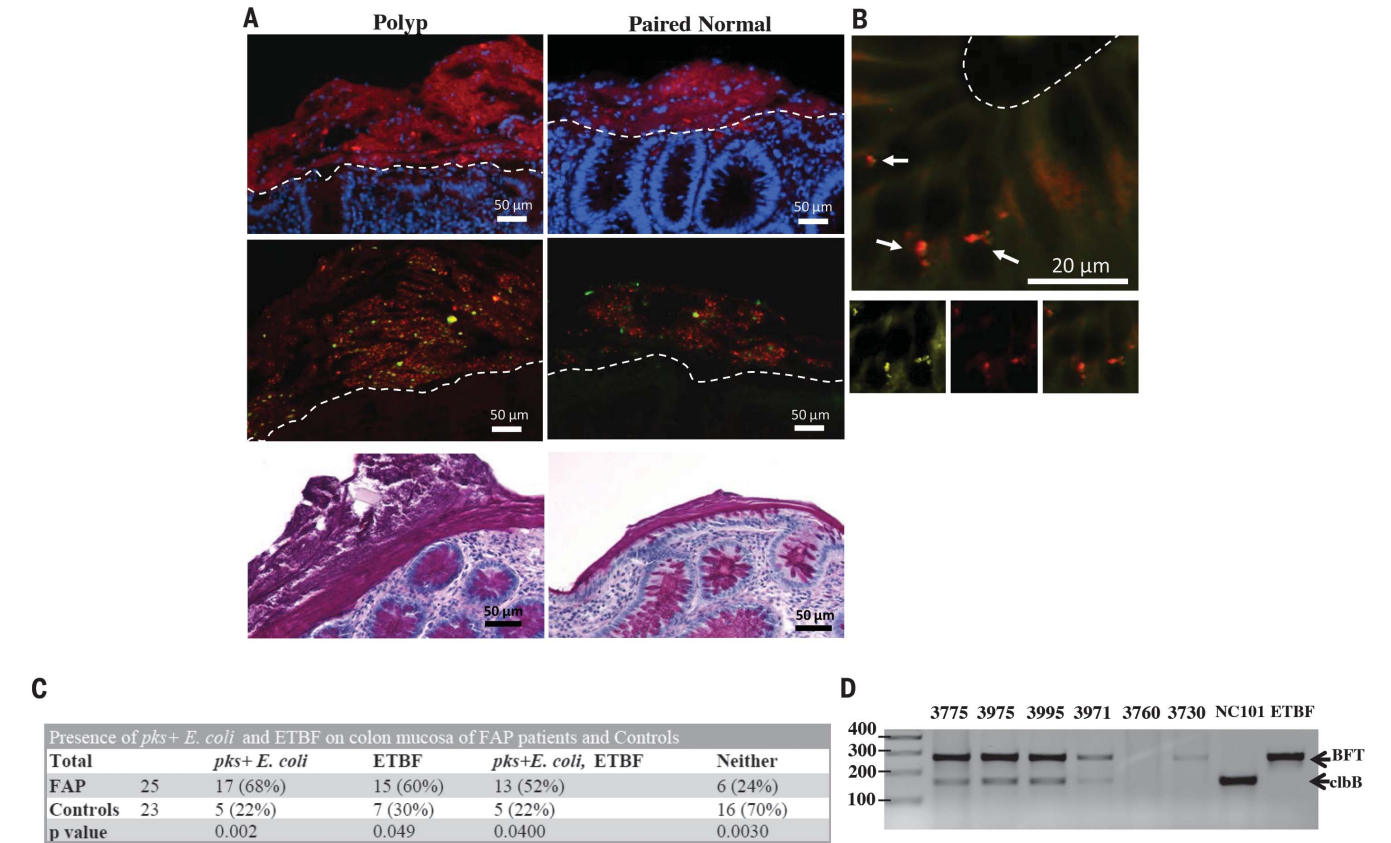
<sup>1</sup>Bloomberg-Kimmel Institute for Cancer Immunotherapy, Johns Hopkins University, Baltimore, MD, USA. <sup>2</sup>Department of Oncology, Johns Hopkins University, Baltimore, MD, USA. <sup>3</sup>Department of Medicine, Johns Hopkins University, Baltimore, MD, USA. <sup>4</sup>Department of Environmental Health Sciences, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, USA. <sup>5</sup>Department of Pathology, Radboud University Medical Center, Postbus 9101, 6500 HB Nijmegen, Netherlands. <sup>6</sup>Department of Surgery, Johns Hopkins University, Baltimore, MD, USA. <sup>7</sup>Department of Molecular and Comparative Pathobiology, Johns Hopkins University, Baltimore, MD, USA. <sup>8</sup>Department of Pathology, Johns Hopkins University, Baltimore, MD, USA.

\*Present address: 10903 New Hampshire Avenue, WO22 RM 5171, Silver Spring, MD 20993, USA. †Present address: Vanderbilt University School of Medicine, 340 Light Hall, Nashville, TN 37232, USA. ‡Present address: Arkansas College of Osteopathic Medicine, 7000 Chad Colley Boulevard, Fort Smith, AR 72916, USA. §Mount Sinai Hospital, Department of Surgery, 5 East 98th Street, New York, NY 10029, USA. ||Deceased. ¶Present address: Department of Surgery, University of California, 513 Parnassus Avenue, S 549, San Francisco, CA 94143, USA. #Corresponding author. Email: csears@jhmi.edu

to healthy subject mucosa (22% *pks+* *E. coli* and 30% ETBF) (Fig. 1C). There was no preferential association of ETBF or *pks+* *E. coli* with polyp or normal mucosa from FAP patients. Typically, mucosal samples from individual patients were concordant for *pks+* *E. coli* or ETBF (73% for *pks+* *E. coli*, 59% for ETBF), similar to results for mucosal *bft* detection in sporadic CRC patients (13). Notably, *pks+* *E. coli* and ETBF mucosal co-association occurred at a higher rate (52%) than expected to occur randomly (40.8%) given the frequencies for the individual species (Fig. 1C). Increased mucosal coassociation also occurred in healthy control subjects (22% observed versus 6.6% expected) (Fig. 1C). Laser capture microdissection of mucosal biofilms from our initial FAP patients (fig. S2 and table S1) contained both *bft* and *clbB* as determined by polymerase chain reaction (PCR) analysis, indicating that the car-

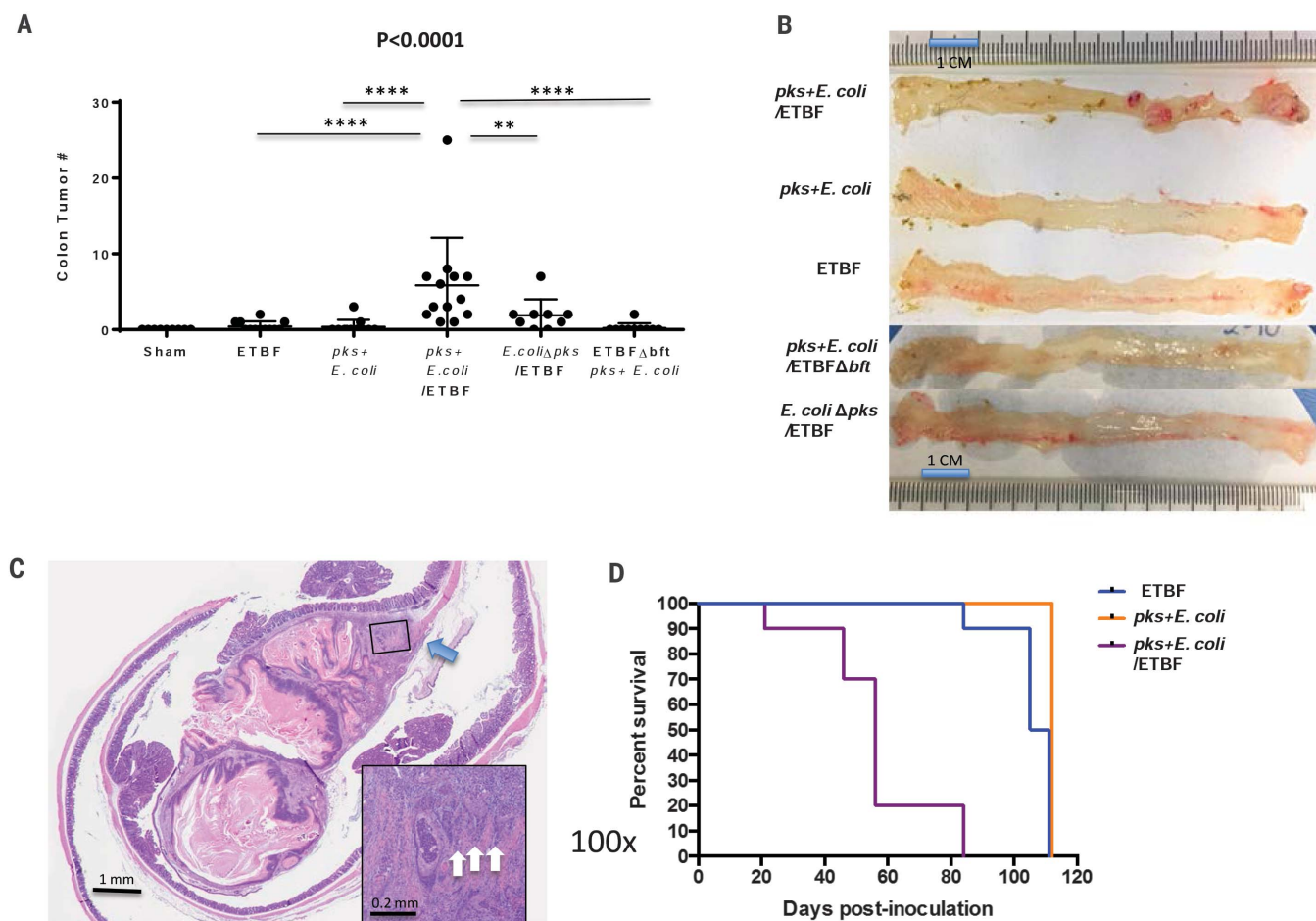
cinogetic subtypes of *B. fragilis* and *E. coli*, respectively, were present in the mucus layer in direct contact with the FAP epithelium (Fig. 1D). In contrast, neither virulence gene was detected in the mucus layer of control subject 3760 whereas *bft* was detected in the mucus layer of control subject 3730, consistent with our prior reported culture analysis of this sample (Fig. 1D) (13). The high frequency of *pks+* *E. coli* and ETBF cocolonization in FAP colons highlights the importance of understanding the potential effects of simultaneously harboring these two carcinogenic bacteria. Consequently, we used two murine models, AOM treatment without dextran sodium sulfate (see materials and methods) and *Apc*<sup>MinΔ716/+</sup> mice to test the hypothesis that *pks+* *E. coli* and ETBF cocolonization enhances colon tumorigenesis compared to monocolonization with either bacterium. The rate of spontaneous

colon tumorigenesis is very low in both model systems. Specific pathogen-free wild-type mice were treated with the carcinogen AOM and monoinoculated or coinoculated with canonical strains of *pks+* *E. coli* (the murine adherent and invasive strain, NC101) and ETBF (strain 086-5443-2-2) (9, 10). Fecal ETBF or *pks+* *E. coli* colonization was similar under monocolonization or cocolonization conditions, persisting until colon tumor formation was assessed at 15 weeks after colonization (fig. S4). Monocolonized (*pks+* *E. coli* or ETBF) mice displayed few to no tumors. However, pronounced tumor induction occurred in colonized mice, including an invasive cancer, suggesting the requirement for both bacteria to yield oncogenesis (Fig. 2, A to C). Tumorigenesis required the presence of BFT and the colibactin genotoxin as in-frame deletions of



**Fig. 1. Fluorescent in situ hybridization (FISH) and microbiology culture analysis of FAP mucosal tissues.** (A) Top panels: Representative FISH images of bacterial biofilms (red) on the mucosal surface of a FAP polyp and paired normal tissues counterstained with DAPI (4',6-diamidino-2-phenylindole) nuclear stain (blue). Middle panels: Most of the biofilm composition was identified as *B. fragilis* (green) and *E. coli* (red) by using species-specific probes. Bottom panels: PAS (periodic acid–Schiff)–stained histopathology images of polyp and paired normal mucosal tissues demonstrating the presence of the mucus layer. Images were obtained at 40× magnification; scale bars, 50 μm. Dotted lines delineate the luminal edge of the colonic epithelial cells. Images are representative of *n* = 4 to 23 tissue samples per patient screened (at least 10 5-μm sections screened per patient). (B) *Enterobacteriaceae* (yellow) and *E. coli* (red) FISH probes on paired normal FAP tissue (100×

magnification) revealing invasion into the epithelial cell layer at the base of a crypt (arrows). Bottom panels with insets of *Enterobacteriaceae* (bottom left panel) in yellow, *E. coli* (bottom middle panel) in red, and overlay (bottom right panel) confirming identification of the invasive species. Scale bar, 20 μm. Images are representative of *n* = 5 to 16 tissue samples per patient screened (at least 10 5-μm sections screened per patient). (C) FAP and control prevalence of *pks+* *E. coli* and enterotoxigenic *Bacteroides fragilis* (ETBF). Chi-square *P*-values are shown that represent the difference in probability of detection of each bacterium in FAP versus control patients. (D) PCR detection of *clbB* (a gene in the *pks* island) and *bft* within laser-captured biofilms containing *E. coli* and *B. fragilis* from designated FAP patients (table S1) and controls (table S2; materials and methods). Data show a representative image from two independent experiments with two or three replicates per experiment performed.



**Fig. 2. Cocolonization by *pks+* *E. coli* and ETBF increases colon tumor onset and mortality in murine models of CRC.** (A) Total colon tumor numbers detected in sham (n = 9), ETBF monocolonized (n = 12), *pks+* *E. coli* monocolonized (n = 11), *pks+* *E. coli*/ETBF cocolonized (n = 13), *E. coli*Δ*pks*/ETBF (n = 9), or *pks+* *E. coli*/ETBFΔ*bft* (n = 10) AOM mice at 15 weeks after colonization. Data indicate mean ± SEM. Overall significance was calculated with the Kruskal-Wallis test, and the overall P value is shown; Mann-Whitney U was used for two-group comparisons; \*\*P = 0.016, \*\*\*\*P < 0.0001. (B) Representative colons of monocolonized (ETBF or *pks+* *E. coli*), cocolonized (ETBF/*pks+* *E. coli*), *E. coli*Δ*pks*/ETBF, and *pks+* *E. coli*/ETBFΔ*bft* mice at 15 weeks after colonization of AOM-treated mice. Images are representative of n = 9 to 13 mice for each group.

(C) H&E (hematoxylin and eosin) histopathology of an invasive adenocarcinoma in a cocolonized (*pks+* *E. coli*/ETBF) AOM mouse at 15 weeks. Main image, 10× magnification; scale bar, 1 mm. Inset image, 100× magnification; scale bar, 0.2 mm. Blue arrow depicts the disruption of the muscularis propria by the invasive adenocarcinoma, and white arrows (inset) identify invading clusters of adenocarcinoma epithelial cells. (D) Kaplan-Meier survival plot of *Apc*<sup>Δ716Min/+</sup> mice (n = 30) colonized with either ETBF (blue; n = 10), *pks+* *E. coli* (orange; n = 10), or cocolonized with *pks+* *E. coli* and ETBF (purple; n = 10). Cocolonization significantly (P < 0.0001) increased the mortality rate. Statistics were analyzed with the log-rank test. All surviving mice (n = 19) were harvested at 110 days.

the *bft* gene and the *pks* virulence island significantly decreased tumors (Fig. 2A).

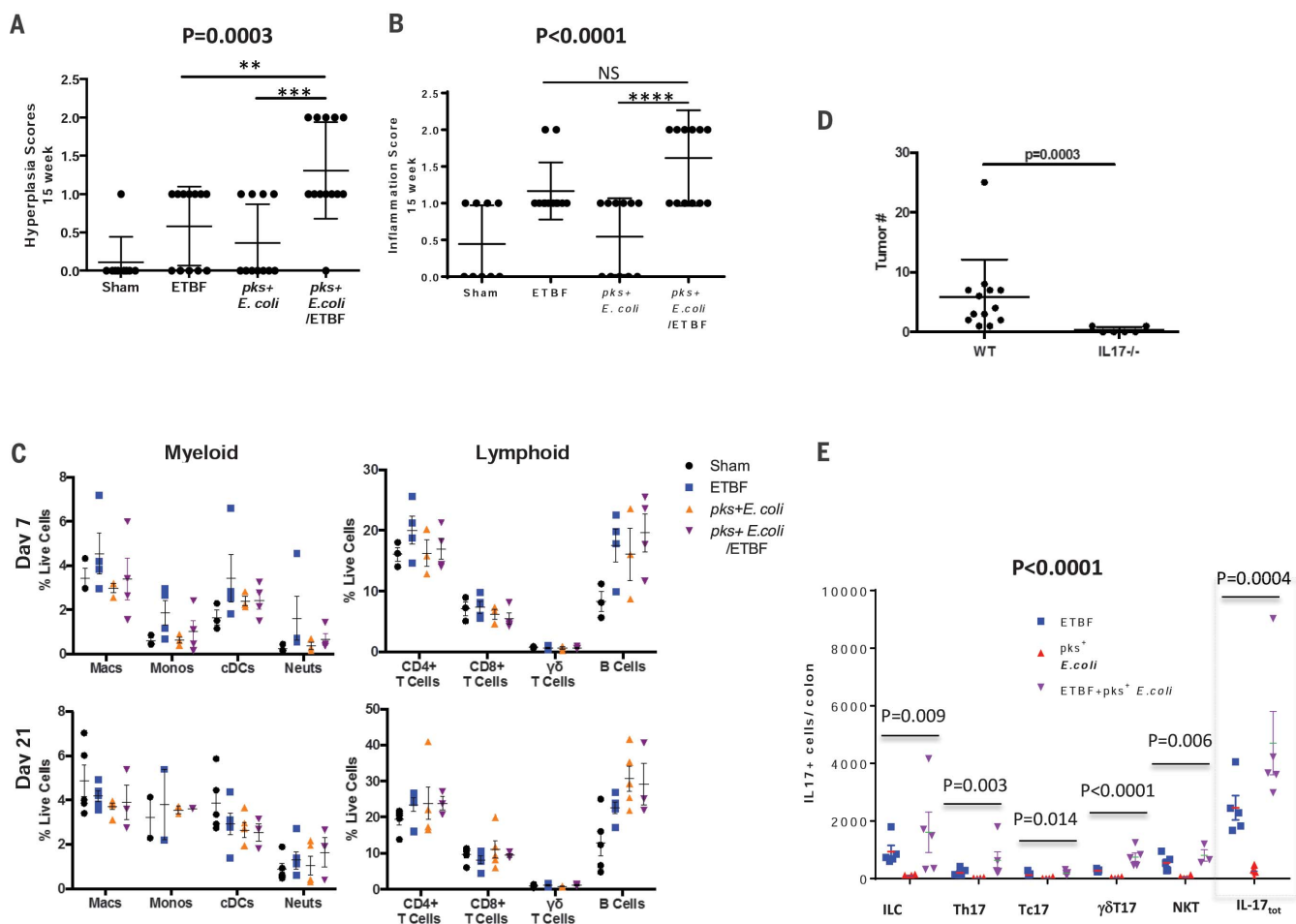
*Apc*<sup>MinΔ716/+</sup> mice cocolonized with ETBF and *pks+* *E. coli* exhibited enhanced morbidity with rapid weight loss and significantly increased mortality (P < 0.0001) [loss of 80% of the mice (n = 8) by 8 weeks and the remaining 20% (n = 2) at 12 weeks after colonization]. In contrast, 90% (n = 9) and 100% (n = 10) of mice monocolonized with ETBF or *pks+* *E. coli*, respectively, survived 15 weeks after colonization (Fig. 2D). The robust tumorigenesis of ETBF alone (at 15 weeks) and cocolonized mice (majority deceased by 8 weeks after colonization) was similar, whereas tumor numbers were significantly increased in the cocolonized cohort compared to *pks+* *E. coli* alone (fig. S5). Notably, at early time points, inflam-

mation was increased in the cocolonized cohort compared to either ETBF or *pks+* *E. coli* alone (fig. S5). Together these results suggest that the significant increase in colon inflammation and early tumorigenesis in the cocolonized mice contributed to their earlier mortality in the *Apc*<sup>Min/+</sup> mouse model.

Consistent with enhanced tumorigenesis, histopathological analysis revealed significantly increased colon hyperplasia and mucosal microadenomas in cocolonized AOM-treated mice compared to monocolonized mice (Fig. 3A and fig. S6A). However, histopathology scoring revealed modest differences in inflammation over time (4 days to 15 weeks) in mono- and cocolonized AOM mice (Fig. 3B and fig. S6B). Thus, overall inflammation did not seem to explain

differential tumor induction. To determine if the type of inflammation contributed to differences in tumorigenesis, we analyzed lamina propria immune-cell infiltrates of monocolonized and cocolonized wild-type AOM mice by flow cytometry. Our general lymphoid panel revealed a marked B cell influx across all colonization groups (Fig. 3C) but no significant differences in the proportion of infiltrating T cells (CD4, CD8, or γδ T cells) and myeloid populations between monocolonized and cocolonized AOM mice (Fig. 3C) either at the acute (1-week) or chronic (3-week) stage of infection.

Of particular interest was IL-17, as the tumorigenic potential of ETBF in *Apc*<sup>MinΔ716/+</sup> mice has been attributed, in part, to IL-17 (9). Because *bft* was necessary for synergistic tumor induction



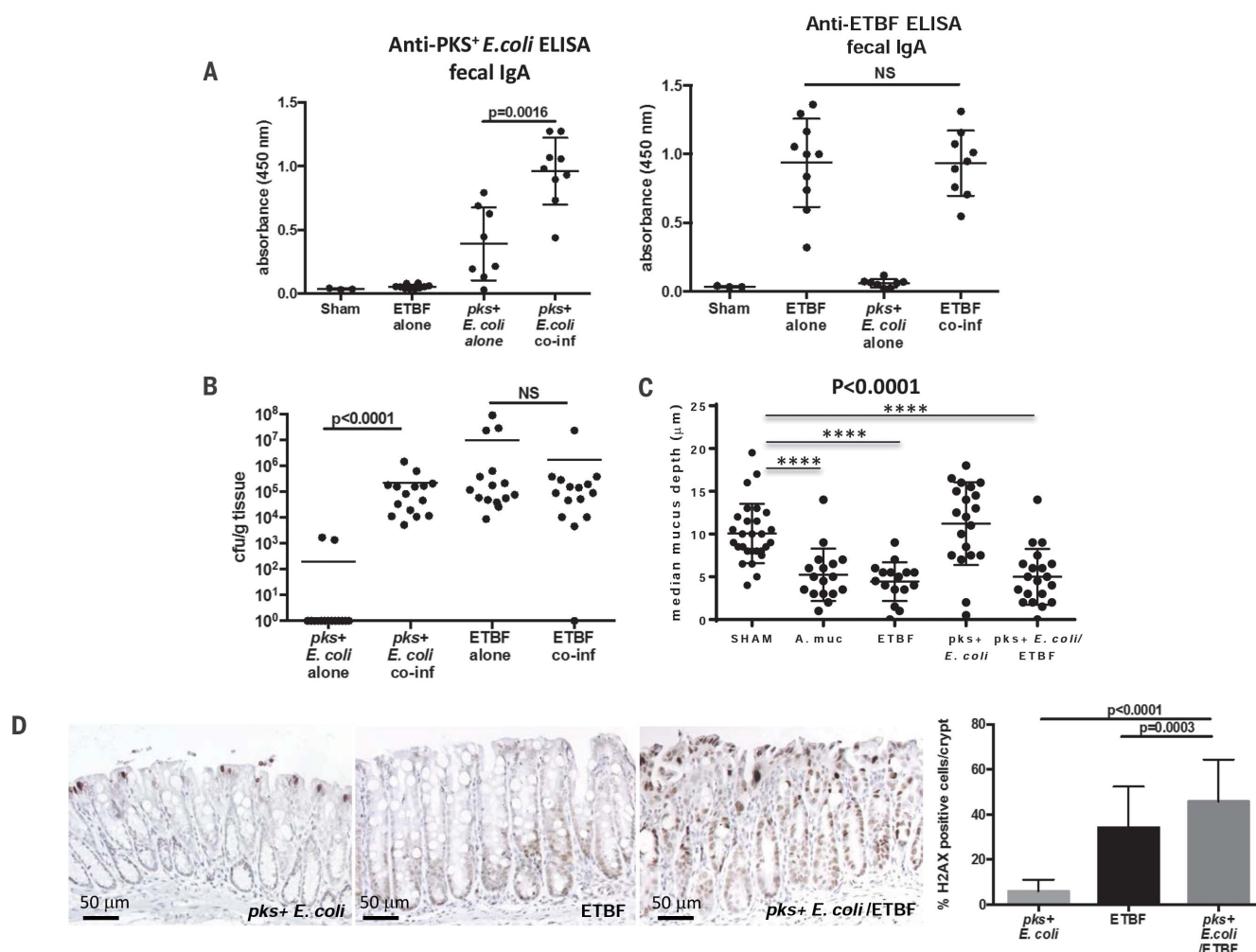
**Fig. 3. IL-17-induced inflammation is necessary for bacterial-driven tumorigenesis.** (A) Histologic hyperplasia and (B) inflammation scores of 15-week AOM sham ( $n = 9$ ), ETBF monocolonized ( $n = 12$ ), *pks+* *E. coli* monocolonized ( $n = 11$ ), or *pks+* *E. coli*/ETBF cocolonized ( $n = 13$ ) mice. Data represent mean  $\pm$  SEM of three independent experiments. For (A) and (B), overall significance was calculated by using the Kruskal-Wallis test, and the overall  $P$  value is shown; Mann-Whitney U was used for two-group comparisons;  $**P = 0.01$ ,  $***P = 0.0014$ ,  $****P = 0.0006$ ; NS, not significant. (C) Myeloid and lymphoid lamina propria immune cell infiltrates plotted as percentage of live cells in AOM mice at day 7 (top panels) and day 21 (bottom panels) after colonization. Data represent mean  $\pm$  SEM of three independent experiments (total three to five mice per group). (D) Total tumor numbers detected in IL-17-deficient AOM-treated mice (IL17<sup>-/-</sup>) versus wild-type AOM mice (WT). Both mouse strains were cocolonized with *pks+* *E. coli* and ETBF and tumors assessed

at 15 weeks. Data represent mean  $\pm$  SEM of two or three independent experiments (total 6 to 13 mice per group). Significance calculated by the Mann-Whitney U test represents differences between the non-normally distributed colon tumors in the independent mouse groups. (E) IL-17-producing cell subsets and total number of IL-17-producing (IL-17<sub>tot</sub>) cells per colon harvested from germ-free C57BL/6 mice monocolonized with *pks+* *E. coli* or ETBF or cocolonized with *pks+* *E. coli* and ETBF for up to 60 hours. Data represent mean  $\pm$  SEM of two independent experiments (total 3 to 5 mice per group). Overall significance across IL-17-producing cell types was calculated by using two-way analysis of variance testing based on log-transformed data (bold  $P$  value). For each cell subset and total number of IL-17-producing cells (gray dotted line box), the overall  $P$  value is shown and was calculated by using the Kruskal-Wallis test. Two-group cell subset and total number of IL-17-producing cell comparisons were analyzed by Mann-Whitney U test and are reported in table S7.

under cocolonization conditions (Fig. 2A), we tested the role of IL-17 in the cocolonized AOM model. Although IL-17 expression analysis by quantitative PCR revealed no significant difference in overall mucosal IL-17 mRNA levels between 15-week ETBF monocolonized and ETBF and *pks+* *E. coli* cocolonized mice (fig. S7), cocolonization of IL-17-deficient AOM mice ablated tumorigenesis (Fig. 3D). To specifically test whether ETBF and *pks+* *E. coli* cocolonization affected early colon mucosal IL-17 production, germ-free C57BL/6 mice were mono- or cocolonized and innate and adaptive lymphocyte sub-

sets analyzed by flow cytometry. Germ-free mice cocolonized with ETBF and *pks+* *E. coli* displayed a trend toward increase in total mucosal IL-17-producing cells when compared to monocolonized ETBF or *pks+* *E. coli* mice, driven by both adaptive [T helper 17 (T<sub>H</sub>17)] and innate (particularly  $\gamma\delta$ T17) cells (Fig. 3E and table S7). Although necessary for tumorigenesis (Fig. 3D), IL-17 alone appears insufficient to explain synergistic tumorigenesis in cocolonized mice because robust IL-17 induction by ETBF monocolonization (fig. S7) induces only meager colon tumorigenesis in AOM mice (Fig. 2A).

Because our general lymphoid panel revealed a marked B cell influx across all colonization groups (Fig. 3C), we profiled the secretory immunoglobulin A (IgA) response by IgA enzyme-linked immunosorbent assay (ELISA) using stool collected 4 weeks after colonization from AOM mice. Cocolonized mice had a significantly more robust IgA response to *pks+* *E. coli* than mice monocolonized with *pks+* *E. coli*, whereas the fecal anti-ETBF IgA response was similar under mono- and cocolonization conditions (Fig. 4A). Thus, the increased fecal IgA response was specific to *pks+* *E. coli* in mice cocolonized with ETBF,



**Fig. 4. ETBF enhances *pks+ E. coli* colonization and colonic epithelial cell DNA damage.** (A) ELISA results showing anti-*pks+ E. coli* (NC101) IgA and anti-ETBF (86-5443-2-2) IgA present in fecal supernatants from wild-type AOM mice under the designated colonization conditions for 4 weeks. Data represent mean  $\pm$  SEM of three independent experiments (total 3 to 10 mice per group). (B) Colonization of distal colon mucosae by *pks+ E. coli* and ETBF under mono- and cocolonization conditions at 4 weeks in AOM mice. Data represent mean of three independent experiments (total of 15 mice per group). (C) Mucus depth ( $\mu$ m) of HT29-MTX-E12 monolayers under the designated colonization conditions. Data represent mean  $\pm$  SEM

of three independent experiments. A. muc, *Akkermansia muciniphila*. (D) Representative images of  $\gamma$ -H2AX immunohistochemistry of distal colon crypts from AOM mice (five mice per condition) mono- or cocolonized with *pks+ E. coli* and ETBF for 4 days with quantification (right panel) of  $\gamma$ -H2AX-positive cells displayed as percentage positive per crypt (see materials and methods). Data represent mean  $\pm$  SEM of three independent experiments. For (A), (B), and (D), significance was calculated with the Mann-Whitney U test for two-group comparisons; for (C), overall significance was calculated with the Kruskal-Wallis test and the overall *P* value is shown; Mann-Whitney U was used for two-group comparisons; \*\*\*\**P* < 0.0001.

suggesting that cocolonization enhanced mucosal exposure to *pks+ E. coli*.

Although fecal colonization of both *pks+ E. coli* and ETBF was equivalent under both mono- and cocolonization conditions (fig. S4), quantification of mucosal-adherent ETBF and *pks+ E. coli* revealed a marked increase in mucosal-adherent *pks+ E. coli* under cocolonization conditions compared to *pks+ E. coli* monocolonization (Fig. 4B). Hence, under monocolonization conditions, *pks+ E. coli* is largely cultivatable only from the colon lumen whereas in the presence of ETBF, *pks+ E. coli* colonizes the mucosa at high levels (10<sup>3</sup> to 10<sup>6</sup> colony-forming units per gram of tissue). Using Muc-2-producing HT29-MTX-E12 monolayers in vitro, we tested the impact of

*pks+ E. coli* and ETBF on mucus. Although *pks+ E. coli* colonization alone had no impact on mucus depth, monolayer colonization with ETBF alone or cocolonized with *pks+ E. coli* significantly reduced mucus depth similar to colonization with *A. muciniphila* a known human colonic mucin-degrading bacterium (Fig. 4C). These results suggest that mucus degradation by ETBF promotes enhanced *pks+ E. coli* colonization. Such a shift in the bacterial niche of *pks+ E. coli* would facilitate the delivery of colibactin, the DNA-damaging toxin released by *pks+ E. coli*, to colon epithelial cells. Consistent with this hypothesis,  $\gamma$ -H2AX immunohistochemistry revealed significantly enhanced DNA damage in the colon epithelial cells of AOM

mice cocolonized with *pks+ E. coli* and ETBF compared to monocolonized (*pks+ E. coli* or ETBF) mice (Fig. 4D). Further, mice cocolonized with ETBF and *E. coli*Δ*pks* displayed similarly enhanced mucosal colonization with the *E. coli* strain (fig. S8) but reduced tumors and no increase in DNA damage or IL-17 (Fig. 2A and fig. S9, A and B, respectively). Lastly, persistent cocolonization of AOM-treated mice with the mucin-degrading *A. muciniphila* and *pks+ E. coli* did not enhance, but rather reduced, the modest colon tumorigenesis (fig. S10, A and B) induced by *pks+ E. coli* monocolonization. These results suggest that mucus degradation alone was insufficient to promote *pks+ E. coli* colon carcinogenesis in AOM mice.

Taken together, these data suggest that cocolonization with ETBF and *pks+* *E. coli*, found in more than half of FAP patients (in contrast to less than 25% of controls), promotes enhanced carcinogenesis through two distinct but complementary steps: (i) mucus degradation enabling increased *pks+* *E. coli* adherence, inducing increased colonic epithelial cell DNA damage by colibactin (Fig. 4D and fig S9); and (ii) IL-17 induction promoted, primarily, by ETBF with early augmentation by *pks+* *E. coli* cocolonization (Fig. 3, D and E, and table S7). We propose that together these mechanisms yield cooperative tumor induction in AOM mice cocolonized with ETBF and *pks+* *E. coli*.

ETBF and *pks+* *E. coli* commonly colonize young children worldwide. Thus, our results suggest that persistent cocolonization in the colon mucosa from a young age may contribute to the pathogenesis of FAP and potentially even those who develop sporadic CRC because *APC* loss or mutation occurs in the vast majority of sporadic CRC. We note that *pks+* *E. coli* are phenotypic and genotypic adherent and invasive *E. coli* (AIEC) (14). Despite this designation, derived primarily from in vitro cell culture experiments, the canonical *pks+* *E. coli* strain (NC101) used in our experiments was only cultivatable from the colon lumen in the absence of concomitant ETBF colonization in our mouse model. This ETBF-dependent shift to marked mucosal *pks+* *E. coli*

colonization is consistent with our observations that ETBF and *pks+* *E. coli* cocolonize FAP colon biofilms, where both bacteria invade and cocolonize the mucus layer throughout the FAP colon. These findings suggest that analysis of coexpression of *bft* and *clbB* may have value in general screening and potential prevention of CRC.

#### REFERENCES AND NOTES

1. E. R. Fearon, B. Vogelstein, *Cell* **61**, 759–767 (1990).
2. F. M. Giardiello *et al.*, *Gastroenterology* **106**, 1542–1547 (1994).
3. C. Dejea, E. Wick, C. L. Sears, *Future Microbiol.* **8**, 445–460 (2013).
4. A. Swidsinski *et al.*, *Gut* **56**, 343–350 (2007).
5. A. Swidsinski, V. Loening-Baucke, A. Herber, *J. Physiol. Pharmacol.* **60** (suppl. 6), 61–71 (2009).
6. C. M. Dejea *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 18321–18326 (2014).
7. C. H. Johnson *et al.*, *Cell Metab.* **21**, 891–897 (2015).
8. J. S. Son *et al.*, *PLOS ONE* **10**, e0127985 (2015).
9. S. Wu *et al.*, *Nat. Med.* **15**, 1016–1022 (2009).
10. J. C. Arthur *et al.*, *Science* **338**, 120–123 (2012).
11. T. P. Prindiville *et al.*, *Emerg. Infect. Dis.* **6**, 171–174 (2000).
12. M. Prorok-Hamon *et al.*, *Gut* **63**, 761–770 (2014).
13. A. Boleij *et al.*, *Clin. Infect. Dis.* **60**, 208–215 (2015).
14. M. Martinez-Medina *et al.*, *J. Clin. Microbiol.* **47**, 3968–3979 (2009).

#### ACKNOWLEDGMENTS

We thank K. Kinzler and B. Vogelstein for valuable discussions; K. Romans and L. Hyland for assistance with patient enrollment; and S. Besharati for assistance with histopathologic analyses. The data presented in this manuscript are tabulated in the main text and supplementary materials and methods. This work was supported

by the Bloomberg Philanthropies and by NIH grants R01 CA151393 (to C.L.S., D.M.P.), K08 DK087856 (to E.C.W.), 5T32CA126607-05 (to E.M.H.), P30 DK089502 (Johns Hopkins University School of Medicine), P30 CA006973 (Johns Hopkins University School of Medicine), and P50 CA62924 (Johns Hopkins University School of Medicine). Funding was also provided through a research agreement with Bristol-Myers Squibb Co-International Immunology Network-IION Resource Model, 300-2344 (to D.M.P.); Alexander and Margaret Stewart Trust (Johns Hopkins University School of Medicine); GSRRIG-015 (American Society of Colon and Rectal Surgeons to E.M.H.); The Netherlands Organization for Scientific Research (NWO 825.11.03 and 016.166.089 to A.B.); and a grant from the Institute Mérieux (to C.L.S. and D.M.P.). D.M.P. discloses consultant relationships with Aduro Biotech, Amgen, Astra Zeneca, Bayer, Compugen, DNatrix, Five Prime, GlaxoSmithKline, ImmuneXcite, Jounce Therapeutics, Neximmune, Pfizer, Rock Springs Capital, Sanofi, Tizona, Janssen, Merck, Astellas, Flx Bio, Ervaxx, and DNAX. D.M.P. receives research support from Bristol-Myers Squibb, Compugen, Ervaxx, and Potenza. D.M.P. is a scientific advisory board member for Immunomic Therapeutics. D.M.P. shares intellectual property with Aduro Biotech, Bristol-Myers Squibb, Compugen, and Immunomic Therapeutics. All other authors declare no competing interests. C.L.S., D.M.P., C.M.D., and E.C.W. are inventors on patent application PCT/US2014/055123 submitted by Johns Hopkins University that covers use of biofilm formation to define risk for colon cancer.

#### SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/359/6375/592/suppl/DC1  
Materials and Methods  
Figs. S1 to S11  
Tables S1 to S7  
References (15, 16)

15 June 2016; resubmitted 28 September 2017  
Accepted 28 December 2017  
10.1126/science.aah3648



Check out our exciting research Prize!



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

The Prize is a new highly competitive international prize which honors scientists for their excellent contributions to neuromodulation research. For purposes of the Prize, neuromodulation is any form of alteration of nerve activity through the delivery of physical (electrical, magnetic, optical) stimulation to targeted sites of the nervous system.”

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

**[www.sciencemag.org/prizes/pins](http://www.sciencemag.org/prizes/pins)**

**Submission Deadline: March 15, 2018**

**Science**  
AAAS



**PINS 品驰**

# 2018 GCHERA WORLD AGRICULTURE PRIZE

## CALL FOR NOMINATIONS



The World Agriculture Prize seeks Nominations for the 2018 prize to be submitted no later than 17:00 GMT on 30 April 2018, to submit a nomination check terms & conditions for the Prize, and complete the online submission at [www.gchera-wap.org](http://www.gchera-wap.org)

### The Call

**WORLD AGRICULTURE PRIZE** is an international award of the **Global Confederation of Higher Education Associations for the Agricultural and Life Sciences (GCHERA)**, <http://www.gchera.com/>). The prize aims to promote the global development of the mission of its member institutions in education, research, innovation and outreach in the agricultural and life sciences. It recognizes the distinguished contribution of an individual to this mission. *A nominee's achievements should be demonstrated locally and in the wider geographical region of the institutions in which the nominee has served, but not necessarily globally.* From 2018, the prize is doubled to 100,000 USD and two recipients awarded, with at least one recipient NOT coming from a country classified as having a developed economy. For the list of those countries with a developed economy see [www.gchera-wap.org](http://www.gchera-wap.org).

### The Prize

Formally proposed by both Prof. Zhai Huqu, the former Vice President of the Chinese Academy of Agricultural Sciences (CAAS) and Prof. Zhou Guanghong, the President of NJAU and Chairman of Education Professional Committee (EPC) of the Chinese Association of Agricultural Science Societies (CAASS) on October 20, 2012 in the occasion of NJAU's 110th anniversary and was approved at the GCHERA steering committee meeting in Uruguay on Oct. 29, 2012. Generously sponsored by the Education Development Foundation of Nanjing Agricultural University and Da Bei Nong Group, the prizes are awarded annually.

### 2018 AWARDING CEREMONY

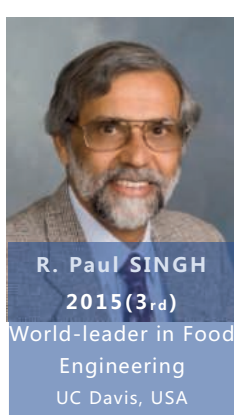
The 2018 GCHERA World Agriculture Prize Awarding Ceremony will be held on **28 October 2018** in Nanjing Agricultural University, Nanjing, China.

For more information, please visit China Coordinating Office for GCHERA World Agriculture Prize: <http://wapcn.njau.edu.cn/>



GCHERA WAP Awarding Ceremony

## LAUREATES



# Did you get the message?

In recent years, the discovery of new classes and modifications of RNA has ushered in a renaissance of RNA-focused research. Did you know that NEB<sup>®</sup> offers a broad portfolio of reagents for the purification, quantitation, detection, synthesis and manipulation of RNA? Experience improved performance and increased yields, enabled by our expertise in enzymology.



## FEATURED PRODUCTS *for* RNA RESEARCH:

### **NEW** Monarch<sup>®</sup> Total RNA Miniprep Kit –

rapidly purify up to 100 µg of high quality total RNA from multiple sample types

### Luna<sup>®</sup> One-Step RT-qPCR kits –

improve RT-qPCR performance with our novel, Luna WarmStart<sup>®</sup> Reverse Transcriptase

### **NEW** NEBNext<sup>®</sup> Ultra<sup>™</sup> II RNA Library Prep kits –

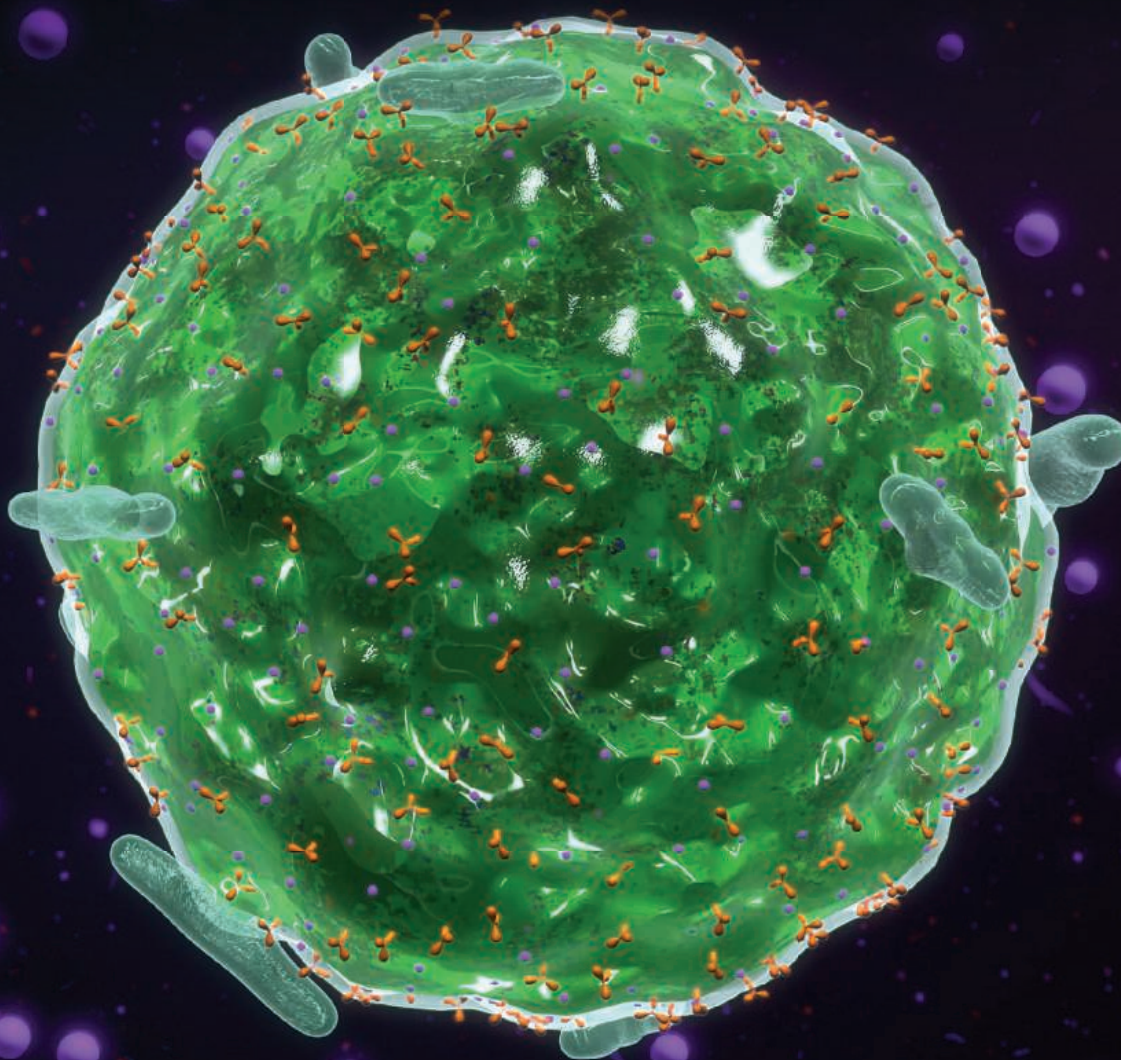
generate high quality RNA-seq libraries, even with limited amounts of RNA

### HiScribe<sup>®</sup> *in vitro* transcription kits –

rapidly synthesize high yields of high quality RNA

Let NEB help streamline your **RNA-related workflows**.  
Get started at **NEBrna.com**

Submit your high-impact research  
to ***Science Immunology***



*Science Immunology* publishes original, peer-reviewed, science-based research articles that report critical advances in all areas of immunological research, including important new tools and techniques. Share your research with *Science Immunology* global readership and submit your manuscript today!

What will your discovery be?

Submit your manuscript today at  
**ScienceImmunology.org**

**Science**Immunology

AAAS

# Serum University: These 5 key subjects will provide you with a strong educational foundation in the use of sera for cell culture and other applications

## A lab commodity whose quality may often be taken for granted, serum can make—or break—the outcome of all of your cell-based assays

Biologists who culture cells or tissue know that serum products, including fetal bovine serum, or FBS, (also sometimes called FCS, or fetal calf serum) constitute an essential component of culture media for most applications. Serum is often one of the most abundant constituents of cell culture media (typically 2-15% by volume), so its quality can have significant impact on culture success. FBS in particular is the most widely-used media supplement for the culture of mammalian cells due to its ready availability, along with its rich supply of growth factors that promote the development and expansion of cells *in vitro*. Serum also acts as a buffer to the cell culture system against a variety of toxic effects that can disrupt cell growth such as pH change, proteolytic activity, or the presence of endotoxin.

### General Biology: how to assess and handle serum in the lab

Anyone who has ever ordered serum for the lab knows that delivery timing is important, as bottles will typically arrive frozen from the supplier. This is because the activity of critical macromolecular factors in serum are best protected by storing serum at -5 to -20°C, and maintaining serum in the frozen state until use will also reduce the incidence of events like the breakdown of lipids. Serum is thawed overnight at 2-8°C to prevent degradation; alternatively, it may be thawed at room temperature with periodic agitation to resuspend the contents. Thawed serum should be mixed thoroughly before being added to cell culture medium. Finally, temperatures above 37°C may compromise serum performance by accelerating the degradation of critical serum factors, and should be avoided in the lab.

Serum can be sensitive to environmental conditions beyond temperature. For example, high intensity/high energy wavelengths of light may degrade the activity of growth

factors, which is important to keep in mind when using serum near UV lighting intended to promote sterility in some biosafety cabinets. Typical laboratory lighting conditions will have little to no effect on serum integrity or performance.

Physicians may interpret a cloudy patient serum sample as indicative of lipemia, or excess fats in the blood. The same may be true of that cloudy serum in the lab—if bacterial contamination can be ruled out by culturing a sample, cloudy serum is most likely explained by the precipitation of lipids and lipoproteins. This can be minimized by aliquoting bulk stock into single-use tubes to avoid repeated freeze-thaw cycles. Although the presence of the clotting protein fibrin is sometimes blamed for serum cloudiness, serum is prepared from the post-coagulation fraction of blood, and therefore would not typically contain significant amounts of fibrin or other clotting factors.

It's also useful to realize that lipid content contributing to cloudiness in serum can vary significantly among species. For example (and perhaps not surprisingly), porcine serum is higher in lipid content than its bovine counterpart. So, serum from diverse biological sources requires unique handling and additional processing steps at the manufacturing facility in order to prepare the final product for use in the lab. Suppliers that offer sera from multiple species will therefore have differential protocols in place for processing raw materials in order to provide a consistent, appropriate serum product to the end user, regardless of lipid biology.

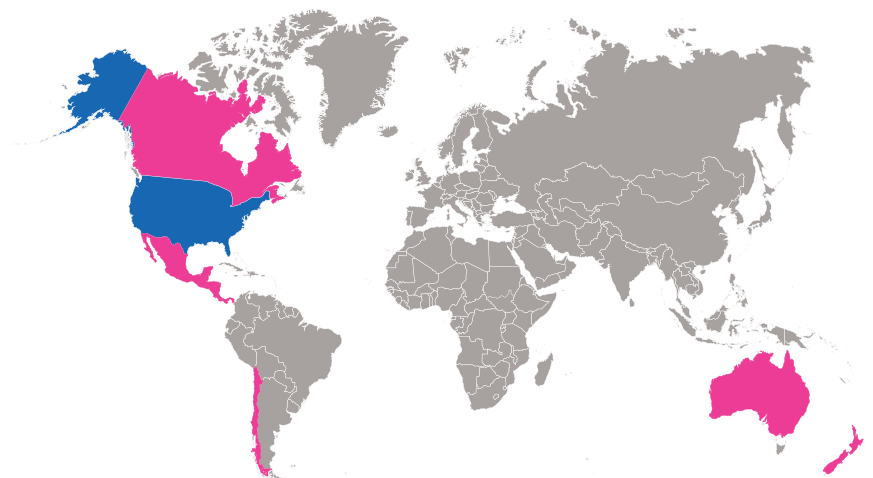
### World Geography: understanding sourcing and countries of origin

Because serum is a biological product, it is subject to variations that arise from the health of the source animals, which is in turn influenced by dietary, environmental, and global disease factors. One of the most important attributes knowledgeable serum product users therefore seek is the country of origin, or source.

Although the incidence and impact of various species-specific pathogens such as bovine spongiform encephalopathy (BSE) on serum are still being studied, the USDA maintains lists of countries approved for importation of serum that are based in part on up-to-date pathogen incidence data. Serum consumers should therefore always expect suppliers to provide source information for their sera and other products of animal origin.



Good practices such as aliquoting serum into single use tubes for frozen storage can reduce the potential for repeated exposure to contaminants. Aliquoting also helps to prevent cloudiness that can arise from the breakdown of lipids and lipid-associated macromolecules as a consequence of repeated freeze-thaw cycles.



The USDA maintains a list of countries approved for importation of serum into the U.S. (guidelines revised annually). This differs from U.S.-origin serum, which is collected from facilities in the U.S. that undergo continuous inspection by the USDA.

### International Regulatory Practices

Reputable suppliers will typically employ consensus testing methods specified by the serum industry to ensure that serum supplied to laboratories meets standards for levels of common contaminants that may impact its performance in cell culture. Some common tests are:

Contaminant/Parameter	Test	Standard (if specified)
<b>Endotoxin</b>	Limulus ameocyte lysate (LAL)	Test and record
<b>Bacteria/fungi</b>	Broth or solid culture of a serum sample	Not detected
<b>Viruses (various)</b>	PCR, anti-viral antibody screen, cytopathology, hemadsorption	Not detected
<b>mycoplasma</b>	LookOut® kit, mycoplasma detection kits/reagents	Not detected
<b>pH</b>	meter or standard laboratory pH measurement	6.5 – 8.5 (rec: 7.0-8.0)

Standards for the testing of general serum properties and components have also been established to enable consistent evaluation for osmolality, total protein, hemoglobin, and general performance in culture applications.



*Limulus polyphemus*, the Atlantic horseshoe crab, is particularly sensitive to the presence of endotoxins produced by the cell wall of Gram-negative bacteria that may also contaminate serum. Suppliers that follow industry guidelines will test for endotoxin in serum using the LAL (Limulus Ameocyte Lysate) assay.

### History: the importance of traceability and documentation in the serum industry

While “origin” indicates the country in which serum was collected, “traceability” refers to how well the journey of a serum product—from collection at site of origin through manufacturing and distribution—has been documented.

Serum delivered frozen to a processing facility is thawed and tested for the presence of bacteria, virus, endotoxin, and for hemoglobin content. The best quality serum is then filtered through a sequence of membrane filters, terminating with a 0.1 µm pore filtration under aseptic conditions. After filtration, aseptic procedure is maintained as serum is dispensed into sterile bottles, quickly frozen to -20°C and quarantined until all lots have met or exceeded quality control standards. A certificate of analysis for each serum lot should always be available upon request from the supplier.

Traceability for biologics is critical to ensure that label descriptions are an accurate representation of product history and contents. Complete documentation ensures that both the customer and regulatory agencies can audit the processing of serum products, and can uncover whether suppliers are engaging in serum pooling to offset local supplier cost fluctuations—a practice which can compromise performance.

### Economics 101: factors that may affect serum pricing

Serum is a commodity product for which price varies among countries of origin. For example, the cost of serum originating from Australia and New Zealand may be higher than that from other countries. The OIE’s current Geographical Risk Assessment rating for these countries is high because they continue to be free of BSE and FMD (foot and mouth disease), in part due to geographical isolation. This, coupled with strict agricultural regulatory practices, may increase demand for serum products from the ANZ region, which in turn can drive up price.

The availability of FBS and other animal sera is inherently subject to fluctuations in global animal disease states, climate, and food and water supply. These factors make industry-wide serum pricing sensitive in ways that pricing for non-biological reagents is not. And last—but certainly not least—as cell culture becomes more prevalent, particularly in the pharmaceutical sector, serum pricing is sure to reflect the changes in supply this demand may precipitate.

**For additional information**  
about serum for cell and tissue  
culture, please visit  
[sigmaaldrich.com/serumFAQs](https://sigmaaldrich.com/serumFAQs)

To request a sample, go to  
[sigmaaldrich.com/serumrequest](https://sigmaaldrich.com/serumrequest)



### Apoptotic Cell-Free DNA Amplification Kit

The TruePrime apoptotic cell-free DNA amplification kit from SYGNIS enables accurate DNA amplification using cell-free DNA (cfDNA) obtained from plasma, serum, urine, cerebrospinal fluid, and many other bodily fluids.

It uses a novel multiple displacement amplification method to overcome the common limitations with analysis of cfDNA from bodily fluids. By exponentially amplifying cfDNA derived from apoptosis (the cell-death mechanism), TruePrime kits provide users with excellent sensitivity, error-free amplification in high yields, and a streamlined workflow for reduced hands-on time.

#### SYGNIS

For info: +49-(0)-6221-3540-120  
www.sygnis.com

### Tissue Fixatives

Poly Scientific R&D offers a line of specialty tissue fixatives. STU-IHC is a tissue preservative designed to give researchers gently fixed tissues that provide unimpeded access to epitopes for immunohistochemical visualization. STUMol is designed for researchers seeking maximum recovery of DNA, RNA, and proteins from source tissues; it is an excellent fixative to preserve tissue for cell harvest by laser-capture microdissection. STU-NGS is designed for those performing NGS studies of tumor tissue; yield and quantity of recovered DNA from STU-NGS fixed tissue exceeds that of formalin. STU-ONE is an all-purpose nonformaldehyde tissue fixative; unlike formalin or paraformaldehyde, STU-ONE does not denature DNA, RNA, and proteins. Poly Scientific R&D will manufacture any stain or reagent to your specifications and ship directly to your laboratory, not to a distributor's shelf.

#### Poly Scientific R&D

For info: 800-645-5825  
www.polyrnd.com

### Flow Cytometer

Designed and developed to remove common barriers associated with flow cytometry, the Attune NxT flow cytometer offers adaptable optical configuration options so you can get the most out of your multicolor analysis. The innovative combination of acoustic focusing technology, coupled with traditional hydrodynamic focusing, results in acquisition speeds up to 10X faster than those of other flow cytometers. Attune NxT is ideal for immunophenotyping and signaling studies, cell-cycle analysis,

detection of rare events, stem-cell analysis, cancer and apoptosis studies, microbiological assays, and more. Its modular design offers the flexibility to meet your future requirements. Purchase what you need now and easily upgrade later. The instrument's small footprint requires minimal bench space—it can even be used in a cell culture hood.

#### Thermo Fisher Scientific

For info: 800-955-6288  
www.thermofisher.com

### Mycoplasma Test Kit

The EZ-PCR Mycoplasma Test Kit is a highly sensitive, specific PCR-based assay designed to detect a broad range of mycoplasma species in cell cultures and other biological materials. This kit contains an optimized, complete PCR mix, including a variety of mycoplasma-specific primers, deoxyribonucleotide triphosphate (dNTP) mix, and *Taq* polymerase. With the kit's simple protocol and ready-to-use format, samples can be prepared and run in about 10 minutes, with accurate results seen in only a few hours.

#### Biological Industries

For info: 860-316-2702  
www.bioind.com/ez-pcr-mycoplasma-test-kit

### Tissue Clearing System

The X-CLARITY Tissue Clearing System is an all-in-one, easy-to-use solution for electrophoretic tissue clearing. Its unique design accelerates the removal of lipids from tissues while preserving the structural integrity of the sample. The X-CLARITY system can clear a whole mouse brain in just 6 hours, an astounding eight times faster than the original technology. Moreover, it successfully produces transparent tissues that are efficiently penetrated by and labeled with macromolecules such as antibodies or oligonucleotides. This capability allows for 3D imaging of large tissues at single-cell resolution.

#### Logos Biosystems

For info: +82-(31)-478-4185  
www.logosbio.com

### Microplate for Microbial Research

The novel shape of the FlowerPlate microplate ensures turbulent mixing for biological suspensions and broths when incubated and shaken on a BioLector or BioLector Pro system. The flower shape acts similarly to baffles in shake flasks, and increases mixing and gas/liquid mass transfer. In addition to the unusual well shape—which is proprietary to this plate—the underside features a clear polystyrene base through which real-time physical measurements can be made during the shaking and incubation cycle, giving valuable kinetic information about these processes as they occur. The plates for online cultivations have a clear bottom and are available without optodes, with pH or dissolved oxygen (DO) optodes, or with both pH and DO optodes. The plates for offline cultivation are nontransparent.

#### m2p-labs

For info: 49-(0)-2401-80-53-30  
www.m2p-labs.com/bioreactors/microtiter-plates/flowerplate

Electronically submit your new product description or product literature information! Go to [www.sciencemag.org/about/new-products-section](http://www.sciencemag.org/about/new-products-section) for more information.

Newly offered instrumentation, apparatus, and laboratory materials of interest to researchers in all disciplines in academic, industrial, and governmental organizations are featured in this space. Emphasis is given to purpose, chief characteristics, and availability of products and materials. Endorsement by *Science* or AAAS of any products or materials mentioned is not implied. Additional information may be obtained from the manufacturer or supplier.



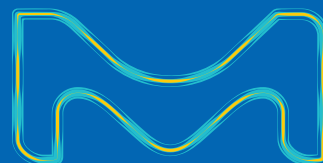
## AAAS MEMBERSHIP. MAKE THE CONNECTION.

**Join AAAS Sections.** They are the foundation of your AAAS membership

- **Be a subject-matter expert.**
- **Represent your discipline.**
- **Network with leaders in your field.**

**[aaas.org/sections](https://aaas.org/sections)**

AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE



# REFINED WITH YOU IN MIND

## Stericup® Quick Release Filtration Systems

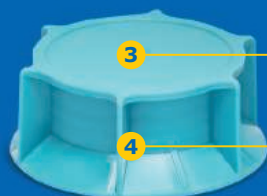
*Work With Ease. Filter With Confidence.*

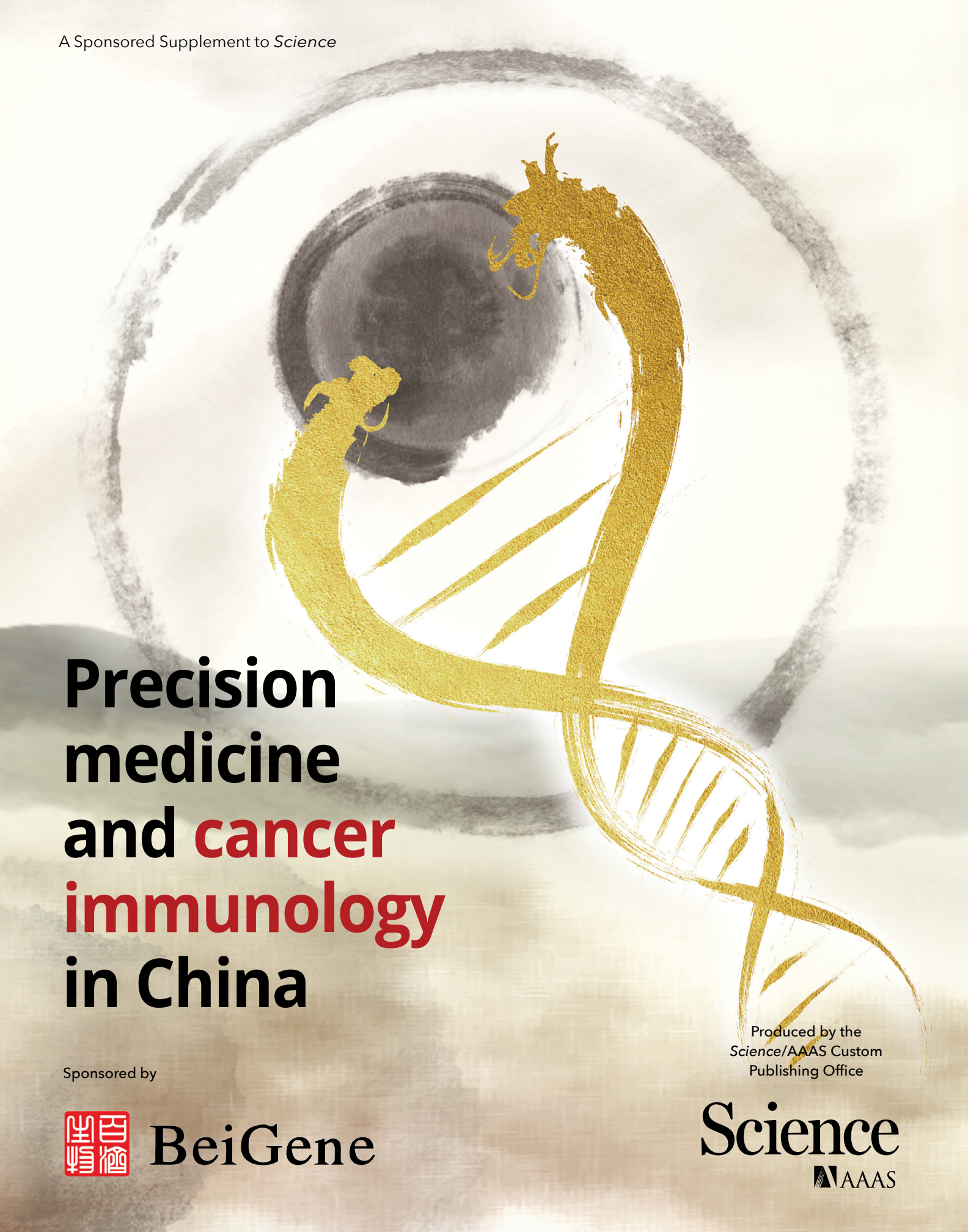
Stericup® Quick Release Filtration Systems streamline your workflow with ergonomic design updates and safeguard your results with the proven performance of Millipore® membranes.

- 1 Quarter-Turn Quick Release Funnel Removal
- 2 Frosted Writing Surface
- 3 Lighter Color for Legibility
- 4 Click-Seal Confidence Cap

### Explore New Features

[SigmaAldrich.com/stericupquickrelease](https://SigmaAldrich.com/stericupquickrelease)





# Precision medicine and **cancer** **immunology** in China

Sponsored by



BeiGene

Produced by the  
Science/AAAS Custom  
Publishing Office

Science  
AAAS

# Step up your job search with *Science* Careers



- Access thousands of job postings
- Sign up for job alerts
- Explore career development tools and resources



Search jobs on **ScienceCareers.org** today

# Precision medicine and **cancer** **immunology** in China



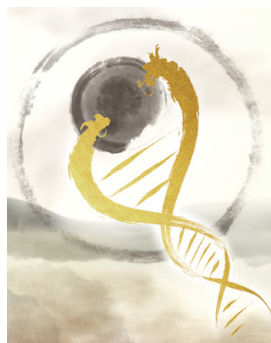
## Introductions

- 3 Foundational biomarkers for precision medicine  
Sean Sanders, Ph.D.  
Jackie Oberst, Ph.D.  
*Science/AAAS*
- 4 A global oncology company rooted in China  
Xiaodong Wang, Ph.D.  
Founder & Chairman of Scientific Advisory Board, BeiGene  
  
John V. Oyler  
Founder & CEO, BeiGene

## Articles

- 5 Targeted therapy for liver cancer: Challenges and opportunities  
Shuzhen Chen, Jing Fu, and Hongyang Wang
- 9 The challenges of radiation oncology in the era of precision medicine  
Ligang Xing and Jinming Yu
- 12 The role of multidisciplinary efforts in precision medicine and immunology for clinical oncology  
Jianzhen Lin, Anqiang Wang, Junyu Long *et al.*
- 14 Current status of immunotherapy in advanced HCC  
Shukui Qin
- 18 Challenges and prospects for precision cancer immunotherapy in China  
Zhihao Lu, Jianling Zou, Shuang Li *et al.*
- 24 The rise of engineered T-cell therapy in China  
Jianshu Wei, Yi Zhang, and Weidong Han

# Precision medicine and **cancer** **immunology** in China



About the cover: An artist's depiction of two traditional Chinese dragons surrounding a black pearl, symbolizing how precision medical treatment can overcome tumor cells using genetic research.

Cover: © 2018 Haitao Zhao (PUMCH)  
Design company: Jzhmed

This supplement was produced by the *Science*/AAAS Custom Publishing Office and sponsored by BeiGene, Ltd.

Editors: Sean Sanders, Ph.D.; Jackie Oberst, Ph.D.  
Proofreader/Copyeditor: Bob French  
Designer: Amy Hardcastle

Materials that appear in this supplement have not been peer-reviewed nor have they been assessed by *Science*. Articles can be cited using the following format: [AUTHOR NAME(S)] [CHAPTER TITLE] in Precision medicine and cancer immunology in China. (*Science*/AAAS, Washington, DC, 2018), p. [xx-xx].

## **Xiaoying Chu**

Director, Global Collaboration  
and Business Development, Asia  
xchu@aaas.org  
+86-131-6136-3212

## **Danny Zhao**

Regional Sales Manager, Asia  
dzhao@aaas.org  
+86-131-4114-0012

© 2018 by The American Association for the  
Advancement of Science. All rights reserved.  
2 February 2018

## Articles continued

- 29 Precision cancer medicine and immunology  
in China  
Xu-Chao Zhang and Yi-Long Wu
- 35 From big data to knowledge in precision  
medicine  
Dechao Bu, Shaoliang Peng, Haitao Luo *et al.*
- 38 The role of circulating cell-free DNA in  
the management of cancer in China  
Ying Hu, Yanhui Chen, Lei Zhang *et al.*
- 44 Next-generation sequencing-based  
testing for cancer precision medicine  
in China: A review of technologies  
and validation procedures  
Weifeng Wang, Weiwei Shi, Ming Yao *et al.*
- 49 ctDNA-NGS: The key to unlocking a  
molecular diagnostic revolution in the  
heart of Asia  
Ying Hou and Kang Ying
- 52 Adoptive cell transfer therapy: A  
strategic rethinking of combination  
cancer therapy  
Minghui Zhang



## Foundational biomarkers for precision medicine

Applications of this technology in the clinic are bringing researchers closer to real-time biomarker tracking.

One might argue that the concept of personalized medicine—in which a treatment is tailored for a specific individual based on their unique physiology as well as their specific disease and drug tolerance—is a foundational aspect of the ancient art of traditional Chinese medicine, practiced for centuries. Today, China’s medical practitioners depend less on ancient remedies and more on evidence-based practice. They bring with them an appreciation for the benefits and rationale of personalizing treatments to each patient. Therefore, to them, the shift from generalized therapies to precision medicine is perhaps an easier and more logical one to make than for those trained in westernized settings.

In order for precision medicine—the term that now appears to have dislodged “personalized medicine”—to be successful, accurate characterization of the patient is necessary. Various biomarkers provide the necessary data, collected through a variety of ‘omics techniques including next-generation DNA sequencing (genomics), analysis of protein levels in blood or tissues (proteomics), or determination of RNA levels (transcriptomics). However, identifying and characterizing biomarkers that accurately reflect a physiological state (normal or diseased), or response to a particular drug or therapy, has turned out to be challenging. Add to this the complication that biomarkers may differ between population groups, or indeed between individuals, and that tracking these biomarkers as the patient’s status changes can be onerous, and the future of precision medicine could be described as bleak.

This pessimistic outlook has not stopped researchers from pushing forward in their search for accurate and robust biomarkers. Big data analysis is helping, by providing a means to crunch millions of datapoints to yield associations that are not at first obvious. It is hoped that these associations will point to the presence of predictive biomarkers or potential targets for therapy, and also help to predict the risk of disease, ascertain the probability of positive clinical outcomes, and evaluate therapeutic efficacy. Such biomarkers are also the ultimate goal of many next-generation sequencing studies being performed on a range of samples, including tumor tissue and circulating cell-free DNA. Applications of this technology in the clinic are bringing researchers closer to real-time biomarker tracking, with implications for cancer detection and the development of safe, effective treatments.

New immunotherapy treatment modalities, such as the use of checkpoint inhibitors, cytokines, and chimeric antigen receptors, are being developed at an increasingly rapid pace, and the success of such therapies depends heavily on extensive knowledge of individual patients, for which high-quality biomarkers are especially important.

The articles presented in this booklet cover many of the topics above, with a focus on precision medicine research currently being performed in China. Researchers there are determined to overcome every obstacle to detecting and exploiting genomic and proteomic biomarkers in a clinical setting for the benefit of their patients. They also hope their insights will advance the practice of precision medicine both domestically and worldwide.

Sean Sanders, Ph.D.  
Jackie Oberst, Ph.D.  
Science/AAAS Custom Publishing Office



## A global oncology company rooted in China

Building on its scientific roots and research foundation in China, BeiGene has established global clinical development capabilities with a significant presence in the United States, China, and Australia.

**B**eiGene, Ltd. (NASDAQ: BGNE) is a global, commercial-stage biotechnology company focused on molecularly targeted and immuno-oncology cancer therapies. With a team of over 800 employees in China, the United States, and Australia, BeiGene is advancing a pipeline of novel small molecules, monoclonal antibodies, and combination therapies for cancer treatment. BeiGene also markets ABRAXANE (nanoparticle albumin-bound paclitaxel), REVLIMID (lenalidomide), and VIDAZA (azacitidine) in China under a license from Celgene Corporation.

BeiGene was founded in 2010 based on the premise that the confluence of two major developments—the revolutionary scientific breakthroughs in cancer medicine, and the emergence of the pharmaceutical market in China, where nearly a quarter of the world’s cancer population has limited access to innovative therapies—may allow new biotech leaders to emerge. With Beijing-based R&D, BeiGene recruits from China’s strong scientific talent pool and has developed a drug discovery platform incorporating tumor samples through local hospital collaborations. Its scientific advisory board consists of world-renowned scientists and clinicians and is chaired by Dr. Xiaodong Wang, cofounder of BeiGene, founding director and architect of China’s National Institute of Biological Sciences, and a member of the Chinese Academy of Sciences and the U.S. National Academy of Sciences.

Over the past seven years, BeiGene has discovered and advanced into clinical development four investigational drug compounds: Bruton’s tyrosine kinase (BTK) inhibitor zanubrutinib (BGB-3111), PD-1 antibody tislelizumab (BGB-A317), PARP inhibitor pamiparib (BGB-290), and RAF dimer inhibitor lifirafenib (BGB-283). Zanubrutinib is in registrational trials both globally and in China, and its global registration program includes a phase 3 head-to-head trial comparing BGB-3111 to ibrutinib, a currently approved BTK inhibitor, with the aim of demonstrating superior depth of response. Tislelizumab is the subject of a strategic collaboration with Celgene and is in registrational trials in China. BeiGene is also testing tislelizumab in combination with pamiparib and zanubrutinib, respectively. The company plans to initiate additional registrational trials of its assets, both in China and globally, and to advance additional preclinical assets into the clinic.

Building on its scientific roots and research foundation in China, BeiGene has established global clinical development capabilities with a significant presence in the United States, China, and Australia. In addition, the company has domestic manufacturing capabilities, including a multipurpose manufacturing facility in Suzhou and a commercial-scale biologics manufacturing facility under construction in Guangzhou, established through a joint venture with the Guangzhou Development District. Through its strategic collaboration with Celgene, BeiGene also recently acquired Celgene’s commercial operations in China and gained exclusive rights to commercialize Celgene’s three approved therapies there, which is expected to help BeiGene prepare for the potential future commercialization of its internally developed compounds and any additional in-licensed compounds in China. BeiGene aspires to be a global biotech leader and is committed to bringing new, potentially life-altering treatments to patients worldwide.

**Xiaodong Wang, Ph.D.**

Founder & Chairman of Scientific Advisory Board, BeiGene

**John V. Oyler**

Founder & CEO, BeiGene

*Certain statements found herein may constitute forward-looking statements that involve numerous risks and uncertainties that are described in BeiGene’s filings with the Securities and Exchange Commission, and are made only as of the date of this publication.*

# Targeted therapy for liver cancer: Challenges and opportunities

Shuzhen Chen<sup>1,2†</sup>, Jing Fu<sup>1,2†</sup>,  
and Hongyang Wang<sup>1,2\*\*</sup>

**L**iver cancer is the sixth most prevalent cancer and the second leading cause of cancer-related death worldwide (1). China alone accounts for over half of the new cases and deaths. It is estimated that in 2015 alone, 466,100 new cases of liver cancer were diagnosed in China and 422,100 deaths occurred there (2). Of all the cancers, the survival rate of liver cancer is the poorest—the age-standardized five-year relative survival rate is only 10.1% (3). Due to difficulties in early diagnosis, most liver cancer patients are diagnosed at an advanced stage, losing the opportunity for curative treatments such as liver resection or ablative procedures.

Fortunately, the development of innovative technology such as next-generation DNA sequencing has enabled a rapid and dramatic increase in our understanding of the genetic, molecular, and morphological changes occurring in individual cancer patients, laying the foundation for the emergence of targeted therapy. Although targeted therapies such as sorafenib treatment have raised hope for advanced liver cancer patients, their clinical benefits remain modest at best (4, 5). It is hoped that targeted therapy will provide functional and even structural corrections at the molecular level, or at least offer a valid alternative to conventional treatment. However, liver cancer is an extraordinarily heterogeneous disease, which makes it difficult to properly stratify patients for optimal targeted treatment and increases the risk of side effects, leading to the persistent failure of targeted therapy (6). In this review, we summarize the progress made in targeted therapy for liver cancer treatment in China and focus on the challenges and opportunities thereof.

## Sorafenib

Sorafenib is the first small-molecule targeted drug that has demonstrated a survival benefit in advanced

hepatocellular carcinoma (HCC) patients (5, 7). It is a multikinase inhibitor of several tyrosine protein kinases, including vascular endothelial growth factor receptor (VEGFR) and platelet-derived growth factor receptor (PDGFR). Sorafenib can also target intracellular serine and threonine kinase signaling such as RAF proto-oncogene kinase, including the C-Raf and B-Raf pathways (8–10). Sorafenib was approved as the only standard systemic treatment for HCC mainly on the basis of two studies: the Sorafenib HCC Assessment Randomized Protocol (SHARP) phase 3 trial (conducted in Europe, North America, South America, and Australasia), and a phase 3 randomized trial conducted in the Asia-Pacific region. According to these two studies, the sorafenib treatment group showed prolonged median survival: 10.7 months in the SHARP study compared to 7.9 months for the placebo group (5), and 6.5 months in the Asia-Pacific study compared to 4.2 months for the placebo (7). However, further analysis of these two trials and results from other studies have shown undesirable tolerability of sorafenib caused by its severe adverse events, including gastrointestinal, dermatologic, hematologic, cardiovascular, and nervous system side effects (11–16), making patients reluctant to continue the treatment.

Drug resistance is another bottleneck issue for sorafenib treatment. Numerous studies have revealed significantly differing responses to sorafenib due to tremendous variability in the way HCC progresses (17, 18). Chinese researchers have made serious efforts to decipher the mechanism underlying resistance to sorafenib and to identify potential biomarkers predictive of sorafenib response. A study by our team demonstrated that HCC patients with high 26S proteasome non-ATPase regulatory subunit 10 (PSMD10) expression had worse prognosis and a poor response to sorafenib therapy. The overall survival time of sorafenib-treated HCC patients with high levels of PSMD10 was much shorter than those with low PSMD10 ( $p=0.0099$ ), with the median survival time reduced by more than 40 months ( $p=0.0099$ ). These results suggest that PSMD10 may be a potential molecular marker in the classification of HCC patients who may not respond effectively to sorafenib (17). In another study, our clinical investigation revealed that HCC patients with low Src-homology 2 domain-containing phosphatase 2 (Shp2) expression benefited from sorafenib administration after surgery. This study showed that Shp2 could promote liver cancer stem cell expansion by augmenting  $\beta$ -catenin signaling and might be a useful indicator when determining chemotherapeutic strategies (18).

<sup>†</sup>International Co-operation Laboratory on Signal Transduction, Eastern Hepatobiliary Surgery Institute, Second Military Medical University, Shanghai, China

<sup>‡</sup>National Center for Liver Cancer, Shanghai, China

<sup>†</sup>Joint first author

<sup>\*</sup>Corresponding author: hywangk@vip.sina.com

The varied cellular metabolic phenotypes of tumor cells may also affect the efficacy of sorafenib. Investigation of the metabolic characteristics of tumor samples from 63 HCC cases showed huge variation in lipid content and glucose uptake. This study found that the rate-limiting enzyme acetyl-coenzyme A carboxylase alpha (ACC $\alpha$ ) enhanced glucose-derived de novo fatty acid synthesis (FAS) and promoted tumor cell survival under energy stress, which contributed to the heterogeneity of metabolic patterns in HCC. Inhibition of ACC $\alpha$ -driven FAS using a specific inhibitor, orlistat, improved the efficacy of sorafenib in xenograft-bearing mice, suggesting that interfering with ACC $\alpha$ -driven FAS could sensitize HCC cells to sorafenib (19). Another study has reported that blocking interleukin-6/signal transducer and activator of transcription 3 (STAT3)-mediated preferential glucose uptake could sensitize liver tumor-initiating cells to sorafenib treatment and enhance its therapeutic efficacy in vivo (20). These findings suggest that a combination of sorafenib and inhibitors of certain metabolic pathways could be a promising approach for some HCC patients.

### EGFR inhibitors

Epidermal growth factor receptor (EGFR) is overexpressed in 40%–70% of human HCCs, a factor that has been proven to be closely linked to the formation and growth of tumors. But EGFR inhibitors have shown disappointing results in clinical trials with unselected patients (21). A study was conducted in Taipei to evaluate the efficacy and safety of vandetanib, an oral inhibitor of both VEGFR and EGFR, in patients with inoperable advanced HCC. The study observed no significant difference in the rate of tumor stabilization or vascular change between the vandetanib group and the placebo group, suggesting that vandetanib had limited clinical activity in HCC (22). Other clinical trials with erlotinib, gefitinib, or cetuximab showed only limited effects in advanced stage HCC or modest effects at most in phase 2 trials (21). A better understanding of the mechanisms underlying how EGFR signaling influences HCC progression is therefore needed.

A study of the role of EGFR in HCC formation showed that the absence of EGFR in macrophages impaired the development of HCC in mice, whereas mice lacking EGFR in hepatocytes unexpectedly developed more HCCs due to increased compensatory proliferation after cell damage. Following inflammatory stimulation, EGFR induces interleukin-6 expression in liver macrophages, triggering hepatocyte proliferation and the development of HCC (23). This study demonstrated that EGFR has different

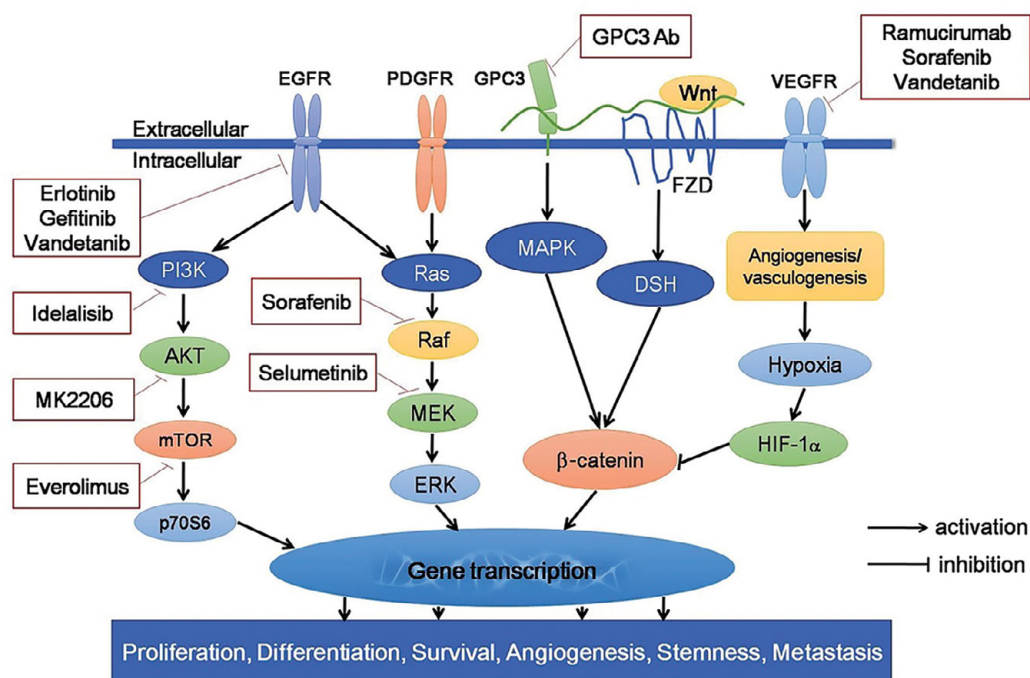
roles in tumor cells than in nontumor cells, providing some explanation of the disappointing results of anti-EGFR agents in HCC treatment. Other recent research by our team indicates that levels of choline kinase alpha (CHKA) are higher in human HCCs than in nontumor tissues, and that CHKA is associated with tumor aggressiveness and reduced overall survival. Further study has revealed that CHKA could facilitate a functional interaction between EGF and mammalian target of rapamycin complex 2 (mTORC2), which could contribute to HCC metastasis by promoting AKT (Ser473) activation. In this way, overexpression of CHKA promotes resistance to EGFR-targeted drugs (gefitinib and erlotinib) in HCC, suggesting that dual inhibition of CHKA and mTORC2 could be a way to overcome the resistance of HCC cells to EGFR-targeted therapies (24).

## Immunotherapy

### GPC3-based immunotherapy

Glypican-3 (GPC3) can be detected in 72% of HCCs, but could not be detected in normal hepatocytes, cirrhotic liver, or benign liver lesions (25). In addition to being a marker for HCC, GPC3 plays a role in the progression of the disease. It activates Wnt signaling and stimulates cell cycle progression and cell survival (26), indicating that anti-GPC3 therapy could be a therapeutic strategy for HCC treatment. The potential usage of GPC3-derived antibody or peptide vaccines has been explored in HCC immunotherapy (27–29). Disappointingly, in clinical trials these agents showed only limited curative effect (30).

Chimeric antigen receptor T (CAR-T) cells have been heralded as a breakthrough technology due to the substantial benefits observed in patients with relapsed or refractory B-cell malignancies. More than 200 CAR-T cell clinical trials have been initiated so far, most of which are CD-19 specific CARs aimed at treating lymphoma or leukemia (31). Researchers interested in HCC have mainly explored the possibility of redirecting T cells to recognize GPC3 for the treatment of HCC. T cells with CARs or high-affinity T-cell receptors (TCRs) targeting GPC3 were therefore engineered. Such targeted cells can efficiently recognize and destroy GPC3-positive human HCC cells in vitro and in vivo (32, 33). In a recent study, Li and colleagues developed T cells carrying two complementary CARs—against both GPC3 and asialoglycoprotein receptor 1 (ASGR1)—to reduce the risk of on-target, off-tumor toxicity, while maintaining relatively potent antitumor activity (34). These preclinical studies suggested that



**FIGURE 1.** Schematic representation of the signaling network and agents involved in targeted therapy for liver cancer. Erlotinib and gefitinib are tyrosine kinase inhibitors targeting EGFR. The PI3K/AKT/mTOR/p70S6 and Ras/Raf/MEK/ERK pathways are involved when EGFR signaling is activated, and can be blocked by independent inhibitors. Vandetanib and sorafenib are multikinase inhibitors of several tyrosine protein kinases such as VEGFR and PDGFR. GPC3 increases the binding of Wnt to FZD, which results in the stimulation of β-catenin transcriptional activity. Approaches to targeting GPC3 in HCC such as anti-GPC3 antibody or GPC3-targeted chimeric antigen receptor have shown some promise in different studies. EGFR, epidermal growth factor receptor; PDGFR, platelet-derived growth factor receptor; GPC3, glypican-3; Ab, antibody; VEGFR, vascular endothelial growth factor receptor; PI3K, phosphatidylinositol-3 kinase; mTOR, mammalian target of rapamycin; p70S6, ribosomal protein S6 kinase; MEK, MAPK/ERK kinase; MAPK, mitogen-activated protein kinase; ERK, extracellular signal-regulated kinase; DSH, dishevelled; FZD, frizzled; HIF-1α, hypoxia-inducible factor 1-alpha.

an open-label, noncomparative, phase 1/2 dose escalation and expansion trial. The study showed that nivolumab produced durable objective responses in long-term survival rates in patients with advanced HCC (37).

However, it is important to recognize that in previous studies the response rate to anti-PD-1 as a stand-alone therapy was 10%–30% overall, which included immunogenic tumors such as malignant melanoma that have a much higher response rate (36). One possible explanation for the low response rate might be the influence of molecules involved in immune escape other than PD-1 and PD-L1/2.

adoptive transfer of GPC3-specific T cells presents a promising therapeutic strategy for treating HCC. Anti-GPC3 CAR-T therapy is now undergoing clinical trials in China.

### Anti-PD-1/L1

Programmed cell death 1 (PD-1) is an immune coinhibitory receptor expressed on immune cells such as T cells, B cells, and natural killer (NK) cells. PD-1 suppresses antigen-specific T-cell activation through interaction with its ligand, PD-L1, which has been observed to be upregulated in tumor cells (35). Clinical trials of antibodies targeting PD-1 or PD-L1 for the treatment of HCC have shown some promising results (36). A recent report in *The Lancet* evaluated the safety and efficacy of PD-1 inhibitor nivolumab in patients with advanced HCC in

Therapy that combines antibodies to block three inhibitory immune checkpoint molecules—PD-1, T cell immunoglobulin and mucin domain 3 (TIM-3), and lymphocyte-activation gene 3 (LAG-3)—has already been reported to restore the immune response to tumor antigens of HCC-derived T cells (38). However, there is still a long way to go before such combined antibody therapy can be established and tested in clinical trials.

### Potential targeted therapies

Recent findings have also shed light on other potential targets for HCC treatment. The development of HCC is a multistep process with high intratumoral heterogeneity, including alterations in tumor microenvironment, signaling pathways, and energy metabolism patterns. In

previous reports, it has been noted that adenosine monophosphate (AMP)-activated protein kinase (AMPK) serves as an energy sensor in eukaryotic cells and plays a role in linking metabolism and cancer development (39–41). However, our recent research has demonstrated that activation of AMPK by the first-line medication metformin for the treatment of type 2 diabetes, not only inhibited HCC cell growth in vivo, but also augmented the growth inhibition induced by the chemotherapy drug cisplatin in these cells (42). Another study observed an imbalance of gut microflora as well as intestinal inflammation in chronic treatment of rats with the carcinogen diethylnitrosamine. Modulation of gut microbiota by probiotics dramatically mitigated liver tumor growth and spread in vivo (43). These studies indicate that an intervention strategy based on studies of HCC heterogeneity may present a new avenue for therapeutic intervention to treat the disease.

## Perspectives

An improved understanding of the molecular pathways that drive development of HCC has led to the identification of various biomarkers and the evaluation of several agents specifically targeted to tumor cells with particular molecular features. However, clinical trials undertaken worldwide have documented only occasional positive responses to such treatments. To date, no single agent or single targeted therapy has been formally found to be a cure for HCC in clinical trials. An increasing number of studies have demonstrated that intratumoral heterogeneity in individual patients is a roadblock for HCC targeted therapy. Therefore, the efficacy of targeted therapy requires a thorough understanding of the tumor microenvironment, metabolism, and gut microbiota of an individual. Meanwhile, combined therapies may be more effective than the administration of a single agent. IL-6 and PD-L1 blockade, or sorafenib combined with anti-PD-L1 monoclonal antibody, have demonstrated better efficacy than a single inhibitor in mouse models (44, 45). It is hoped that the establishment of combined therapies can offer a way to successfully manage HCC patients in the future.

## References

1. L. A. Torre et al., *CA Cancer J. Clin.* **65**, 87–108 (2015).
2. W. Chen et al., *CA Cancer J. Clin.* **66**, 115–132 (2016).
3. H. Zeng et al., *Int. J. Cancer* **136**, 1921–1930 (2015).
4. V. Di Marco et al., *Ann. Oncol.* **24** (Suppl. 2), ii30–ii37 (2013).
5. J. M. Llovet et al., *New Engl. J. Med.* **359**, 378–390 (2008).
6. A. Pribluda, C. C. de la Cruz, E. L. Jackson, *Clin. Cancer Res.* **21**, 2916–2923 (2015).
7. A. L. Cheng et al., *Lancet Oncol.* **10**, 25–34 (2009).
8. K. S. Smalley et al., *Oncogene* **28**, 85–94 (2009).
9. S. M. Wilhelm et al., *Mol. Cancer Ther.* **7**, 3129–3140 (2008).
10. G. M. Keating, A. Santoro, *Drugs* **69**, 223–240 (2009).
11. M. Kudo, *Liver Cancer* **1**, 62–70 (2012).
12. G. H. Choi et al., *Radiology* **269**, 603–611 (2013).
13. M. Iavarone et al., *Hepatology* **54**, 2055–2063 (2011).
14. S. Amar, K. J. Wu, W. W. Tan, *Mayo Clin. Proc.* **82**, 521 (2007).
15. J. Autier, B. Escudier, J. Wechsler, A. Spatz, C. Robert, *Arch. Dermatol.* **144**, 886–892 (2008).
16. D. T. Alexandrescu, R. McClure, H. Farzanmehr, C. A. Dasanu, *J. Clin. Oncol.* **26**, 4047–4048 (2008).
17. T. Luo et al., *Autophagy* **12**, 1355–1371 (2016).
18. D. Xiang et al., *Hepatology* **65**, 1566–1580 (2017).
19. M. D. Wang et al., *Hepatology* **63**, 1272–1286 (2016).
20. H. L. Zhang et al., *Cancer Lett.* **388**, 1–11 (2017).
21. S. Whittaker, R. Marais, A. X. Zhu, *Oncogene* **29**, 4989–5005 (2010).
22. C. Hsu et al., *J. Hepatol.* **56**, 1097–1103 (2012).
23. H. Lanaya et al., *Nat. Cell Biol.* **16**, 972–977 (2014).
24. X. M. Lin et al., *Gastroenterology* **152**, 1187–1202 (2017).
25. M. Capurro et al., *Gastroenterology* **125**, 89–97 (2003).
26. M. I. Capurro, Y. Y. Xiang, C. Lobe, J. Filmus, *Cancer Res.* **65**, 6245–6254 (2005).
27. M. Feng et al., *Proc. Natl. Acad. Sci. U.S.A.* **110**, E1083–E1091 (2013).
28. W. Gao et al., *Nat. Commun.* **6**, 6536 (2015).
29. Y. Sawada et al., *Clin. Cancer Res.* **18**, 3686–3696 (2012).
30. G. K. Abou-Alfa et al., *J. Hepatol.* **65**, 289–295 (2016).
31. J. Hartmann, M. Schussler-Lenz, A. Bondanza, C. J. Buchholz, *EMBO Mol. Med.* **9**, 1183–1197 (2017).
32. H. Gao et al., *Clin. Cancer Res.* **20**, 6418–6428 (2014).
33. C. Dargel et al., *Gastroenterology* **149**, 1042–1052 (2015).
34. C. Chen et al., *Cancer Immunol. Immunother.* **66**, 475–489 (2017).
35. Q. Gao et al., *Clin. Cancer Res.* **15**, 971–979 (2009).
36. M. Kudo, *Oncology* **92** (Suppl. 1), 50–62 (2017).
37. A. B. El-Khoueiry et al., *Lancet* **389**, 2492–2502 (2017).
38. G. Zhou et al., *Gastroenterology* **153**, 1107–1119 (2017).
39. D. G. Hardie, D. Carling, M. Carlson, *Annu. Rev. Biochem.* **67**, 821–855 (1998).
40. R. G. Jones, C. B. Thompson, *Genes Dev.* **23**, 537–548 (2009).
41. M. M. Mihaylova, R. J. Shaw, *Nat. Cell Biol.* **13**, 1016–1023 (2011).
42. L. Zheng et al., *Clin. Cancer Res.* **19**, 5372–5380 (2013).
43. H. L. Zhang et al., *J. Hepatol.* **57**, 803–812 (2012).
44. Y. Wang et al., *Tumor Biol.* **36**, 1561–1566 (2015).
45. H. Liu, J. Shen, K. Lu, *Biochem. Biophys. Res. Commun.* **486**, 239–244 (2017).

# The challenges of radiation oncology in the era of precision medicine

Ligang Xing and Jinming Yu\*

**R**adiotherapy is an essential treatment in the management of cancer. Seventy percent of all cancer patients will get radiotherapy as at least part of their treatment (1). Advances in physics, mathematics, computer science, electrical engineering, and radiobiology have significantly improved the safety, precision, and efficacy of radiotherapy, the concomitant control of tumor growth, and the probability of a cure for many cancer sufferers. In recent years, the sequencing of the human genome has paved the way for precision medicine, which aims to deliver “the right treatment to the right patient at the right time.” Although discoveries arising from studying the genome have affected the delivery of chemotherapy and targeted biological agents (2), they have yet to impact the clinical use of radiotherapy. In this new era of precision medicine, radiotherapy poses both great challenges and great opportunities for physicians and researchers.

## Progress in modern radiotherapy techniques

New technologies have been the main driving force behind innovations in radiation oncology over the last two decades. Technological advances have been put to clinical use, leading to better localization of the radiation dose and less damage to healthy tissue. These methods include technologies such as three-dimensional conformal radiotherapy (3D-CRT), intensity-modulated radiotherapy (IMRT), 3D brachytherapy, stereotactic radiotherapy, image-guided and adaptive radiotherapy (IGRT and ART), and charged particle radiotherapy using protons and carbon ions. Using 3D-CRT and IMRT, we can make the radiation conform to the shape of the target volume, solving the problem of irradiating complex targets that lie close to critical healthy structures. 3D-CRT and IMRT are routinely and

successfully applied for head and neck cancers, prostate cancer, and many other common cancers including those of the lung, liver, esophagus, and breast (3, 4). In the past 10 years, there has been rapid growth in clinical research on and application of stereotactic ablative body radiotherapy (SABR), also known as stereotactic body radiation therapy (SBRT), for cancers such as lung, liver, spine, pancreas, and prostate. For example, by accurately delivering high-dose radiation to the tumor, SBRT has emerged as the standard of care for medically inoperable stage I non-small cell lung cancer (NSCLC) and may even outperform surgery in operable patients (5, 6).

## Genomics for personalized radiotherapy

From 3D-CRT and IMRT to SBRT, two factors limit the efficacy of radiotherapy: (1) defining the target, or differentiating between tumor and normal tissue; and (2) determining the appropriate total dose of radiation and its “fractions”—the number of separate treatments into which the total dose is divided. In clinical practice, radiotherapy dosages and fractions have been determined empirically, which has resulted in reasonable disease control and acceptable levels of toxicity. However, results suggest that current radiotherapy dosing protocols can be further optimized with a modern precision oncology approach, such as gene profiling to detect relevant biomarkers or biomarker signatures that would inform clinicians of a particular patient’s sensitivity or resistance to radiotherapy. Existing tests that predict a patient’s sensitivity to radiation can be grouped into three categories: those that determine intrinsic radiosensitivity, those that determine tumor oxygen levels, and those that determine a tumor’s chance of growing (7). Unfortunately, none of these approaches is practical for clinical application. Radiotherapy is used in different settings depending on the site of the disease, so the clinical utility of a molecular biomarker signature indicating sensitivity to radiation would vary depending on the clinical application. The development of clinically relevant radiosensitivity molecular signatures is therefore challenging (8).

Recently, Scott and colleagues identified 10 genes that could index radiosensitivity (9). This could allow the radiation dose to be individually tuned to a tumor’s radiosensitivity and provide a framework for designing genomically guided clinical trials in radiation oncology. Being able to increase radiation dosage for more resistant tumors and lower it for more sensitive tumors would also

Department of Radiation Oncology and Shandong Key Laboratory of Radiation Oncology, Shandong Cancer Hospital Affiliated with Shandong University, Shandong Academy of Medical Science, Shandong, China  
\*Corresponding author: sdyujinming@126.com

lower the risk of complications from the therapy. It should be emphasized that tumor genomic data gives information only about a tumor's intrinsic radiosensitivity. Additional biological insights about the tumor and its microenvironment, as well as information about the patient, are also important for optimizing radiotherapy dosing.

### **Combining radiotherapy with targeted therapy**

Biomarkers that can predict tumor sensitivity to therapy are considered the gatekeepers necessary to develop precise and personalized medicine (10). These biomarkers are therapy-specific, and can therefore aid in therapeutic decision-making. For example, mutations in the epidermal growth factor receptor (EGFR) gene have been shown to predict the benefit derived from using tyrosine kinase inhibitors (TKIs), while anaplastic lymphoma kinase (ALK) gene rearrangements have been shown to predict the efficacy of ALK inhibitors in treating NSCLC (11, 12). Building on the preclinical rationale that inhibitors of EGFR function create a strong sensitivity to radiation (13), serial clinical trials have been conducted to test combinatorial treatments using EGFR inhibitors plus radiotherapy. In one pivotal phase 3 trial, adding the chemotherapy drug cetuximab to radiation improved localization of therapy in locally advanced head and neck squamous cell cancer and improved overall survival (14). However, phase 3 trials that evaluated cetuximab in combination with chemoradiotherapy for NSCLC and esophageal cancer all failed to improve overall survival in an unselected patient population (15, 16).

Phase 1 and 2 trials of EGFR TKIs in combination with radiotherapy for locally advanced NSCLC or metastatic NSCLC have shown a favorable safety profile and some encouraging outcomes (17, 18). However, these trials were all performed in patients without information on whether their tumors carry the EGFR mutation, making the results less informative. These studies highlight the need for predictive biomarkers in cases where targeted therapy is combined with radiotherapy.

### **Radiotherapy combined with modern immune-targeted therapy**

Understanding of the interaction between the immune system and tumor growth has led to the development of modern cancer immunotherapies. These include cancer vaccines, chimeric antigen receptor T-cell (CAR-T) therapy (in which immune system T cells are reengineered to act against a cancer), and immune checkpoint inhibitors,

which interfere with proteins that prevent T cells from responding to cancer. Immune checkpoint inhibitors targeting several types of proteins, including cytotoxic T-lymphocyte antigen-4, programmed death-1 (PD-1), or programmed death ligand-1 (PD-L1), have demonstrated clinical efficacy against a broad spectrum of tumor types—a significant step for both science and medicine (19). Early studies revealed that radiotherapy could provoke an immune response not only at the irradiated site, but also at remote, nonirradiated tumor locations—the so-called “abscopal effect.” Cell death in the irradiated tumor can enhance antitumor immunity by inducing the expression of certain antigens on tumor cells and by activating lymphocytes to attack the tumor. Preclinical and clinical studies have demonstrated the efficacy and safety of radiotherapy combined with immunotherapy (20, 21). Currently, clinical trials of such combined treatments for a variety of tumor types are underway.

Most modern immunotherapies are not yet cost-effective, especially for patients in China, and immune checkpoint inhibitor therapy is not sufficiently precise yet. A crucial step in refining these therapies is the identification of biomarkers that can predict a tumor's response to checkpoint blockades (22). The overexpression of the PD-L1 antigen, the presence of tumor-infiltrating immune cells, or a variety of molecules in the tumor's microenvironment may be important predictive biomarkers and are being extensively explored. However, they are not yet sufficiently predictive to allow them to be used to routinely stratify patients. Gene analysis is a new approach for judging the potential clinical benefit of checkpoint inhibitors (23), but further preclinical and clinical studies are necessary before it can be applied in clinical practice. In order to move the strategy successfully into the clinic, it is also critical to clarify the appropriate fractions and doses of radiotherapy and the suitable combinations of radiotherapy and immunotherapy (24).

### **Molecular image-guided precision radiotherapy**

Imaging plays a critical role in precision medicine, from screening and early diagnosis to guiding treatment, evaluating responses to therapy, and assessing the likelihood of disease recurrence (25). Rapid advances in imaging technologies permit better anatomical resolution and provide noninvasive measurements of functional and physiological properties of tissues and lesions at the molecular

level. The development and application of molecular imaging techniques brings new opportunities for creating more precise treatment.

Novel molecular imaging approaches are being developed and validated in many critical molecular pathways, such as glucose and amino acid metabolism, cell proliferation, hypoxia, angiogenesis, and receptor expression. The concept of “biological target volume” has been introduced as a factor when determining the intensity of radiotherapy needed for treatment (26). We are looking for the best way to use molecular imaging to guide radiotherapy for certain cancers, either to help define the target volume to be irradiated, or to aid in patient stratification. It was recently reported that an escalated radiation dose to treat a particular type of lung tumor detected by a mid-treatment positron emission tomography (PET) scan allowed clinicians to deliver higher-dose radiation to the more aggressive areas of the tumor and improve local control of tumor growth without increasing radiotherapy-induced lung toxicity (27). PET and computed tomography scans could also identify and delineate hypoxic areas that could be targeted for elevated dosing in lung cancer patients (28).

The role of cancer imaging in precision medicine is being explored from another angle as “radiomics,” which assesses a large number of imaging features that characterize the observable properties of a tumor, using descriptors beyond simply its size to predict clinical outcomes with increased prognostic power, or even correlate with gene expression profiles (29). This approach could be important in helping to stratify patients who are at risk of disease recurrence (30, 31).

In summary, precise radiation therapy is being explored at four different levels: (1) clinical features such as the molecular structure of the tissue, cancer stage, and tumor volume and location(s); (2) adaptive radiotherapy based on images collected during treatment; (3) biomarker-guided therapy; and (4) personalized radiotherapy delivery schedule (32). We believe that with multidisciplinary guidance, the strong support of science and technology, and an eye to cost-effectiveness, precision radiotherapy that incorporates radiobiology, bioinformatics, and molecular imaging will eventually be realized.

## References

1. S. M. Bentzen *et al.*, *Radiother. Oncol.* **75**, 355–365 (2005).

2. N. Roper, K. D. Stensland, R. Hendricks, M. D. Galsky, *Cancer Treat. Rev.* **41**, 385–390 (2015).
3. S. G. Chun *et al.*, *J. Clin. Oncol.* **35**, 56–62 (2017).
4. V. Gregoire *et al.*, *Oncologist* **12**, 555–564 (2007).
5. R. Timmerman *et al.*, *JAMA* **303**, 1070–1076 (2010).
6. J. Y. Chang *et al.*, *Lancet Oncol.* **16**, 630–637 (2015).
7. J. F. Torres-Roca, C. W. Stevens, *Cancer Control* **15**, 151–156 (2008).
8. J. F. Torres-Roca, *Pers. Med.* **9**, 547–557 (2012).
9. J. G. Scott *et al.*, *Lancet. Oncol.* **18**, 202–211 (2017).
10. W. S. Dalton, S. H. Friend, *Science* **312**, 1165–1168 (2006).
11. M. Maemondo *et al.*, *New. Engl. J. Med.* **362**, 2380–2388 (2010).
12. A. T. Shaw *et al.*, *Lancet. Oncol.* **18**, 874–886 (2017).
13. M. Wang *et al.*, *Cancer Res.* **71**, 6261–6269 (2011).
14. J. A. Bonner *et al.*, *New. Engl. J. Med.* **354**, 567–578 (2006).
15. M. Suntharalingam *et al.*, *JAMA Oncol.* **3**, 1520–1528 (2017).
16. J. D. Bradley *et al.*, *Lancet Oncol.* **16**, 187–199 (2015).
17. R. Komaki *et al.*, *Int. J. Radiat. Oncol. Biol. Phys.* **92**, 317–324 (2015).
18. N. Ready *et al.*, *J. Thorac. Oncol.* **5**, 1382–1390 (2010).
19. X. Meng *et al.*, *Cancer Lett.* **405**, 29–37 (2017).
20. M. B. Bernstein, S. Krishnan, J. W. Hodge, J. Y. Chang, *Nat. Rev. Clin. Oncol.* **13**, 516–524 (2016).
21. R. C. Walshaw, J. Honeychurch, T. M. Illidge, *Br. J. Radiol.* **89**, 20160472 (2016).
22. X. Meng, Z. Huang, F. Teng, L. Xing, J. Yu, *Cancer Treat. Rev.* **41**, 868–876 (2015).
23. D. T. Le *et al.*, *Science* **357**, 409–413 (2017).
24. F. Teng, L. Kong, X. Meng, J. Yang, J. Yu, *Cancer Lett.* **365**, 23–29 (2015).
25. A. Giardino *et al.*, *Acad. Radiol.* **24**, 639–649 (2017).
26. C. C. Ling *et al.*, *Int. J. Radiat. Oncol. Biol. Phys.* **47**, 551–560 (2000).
27. F. M. Kong *et al.*, *JAMA Oncol.* **3**, 1358–1365 (2017).
28. P. Vera *et al.*, *J. Nucl. Med.* **58**, 1045–1053 (2017).
29. E. Segal *et al.*, *Nat. Biotechnol.* **25**, 675–680 (2007).
30. E. Huynh *et al.*, *Radiother. Oncol.* **120**, 258–266 (2016).
31. X. Dong *et al.*, *Oncotarget* **8**, 14969–14977 (2017).
32. J. Y. Jin, F. M. Kong, *Adv. Exp. Med. Biol.* **890**, 175–202 (2016).

## Acknowledgments

This review was supported by grants from the National Health and Family Planning Commission of China (201402011), the National Key Research and Development Project of China (2016YFC0904700), the National Natural Science Foundation of China (81572970), and the Innovation Project of the Shandong Academy of Medical Science.

# The role of multidisciplinary efforts in precision medicine and immunology for clinical oncology

Jianzhen Lin<sup>1</sup>, Anqiang Wang<sup>2</sup>,  
Junyu Long<sup>1</sup>, Haohai Zhang<sup>1</sup>, Yi Bai<sup>1</sup>,  
Xiaobo Yang<sup>1</sup>, Jie Pan<sup>3</sup>, Ke Hu<sup>4</sup>, Lin Zhao<sup>5</sup>,  
Xinting Sang<sup>1</sup>, and Haitao Zhao<sup>1,6\*</sup>

**C**ancer is a group of highly heterogeneous diseases in which dysregulation of key cellular processes leads to neoplastic transformation. Tumor immunity and suppression of the immune response within the tumor microenvironment support its progression. In the past few years, dramatic advances in DNA sequencing technologies have facilitated key insights into the genomic alterations and somatic mutations that enable cancer formation and progression, allowing for significant advances in precision medicine. The broad lexicon of precision medicine describes the narrowing of medical care to the characteristics of an individual patient. It improves upon the current approach of stratifying patients into treatment groups based only on phenotypic biomarkers, and instead makes use of a patient's molecular information (including genomics and proteomics) to inform diagnosis, prognosis, treatment, and disease prevention for that individual.

Oncology is at the frontline of precision medicine. It is evolving from the previous model of administering cancer therapeutics through a unified treatment regimen based on cursory tumor classification, to one that applies the precise molecular profile of an individual's cancer genome to optimize personalized treatment. Another new and highly successful therapeutic approach to cancer, immunotherapy (*Science's* 2013 Breakthrough of the Year), has created great excitement for clinicians, patients, researchers, and scientific journals. Current strategies in cancer immu-

notherapy include chimeric antigen receptor (CAR) T-cell therapy, immune checkpoint inhibitors, and cancer vaccines. These treatments have demonstrable clinical benefits and show great promise. However, there are still some challenges that precision medicine and immuno-oncology must confront.

For precision medicine the most pressing are:

## 1. Tumor heterogeneity and molecular evolution

Tumors exhibit tremendous genetic heterogeneity, both among different kinds of tumors and within a single tumor. This creates phenotypic variation, posing a significant challenge to personalized cancer medicine. Intratumor heterogeneity increases the complexity of cancer prognosis and likely contributes to tumor metastasis under therapeutic pressure. Exploring genomic alterations by time series analysis can identify the factors that drive a tumor's evolution during treatment, which may identify the molecular targets of resistance or tumor progression. Rapid advances in technology for studying tumor heterogeneity would help us understand the tumor genomic landscape. Among such technologies are liquid biopsy assays of circulating tumor DNA, detection and analysis of circulating cancer stem cells, and multiregion next-generation sequencing.

## 2. Drug resistance

Understanding the clonal make-up of tumors, their molecular evolution, and their response and resistance to drugs poses perhaps the greatest challenge, not only to the application of traditional therapies, but also to personalized cancer medicine. To improve an individual's treatment response and clinical outcomes, clinicians are obliged to inspect the evolving nature of the cancer genome. Moreover, combinational cancer treatments that target complementary signaling pathways, or harness the immune system through immunotherapy, have promise to overcome resistance while improving efficacy.

## 3. Shortage of agents that target specific genomic aberrations

Precise targeted therapy, which is based on the genomic characteristics of an individual patient and the mutations identified in their particular cancer, is limited by a shortage of biochemical agents that target specific genomic aberrations. In clinical practice, it is common to detect significant gene mutations in a cancer patient but to have no existing drug that can target those mutations. Or, as also happens, the targeted agent used has little antitumor efficacy. This may be a result of tumor heterogeneity (in which clones of resistant tumor cells survive and proliferate)

<sup>1</sup>Department of Liver Surgery, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College (PUMC), Beijing, China

<sup>2</sup>Department of Gastrointestinal Surgery, Peking University Cancer Hospital & Institute, Beijing, China

<sup>3</sup>Department of Radiology, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and PUMC, Beijing, China

<sup>4</sup>Center of Radiotherapy, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and PUMC, Beijing, China

<sup>5</sup>Department of Medical Oncology, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and PUMC, Beijing, China

<sup>6</sup>Center of Translational Medicine, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and PUMC, Beijing, China

\*Corresponding author: zhaoh@pumch.cn

and the variability in the drug's effect on the targeted genes, or the presence of nonidentical molecular pathways in different patients due to inherent physiological and biochemical differences.

For cancer immunotherapy treatments, the most pressing challenges are:

### **1. Lack of a definitive response to treatment**

In many immunotherapy trials, durable responses and extended, long-term survival are seen only in a subset of patients, and it has been reported in some clinical cases that immunotherapy treatment could even promote cancer progression. These challenges highlight the importance of identifying distinguishing clinical and molecular characteristics that can explain the differential response. Clues to what factors might predict a patient's response to therapy must be sought, and clinicians must be cognizant of overprescription in refractory patients.

### **2. Identification of suitable patients**

The limitations of immunotherapy approaches have aroused researchers' attention. CAR-T therapy, for example, has had clinical success against lymphocytic malignancies, but not against more common solid tumors. Even though antibodies that modify immune function can help patients defend against tumors, overall response rates are still relatively low, ranging from 10% to 30%. Thus, precisely identifying which patients are likely to benefit from immunotherapy is an important issue to address. Identifying unique clinical, genomic, and molecular patient characteristics will help researchers identify those factors that might predict response, and presents an opportunity to heighten response in refractory patients. Biomarkers are crucial to identifying suitable patients, but the complexity of the immune system makes their development a challenge.

### **3. Improving efficacy using combinational therapies**

There is considerable current interest in combining treatment modalities to improve the effectiveness of immunotherapies in a broad array of patients. This includes immunotherapies used in conjunction with chemotherapy, radiotherapy, targeted therapy, and even other immunotherapies. But caution is recommended. Our understanding of biological mechanisms leads us to expect additive or synergistic responses from such combined treatments. However, the complexity of the immune response could lead to unexpected, even serious, adverse events. Crucially, immunomodulatory therapies carry distinct risks, including autoimmune reactions that can be fatal if doctors are not alert and ready with proper treatment.

To overcome these tough challenges, precision medicine and immunology for clinical oncology require multidisciplinary efforts. Large-scale genomic data needs to be integrated with clinical data, analyzed, and translated into information that can guide clinical decisions. Multidisciplinary efforts can also ensure the rational application of cancer immunotherapy and combinational therapies by using multi-omics data to improve the diagnosis, prognosis, and treatment of cancer.

The Department of Liver Surgery and the Center for Translational Medicine at Peking Union Medical College Hospital have been focused for several years on researching precision tumor medicine and immunotherapy and their translation to clinical use. As personalized cancer medicine moves to the clinic, conceiving of cancer as a systemic, highly heterogeneous and complex disease becomes even more apt. Quality cancer care demands that we form a multidisciplinary team (MDT) of highly qualified health care professionals, with medical oncologists at its core.

We see the characteristics of our MDT as including:

#### **1. A multidisciplinary approach**

Our MDT members come from specialties including hepatobiliary surgery, radiology, radiotherapy, medical oncology, and pathology. Also on the team are experts in cancer biology and bioinformatics who can help interpret information on genomic aberrations. We strive to translate dispersed knowledge into an integrated, coherent, and personalized treatment regimen.

#### **2. A patient-centric model**

Patients' perceptions of their care are created by the quality of the care, the outcome of the treatment, the empathy displayed by the physician and health care team during their interactions, and each patient's individual world view. Therefore, the consulting model of our MDT practice is patient-centric. This means that all decisions consider the relationship between the patient, his or her family, and our health care team. We provide face-to-face counseling for every patient, and guarantee no less than a half-hour consultation for deciding on a personalized therapeutic regimen and answering patient questions.

#### **3. Guideline-first decision making**

A core principle of our MDT decision-making is "guideline-first," which means that every decision is evidence-based, and that each patient is treated according to clinically accepted and approved guidelines. The goal is to provide appropriate therapy for cancer patients to prolong their overall survival.

#### 4. A personalized therapeutic regimen

Personalizing cancer care implies using MDTs for clinical decision-making, since personalized care requires input from a range of different scientific and care domains. In our teams, clinical information and genomic profiling results are reviewed by professionals to identify clinically significant results. Patients for whom traditional therapies such as surgery, chemotherapy, or radiotherapy have failed, or who have elected to give up such treatments to try targeted therapy, are matched with a targeted therapy or immunotherapy regimen if one is available. These treatments might be administered through a clinical trial or using a drug approved by the China Food and Drug Administration or the U.S. Food and Drug Administration. The team also creates personalized therapeutic regimens using cancer immunotherapy or combinational therapy based on individual clinical characteristics, genomic aberrations, and the specific microenvironment of each tumor.

#### 5. Public welfare

The operation of our MDTs is supported by the Chinese Precision Medicine and Immunotherapy for Clinical Oncology Fund, which is a public welfare organization affiliated with the China Social Welfare Foundation. We primarily provide free services to hepatobiliary cancer patients, including multidisciplinary consultation, tumor genomic sequencing, and protein expression analysis to identify immune system biomarkers. Our mission is firstly to raise awareness of both current achievements using targeted therapies and immunotherapy, and of the potential and limitations of these therapies; and secondly to guide patients in seeking out clinical trials where their tumors can be better profiled and they can gain access to novel treatments. We also plan to develop more clinical trials to implement personalized therapeutic regimens.

Despite some obvious challenges, the largely encouraging clinical results from targeted therapeutics and biomarker-guided clinical trials are fueling further technological advancements in next-generation sequencing technology, data interpretation, and the associated preclinical and clinical cycles of drug development. All of this translates into a major shift in clinical practice. Moreover, cancer immunotherapy, especially using single-antibody immune checkpoint inhibitors, has shown promising efficacy against many solid cancers, and significantly improves the response rate against several solid cancers when used in combination with targeted drugs. These advances will improve the survival and quality of life of many cancer patients in the near future.

## Current status of immunotherapy in advanced HCC

Shukui Qin\*

Primary liver cancer, consisting mainly of hepatocellular carcinoma (HCC), is one of the most common cancers worldwide, with especially high prevalence in China and a growing incidence: 782,000 new cases were reported worldwide in 2012 (1, 2). As one of the leading causes of cancer-related deaths, HCC was responsible for an estimated 746,000 deaths worldwide in 2012 (2). China alone accounted for more than 50% of both new HCC cases and HCC-related deaths globally (3). Chronic infection due to hepatitis B virus (HBV) or hepatitis C virus (HCV) contributes to an estimated 75% of all HCC cases (4). In the Asia-Pacific region, more than 75% of cases are associated with HBV infection (5).

Surgical resection or orthotopic liver transplant (OLT) offer the best chance for successful treatment of HCC (6). However, surgical resection, liver transplantation, and radiofrequency ablation (RFA) are only applicable to a small portion of patients with well-preserved liver function who have early-stage or localized HCC (7). Due to difficulties in early diagnosis, most HCC cases are locally advanced or show distant metastases at the time of diagnosis. These patients generally have a poor prognosis, with median survival of 6 to 20 months (8), and a five-year survival rate of less than 16% (9).

For local advanced or metastatic HCC, systemic therapy is often used as an important palliative treatment. Various conventional systemic chemotherapy regimens such as doxorubicin have been used clinically, although there are few well-controlled studies demonstrating the efficacy of systemic chemotherapy in the management of HCC (10). In 2012, our group was the first to demonstrate that a regimen of "FOLFOX4" (oxaliplatin, 5-fluorouracil, and leucovorin) significantly improved the objective response rate (ORR) and prolonged overall survival (OS) compared with doxorubicin alone in a randomized phase 3 clinical trial in Chinese HCC patients (EACH

People's Liberation Army Cancer Center of Bayi Hospital affiliated with Nanjing University of Chinese Medicine, Nanjing, Jiangsu, China  
\*Corresponding author: qinsk@csco.org.cn

study) (11). Based on this trial, the FOLFOX4 regimen was approved by the Chinese Food and Drug Administration (CFDA) for use in first-line treatment of advanced or metastatic HCC patients in China.

Besides chemotherapy, molecular target therapies have also been evaluated in clinical practice globally. In 2007, sorafenib, a tyrosine kinase inhibitor against multiple targets including RAF, vascular endothelial growth factor receptor, and platelet-derived growth factor receptor, was shown to extend median overall survival in patients with advanced HCC, and was approved by the U.S. Food and Drug Administration (FDA) and the European Medicines Agency (EMA) as a first-line treatment for advanced HCC. However, its use in the clinic has been limited due to low ORR, limited survival benefit, toxicity, and cost. After sorafenib, a series of molecular-targeted drugs including sunitinib, brivanib, lapatinib, linifanib, everolimus, axitinib, pazopanib, tivantinib, and ramucirumab were studied. Unfortunately, all failed in phase 3 trials. Additionally, trials combining either sorafenib with erlotinib or sorafenib with transarterial chemoembolization were also unsuccessful. In the decade since the approval of sorafenib, only two phase 3 studies, both involving multikinase inhibitors, generated positive results: regorafenib versus placebo as a second-line treatment (RESORCE study, ClinicalTrials.gov: NCT01774344) (12) and lenvatinib versus sorafenib as a first-line treatment (REFLECT study, ClinicalTrials.gov: NCT01761266) (13). Most recently, the U.S. FDA approved regorafenib for advanced HCC patients who have been previously treated with sorafenib. Lenvatinib is currently under regulatory review by FDA and EMA for first-line treatment of advanced HCC patients.

Although there are a growing number of molecular target therapies and a few cytotoxic drugs under rapid development, clinical outcomes have been unsatisfactory to date. There is a great unmet medical need for novel therapies with better response rates and survival.

### Immunotherapies for HCC

Tumor immunotherapies, especially immune checkpoint inhibitors, have recently been established as effective treatments for several types of cancers. Immune checkpoint inhibitors release the “brakes” on the immune system and restore the ability of immune cells to eliminate cancer cells. Several immune checkpoint targets such as programmed cell death protein 1 (PD-1), programmed death ligand 1 (PD-L1), and cytotoxic T-lymphocyte

antigen 4 (CTLA-4) are being investigated extensively in various types of cancers. In the past few years, many clinical studies have shown immunotherapy to be a promising treatment option for many solid tumors and hematologic malignancies. Some of these agents are now approved for the treatment of melanoma, non-small cell lung cancer, renal cell carcinoma, head and neck cancer, bladder cancer, liver cancer, gastric cancer, mismatch repair (MMR)-deficient cancer, Merkel cell carcinoma, and Hodgkin’s lymphoma.

The human liver has a unique immunobiology whereby multiple regulatory mechanisms are in place to maintain an immunosuppressive environment. Normal liver tissue is inherently tolerogenic to environmental autoantigens and toxins in order to prevent aberrant immunity to pathogens encountered through arterial circulation and from the gut (14, 15). Clinical and nonclinical data showed increased numbers of immunosuppressive cells in HCC, including regulatory T cells and myeloid-derived suppressor cells, as well as increased expression of inhibitory signaling molecules such as CTLA-4 and PD-1 (16–18). HBV and HCV infections have also been associated with the increased proliferation of regulatory T cells and upregulation of PD-L1/PD-1 expression, thereby suggesting a role for this pathway in HBV or HCV-mediated hepatocarcinogenesis (15, 18–21). Overexpression of PD-L1 or PD-L2 has also been shown to be associated with tumor aggressiveness, disease progression, and high mortality in HCC patients (17, 22). Therefore, therapeutic agents that target immune checkpoints may potentially improve clinical outcomes by reversing the immunosuppressive nature of the HCC tumors and stimulating host immunity against the tumor cells. Interestingly, encouraging antitumor activity has now been reported in several clinical studies with anti-CTLA-4, anti-PD-1, and anti-PD-L1 antibodies.

Tremelimumab is a fully humanized immunoglobulin G2 (IgG2) monoclonal antibody that blocks CTLA-4. In a phase 2 study reported in 2013, of 21 HCV-positive patients, 17 were assessable for tumor response (23). The tumors shrank or disappeared in three patients (17.6%), and the disease control rate was 76.4%, with a median time to progression of 6.48 months. This study provided the first evidence of the antitumor activities of immune checkpoint inhibitors in HCC, and greatly encouraged further investigations and clinical trials.

Several anti-PD-1 antibodies are currently being evaluated in HCC patients globally. In September

2017, the U.S. FDA approved nivolumab for the treatment of patients with HCC who have been previously treated with sorafenib. Accelerated approval for this indication was granted based on the tumor response rate and durability of response observed in the CheckMate-040 trial (24). In this trial, 14.3% [95% confidence interval (CI): 9.2-20.8; 22/154] of patients responded to treatment with nivolumab. The percentage of patients with a complete response was 1.9% (3/154), and the percentage with a partial response was 12.3% (19/154). Among responders (n=22), duration of response ranged from 3.2 to 38.2+ months; 91% of those patients had responses lasting six months or longer and 55% had responses lasting 12 months or longer. The median time to response was 2.8 months (range: 1.2-7.0). The overall response rate (based on modified RECIST) criteria was 18.2% (95% CI: 12.4-25.2; 28/154). Complete response rate was 3.2% (5/154) and partial response rate was 14.9% (23/154), also based on modified RECIST criteria. A phase 3 randomized trial (CheckMate-459) is now ongoing to assess the clinical activity of nivolumab versus sorafenib in first-line HCC treatment (ClinicalTrials.gov: NCT02576509) (25). Another anti-PD-1 antibody, pembrolizumab, is also under investigation for treatment of HCC patients. The KEYNOTE-240 study is a phase 3 study to assess pembrolizumab versus placebo plus best support care as a potential second-line therapy in patients with previously systemically treated advanced HCC (ClinicalTrials.gov: NCT02702401) (26). This study is currently ongoing.

In addition to anti-PD-1 antibodies, anti-PD-L1 antibodies are also being evaluated in HCC patients. A phase 1/2 clinical trial of durvalumab in predominantly sorafenib-pretreated HCC patients showed an ORR of 10%, with a median OS of 13.2 months and a well-tolerated safety profile (ClinicalTrials.gov: NCT01693562) (27). A separate phase 1/2 study of durvalumab in combination with tremelimumab in unresectable HCC showed an ORR of 25%, with no unexpected safety signals (ClinicalTrials.gov: NCT02519348) (28). Overall, the trial demonstrated that a regimen of durvalumab plus tremelimumab was well-tolerated in this unresectable HCC patient population.

Many combinations of immunotherapies with molecularly targeted therapies are being evaluated in advanced HCC patients. Examples include pembrolizumab/lenvatinib (ClinicalTrials.gov: NCT03006926) (29), nivolumab/galunisertib (ClinicalTrials.gov: NCT02423343) (30), nivolumab/

yttrium Y 90 glass microspheres (ClinicalTrials.gov: NCT02837029) (31), and nivolumab/cabozantinib (ClinicalTrials.gov: NCT 01658878) (32).

In addition to checkpoint inhibitors, chimeric antigen receptor T cell (CAR-T) therapy has also made significant progress in the field of cancer immunotherapy. Remarkable clinical outcomes have been shown with CAR-T treatment, especially CD-19-directed CAR-T in various hematologic malignancies (33). However, to date, the clinical activity of adoptive CAR-T treatment in solid tumors has been less impressive. In HCC, CAR-T therapies targeting various antigens, including CEA, MUC1, GPC3, EGFR, EpCAM, and CD133, are under investigation in early stage clinical trials (34).

### Immunotherapies for HCC in China

HCC remains a cancer with a high mortality rate and a clear unmet medical need in China, despite the fact that multiple HCC prevention programs supported by the Chinese government have been in place for many years (35, 36). Based on data retrieved from the China National Central Cancer Registry, estimated new cases of liver cancer numbered about 356,000 in China in 2011, and the incidence rate was 26.39 per 100,000. There was also an increasing trend in the incidence rate of liver cancer in China from 2000 to 2011. As a result, the burden of liver cancer is still very high in China (37). Development of effective therapies for HCC patients remains a big challenge and an important unmet medical need.

With the boom in the Chinese biotech industry over the past five years, several China-based biopharmaceutical companies have been actively developing innovative immuno-oncology drugs for the treatment of cancers, including HCC. BGB-A317, developed by BeiGene, is an anti-PD-1 antibody engineered to minimize Fc-gamma receptor binding on macrophages, with the aim of abrogating antibody-dependent phagocytosis, a potential mechanism of T-cell clearance. Results from its ongoing phase 1 study were recently reported (38). At the time of data cutoff, 40 patients with unresectable HCC had been enrolled. A majority (85%) of these patients were infected with HBV. The median treatment duration was 64 days, with a range of 1 to 471 days. BGB-A317 was well tolerated: Twenty evaluable patients remained on the treatment at the data cutoff date; partial responses were observed in three patients (all HBV-positive) and nine patients achieved stable diseases, some of whom also had significant reductions in  $\alpha$ -fetoprotein (AFP) levels, which indicates a positive

treatment response. Based on these initial results, a randomized, open-label, multicenter phase 3 study has been initiated to assess the efficacy and safety of BGB-A317 versus sorafenib in first-line patients with unresectable HCC.

In addition, SHR-1210, a humanized anti-PD-1 IgG4 antibody developed by Jiangsu Hengrui Pharmaceutical Co., is under investigation either as a single agent or in combination with apatinib or FOLFOX4 in HCC. There are also at least 15 other anti-PD-1 and anti-PD-L1 antibodies from different pharmaceutical companies that have either received authorization from CFDA or are currently under review to initiate clinical studies in China (39).

Glypican 3 (GPC3), a member of the glypican family of heparan sulfate proteoglycans, is highly expressed on the surface of liver cancer cells with minimal expression on normal tissues (40). A clinical trial was initiated by Chinese investigators from Renji Hospital in collaboration with CARsgen Therapeutics to investigate GPC3-targeted CAR-T in Chinese patients with GPC3-positive refractory or relapsed HCC. Its preliminary phase 1 results showed that GPC3-targeted CAR-T was well tolerated in 13 Chinese patients with GPC3-positive HCC (41). Of these 13 patients, eight were treated with lymphodepleting conditioning (LDC) at baseline. Among these eight patients, there was one partial response, three showed stable diseases, two had progressive diseases, and two were unevaluable. Taken together, GPC3-positive CAR-T therapy showed encouraging preliminary clinical activity in Chinese patients with advanced HCC.

## Conclusions

HCC remains one of the most devastating malignancies in the world, and a disease with significant unmet medical need, particularly in China. Immunotherapies, especially immune checkpoint inhibitors, have demonstrated preliminary but promising clinical activity in patients with HCC, and several large randomized phase 3 clinical trials of these treatments are currently underway. Despite the limited progress and challenges in the development of HCC treatment in the past decades, there is hope on the horizon. With the joint efforts of pharmaceutical companies and academic institutions, we look forward to the continuous development of safer, more effective treatments for patients with advanced or metastatic HCC.

## References

1. J. Ferlay *et al.*, *Int. J. Cancer* **127**, 2893-2917 (2010).
2. L. A. Torre *et al.*, *CA Cancer J. Clin.* **65**, 87-108 (2015).
3. International Agency for Research on Cancer, Globocan 2012, "Liver Cancer Estimated Incidence, Mortality and Prevalence Worldwide in 2012"; available at <http://globocan.iarc.fr/old/factsheets/cancers/liver-new.asp>.
4. H. B. El-Serag, K. L. Rudolph, *Gastroenterology* **132**, 2557-2576 (2007).
5. K. A. McGlynn, L. Tsao, A. W. Hsing, S. S. Devesa, J. F. Fraumeni, Jr., *Int. J. Cancer* **94**, 290-296 (2001).
6. Chinese Society of Liver Cancer, Chinese Society of Clinical Oncology, Liver Cancer Society of the Chinese Medical Association, *Chin. Clin. Oncol.* **14**, 259-269 (2009).
7. L. P. Waller, V. Deshpande, N. Pyrsopoulos, *World J. Hepatol.* **7**, 2648-2663 (2015).
8. [No authors listed], *Hepatology* **28**, 751-755 (1998).
9. T. F. Greten, A. G. Duffy, F. Korangy, *Clin. Cancer Res.* **19**, 6678-6685 (2013).
10. G. L. Deng, S. Zeng, H. Shen, *World J. Hepatol.* **7**, 787-798 (2015).
11. S. Qin *et al.*, *Oncologist* **19**, 1169-1178 (2014).
12. J. Bruix *et al.*, *Lancet* **7**, 56-66 (2017).
13. A. Cheng *et al.*, *J. Clin. Oncol.* **35**, Abstract 4001 (2017), doi: 10.1200/JCO.2017.35.15\_suppl.4001.
14. L. Rimassa *et al.*, *J. Clin. Oncol.* **35**, Abstract 4000 (2017), doi: [http://ascopubs.org/doi/abs/10.1200/JCO.2017.35.15\\_suppl.4000](http://ascopubs.org/doi/abs/10.1200/JCO.2017.35.15_suppl.4000).
15. A. D. Pardee, L. H. Butterfield, *Oncoimmunology* **1**, 48-55 (2012).
16. S. Zongwen *et al.*, *Cell Mol. Immunol.* **13**, 354-368 (2016).
17. Q. Gao *et al.*, *Clin. Cancer Res.* **15**, 971-979 (2009).
18. T. Hato, L. Goyal, T. F. Greten, D. G. Duda, A. X. Zhu, *Hepatology* **60**, 1777-1782 (2014).
19. C. Miroux, T. Vausselin, N. Delhem, *Expert Opin. Biol. Ther.* **10**, 1563-1572 (2010).
20. L. Golden-Mason *et al.*, *J. Virol.* **81**, 9249-9258 (2007).
21. G. Peng *et al.*, *Immunology* **123**, 57-65 (2008).
22. H. I. Jung *et al.*, *Cancer Res. Treat.* **49**, 246-254 (2017).
23. B. Sangro *et al.*, *J. Hepatol.* **59**, 81-88 (2013).
24. A. B. El-Khoueiry *et al.*, *Lancet* **389**, 2492-2502 (2017).
25. B. Sangro *et al.*, *J. Clin. Oncol.* **34**, Abstract TPS4147 (2016), doi: 10.1200/JCO.2016.34.15\_suppl.TPS4147.
26. R. S. Finn *et al.*, *J. Clin. Oncol.* **35**, Abstract TPS503 (2017), doi: 10.1200/JCO.2017.35.4\_suppl.TPS503.
27. Z. A. Wainberg *et al.*, *J. Clin. Oncol.* **35**, Abstract 4071 (2017), doi: 10.1200/JCO.2017.35.15\_suppl.4071.
28. G. K. Abou-Alfa *et al.*, *J. Clin. Oncol.* **34**, Abstract TPS3103 (2016), doi: 10.1200/JCO.2016.34.15\_suppl.TPS3103.
29. ClinicalTrials.gov, "A Trial of Lenvatinib Plus Pembrolizumab in Subjects With Hepatocellular Carcinoma"; available at <https://clinicaltrials.gov/ct2/show/NCT03006926>.
30. S. C. Guba *et al.*, *Ann. Oncol.* **27**, 1102 (2016), doi: <https://doi.org/10.1093/annonc/mdw378.55>.

31. A. B. El-Khoueiry *et al.*, *Lancet* **389**, 2492–2502 (2017).
32. J. J. Knox *et al.*, *J. Clin. Oncol.* **33**, 1835–1844 (2015).
33. M. H. Kershaw, J. A. Westwood, P. K. Darcy, *Nat. Rev. Cancer* **13**, 525–541 (2013).
34. Q. Zhang *et al.*, *Oncoimmunology* **5**, e1251539 (2016).
35. M. Chen, T. Therneau, L. S. Orsini, Y. L. Qiao, *BMC Gastroenterol.* **11**, 53 (2011).
36. J. G. Chen, S. W. Zhang, W. Q. Chen, *Zhonghua yu fang yi xue za zhi [Chinese Journal of Preventive Medicine]* **44**, 383–389 (2010).
37. T. Zuo, *et al.*, *Zhonghua zhong liu za zhi [Chinese Journal of Oncology]* **37**, 691–696 (2015).
38. B. M. Chia-Jui Yen *et al.*, *Ann. Oncol.* **28**, mdx261.139 (2017), doi: <https://doi.org/10.1093/annonc/mdx261.139>.
39. J. Huang, *J. Clin. Oncol.* **35**, Abstract e15572 (2017), doi: 10.1200/JCO.2017.35.15\_suppl.e15572.
40. D. Baumhoer *et al.*, *Am. J. Clin. Pathol.* **129**, 899–906 (2008).
41. B. Zhai, *J. Clin. Oncol.* **35**, Abstract 3049 (2017), doi: 10.1200/JCO.2017.35.15\_suppl.3049.

## Challenges and prospects for precision cancer immunotherapy in China

Zhihao Lu, Jianling Zou, Shuang Li, and Lin Shen\*

With increasing incidence and mortality, cancer has become a major public health problem in China (1). Surgery, radiotherapy, chemotherapy, and targeted therapy remain standards of care for cancer patients. However, the effectiveness of these treatments is not completely satisfactory. The success of immunotherapy, especially anti-PD-1/PD-L1 treatment, which blocks the ability of tumor cells to shield themselves from attack by the immune system, has led to some important strides in cancer care in recent years. At the same time, immunotherapy has become a hot area in cancer research and treatment in China. Here, we discuss the challenges and prospects for precision immunotherapy in the treatment of cancer, based on the unique characteristics of the Chinese population.

### Current status of cancer immunotherapy

Chinese pharmaceutical companies and researchers are highly motivated to explore new immunotherapy agents and therapies. Even though no PD-1 or PD-L1 inhibitor has yet been approved by the Chinese Food and Drug Administration (CFDA), the Chinese government is making earnest efforts to improve the overall clinical research environment, and the CFDA is actively reforming the regulatory framework to accelerate approval of novel agents (2). By July 2017, Chinese pharmaceutical companies had developed 10 anti-PD-1/PD-L1 inhibitors, eight of which have been approved by CFDA for phase 1 to phase 3 clinical trials in patients with advanced solid tumors (Table 1). In November 2016, clinical trials of PD-1 inhibitor KN035 were approved by the U.S. Food and Drug Administration (FDA) (3). In the past five years, around 1,000 international clinical trials of PD-1/PD-L1 therapies for solid tumors have been registered in the ClinicalTrials.gov database, of which 100 involved Chinese sites (2). Thus, Chinese

Department of Gastrointestinal Oncology, Key Laboratory of Carcinogenesis and Translational Research (Ministry of Education), Peking University Cancer Hospital and Institute, Beijing, China

\*Corresponding author: [linshenpku@163.com](mailto:linshenpku@163.com)

Agent	Target	Pharmaceutical company	Phase	Condition	ClinicalTrials.gov Identifier
JS001	PD-1	Shanghai Junshi Biosciences	Phase 1, 2	Melanoma, urological cancer	NCT02836795, NCT03013101
SHR-1210	PD-1	Jiangsu Hengrui Medicine	Phase 1, 2, 3	Breast cancer, gastrointestinal cancer, hepatocellular cancer	NCT02742935, NCT02989922
BGB-A317	PD-1	BeiGene	Phase 1, 2	Various malignant solid tumors	NCT02407990, NCT03209973
Genolimzumab	PD-1	Genor BioPharma	Phase 1	Various malignant solid tumors	NCT03053466
IBI308	PD-1	Innovent Biologics	Phase 1, 2	Various malignant solid tumors	NCT02937116, NCT03116152
Recombinant humanized anti-PD-1 monoclonal antibody	PD-1	Harbin Gloria Pharmaceuticals	Phase 1	Various malignant solid tumors	NA
Recombinant humanized anti-PD-1 monoclonal antibody	PD-1	Taizhou Hanzhong	IND	Various malignant solid tumors	NA
Recombinant humanized anti-PD-1 monoclonal antibody	PD-1	Bio-Thera Solutions	IND	Various malignant solid tumors	NA
Recombinant humanized PD-L1 single-domain antibody Fc fusion protein	PD-L1	3D Medicines/ Alphamab	Phase 1	Various malignant solid tumors	NA
Recombinant fully human anti-PD-L1 monoclonal antibody	PD-L1	CStone Pharmaceuticals	Phase 1	Various malignant solid tumors	NA

TABLE 1. PD-1/PD-L1 blockades in clinical trials in China. NA, none available; IND, investigational new drug.

clinical trials are developing rapidly and catching up with those in other parts of the world.

However, there are some problems with the current situation. The promise of these drugs is leading to a crowded market. Numerous companies are jockeying for position, while many others are hoping for profitable out-licensing deals. All of this makes quality control a serious challenge. Fur-

thermore, whether in pharmaceutical development or regimen design, China has been producing too many “me-too” drugs—innovation has been lacking. Finally, because of overreaching pragmatism and the current system for evaluating research (which places great importance on publication numbers), it is possible that applied research could impede the progress of pure science.

### **Challenges facing cancer immunotherapy**

Anti-PD-1/PD-L1 treatment is one of the most successful immunotherapies against solid tumors, but it faces many difficulties in China. Several of these relate to biomarkers:

#### ***Lack of predictive biomarkers for treatment response***

Multiple tumor types have been shown in early phase trials to have a long-lasting response to anti-PD-1/PD-L1 treatment (4, 5). However, the overall response rate to this treatment is only 10%–44% (6, 7). Therefore, better biomarkers are needed to guide patient selection and provide early indicators of treatment response.

#### ***Lack of predictive biomarkers for adverse events***

Adverse effects, including lethal complications, have been reported during anti-PD-1/PD-L1 treatment (8). Most clinical trials focus on potential predictive biomarkers of treatment response, but no reliable biomarkers for adverse events have been identified. However, they are equally necessary in the clinic.

#### ***Lack of predictive biomarkers for hyperprogressive disease***

Champrat and colleagues reported for the first time instances in which a cancer progressed unusually rapidly in patients receiving immune checkpoint inhibitors like anti-PD-1/PD-L1, described as “hyperprogressive disease.” The incidence varied from 9% to 29% (9, 10) of patients receiving the immune checkpoint therapy. As this study indicates, it is necessary to identify biomarkers for hyperprogression in order to selectively avoid those patients who might be harmed by immune checkpoint inhibitors. As of now, unfortunately no such explicit biomarkers have been found.

In addition to these difficulties, Chinese patients have some special characteristics that need to be considered in designing immune therapies for them. They have relatively high rates of hepatitis B virus infection, for example, and more frequently use Chinese traditional medicine. Chinese patients also have more “driver gene” mutations that are linked to cancer development, better clinical responses to chemotherapy, and different toxicity profiles (11–14), which may have an effect on anti-PD-1/PD-L1 therapy. These factors have rarely been considered in trial designs. Furthermore, Chinese patients have limited access to immune

checkpoint inhibitors such as antibodies against PD-1 and PD-L1, so there has been only a small patient pool in which to explore biomarkers in the Chinese population.

### **Toward precision cancer immunotherapy in China**

Because of the dynamic nature of the immune system and the multiple elements involved in the complex immune response to cancer, it is much more challenging to develop biomarkers for immunotherapy than for other treatments. Thus, it is imperative to establish a paradigm for precision immunology to steer patients through immunotherapy. Some strategies for precision cancer immunotherapy include the following:

#### **1. Mechanism-based predictive biomarkers for patient selection**

##### ***Mismatch repair deficiency***

DNA mismatch repair (MMR) describes a specific genetic aberration that results in a many-fold increase in tumor mutational burden in various tumor types and has been explored for responsiveness to anti-PD-1 (15). The high level of microsatellite instability (MSI-H)—a predisposition to mutation that is characteristic of MMR-deficient (dMMR) cancers—or the absence of one of the MMR proteins, is used to identify this genetic subset of patients.

In 2017, the U.S. FDA granted accelerated approval to pembrolizumab for the treatment of adult and pediatric patients with unresectable or metastatic solid tumors that have been identified as MSI-H or dMMR. The approval covers patients whose disease has progressed after prior treatment and who have no satisfactory alternative treatment options. However, only about 4%–5% of patients with advanced solid tumors have been identified as MSI-H or dMMR, while their therapeutic response is approximately 40% (16). Thus, further exploration of novel biomarkers is needed to refine anti-PD-1 treatment strategies.

##### ***PD-L1 expression***

The protein PD-L1, detected on patient tumor cells using immunohistochemistry (IHC), is currently the most common clinically detected biomarker for predicting patient response to anti-PD-1/PD-L1 therapy (17).

A recently published randomized phase 3 trial in advanced non-small cell lung cancer (NSCLC) revealed a correlation between high PD-L1

expression and longer progression-free survival and overall survival in immunotherapy treatment regimens (18). Nevertheless, a recent review of immune biomarker utility in NSCLC therapy indicated a lack of consensus on how to use PD-L1 expression as a biomarker to help in selecting treatment regimens. In fact, patients with little or no detectable expression of this biomarker can also benefit from anti-PD-L1 therapy (19). A similar scenario has been observed in other tumors such as gastric cancer (20), renal cell carcinoma (21), and urothelial cancer (22). However, due to the complexities of PD-L1 IHC, further studies are needed to carefully validate these observations and other characteristics of the tumor microenvironment.

### **Mutational landscape and mutational burden**

Various studies have shown how the number of mutations in a tumor's genome—the mutational burden—correlates with greater efficacy of diverse anti-PD-1/PD-L1 drugs (23, 24). Analyzing the genetic signatures of tumors might therefore identify patients who have a better chance of responding to checkpoint inhibition (25). NSCLC patients with a high nonsynonymous mutational burden showed long-lasting clinical benefits from treatment with pembrolizumab. Furthermore, another study showed that almost all patients with a low nonsynonymous mutational burden did not benefit from the same treatment (24). These results make the mutational landscape of the tumor one of the most promising predictive factors both at the beginning of treatment and after an initial favorable response.

### **Other predictive biomarkers**

It is well known that there is a relationship between inflammatory cells, cancer, and proinflammatory proteins such as chemokines. Some findings demonstrate that the genetic signatures of genes related to the cytokine interferon  $\gamma$  appear to be reasonable tumor biomarkers that could soon be incorporated into predictive models of how patients will respond to checkpoint inhibitors (26, 27).

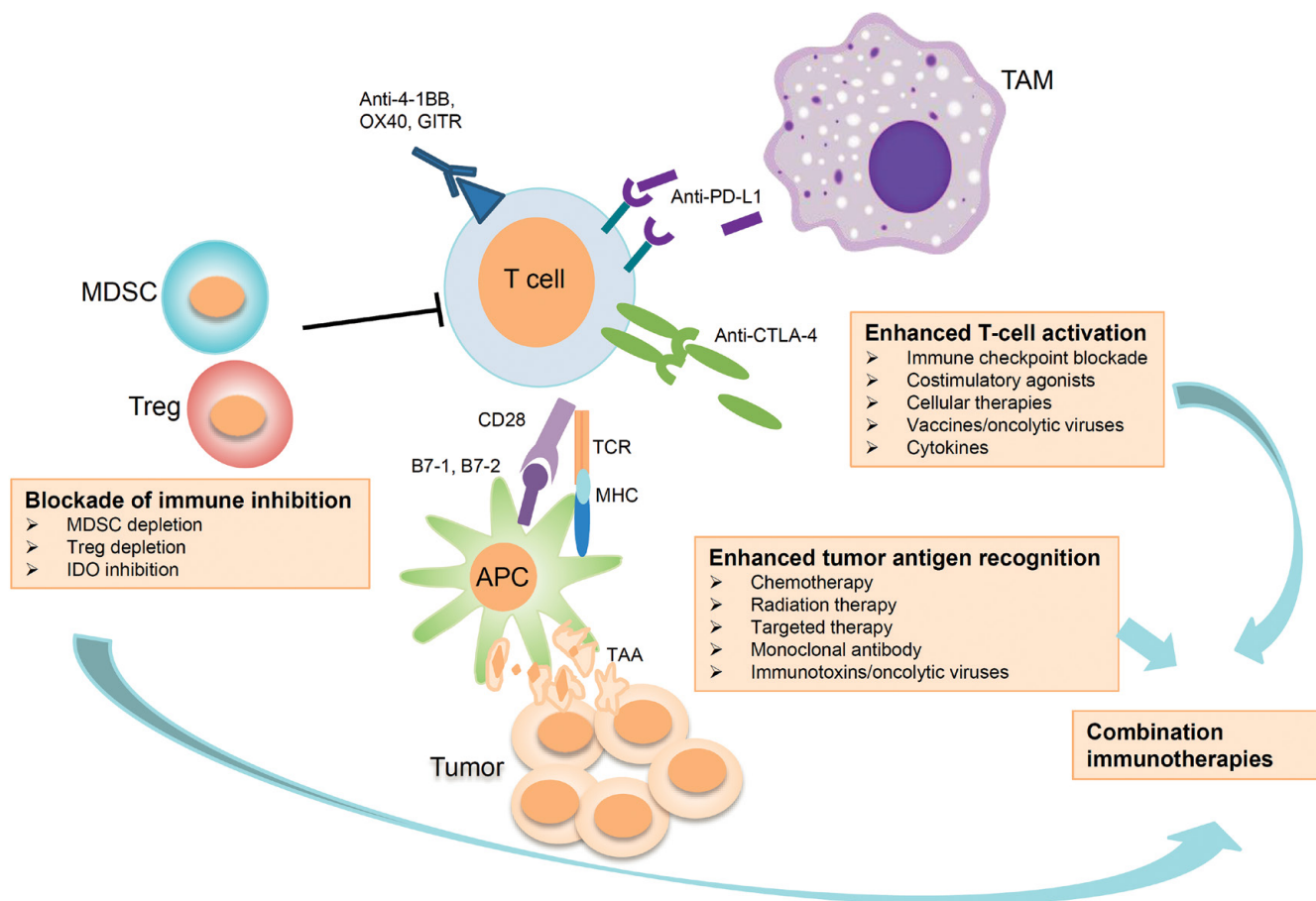
Recent studies have also shed light on the previously unknown effect of the gut microbiota on immunotherapy for patients being treated with checkpoint blockade inhibitors (28, 29). Future studies of the gut microbiota of cancer patients who respond to checkpoint inhibitors could provide invaluable information to aid in patient selection.

Findings from 2017 show that CKLF-like MARVEL transmembrane (CMTM) domain-containing proteins CMTM6 and CMTM4 are associated with the PD-L1 protein and enhance the ability of tumor cells that produce PD-L1 to inhibit T cells, suggesting a potential therapeutic target in the effort to overcome immune evasion by tumor cells (30, 31).

Results of completed trials indicate that numerous biomarkers that predict toxicity to antibodies against immune checkpoint inhibitor cytotoxic T-lymphocyte antigen 4 (CTLA-4) should be considered for further evaluation as an anticancer treatment in larger prospective trials. Of these promising biomarkers, neutrophil activation markers and several immunoglobulin genes warrant further investigation as biomarkers predictive of toxicity to anti-CTLA-4 treatment. There is currently no predictive biomarker for toxicity to anti-PD-1 treatment. Evaluating the gene expression profile of patients who develop toxicities to anti-PD-1/PD-L1 treatment may also be an interesting landscape to explore in the future.

## **2. Predictive biomarkers related to hyperprogression**

Champiat and colleagues at Université Paris Saclay found a greater incidence of hyperprogression in patients 65 years old or older than in those younger than 65 (9). That may be because the functioning of immune cells, chemotaxis, phagocytosis, and intracellular killing of pathogens is decreased in older patients (32), but the mechanism by which this happens is not clear. Saada-Bouazid and colleagues (10) found that hyperprogression correlated more significantly with the presence of a recurrence near the original tumor than with a recurrence farther away (90% versus 37%) in patients with recurrent or metastatic head and neck squamous cell cancer. Kato and colleagues (33) investigated potential genomic biomarkers associated with hyperprogression after immunotherapy. Their results showed that murine double minute 2 (MDM2) family amplification or epidermal growth factor receptor (EGFR) alterations were correlated with hyperprogression in patients previously treated with PD-1/PD-L1 inhibitor, but not CTLA-4. However, the exact mechanism linking MDM2 family amplification and hyperprogression is unclear. Pretreatment and early postimmunotherapy biopsies would be valuable in exploring the underlying mechanism in the future.



**FIGURE 1. Opportunities for precision cancer immunotherapy and combination therapy strategies.** Multiple approaches that target different pathways are required for successful anticancer treatment, including T-cell activation, blockade of immune inhibition, and correct recognition of tumor antigens. The combination immunotherapies represented here offer the possibility for achieving improved efficiencies in cancer treatment. TAM, tumor-associated macrophage; MDSC, myeloid-derived suppressive cell; Treg, regulatory T cell; APC, antigen-presenting cell; GITR, glucocorticoid-induced tumor necrosis factor-related protein; CTLA-4, cytotoxic T-lymphocyte antigen 4; TCR, T-cell receptor; MHC, major histocompatibility complex; IDO, indoleamine 2,3-dioxygenase; TAAs, tumor-associated antigens.

### 3. Combination immunotherapy

In recent years, there has been an increase in the number of novel combination therapies used in cancer immunotherapy treatments. The goal of combination approaches is to expand the spectrum of patients who would benefit from immunotherapy, to enhance its curative effects, and to reduce the number of side effects compared with single therapy regimens (Figure 1).

Several hundred clinical trials are currently exploring the effects of anti-PD-1/PD-L1 used in combination with experimental immune modulators, such as other monoclonal antibody checkpoint inhibitors, cancer vaccines, epigenetic

drugs, and drugs targeting metabolic pathways, and also with surgery, radiotherapy, and chemotherapy. Preliminary evidence [for example using a combination of PD-1 and CTLA-4 in treatment of melanoma (34) and NSCLC (35)], has highlighted the potential to enhance the clinical benefits of monotherapies by combining them with agents that have synergistic mechanisms of action. Initial analyses from the international, multicohort, phase 2 KEYNOTE-059 study showed that objective response rate (includes both partial and complete responses) was 60% across all patients (36).

Despite the promise of combination therapies, several issues need to be further investigated:

**1. Combination agents.** Although there are dozens of clinical trials of immune modulators used in combination with anti-PD-1/PD-L1 treatment, there is little research in China on the rationale, dosages, schedules of treatment, and configuration of the specific combinations used.

**2. Predictive biomarkers for combination immunotherapy.** Although studies have found the rate of response to combination therapies to be much higher than that of monotherapies, there are still a proportion of patients who do not survive. Finding predictive biomarkers for combination immunotherapy is therefore of prime importance, and will probably require dynamic monitoring of the immune response and a fuller understanding of the changes in tumor microenvironments before, during, and after treatment.

**3. Predictive biomarkers for adverse events.** Combination immunotherapy usually brings more instances of adverse events. In the KEYNOTE-059 study, grade 3–4 treatment-related adverse events (TRAEs) including neutropenia, stomatitis, anemia, thrombocytopenia, anorexia, and fatigue, occurred in 76% of patients—a much higher frequency than with chemotherapy or targeted therapies. We should therefore pay more attention to TRAEs associated with combination therapy, and attempt to identify novel predictive biomarkers for them.

## Conclusions

The ultimate goal of precision immunology for cancer is to select a subset of patients with a common biological basis of cancer: the patients who are most likely to benefit from a particular immunotherapy. Hopefully, with the development of new technologies to dynamically monitor the cancer-immune system interaction and by considering the special features of different populations, precision cancer immunology will be applied more widely, improving diagnosis, stratification, and treatment of cancer patients in China.

## References

1. W. Chen *et al.*, *CA Cancer J. Clin.* **66**, 115–132 (2016).
2. Q. Zhou *et al.*, *Nat. Rev. Clin. Oncol.* **14**, 577–583 (2017).
3. S. Y. Liu, Y. L. Wu, *J. Hematol. Oncol.* **10**, 136 (2017).
4. E. B. Garon *et al.*, *N. Engl. J. Med.* **372**, 2018–2028 (2015).
5. C. Robert *et al.*, *N. Engl. J. Med.* **372**, 2521–2532 (2015).
6. F. S. Hodi *et al.*, *J. Clin. Oncol.* **32**, 9002–9002 (2014).
7. Y. Iwai *et al.*, *J. Biomed. Sci.* **24**, 26 (2017).
8. A. Sgambato *et al.*, *Curr. Drug Saf.* **11**, 62–68 (2016).
9. S. Champiat *et al.*, *Clin. Cancer Res.* **23**, 1920–1928 (2017).
10. E. Saada-Bouazid *et al.*, *Ann. Oncol.* **28**, 1605–1611 (2017).
11. L. Y. Gou, Y. L. Wu, *Lung Cancer (Auckl.)* **5**, 1–9 (2014).
12. T. Kohno *et al.*, *Transl. Lung Cancer Res.* **4**, 156–164 (2015).
13. S. Lu *et al.*, *Clin. Lung Cancer* **17**, e103–e112 (2016).
14. Q. Zhang *et al.*, *Sci. Rep.* **5**, 17413 (2015).
15. S. Y. Liu, T. Mok, Y. L. Wu, *Cancer* **121** (Suppl. 17), 3089–3096 (2015).
16. M. Bupathi, C. Wu, *J. Gastrointest. Oncol.* **7**, 713–720 (2016).
17. S. P. Patel, R. Kurzrock, *Mol. Cancer Ther.* **14**, 847–856 (2015).
18. M. Reck *et al.*, *N. Engl. J. Med.* **375**, 1823–1833 (2016).
19. Y. K. Chae *et al.*, *Clin. Lung Cancer* **17**, 350–361 (2016).
20. Y. Y. Janjigian *et al.*, *J. Clin. Oncol.* **34**, 4010–4010 (2016).
21. F. Xu *et al.*, *Int. J. Clin. Exp. Med.* **8**, 14595–14603 (2015).
22. J. E. Rosenberg *et al.*, *Lancet* **387**, 1909–1920 (2016).
23. H. Matsushita *et al.*, *Nature* **482**, 400–404 (2012).
24. N. A. Rizvi *et al.*, *Science* **348**, 124–128 (2015).
25. C. Kandoth *et al.*, *Nature* **502**, 333–339 (2013).
26. J. Gao *et al.*, *Cell* **167**, 397–404.e9 (2016).
27. J. M. Zaretsky *et al.*, *N. Engl. J. Med.* **375**, 819–829 (2016).
28. A. Sivan *et al.*, *Science* **350**, 1084–1089 (2015).
29. M. Vetizou *et al.*, *Science* **350**, 1079–1084 (2015).
30. M. L. Burr *et al.*, *Nature* **549**, 101–105 (2017).
31. R. Mezzadra *et al.*, *Nature* **549**, 106–110 (2017).
32. R. Solana *et al.*, *Semin. Immunol.* **24**, 331–341 (2012).
33. S. Kato *et al.*, *Clin. Cancer Res.* **23**, 4242–4250 (2017).
34. J. Larkin *et al.*, *N. Engl. J. Med.* **373**, 23–34 (2015).
35. A. Somasundaram, T. F. Burns, *J. Hematol. Oncol.* **10**, 87 (2017).
36. Y.-J. Bang *et al.*, *J. Clin. Oncol.* **35**, 4012–4012 (2017).

# The rise of engineered T-cell therapy in China

Jianshu Wei<sup>1</sup>, Yi Zhang<sup>2,3\*</sup>, and Weidong Han<sup>1\*</sup>

**T**here are many treatment methods for controlling or eliminating cancer cells, including tumor vaccines, chemotherapy, radiation therapy, and tyrosine kinase inhibitors. The advancement of basic research into the progression of malignancies has resulted in significant improvements in treatment. However, complete and durable remission for inoperable malignancies remains rare. There is a tremendous effort underway to find novel therapies that can offer better prospects for destroying tumors.

Recently, immunotherapy as a tumor treatment has come to the forefront because of the significant successes of therapies using immune checkpoint antibodies and chimeric antigen receptor modified T (CAR-T) cells. Simply speaking, immune checkpoint antibodies fight tumors by activating the naturally occurring lymphocytes in the body that can recognize and kill the tumor cells. CAR-T cells are the patient's own cells educated in vitro by genetic manipulation to kill tumor cells. The immune system is so sophisticated that the mechanisms underlying cancer immunology are far from fully explained, although there has been steady progress over the past 30 years. As clinical applications have developed, T cell-centered cancer immunotherapy has proven to be an effective antitumor treatment.

In 2017, the U.S. Food and Drug Administration (FDA) approved the drug CTL019 (tisagenlecleucel-T, Novartis), the first CAR-T treatment to get the regulatory nod. It was a milestone in tumor immunotherapy, marking the start of the CAR-T era. With strong support from the government, CAR-T therapies are developing very rapidly in China, where the number of registered clinical trials recently surpassed that in the United States.

Cancer immunotherapy has been in clinical use in China since before 1988, in the form of a method using lymphokine-activated killer (LAK) cells. LAK cells are white blood cells stimulated to grow in vitro

and cultured in the presence of interleukin 2 (IL-2). In 1988, Rosenberg and colleagues first reported a method for preparing LAK cells (1). Since then, LAK cells have been used to treat multiple types of tumors (2).

In 1991, Schmidt-Wolf reported on cytokine-induced killer (CIK) cells (3). CIK cells were more cytotoxic than LAK cells and exhibited better proliferation efficiency. Several clinical trials on CIK cells have been conducted since then. In China, the first research article on the subject was published in 2002 (4). Subsequently, CIK therapy has been widely utilized for a variety of reasons, including its safety, relative technical simplicity, and the support it received from the government. During this period, other treatments using natural killer (NK) cells and dendritic cell-activated CIK cells (DC-CIKs) were also in clinical use, providing a solid foundation for the development of newer CAR-T treatments. In general, the tumor-killing cells utilized during that time were natural lymphocytes—small white blood cells—with no modifications, and the killing effect was weak and depended largely on nonspecific tumor cell recognition mechanisms. As a result, the overall clinical benefits were minimal.

To improve the clinical efficacy of these therapies, Chinese scientists have tried a variety of ways to increase the number of T lymphocytes that recognize a specific tumor, of which CAR-T technology is undoubtedly the most attractive. The generation of the first CAR molecule was reported in 1989 by Gideon Gross (5). This first-generation CAR-T showed poor activity in phase 1 clinical trials. The first domestic research using improved second-generation CAR-T was published in 2009, and involved a human epidermal growth factor receptor 2 (HER2)-specific CAR that was designed for the treatment of breast cancer (6). The first article on CAR-T clinical trials in China was published in 2014 by the Han laboratory of the People's Liberation Army General Hospital (PLAGH) (7). With the government restricting the clinical use of CIK and similar minimally successful immune therapies, the use of natural lymphocytes gradually diminished, and China entered a new era of precision immunotherapies, exemplified by CAR-T.

By early 2015, 12 clinical trials of CAR-T in China were registered on the ClinicalTrials.gov website, of which eight were conducted by PLAGH, two by Beijing Cancer Hospital, and one each by Renji Hospital and the 307 Hospital of the People's Liberation Army. By January 2016, the total number of trials had risen to 26. As of September 2017, 121 CAR-T trials have been registered by domestic

<sup>1</sup>Department of Molecular Immunology, People's Liberation Army General Hospital, Beijing, China

<sup>2</sup>Biotherapy Center, the First Affiliated Hospital of Zhengzhou University, Zhengzhou, Henan, China

<sup>3</sup>School of Life Science, Zhengzhou University, Zhengzhou, Henan, China

\*Corresponding authors: yizhang@zzu.edu.cn (Y.Z.), hanwdrsw69@yahoo.com (W.D.H)

hospitals, academic institutions, and private companies, and China has become the country with the most registered CAR-T trials worldwide.

Of the 121 registered trials, 53 are CAR-T therapies targeting CD19 for treatment of leukemia, and 44 of these are still ongoing. CD19 is a membrane protein found on all B cells and critical for B-cell activation and proliferation. Generally speaking, acute lymphocytic leukemia (ALL) is most responsive to this therapy, known as CART-19. Five of the nine patients enrolled in the first domestic CART-19 trial saw complete remission (CR) for two to nine months (8), and cytokine release syndrome (CRS) was observed in most cases. CRS occurs when large numbers of white blood cells release proinflammatory cytokines, which go on to activate more white blood cells, creating a potentially fatal cascade effect called a “cytokine storm.” Two patients who had previously received hematopoietic stem cell transplantation (HSCT) presented with grade 2–3 graft-versus-host diseases (GVHD) after infusion of donor-derived CAR-T cells, suggesting that HSCT-treated patients without 100% chimerization (i.e., in which the hematopoietic system was not completely reconstructed from donor-derived bone marrow cells) had a higher risk of GVHD when receiving donor-derived CAR-T cells. Overall, the rate of complete remission is a bit lower than that seen in results obtained at the Memorial Sloan Kettering Cancer Center (MSKCC), the University of Pennsylvania, and Novartis. It is worth noting that the patients recruited in this trial were more aggressively treated and in later stages of disease than the participants in the United States. Results from several subsequent domestic CART-19 trials have shown that overall CR rates were comparable to those in the United States. For example, the First Affiliated Hospital of Zhejiang University enrolled 15 patients with relapsed or refractory (R/R) ALL (9), 12 of whom had CR responses one month after CART-19 infusion.

Relapse is the major obstacle to long-term survival of patients hoping for complete remission. Researchers in China have been exploring ways to avoid relapse. The first option is to use bispecific CAR-T cells that are active against two different antigens. A group from the Southwest Hospital in Chongqing has registered a trial in which CART-19 and CART-20 (targeting CD-20) will be infused sequentially into patients with diffuse large B-cell lymphomas (DLBCL). And the PLAGH group has carried out two trials in which tandem CART-19/20 and tandem CART-19/22 were used for patients with R/R B-cell malignancies. Another common approach is to initiate HSCT treatment after obtaining a CR response using

CAR-T therapy. A group at Lu Daopei Hospital in Beijing has treated 27 patients with this regimen (10). Of the 27 CR patients treated with allogeneic HSCT following CART-19, 23 remained disease-free for at least 45 days (median follow-up time was 206 days; range was 45–427 days). This suggested that further use of allogeneic HSCT could effectively reduce the risk of relapse. The Southwest Hospital group has also registered a trial in which HSCT will be given after CART-19 treatment for patients with B cell malignancies. In collaboration with PLAGH, a team led by Huisheng Ai launched the world's first treatment to use coinfusion of haploidentical donor-derived CART-19 cells and mobilized peripheral blood stem cells (PBSCs)—stem cells stimulated to move from the bone marrow into the bloodstream—for a 71-year-old female with R/R ALL (11). Any residual disease was undetectable following treatment, and CRS and GVHD were controlled. This result suggested that coinfusion of allogeneic CAR-T cells and PBSCs may induce full donor engraftment in R/R ALL, offering a novel option for combining CAR-T and HSCT.

CD20 and CD22 have expression patterns similar to those of CD19, which are expressed in all B lymphocytes except B progenitor cells, making these proteins potential targets for managing B-cell malignancies. At present, there are five CD20 and six CD22 CAR-T trials registered in China. The initial article on a CD20 CAR-T trial was published in 2014 by researchers at PLAGH (7), being the first to report results of CART-20 therapy for DLBCL. Seven patients with DLBCL were treated with CART-20 cells, and five of the six evaluable patients showed an objective remission response (ORR).

ALL and B-cell lymphoma are the cancers most sensitive to CAR-T therapy, so they are prime candidates for further research, both in China and abroad. But China's research groups have also done pioneering work on other types of blood-based malignancies.

Multiple myeloma (MM) is a fatal plasma cell cancer and the second most common hematological malignancy. Typically, CD19, CD20, and CD22 are not detectable in MM. Clinical results have suggested that CD138, a transmembrane proteoglycan found in certain hematopoietic cancers and carcinomas, may be an effective target for refractory and progressive MM. In 2015, researchers at PLAGH reported the first results of CART-138 therapy on MM. To date, five patients have been treated, and in four the disease was stabilized for more than three months, while one patient with progressive MM had a significant reduction in myeloma cells in peripheral blood—from 10.5% to less than 3% (12). This study

suggested that CART-138 could be a promising approach for treating MM. There are three CART-138 clinical trials presently registered in China.

B-cell maturation antigen (BCMA), a protein reported to be expressed in B cells, including MM cells, is another potential target for MM treatment. At the 2017 conference of the American Society of Clinical Oncology (ASCO), Nanjing Legend Biotech Co., Ltd. presented data from their trial of BCMA-targeted CAR-T therapy (CART-BCMA) in 35 participants with R/R MM. Thirty-three evaluable patients saw significant clinical responses (defined as a complete response or a very good partial response) within two months following CAR-T infusion. The overall ORR was as high as 100%. Four CART-BCMA clinical trials are currently registered in China.

Acute myeloid leukemia (AML) is the most common acute leukemia affecting adults (13), and CD33 antigen is expressed on leukemic cells in more than 90% of patients with the disease (14). In 2015, a clinical trial was conducted at PLAGH to assess the feasibility and efficacy of CART-33–CAR-T therapy targeting CD33—for the treatment of R/R AML. A marked decrease of myeloid blasts in the bone marrow two weeks after cell infusion was observed, but the disease progressed after nine weeks (15). This was the first demonstration of the potential of CART-33 for the treatment of AML. Five clinical trials for treating AML with CART-33 are currently registered in China.

In Hodgkin's lymphoma (HL), the CD30 protein is a potentially potent target. This year, the first phase 1 clinical trial results on CART-30–CAR-T therapy directed at CD30—were reported by PLAGH researchers. Eighteen patients with progressive R/R HL were enrolled. The data showed that the treatment was well tolerated without severe toxicity. Seven patients achieved partial remission and the disease was stabilized in six patients (16). There are now four CART-30 clinical trials underway in China.

Compared with the performance of CAR-T seen with hematologic malignancies, there is general pessimism about the prospects of this therapy for treatment of solid tumors. Chinese researchers have nonetheless pursued this possibility with enthusiasm and have made many original contributions. Currently, 44 clinical trials of CAR-T cell treatments of solid tumors have been registered in China, covering 14 different targets.

Among them, the glypican-3 (GPC3) protein has attracted the most interest, with nine trials underway. GPC3 is generally overexpressed in multiple solid tumors, including lung cancer and hepatocellular

carcinoma (HCC) (17, 18). A group led by Zonghai Li from Renji Hospital has done several preclinical studies to verify the feasibility of CART-GPC3 in the treatment of lung cancer and HCC (18, 19).

Epidermal growth factor receptor (EGFR) is a glycoprotein that is overexpressed in many malignancies that derive from the epithelium. These include non-small cell lung cancer (NSCLC) and multiple digestive tract cancers (20). A preclinical study of CART-EGFR was carried out in 2013 by Yuquan Wei's laboratory at Sichuan University, which provided an experimental basis for further clinical study (20). The first CART-EGFR clinical results were published by a group at PLAGH in 2016. Patients with EGFR-positive (>50% expression) R/R NSCLC received escalating doses of CART-EGFR. Among 11 evaluable patients, two achieved partial remission and five had a stabilized disease response for two to eight months. Significantly, eradication of EGFR-positive tumor cells was observed in tumor biopsies after the CAR-T infusion, providing the most direct evidence supporting the effectiveness of this treatment (21). However, this result also emphasized that the heterogeneity of solid tumors may present a formidable obstacle to the CAR-T treatment. Six CART-EGFR trials have been registered in China as of now.

Several other domestic clinical trials of CAR-T treatment for solid tumors have been published. In 2016, a group headed by Lin Yang published a case report on CAR-T therapy targeting mucin 1 (MUC1) for a patient with MUC1-positive metastatic seminal vesicle cancer in which significant tumor necrosis was observed (22). In 2017, a group at PLAGH published results from a phase 1 clinical trial using CAR-T cells to target HER2 in patients with advanced biliary tract cancers and pancreatic cancers (23). Cheng Qian's group published results of a phase 1 trial of CAR-T cells targeting carcinoembryonic antigen (CEA) therapy for patients with CEA-positive metastatic colorectal cancers (24). Data from all of these studies demonstrated the safety and feasibility of CAR-T therapy for a variety of cancers.

What can be done to provide patients with greater clinical benefits following CAR-T treatment? To answer this question, Chinese research groups have done some pioneering work. For example, cancer stem cells (CSCs) are believed to play an important role in tumor growth and metastasis. Theoretically, CAR-T therapy targeting CSCs might effectively constrain tumor metastasis. The antigen CD133 is thought to be a specific marker for CSCs, and is also expressed by various carcinomas including cholangiocarcinoma (CCA), a cancer of the bile ducts. A 2017 article reported on a 52-year-old

patient with advanced CCA who was treated with CART-133 after seeing a partial remission following CART-EGFR treatment. A further 4.5-month partial remission was observed (25). Although the killing of CSCs by CART-133 was not conclusively shown, the novel sequential use of the CAR-T cocktail indicated intriguing possibilities. Epithelial cell adhesion molecule (EpCAM)—a transmembrane glycoprotein involved in epithelial cell adhesion—is another CSC marker, and several preclinical studies on CART-EpCAM cells have been published in China (26, 27). Currently, four Chinese CART-EpCAM trials are underway.

CAR-T research in China has been developing rapidly in recent years. However, we cannot overlook the gap in domestic research relative to that seen internationally.

For example, although CAR-T treatment for hematological malignancies is efficient, researchers in the United States continue to make improvements. A group at the Fred Hutchinson Cancer Research Center in Seattle, Washington has been working to optimize the composition of CAR-T treatments. Clinical data has indicated that optimizing the types of T cells used (e.g., including higher percentages of naïve or central memory T cells) or utilizing a defined ratio of CD4-positive to CD8-positive T cells gives better clinical results (28, 29). Along these lines, Xinqiao Hospital in Chongqing has registered a CART-19 clinical trial that will use a CAR-T therapy in which a higher percentage of memory T cells will be used. A case report from Carl June's group at the University of Pennsylvania demonstrated the possibility of treating MM with CART-19 cells, something previously considered to be impossible (30). Such innovative attempts to subvert traditional ideas rarely occur in China.

For CAR-T cell preparation, the vast majority of domestic clinical trials use viral transfection. In contrast, the University of Texas MD Anderson Cancer Center has been promoting the use of the so-called "sleeping beauty" system, a synthetic DNA transposon vector designed to introduce DNA sequences into chromosomes following electroporation (31). These scientists believe this system can better meet future demands for lower cost and greater gene-editing capabilities.

The challenges related to solid tumor treatment have motivated researchers in the United States to develop new strategies and clinical applications that have been rapidly translated into the clinic. One such strategy was developed by a group from City of Hope National Medical Center (CHNMC) in Duarte, California. In order to overcome

the immunosuppressive microenvironment, they designed a PD-L1:CD28 chimeric receptor, converting an inhibitory receptor to a costimulatory receptor by exchanging the transmembrane and cytoplasmic protein domains of PD-L1 with that of CD28. This PD-L1:CD28 chimera exhibited enhanced functional attributes and demonstrated a novel strategy to overcome PD-L1-mediated immunosuppression (32). Another new strategy was developed by a group from MSKCC to avoid "on-target, off-tumor" toxicities; the group designed an antigen-specific inhibitory CAR that could constrain the activation of T cells when a specific antigen was encountered (33). To reduce the risk of off-target toxicity, affinity-tuned CARs that could more precisely distinguish tumor cells from normal cells were tested by Carl June's group (34). Unlike in the United States, original research in this area is lacking in China. However, it is gratifying to note that some innovative clinical trials have also been initiated there. For example, six trials in which CAR-T cells will be modified to express soluble PD-1 or CTLA-4 antibodies to overcome the immunosuppressive microenvironment have been registered.

It is true that clinical trials in the United States are more varied than those in China; China's clinical trials are relatively elementary and mostly similar to one another. Yet the diversity of U.S. trials has brought us valuable insights into how to design effective CAR-T treatments. For example, original U.S. studies have illuminated the benefits of using a preconditioning regimen with the drugs fludarabine and cyclophosphamide to improve the proliferation and persistence of CAR-T cells, and have shown how to optimize CAR-T preparation protocols and the framework of the CAR molecule.

Compared with traditional drugs, CAR-T therapy offers more opportunity for businesses—especially small companies—because of the great diversity of targets and the many possible ways to optimize CAR molecular structures and preparation methods. American businesses have also made outstanding contributions to the industrialization of CAR-T therapy. Of the seven ongoing phase 2 CAR-T trials, six are being conducted by three companies (Kite Pharma, Juno Therapeutics, and Novartis). However, in China, most CAR-T trials are led by institutions or hospitals. Business participation has been mainly limited to providing technical support or supplying viruses. The most effective way to encourage domestic enterprises to participate in the industrialization of CAR-T therapy is a topic deserving further discussion.

The U.S. FDA has issued clear guidelines and regulations for clinical application of CAR-T therapy after extensive input from experts in relevant fields. In China, no formal regulations on CAR-T therapy have yet been issued. This policy uncertainty is a cause for concern for businesses. Achieving consensus within the field and deepening communication with regulators has become a top priority for the industry. In recent years, several related associations and conferences have been established, including the Chinese Research Hospital Association (CRHA), the China Biotherapy Conference (CBC), and the China Disease Biotherapy Conference (CDBC). Meetings over the past few years have effectively promoted communication, cooperation, consensus building, and the standardization of clinical trials. It is gratifying that the regulatory authorities have issued a policy draft on CAR-T therapy, and that formal regulations are expected by the end of 2017. We believe that CAR-T therapy will be regulated as a novel drug type in China as it is in the United States. Furthermore, the government has been actively supporting CAR-T therapy; last year, it announced 14 key research projects, and CAR-T therapy was among them.

### The future of CAR-T therapy

The side effects of a CAR-T cell infusion can include serious consequences, even death. Without proper control, serious clinical accidents could cause significant human and financial losses. For example, a phase 2 CART-19 trial sponsored by Juno Therapeutics was terminated this year because three patients died from cerebral edemas. It is therefore very important to establish more clinical teams with experience in immunotherapy. And it is strongly recommended that the businesses involved in producing the therapies establish close cooperation with experienced clinical teams who can help reduce the risk of clinical accidents.

The preparation of CAR-T cells is complicated and greatly affected by variation among patients. This places a heavy burden on the preparation technologies, which may be an obstacle for some companies wishing to carry out CAR-T clinical trials. To address this problem, some domestic enterprises have begun to develop automated and standardized CAR-T preparation platforms. For example, Sinobiocan (Beijing) Biology Technology is building a one-stop CAR-T preparation apparatus, including an automated cell-preparation platform, and is also providing a steady supply of reliable reagents. As an alternative, "off-the-shelf" universal CAR-T cells can be prepared using the CRISPR/Cas9 system and other gene-editing technologies. Currently, there are

two clinical trials utilizing universal CAR-T technology ongoing in China.

The deficiencies mentioned above point to a clear path for improvements to the current status of CAR-T research in China. We believe that this powerful form of cancer treatment will undergo rapid development with the Chinese government's active support and with our joint efforts.

### References

1. P. Aebbersold *et al.*, *J. Immunol. Methods* **112**, 1-7 (1988).
2. R. O. Dillman *et al.*, *J Clin. Oncol.* **9**, 1233-1240 (1991).
3. I. G. Schmidt-Wolf *et al.*, *J. Exp. Med.* **174**, 139-49 (1991).
4. F. Wang *et al.*, *World J. Gastroenterol.* **8**, 464-468 (2002).
5. G. Gross *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 10024-10028 (1989).
6. H. Wang *et al.*, *Clin. Cancer Res.* **15**, 943-950 (2009).
7. Y. Wang *et al.*, *Clin. Immunol.* **155**, 160-175 (2014).
8. H. Dai *et al.*, *Oncoimmunology* **4**, e1027469 (2015).
9. Y. Hu *et al.*, *Clin. Cancer Res.* **23**, 3297-3306 (2017).
10. J. Pan *et al.*, *Leukemia*, Epub ahead of print (2017), doi: 10.1038/leu.2017.145.
11. B. Cai *et al.*, *J. Hematol. Oncol.* **9**, 131 (2016).
12. B. Guo *et al.*, *J. Cell. Immunother.* **2**, 28-35 (2015).
13. G. Tamamyan *et al.*, *Crit. Rev. Oncol. Hematol.* **110**, 20-34 (2017).
14. A. Ehninger *et al.*, *Blood Cancer J.* **4**, e218 (2014).
15. Q. S. Wang *et al.*, *Mol. Ther.* **23**, 184-191 (2015).
16. C. M. Wang *et al.*, *Clin. Cancer Res.* **23**, 1156-1166 (2017).
17. X. Yu *et al.*, *Genet. Mol. Res.* **14**, 10185-10192 (2015).
18. H. Gao *et al.*, *Clin. Cancer Res.* **20**, 6418-6428 (2014).
19. K. Li *et al.*, *Oncotarget* **7**, 2496-2507 (2016).
20. X. Zhou *et al.*, *Neoplasia* **15**, 544-553 (2013).
21. K. Feng *et al.*, *Sci. China Life Sci.* **59**, 468-479 (2016).
22. F. You *et al.*, *Sci. China Life Sci.* **59**, 386-397 (2016).
23. K. Feng *et al.*, *Protein Cell*, Epub ahead of print (2017), doi: 10.1007/s13238-017-0440-4.
24. C. Zhang *et al.*, *Mol. Ther.* **25**, 1248-1258 (2017).
25. K. C. Feng *et al.*, *J. Hematol. Oncol.* **10**, 4 (2017).
26. Z. Deng *et al.*, *BMC Immunol.* **16**, 1 (2015).
27. W. X. Ang *et al.*, *Oncotarget* **8**, 13545-13559 (2017).
28. D. Sommermeyer *et al.*, *Leukemia* **30**, 492-500 (2016).
29. C. J. Turtle *et al.*, *Sci. Transl. Med.* **8**, 355ra116 (2016).
30. A. L. Garfall *et al.*, *N. Engl. J. Med.* **373**, 1040-1047 (2015).
31. S. N. Maiti *et al.*, *J. Immunother.* **36**, 112-123 (2013).
32. M. E. Prosser *et al.*, *Mol. Immunol.* **51**, 263-272 (2012).
33. V. D. Fedorov, M. Themeli, M. Sadelain, *Sci. Transl. Med.* **5**, 215ra172 (2013).
34. X. Liu *et al.*, *Cancer Res.* **75**, 3596-3607 (2015).

### Acknowledgments

This research was supported by grants from the Science and Technology Planning Project of Beijing City (Z151100003915076 to WDH), the National Key Research and Development Program of China (2016YFC1303501 and 2016YFC1303504 to WDH), and the National Natural Science Foundation of China (81502679 to CL).

# Precision cancer medicine and immunology in China

Xu-Chao Zhang<sup>1,2</sup> and Yi-Long Wu<sup>1\*</sup>

**P**recision medicine in the context of clinical oncology treats patients by devising therapies based on information about genetic and epigenetic changes to the genes in their tumors. Canonical examples of such mutated genes include epidermal growth factor receptor (*EGFR*) and anaplastic lymphoma kinase (*ALK*) genes in lung cancer, human epidermal growth factor receptor 2 (*HER2*) in breast cancer, and the *BCR-ABL* gene fusion in chronic myeloid leukemia (CML), among many others. The term “precision medicine” incorporates personalized medicine, since in a clinical setting it implies choosing the right drug for the right molecular target in the right patient. Precision medicine has dramatically changed clinical practice and is further shaping the landscape of oncology treatment standards.

Chinese investigators have contributed significantly to the establishment of standard treatments for non-small-cell lung cancers (NSCLCs) in which the *EGFR* and *ALK* genes are mutated. In 2009, the world’s first phase 3 study was carried out showing the efficacy of first-generation *EGFR* tyrosine kinase inhibitor (TKI) drugs in patients whose tumors had a mutated *EGFR* gene—a landmark in the study and treatment of lung cancer (1). Since then, there have been further studies of second- and third-generation *EGFR* TKIs in China (2–5). In addition to *EGFR*, several trials of drugs targeting *ALK*, *ROS1*, and other so-called “driver genes” for lung cancer have generated good evidence for the establishment of new standards of care for patients with specific genomic alterations, which also promoted the approval from the China Food and Drug Administration (CFDA) for new drugs to be marketed in China (6, 7). Research into immune therapies has also been making rapid progress, and some have already been approved for use in the United States and other countries. Early clinical trials of these therapies in China are coming to fruition, and it is hoped that the resulting

treatments will be approved by the CFDA in the near future. This area is very competitive in China because of the number of similar drugs available internationally; several local antibody-based drugs are entering the arena, although they are still only in early trials.

Progress in precision medicine has already changed the practice of oncology in China, which is seeing broader use of genotyping and TKI-based treatments. Nonetheless, the availability of targeted therapies is affected by several factors and could benefit from certain improvements. For example, the processes by which Chinese patients are selected to receive immunotherapy need both regulatory and clinical consensus.

In this article, we will talk about the current state of precision medicine in China, specifically in the treatment of lung cancer. We will also address future directions and strategies in this area.

## Precision medicine in lung cancer treatment

Lung cancer is a highly heterogeneous group of diseases. In Chinese lung cancer patients, the major driver genes—including *EGFR*, *ALK*, *ROS1*, *RET*, *MET*, *HER2*, *BRAF*, and *KRAS*—are very similar to those listed in the guidelines published by the National Comprehensive Cancer Network (NCCN), an alliance of American cancer centers (8, 9). However, there are notable differences. The *EGFR* mutation rate, for instance, is around 35% in Chinese patients, and only 15% in Caucasian patients, while the *KRAS* mutation is around 5% to 10% in Chinese lung cancer patients versus 30% in Caucasians. There is no obvious difference in prevalence between populations for other mutated genes that occur with lower frequency (Figure 1, next page).

*EGFR* mutations were identified as a focus for targeted therapy for lung cancer in 2004. With the first report of *EGFR* mutations in the Chinese population and preliminary data about the response rate to TKIs, Chinese investigators Tony Mok, Yi-Long Wu, and others led a milestone phase 3 randomized study, named the Iressa Pan-Asia Study (IPASS), published in the *New England Journal of Medicine* in 2009 (1). The study showed for the first time that an *EGFR* TKI was effective in treating *EGFR* mutant cancers. Based on IPASS and follow-on studies, *EGFR* TKI was established as a first-line standard treatment for patients with *EGFR* mutated tumors, and was incorporated into the treatment guidelines of numerous international cancer organizations including NCCN, the American Society of Clinical Oncology (ASCO),

<sup>1</sup>Guangdong Lung Cancer Institute and <sup>2</sup>Medical Research Center, Guangdong General Hospital and Guangdong Academy of Medical Sciences, Guangzhou, China

\*Corresponding author: syylwu@live.cn

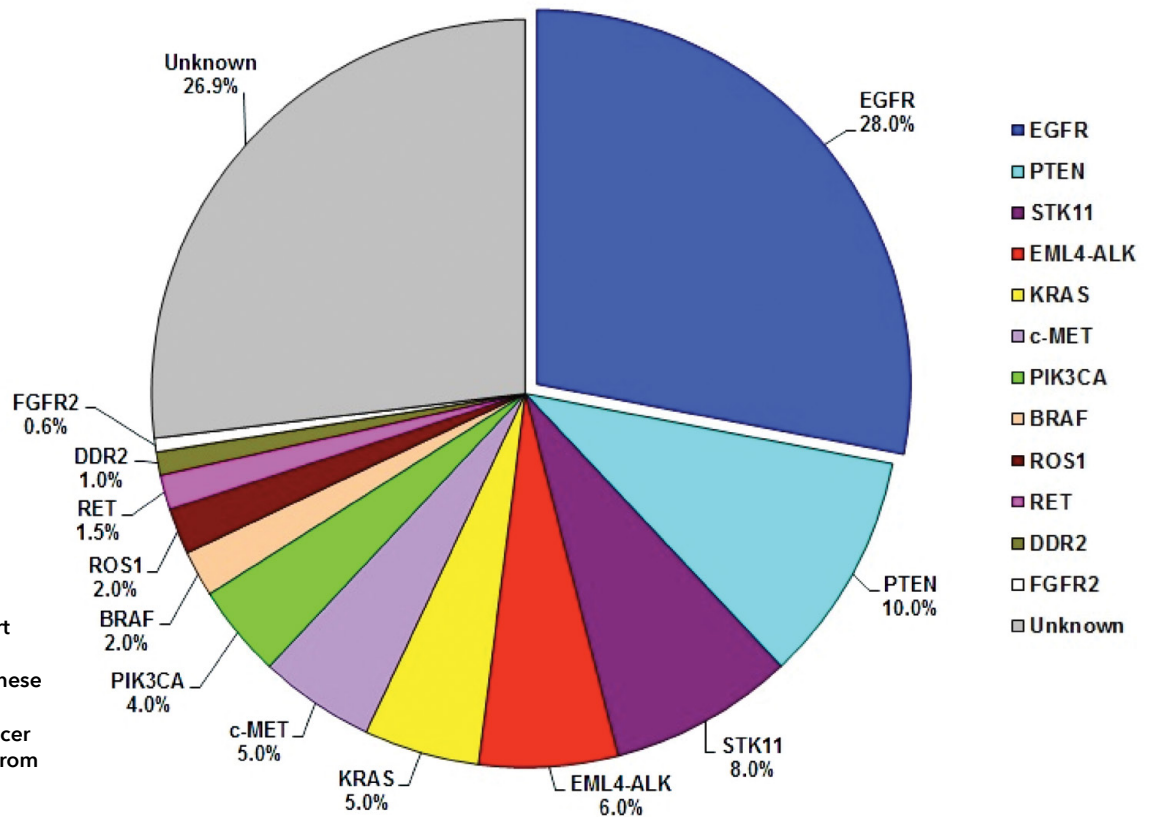


FIGURE 1. Pie chart of the driver gene distribution in Chinese patients with non-small cell lung cancer (NSCLC) (revised from Guangdong Lung Cancer Institute).

and the European Society for Medical Oncology (ESMO). In 2007, oncologist Yi-Long Wu, who is also on the board of directors of the International Association for the Study of Lung Cancer (IASLC) and president of the Chinese Society of Clinical Oncology (CSCO), founded the Chinese Thoracic Oncology Group (CTONG). CTONG now has 31 member hospitals across China and has become an important platform for organizing trials and translational research studies. It contributed significantly to the development of targeted therapies in China, not only participating in several international synchronized trials, but also taking the leading role in dozens of national studies.

All the above studies bolstered the registration and approval of targeted therapies in China as well as the capacity for designing and organizing trials. The CFDA has now approved three first-generation EGFR TKIs (gefitinib, erlotinib, and icotinib), one second-generation EGFR TKI (afatinib), and a third-generation EGFR TKI, osimertinib, for *EGFR* mutant tumors. Crizotinib, aimed at Chinese patients who have *ALK* or *ROS1* mutations, was also approved by the CFDA. Other *ALK* TKIs are in clinical trials led

by investigators from CTONG, while TKIs targeting numerous other mutated genes are in different stages of clinical trials or already in use in clinical settings.

### Novel trial design for oncology

A pragmatic question faces developers of targeted therapies: How can they ensure that a sufficient number of patients with specific genetic alterations are enrolled in the clinical trials for each drug? The idea of simultaneously screening patients for a set of genetic biomarkers linked to the disease and matching the drugs under investigation to each specific genetic or pathway alteration has now been adopted worldwide.

Yi-Long Wu, as the president of CSCO, has advocated this kind of trial design for many years, attempting to push molecular genotyping and targeted therapies into clinical practice. In 2015, Wu launched the CLUSTER 1.0 trial, which is the first multiple-biomarker-driven clinical research study to be conducted in Asia (10). The enrolled patients were assigned to groups receiving different agents according to their genetic

profiles. Five agents targeting different molecular abnormalities are being studied based on the mutation(s) found, making up the five arms of the study. Each arm will be analyzed independently. If the response rate to an agent reaches 40%, a phase 2 trial will be conducted. This multiple-arm, parallel assignment is intended to reduce the overall drug screening failure rate. This study drew much attention from the Center for Drug Evaluation of the CFDA, and is a significant milestone for precision medicine in China. Design of the follow-up CLUSTER 2.0 lung cancer trial is complete, and will be launched in the fourth quarter of 2017. CLUSTER 2.0 is also biomarker-driven and will evaluate novel targeted agents in at least 10 study arms focused on specific genetic or pathway alterations.

During the process of developing new anticancer agents such as icotinib, avitinib, and apatinib, Chinese researchers have already had experience with innovative clinical trial design. One example is a phase 1/2 study that aims to determine the safety profile and recommended phase 2 dose (RP2D) of avitinib in patients carrying a particular *EGFR* mutation. Instead of using only the maximum tolerated dose, as is generally done, in this study RP2D was determined by the pharmacokinetics, preliminary efficacy, and safety profile of the agent. This innovative design established a new paradigm that allows the testing of a wider range of dosages showing indications of efficacy and low toxicity at an early stage. Based on the expedited determination of RP2D, in 2016 researchers applied for approval from the CFDA to conduct a phase 2/3 trial with avitinib—the fastest-ever application for a novel agent in China.

Currently, the CFDA has approved only five *EGFR* TKIs (gefitinib, erlotinib, icotinib, afatinib, and osimertinib) and one *ALK/ROS1* TKI (crizotinib). No drugs are clinically available for lung cancer patients with other common mutations. Therefore, to be enrolled in a new trial offers such patients a way to gain access to precision medicine. Indeed, through these ongoing innovative clinical trials, patients with actionable genomic alterations can quickly access drugs under study internationally, as well as drugs approved in the United States, Japan, or Europe (but not in China).

### **Guidelines for biomarker testing, diagnosis, and treatment**

In precision oncology, molecular diagnosis precedes precise treatment. *EGFR* and *ALK* testing, and their corresponding treatments, have been

included in the guidelines of numerous cancer associations, including NCCN, ASCO, and ESMO. IASLC has released the *EGFR* and *ALK* testing atlas, a series of consensus documents on molecular testing (11).

In China, local guidelines for different tumor types have been formulated. Since 2016, CSCO has taken the lead in compiling the annual *Guidelines for Diagnosis and Treatment* for patients with lung and other cancers. The Ministry of Health and Family Planning Committee, formerly the Ministry of Health, also organized national experts to produce *Clinical Path and Guidelines for Diagnosis and Treatment for Primary Lung Cancer Patients*. Yi-Long Wu played a leading role in compiling both of these guidelines as they pertain to lung cancer. In the CSCO guidelines, due to the limited availability of drug and testing technologies, and disparities and imbalances in resources distribution, only *EGFR* mutations and *ALK* and *ROS1* rearrangements and their targeted drugs are listed for routine clinical practice. Circulating tumor DNA from plasma is regarded in the report as an alternative biomaterial to direct tumor sampling for screening of *EGFR* mutations, but its potential for producing false negatives is emphasized. The national and local guidelines for testing, diagnosis, and treatment are modified annually to account for progress and new evidence from trials.

Several *EGFR* and *ALK* gene testing kits have been approved by CFDA and are also recommended by national or local guidelines. According to one as-yet unpublished report, the overall rate of testing for *EGFR* mutations in patients in China is around 50%, which although higher than the previous rate, is still not on par with the rate in developed Western countries, Japan, Korea, and Singapore.

Many leading hospitals carry out molecular testing in their own laboratories under license from the Ministry of Health and Family Planning Committee, using CDFA-approved kits or tests developed in-house. Some hospitals, especially those with fewer resources, outsource molecular testing to third-party laboratories. As sophisticated next-generation sequencing (NGS) technology comes into wide use, it is reported that many companies are emerging across the country to deliver that service. The CSCO biomarker committee, led by Yi-Long Wu and Xu-Chao Zhang, is paying attention and making efforts to standardize the clinical use of NGS technology and to ensure testing quality for clinical oncology practice. In a draft of *Consensus of the Use of NGS in Clinical Oncology*, to be released in the near

future, NGS technology is regarded as a complex genotyping tool for which quality control is critical. In addition, diagnostic kits based on NGS used to stratify patients with specific gene mutations are currently being developed by many companies. They will be validated as companion diagnostic tools in the CLUSTER 2.0 trial.

For patients with driver genes or genetic alterations other than *EGFR*, *ALK*, or *ROS1*, the CSCO guidelines recommend participation in trials in which investigational drugs are provided. Patients with disease relapse or those who are resistant to the available TKIs are encouraged to undergo additional molecular screening to clarify the resistance mechanism and to participate in clinical trials pertinent to their diagnosis.

Precision medicine in China has made substantial progress in the treatment of lung cancers harboring *EGFR*, *ALK*, and *ROS1* alterations. Chinese investigators have contributed their efforts to the global and national establishment of standard targeted therapies. Yet if precision oncology is to improve, more effort is needed to improve the detection rate of key driver genes like *EGFR*, to develop new drugs, to launch novel trials, and to standardize molecular testing.

### **Immunology of cancer**

Internationally, immune therapy has been recognized as a potentially powerful treatment since the establishment of the field of immuno-oncology more than 100 years ago. In the United States, Europe, and Japan, immune checkpoint inhibitors—such as drugs that target CTLA-4, PD-1, and PD-L1, proteins known to inhibit T cells from fighting cancer—have already been approved by regulatory agencies for a variety of cancer types including melanoma and lung cancer. By the end of December 2016, the U.S. Food and Drug Administration (FDA) had approved therapies that block the PD-1 pathway, including nivolumab, pembrolizumab, and atezolizumab. The working principle of this kind of inhibitor is to remove the brakes on immune activation or block the inhibition of immune attack in a tumor's microenvironment. Antibodies against other immune checkpoints like TIM3 and IDO are in the preclinical or clinical stage of development.

### **Immune therapy for lung cancer**

In China, no PD-1 or PD-L1 inhibitors have been approved by CFDA. The contrast with the United States indicates that Chinese R&D in immuno-

oncology biologics (protein-based drugs) lags behind, as it does in the area of small-molecule drugs. However, various clinical trials are actively investigating both international and domestic drugs in Chinese patients with a range of cancer types including lung cancer.

Between January 1, 2013 and April 6, 2017, ClinicalTrials.gov registered 270 international clinical trials of PD-1/PD-L1 therapies for NSCLC, including trials for nivolumab, pembrolizumab, atezolizumab, and durvalumab. These 270 trials included 61 studies that involved East Asian sites and 14 that involved Chinese sites, 12 of which were multinational trials and 2 that evaluated only Chinese patients (12).

In the 14 international trials that enrolled Chinese patients, six involve first-line/primary therapies, four involve second-line therapies (used if primary treatments fail), two are studies of adjuvant therapies, and two are phase 1 studies using only Chinese patients. There are five studies for atezolizumab, four for durvalumab, three for pembrolizumab, and two for nivolumab. It is noteworthy that PD-L1 expression was set as a criterion for recruitment in some trials, but not in others, and that different diagnostic antibodies and screening platforms were used in this determination.

Nonuniform PD-L1 expression testing and cutoffs make it extremely difficult to compare results from different trials. Nonetheless, it is encouraging that Chinese pharmaceutical companies are working intensively to develop PD-1 and PD-L1 drugs. With the efforts of CFDA to reform the regulatory framework for clinical trial approval of novel agents and to adjust its policies to respond to advances from international clinical trials, immunotherapy trials are now getting underway faster. Chinese pharmaceutical companies had developed eight PD-1 or PD-L1 inhibitors by January 7, 2017, and four drugs have been approved by CFDA for phase 1 trials among patients with advanced solid tumors including NSCLC. Another four drugs are now being considered by CFDA for clinical trial approval. In November 2016, the PD-1 inhibitor KN035, which is administered by subcutaneous injection, received approval from the U.S. FDA for clinical trials.

Notably, the immuno-oncology market is very competitive for domestic Chinese pharmaceutical companies. As of October 2017, there are 17 inhibitors targeting immune checkpoints PD-1,

PD-L1, or CTLA-4 from 14 domestic corporations treating a range of cancers (Table 1).

### **Improving patient selection for immune checkpoint inhibitor therapy**

PD-L1 protein expression on tumor cells or immune cells in the tumor microenvironment has been widely used as a companion diagnostic, or evaluated retrospectively in clinical trials. The overall response rate to PD-L1 inhibitor therapy in patients selected for using PD-L1 expression levels increased

from 20% to 40%–60%, a smaller increase than expected (13). However, testing for PD-L1 protein expression by immunohistochemistry (IHC) was performed using different platforms, antibodies, and scoring methods, making it difficult to compare results from different trials. Hirsch and colleagues reported that in one of the studies, the discrepancies caused by testing for PD-L1 on different platforms could result in misclassification of PD-L1 status (14).

In China, no PD-L1 protein expression tests or kits have been approved by CFDA. No PD-L1 test is recommended in the CSCO guidelines for any cancer type. In contrast, a PD-L1 IHC test has been approved by the U.S. FDA. In ongoing trials in China, a PD-L1 test is taken as one criterion for inclusion in certain trials but not in others. Whether PD-L1 should be used as a biomarker to select Chinese patients for PD-L1 inhibitor treatment requires further investigation.

Are there any potential biomarkers other than PD-L1 to be developed for the Chinese population? Theoretically, molecules in the microenvironment of a tumor and in the circulating blood could potentially be predictive markers. Such molecules could be involved in any one of the many steps between the appearance of an antigen and the launch of an immune attack on tumor cells. Biomarkers such as tumor mutational burden, DNA mismatch repair status, high microsatellite instability, and circulating T-cell characteristics are under extensive investigation in trials using several cancer types.

In ongoing clinical trials in China, investigators have a good opportunity to study these biomarkers by testing biomaterials collected from tumor and blood samples. In one study (with the acronym CHOICE) led by Yi-Long Wu, researchers looked at the immune scores of 250 heavy smokers with lung cancer by analyzing the RNA sequencing data of the immune cells isolated from the tumors. Preliminary data showed a correlation between

Company	MoA	Drug/Asset
Jiangsu Hengrui Medicine	PD-1	SHR-1210
	PD-L1	SHR-1316
Shanghai Junshi Bioscience	PD-1	JS001
	PD-1	GB226
Innovent Biologics	PD-1	GB226
BeiGene	PD-1	BGB-A317
Genor BioPharma	PD-1	IBI308
3DMed	PD-1	KN035
	CTLA-4	KN044
Harbin Gloria Pharmaceuticals	PD-L1	GLS-010
CStone Pharmaceuticals	PD-L1	Anti-PD-L1
Akeso Biopharma	PD-1	AK101, AK103
Bio-Thera	PD-1	Anti-PD-L1
Livzon Pharmaceutical	PD-1	Anti-PD-L1
Lee's Pharma	PD-1	Anti-PD-L1
Sichuan Kelun	PD-L1	KL-A167
Shanghai Henlius Biotech	PD-1	Anti-PD-L1

**TABLE 1. Summary of domestic pharmaceutical companies with R&D in immuno-oncology.** Source: [www.chinadrugtrials.org.cn](http://www.chinadrugtrials.org.cn). MoA, mechanism of action.

immune score and disease prognosis (unpublished data). This RNA-sequencing-based method can be further evaluated in clinical trial patients to see if the immune scores can be predictive for positive response to immunotherapy. It is important to note that some biomarkers require sophisticated technologies like multicolor IHC, whole-exome deep sequencing, or the use of cancer gene panels. One other challenge is how to integrate several different kinds of immune-related biomarkers, for example, combining tumor mutational burden with PD-L1 expression.

Overall, the development of novel biomarkers to predict and guide immune therapy using PD-1/PD-L1 inhibitors needs further exploration, particularly in trial patients who can provide ongoing samples and whose treatment and follow-up data can be monitored.

### Summary and future directions

Over the past 10 years, the practice of clinical oncology in China has gradually incorporated the attributes of precision medicine, as is happening in other countries. There has also been significant progress in clinical research and drug development. Chinese investigators have contributed in important ways to international studies, and have conducted trials based on the particular characteristics of Eastern populations. A series of novel targeted agents with good efficacy has been synthesized and introduced for clinical use. In the arena of immune therapy, dozens of national and international trials are underway in the hope of moving PD-1/PD-L1 inhibitors into clinical use. Eight such inhibitors from local Chinese companies are in the early stages of clinical development for use against a variety of cancer types. And a handful of other small-molecule TKIs and immune checkpoint inhibitors are undergoing validation in clinical trials. National and local guidelines for clinical practice have been written in recent years to standardize the use of these effective new treatments.

Nonetheless, China is lagging behind in some areas, including development of novel drugs, availability of TKIs and antibody drugs that have already been approved in other countries, development of companion diagnostics, and availability of clinical biomarker tests. Furthermore, many domestic novel anticancer agents are structural modifications of agents that are already on the market or are undergoing research; innovative design is rare.

However, there are many possible future directions. The pharmaceutical industry should shift its focus from generic agents to the development of

pioneering agents. New technologies like NGS and liquid biopsies of ctDNA can be effective ways to push innovative biomarker-driven trials forward, especially for those patients with rare genetic alterations in cancer-related genes. Attention should also be paid to the mechanisms that cause resistance to EGFR and ALK TKIs, and to drugs that can overcome this resistance. More translational studies on immunotherapies including checkpoint inhibitors and adoptive cell therapy need extensive exploration. Finally, development of biomarker companion diagnostics and standardization of biomarker tests in a clinical setting should be addressed.

It is well known that much effort is needed to improve drug development by local companies, as well as the competitiveness of those companies. The Chinese government has policies to encourage innovative R&D on new drugs and to incubate pioneering startups in China. Influenced by the Precision Medicine Initiative declared by former American President Obama and the Cancer Moonshot Initiative led by former Vice President Biden, China initiated a set of national science and technology programs designed to stimulate R&D on precision medicine in oncology.

To further drive precision oncology forward, all stakeholders, including the government, academic investigators, pharmaceutical companies, innovative biotechnology companies, patients, and disease advocates should come together and bring their resources to bear. Only this way will we be able to bring precision medicine to the patients who need it most.

### References

1. T. S. Mok *et al.*, *N. Engl. J. Med.* **361**, 947-57 (2009).
2. J. C. Yang *et al.*, *Lancet Oncol.* **16**, 830-838 (2015).
3. J. C. Yang *et al.*, *Lancet Oncol.* **16**, 141-151 (2015).
4. Y. L. Wu *et al.*, *Lancet Oncol.* **15**, 213-222 (2014).
5. T. S. Mok, Y. L. Wu, V. A. Papadimitrakopoulou, *N. Engl. J. Med.* **376**, (1993).
6. B. J. Solomon *et al.*, *N. Engl. J. Med.* **371**, 2167-2177 (2014).
7. J. J. Yang *et al.*, *Clin. Cancer Res.* **20**, 1383-1392 (2014).
8. D. S. Ettinger *et al.*, *J. Natl. Compr. Canc. Netw.* **15**, 504-535 (2017).
9. "NCCN Clinical Practice Guidelines in Oncology," National Comprehensive Cancer Institute; available at [https://www.nccn.org/professionals/physician\\_gls/f\\_guidelines.asp#nsc1](https://www.nccn.org/professionals/physician_gls/f_guidelines.asp#nsc1).
10. S. Su, Y. L. Wu, *J. Hematol. Oncol.* **10**, 147 (2017).
11. N. B. Leighl *et al.*, *J. Clin. Oncol.* **32**, 3673-3679 (2014).
12. S. Y. Liu, Y. L. Wu, *J. Hematol. Oncol.* **10**, 136 (2017).
13. F. R. Hirsch *et al.*, *Lancet* **389**, 299-311 (2016).
14. F. R. Hirsch *et al.*, *J. Thorac. Oncol.* **12**, 208-222 (2016).

# From big data to knowledge in precision medicine

Dechao Bu<sup>1,2†</sup>, Shaoliang Peng<sup>3†</sup>, Haitao Luo<sup>1</sup>,  
Xiyuan Li<sup>1</sup>, Hui Li<sup>1</sup>, Liang Sun<sup>1,4</sup>, Yang Wu<sup>1</sup>,  
Li Xiao<sup>1</sup>, Shuangfang Fang<sup>1</sup>, Chunlong Luo<sup>1</sup>,  
Wenzheng Fang<sup>1</sup>, Deju Kong<sup>1</sup>, Shengyou Liao<sup>1</sup>,  
Shunmin He<sup>5</sup>, Runsheng Chen<sup>5\*</sup>, and Yi Zhao<sup>2\*</sup>

## The explosive growth of medical and 'omics big data

The biomedical field is experiencing an era of significant growth due to the large amount of clinical, biomolecular, and health data being collected. The increased adoption of electronic medical record systems worldwide has enabled large volumes of clinical data to be captured and stored. These data include screening, diagnostic, and treatment results as well as medical and family histories, and data from biochemical testing, imaging, and follow-ups. Furthermore, advanced technologies that collect biomolecular data from studies in genomics, transcriptomics, proteomics, epigenomics, metabolomics, metagenomics, and exposomics—collectively termed “omics”—can provide results within hours or days, dramatically speeding up clinical analysis. “Multi-omics” molecular data are widely used in the field of precision medicine and related clinical applications. These data constitute one of the fastest-growing data types in biomedicine. Moreover, a large amount of daily biological data related to personal health is being continuously collected by wearable devices and smartphones. These multidimensional data require efficient methods of storage and analysis if they are to be useful tools for exploring the mysteries of disease and health.

## Large cohort studies in Chinese precision medicine

As a consequence of major advances in gene sequencing technology, programs promoting

precision medicine have been launched worldwide, with the aim of improving accuracy in screening, diagnosis, intervention, and treatment. Scientists around the world, but particularly in the United States, the United Kingdom, and China, are collecting the genomic data of millions of people to establish a database of genetic variants. In 2016, the Chinese Ministry of Science and Technology founded the Precision Medicine Research Key Special Project (PMRKSP) and announced an investment of billions of Chinese yuan before 2030. At present, three types of large cohort studies have been launched as part of this effort. The first is the study of natural national populations, in which millions of participants from seven main regions covering four municipalities and 24 provinces of China will be analyzed. The second is a study of eight cohorts affected by critical diseases, including cardiovascular, cerebrovascular, respiratory, metabolic, neurological, psychosomatic, and immune system disorders as well as seven common malignant tumors. A primary goal of this study is to collect samples from a total of 700,000 people. The third investigation is a clinical cohort study of 50 rare diseases that aims to collect more than 50,000 patients in a single cohort. In addition, the Chinese Academy of Sciences (CAS) has announced a precision medicine research plan targeting the collection of genetic information from 4,000 Chinese volunteers within four years.

## From big data to knowledge

The meticulous analysis of big data is essential to extract accurate information and ultimately convert them into knowledge (1) that can assist in clinical decision-making and health management (Figure 1, next page). In order to support this transformation, the Chinese government has developed a series of plans that focus on key technologies involved in processing large amounts of data. The 13th Five-Year National Science and Technology Innovation Plan, issued by the State Council, will develop technologies for precision medicine, integrate them into a multilevel knowledge database, and create a national platform to share biomedical big data. PMRKSP has also initiated a series of research projects that focus on the most important technologies for big data integration, storage, use, and sharing. The program has taken three steps toward transforming data into information, and information into knowledge. This program first started to build fundamental data infrastructure (data management, processing systems, and

<sup>1</sup>Bioinformatics Research Group, Key Laboratory of Intelligent Information Processing, Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China

<sup>2</sup>Luoyang Branch of the Computing Technology Institute, Chinese Academy of Sciences, Luoyang, China

<sup>3</sup>School of Computer Science, National University of Defense Technology, Changsha, China

<sup>4</sup>Beijing Zhongke Jingyun Technology Company Ltd., Beijing, China

<sup>5</sup>Chinese Academy of Sciences Key Laboratory of RNA Biology, Institute of Biophysics, Chinese Academy of Sciences, Beijing, China

<sup>†</sup>Contributed equally to this work.

\*Corresponding authors: rschen@ibp.ac.cn (R.S.C.), zhaoyi@gene.ac (Y.Z.)

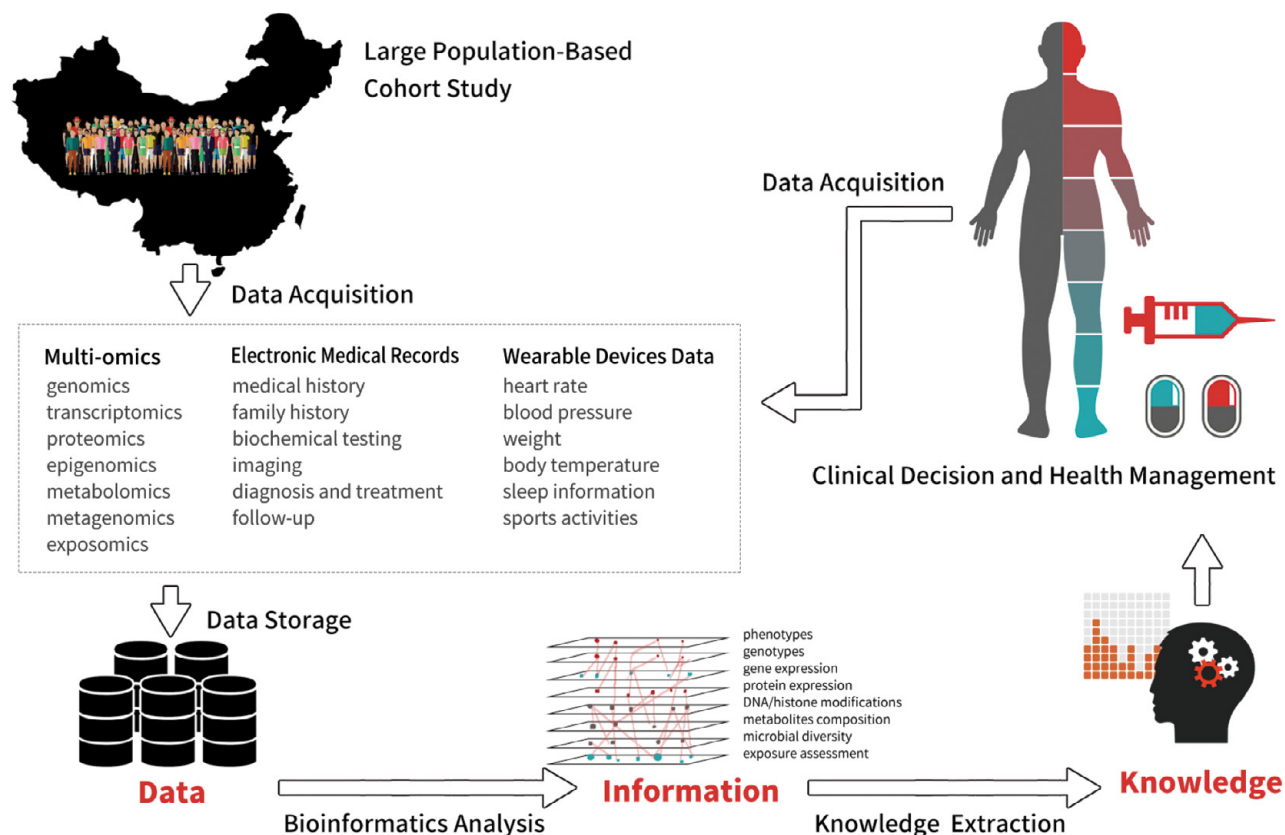


FIGURE 1. Research and applications of big data in precision medicine.

disease databases) in 2016. At the beginning of 2017, scientists began processing the raw data generated from PMRKSP into multilayered biomolecular information repositories, which assist in the integration of detailed standardized information and metadata, including clinical patient information and its associated genomic, proteomic, metabolic, and other 'omics data. Finally, by the start of 2018, knowledge mining will begin on the results of the first two stages of work, with the purpose of discovering important genetic diseases, key mutation sites, and epigenetic information. This project will facilitate the implementation of research in precision diagnosis and the treatment of disease, and will lead to important developments in precision medicine.

### Construction of big data platforms for precision medicine

With the support of a national policy concerning big data platforms, China has been able to construct

a comprehensive infrastructure to successfully implement precision medicine across both state-owned and private health care sectors. Under the guidance of China's National Health and Family Planning Commission, a new company known as China Healthcare Big Data Development Co., Ltd. (CHBDCC) was established. CHBDCC designated two provinces and four cities to participate in a pilot program of national health care big data centers. The China National GeneBank (CNCB), whose establishment was approved by the National Development and Reform Commission, was launched in 2016 and will store and manage the country's unique genetic resources. In addition to the national big data centers, domestic biological big data platforms such as BGI Online ([www.bgionline.com](http://www.bgionline.com)) and JingYun ([www.gene.ac](http://www.gene.ac)) have been built to provide fast and accurate computing services in bioinformatics. Information technology cloud service providers, including Alibaba and Huawei, provide specialized cloud storage and computing services for biomedical

data, which allow information technology and biotechnology to be closely integrated.

### **The potential of medical big data**

Medical and 'omics big data is improving health care efficiency and quality by transforming data into knowledge (2-4). There are three aspects of big data that best describe its potential value. Firstly, the data can describe the entire process of a disease by finding the hidden associations and patterns within complex data, uncovering molecular mechanisms, revealing biomarkers, identifying disease-associated genetic factors, and evaluating therapeutic effectiveness. Secondly, the genomic and clinical data can be used to predict the risk of genetic diseases, the occurrence of cancers, and the clinical outcome of treatments aimed at novel disease targets. Finally, big data can help doctors to prevent disease by prescribing personalized health care regimens using multidimensional data from individuals as well as real-time health and disease information available through the dynamic collection of data acquired daily through personal health monitoring devices.

### **Challenges of big data use in precision medicine**

#### **Data sharing**

To maximize the contributions of big data, generate reproducible results, and improve research practice, it is essential to allow data to be open and shared among different institutions (3). PMRKSP requires that data from clinical studies must be unconditionally shared through a national big data center. However, a mutually beneficial alliance is necessary to encourage research institutions and businesses to share key data generated outside of PMRKSP.

#### **Security and privacy of data**

In the future, a large volume of multi-omics information on populations and electronic medical and health records will be stored in big data platforms. If personal privacy is not guaranteed, ethical issues are bound to arise. To avoid privacy disputes, the government should formulate strong data protection laws. Additionally, methodologies that do not reveal an individual's information, such as multiparty computation, are indispensable and should be established (5).

#### **Standardization of data generation and analysis**

Currently, clinical research based on data from

next-generation sequencing (NGS) lacks standards and specifications for sample collection, storage, handling, and analysis. These issues lead to poor data quality, errors in analysis, and misinterpretation of results. In 2016, the Chinese Society of Clinical Oncology released the first draft of "Consensus on Second Generation Sequencing Technology Applied in Precision Diagnosis and Treatment of Clinical Cancer" in order to promote the standardization of NGS in clinical testing. The following year, experts from West China Hospital of Sichuan University and Peking Union Medical College Hospital published the "Expert Consensus on Second Generation Gene Sequencing in Clinical Molecular Pathology Laboratories." These documents will greatly accelerate standardization of the generation and interpretation of NGS data. However, more informative industry standards are needed to ensure the reliability and reproducibility of precision medical data.

#### **Integration of multidimensional 'omics data**

To overcome the deficiencies of the single 'omics approach, it is crucial to introduce redundancy by generating multi-omics data that allow information on different types of biomolecular interactions to be extracted. In addition, expanding the scope of the methods used to collect data, such as including data from noncoding DNA regions, will further enhance the power of the big data analyses being performed.

#### **Deep integration with information technology**

Medical big data raise significant challenges in storage and computation, which urgently require innovative information technology approaches to provide solutions. A key problem associated with big-data knowledge mining is the rapid rate at which the size and complexity of the information network expands, resulting in bottlenecks in computing speed. One resolution to this issue has recently been offered by the Chinese technology company Hanwuji Intelligence, which improved computing speed by using a dedicated processor that analyzes data through a type of machine learning known as "deep learning" (6). This solution suggests that resolving the unique problems associated with medical big data processing will require a range of cutting-edge information technology, potentially requiring dedicated hardware and software to be embedded in the big data platform.

#### **Future trends in big data for precision medicine**

Future big data technology will be closely inte-

grated with artificial intelligence (AI) technology to promote the development and application of precision medicine. The clinical application of IBM's Watson supercomputer has demonstrated that AI can enable clinical decisions to be made quickly and accurately through the integration and analysis of big data. At present, China has insufficient medical resources and is facing problems with an irrationally structured medical system as well as unbalanced distribution of its medical resources. This situation makes it difficult for citizens to seek medical treatment, as is explained by the State Council in its position paper *Key Tasks to Deepen the Reform of the Medical and Health System in 2016*. To address these issues, the government is establishing diagnosis and treatment guidelines, strengthening basic services, and promoting better allocation of high-quality medical resources for Chinese citizens. The intelligent diagnosis and treatment system constructed with big data, in conjunction with AI, will incorporate the experience of experts, effectively enhancing the quality of grass-roots services available and addressing the deficit in the availability of medical resources in China.

#### References

1. R. Margolis et al., *J. Am. Med. Inform. Assoc.* **21**, 957–958 (2014).
2. T. B. Murdoch et al., *J. Am. Med. Assoc.* **309**, 1351–1352 (2013).
3. V. Marx, *Nature* **498**, 255–260 (2013).
4. E. Bender, *Nature* **527**, S1 (2015).
5. K. A. Jagadeesh et al., *Science* **357**, 692–695 (2017).
6. T. Chen et al., in *Proceedings of the 19th International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS '14)*, Salt Lake City, UT, U.S.A., March 1–5, 2014 (ACM, New York, 2014), pp. 269–284.

## The role of circulating cell-free DNA in the management of cancer in China

Ying Hu<sup>1‡</sup>, Yanhui Chen<sup>1‡</sup>, Lei Zhang<sup>2‡</sup>,  
Haitao Zhao<sup>3</sup>, Hui Zeng<sup>1\*</sup>, and Henghui Zhang<sup>1,2\*</sup>

### Clinical needs in the management of cancer in China

Cancer is a major public health problem in China. Wanqing Chen and colleagues reported that an estimated 4.9 million new cancer cases and 2.8 million cancer deaths occurred in China in 2015, with lung cancer being the most common and the leading cause of cancer death (1). Antitumor treatments currently being developed in China aim to improve overall survival and quality of life for patients.

Some of the most important clinical needs for effectively managing cancer include:

1. Biomarkers that are sensitive and specific to particular kinds of cancer cells, enabling earlier diagnosis and identification of the tissue of origin for unknown metastases and helping scientists distinguish between benign and malignant tumors.
2. Real-time tracking of biomarkers that can assist in clinical decision making, predict the safety and efficacy of a therapy, monitor the patient's response to a therapy, allow early detection of any resistance to treatment, and aid in prognosis.
3. Biomarkers that can define the molecular characteristics of a cancer and enable monitoring of these characteristics as the cancer progresses or shrinks following treatment.

Cell-free DNA (cfDNA)—nucleic acids not bound inside cells but circulating in the plasma—was detected in the bloodstream of patients with cancer as early as 1948 (2). Clinical studies indicated that plasma cfDNA could be used as a biomarker for diagnostic screening, predicting responses to therapy, monitoring the size of the tumor, and diagnosing a relapse at an early stage (3–5).

<sup>1</sup>Institute of Infectious Diseases, Beijing Ditan Hospital, Capital Medical University, Beijing, China

<sup>2</sup>Genecast Precision Medicine Technology Institute, Beijing, China

<sup>3</sup>Department of Liver Surgery, Peking Union Medical College (PUMC) Hospital, Chinese Academy of Medical Sciences and PUMC, Beijing, China

<sup>‡</sup>Contributed equally to this work

\*Corresponding authors: zenghui@ccmu.edu.cn (H. Zeng), zhhbao@ccmu.edu.cn (H. Zhang)

### cfDNA in cancer diagnosis

Circulating cfDNA as a noninvasive biomarker has been used for early diagnosis and monitoring of tumor burden (6, 7). Several studies have shown that levels of cfDNA are higher in cancer patients than in healthy people (8–10). However, the use of circulating cfDNA in early cancer diagnosis has some limitations, since it is thought to originate from both tumor cell DNA and normal cell DNA. In early stage tumors, the low abundance of tumor-derived cfDNA may make detection difficult. Somatic mutations—nonheritable genetic alterations—found in cfDNA may provide information to help distinguish benign from malignant tumors. In practice, some variants identified in gene tests are difficult to classify clinically as pathogenic or benign, making it challenging to clinically distinguish benign from malignant tumors.

### Using cfDNA to identify the tumor origin

Recent studies have demonstrated that epigenetic information embedded in cfDNA can provide information useful for predicting a tumor's tissue of origin (11, 12). Guo and colleagues performed a systematic search of tissue-specific methylation regions across the entire human genome. Using a reference database of known methylation signatures, the authors identified tissue-specific methylation markers and demonstrated accurate determination of tissue of origin from cfDNA from patients with lung or colorectal cancer (11).

### Using cfDNA in selecting patients for targeted therapy

cfDNA has been used in clinical practice to stratify non-small cell lung cancer (NSCLC) patients for targeted therapy. The U.S. Food and Drug Administration (FDA) and the European Medicines Agency (EMA) have approved the use of epidermal growth factor receptor (EGFR) mutations detected in circulating cfDNA to select NSCLC patients with EGFR-actionable mutations for EGFR tyrosine kinase inhibitor (EGFR-TKI) therapy. However, this presents challenges for clinical decision-making, including the detection of discordant mutations in the tumor tissue compared to circulating tumor DNA (ctDNA). A large retrospective study of Chinese NSCLC patients showed that among 2,463 NSCLC patients with matched tumor tissue (T) and ctDNA (C) specimens, 1,017 patients carried the EGFR mutation in tumor tissues and/or ctDNA. Of these patients, 472 received EGFR-TKI treatment and were divided into three groups based on whether they carried the mutation in the tumor, in circulating cfDNA, or both: T<sup>+</sup>/C<sup>+</sup> (n=264),

T<sup>-</sup>/C<sup>+</sup> (n=28), and T<sup>+</sup>/C<sup>-</sup> (n=180). The median progression-free survival across the groups were similar. Further PCR and next-generation sequencing (NGS) validation from microdissected surgical specimens suggested that intratumor heterogeneity and relatively low sensitivity of the mutation detection assay contributed to discordant EGFR mutant status between tissues and ctDNA. In fact, neither tissue nor ctDNA analysis cover all EGFR mutations, suggesting that a combination of data from both tissue and ctDNA might be needed to accurately detect the presence of EGFR mutations. This study suggested that primary screening using ctDNA, followed by tumor tissue analysis in the case of a negative result, may be a preferable approach to determine whether EGFR-TKI therapy is warranted (13).

Previous studies have detected RAS mutations in ctDNA, using this information to stratify metastatic colorectal cancer (mCRC) patients for cetuximab or panitumumab treatment (4, 14). Recently, Xu and colleagues found a novel phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha (PIK3CA) mutation (p.K944N) in patients with resistance to cetuximab. These results indicated that patients harboring PIK3CA or RAS mutations in ctDNA showed shorter progression-free survival than those with wild-type sequences (15).

### Role of cfDNA in monitoring clonal evolution

The early detection of relapse following primary treatment for cancer and characterization of emerging molecular subsets might offer new therapies to limit tumor recurrence. ctDNA analysis is a potentially powerful method to noninvasively track tumor clonal evolution. Results from the TRACERx [TRACKing Cancer Evolution through therapy (Rx)] lung cancer study showed that phylogenetic cfDNA profiling closely tracks the subclonal nature of lung cancer relapses and metastases. It also found that the tumor volume was correlated with the variant allele frequency determined from plasma cfDNA. cfDNA profiling can allow for tracking of tumor evolutionary dynamics and detection of adjuvant chemotherapy resistance in NSCLC (5). Recently, Jiang *et al.* reported that the mutational landscape of cfDNA is associated with therapeutic response to first-line, platinum-based doublet chemotherapy in patients with advanced NSCLC (16).

The CRC patients with wild-type RAS could be stratified based on the tumor tissue genotype (17). However, RAS mutations often occur after EGFR blockade treatment, and it is challenging to obtain repeated tumor biopsy tissue to track tumor

evolution. Siravegna *et al.* observed RAS clonal evolution in circulating cfDNA of mCRC patients. Interestingly, they found that the abundance of mutated KRAS clones in plasma increased during EGFR blockade treatment, but decreased following withdrawal of EGFR-specific antibodies (4).

### cfDNA in immune checkpoint therapies

Recent clinical data support the use of new immune checkpoint inhibitors in several types of cancer. These drugs interfere with proteins that prevent the immune system from attacking cancer cells, and include such drugs as anti-PD-1 and anti-PD-L1 antibodies (18–22). A number of biomarkers have been developed to stratify patients likely to benefit from these immune checkpoint blockers. Their use has improved the survival of cancer patients and reduced treatment costs. Among these biomarkers, the PD-L1 protein is the most well known. Numerous clinical trials suggest that patients expressing PD-L1 are more likely to benefit from treatment with PD-1/PD-L1 inhibitors (23–25). However, these studies showed that approximately 10% of PD-L1-negative patients can also benefit from PD-1/PD-L1 inhibitors, suggesting that there might be other biomarkers that can better predict drug efficacy (23–25).

Several studies have reported other biomarkers that can also predict drug efficacy. A deficiency in DNA mismatch repair (dMMR) is a classical biomarker of colorectal and other cancers, and is associated with the efficacy of chemotherapy. High microsatellite instability (MSI-H) or dMMR leads to the accumulation of genetic mutations that may increase the risk of immune system identification and destruction of the tumor, suggesting that patients with dMMR progressive tumors can benefit from treatment with pembrolizumab or other PD-1 inhibitors (20, 26). Recent studies identified the mechanisms of acquired immune resistance after anti-PD-1 monoclonal antibody treatment and found two specific genetic variants in JAK1/JAK2 and beta-2-microglobulin to be involved (27, 28). A good response to immune checkpoint inhibitor treatment was positively associated with significantly increased tumor-infiltrating CD8 T cells or T-cell receptor (TCR) diversity (29, 30). In these studies, comparison of TCR clonality at baseline and postdosing biopsies showed that samples from responders had over 10 times as many clones expand after anti-PD-1 therapy; peripheral blood TCR diversity increased in responders after anti-CTLA-4 therapy. Peripheral blood T cells are relatively easy to obtain from

cancer patients, so TCR diversity testing might be a good predictor of response to immune checkpoint inhibitors. In NSCLC, the amount of tumor variant load or tumor mutational burden (TMB) can impact the efficacy of immune checkpoint inhibitors (such as anti-PD-1 antibody) (31–36).

Biomarkers such as those mentioned above, which can predict the efficacy and prognosis of immune checkpoint inhibitor treatments, are found in many tumor tissues. It is well known that the predictive efficacy of biomarkers from tumor tissue is limited because of the difficulty of obtaining samples from advanced-stage cancer patients and because of tumor heterogeneity (37). Clinical studies have shown that cfDNA can effectively reflect the tumor's size, malignant state, and ability to metastasize. It can also provide real-time information on causative mutations, providing more comprehensive gene mutation data than can be derived directly from tumor tissues (38). Researchers at the 2017 American Society for Clinical Oncology conference reported that TMB data obtained from sequencing of cfDNA could effectively predict those lung cancer patients who would benefit most from anti-PD-1 antibody treatment (39). In this study, ctDNA TMB was associated with a significantly higher number of DNA repair mutations. Additionally, smoking was associated with a higher TMB score. However, lower TMB detected in ctDNA predicted that a subset of patients would respond better to checkpoint inhibitors, in contrast to results from other studies (31–36). Potential reasons for this result include the small sample size, the possibility that ctDNA does not accurately reflect tumor burden, and the limited length of DNA sequenced (78,000 bp–138,000 bp). Larger prospective studies are necessary to validate these findings.

In the field of commercial gene testing, NGS-based MSI-H detection in cfDNA samples has been used to guide treatment choices, such as whether to use anti-PD-1/PD-L1 antibodies. In China, lung cancer, colorectal cancer, and cholangiocarcinoma are highly prevalent, providing a good opportunity to apply MSI-H and TMB detection assays. In addition, since the mechanism of tumor development in Chinese cancer patients has unique characteristics, research into tumor heterogeneity based on cfDNA analysis may identify biomarkers more suitable for these patients.

Recently, Ghoneim and colleagues reported that *de novo* DNA methylation plays a role in establishing the PD-1 blockade-nonresponsive

Region	Recruitment status	Count	Cancer type (number of studies)
Chinese mainland	Recruiting	20	Lung cancer (24) Ovarian cancer (3)
	Not yet recruiting	10	Gastric cancer (2) Hepatocellular carcinoma (2) Lymphoma (2)
	Active, not recruiting	4	Glioma (2) Breast cancer (1) Colon cancer (1)
	Completed	4	Biliary tract cancer (1) Prostate cancer (1)
	Status unknown	3	Bladder cancer (1) Nasopharyngeal cancer (1)
Hong Kong	Recruiting	1	Lung cancer (1)
Taiwan	Recruiting	7	Lung cancer (4) Hepatocellular carcinoma (4)
	Active, not recruiting	2	Lymphoma (2) Breast cancer (1)
	Enrolling by invitation	1	Head and neck cancer (1) Colon cancer (1)
	Completed	2	Glioma (1) Leukemia (1)
	Status unknown	5	Others (2)

**TABLE 1. Summary of clinical trials using ctDNA currently underway in China.** Data from ClinicalTrials.gov (<https://www.clinicaltrials.gov>).

reliably predict overall survival in response to immune checkpoint blockers. By contrast, ctDNA could be a promising means to monitor treatment efficacy. In one prospective proof-of-principle study, lack of detection of ctDNA at week 8 following treatment was a significant predictor of progression-free survival and overall survival in patients with lung cancer, uveal melanoma, and some forms of colorectal cancer following treatment with nivolumab or pembrolizumab (41).

Some Chinese companies, including Innovent Biologics (Suzhou) Co., Ltd., Jiangsu Hengrui Medicine

Co., Ltd., and Shanghai Junshi Biosciences Co., Ltd., are developing new PD-1/PD-L1 inhibitors that are currently in clinical trials. The authors believe that the biomarkers described above would be very helpful in these trials.

### Role of cfDNA in T-cell transfer immunotherapy

T-cell immunotherapy is a potentially life-saving treatment for patients with advanced stage tumors. As with immune checkpoint inhibitor treatments, there are still no good biomarkers that can identify responders or predict patient response. Since changes in ctDNA can be used to monitor the course of a disease, treatment responses, and recurrence (42–45), ctDNA levels might be useful

state of CD8 T-cell exhaustion, and that administration of DNA-demethylating agents (such as decitabine) prior to PD-1 blockade therapy may enhance the reinvigoration of antitumor CD8 T cells. This suggests that combining a PD-1 blockade drug with a DNA-demethylating agent may improve the therapeutic response (40).

In evaluating the efficacy of immunotherapy, it is worth noting that changes in cfDNA may occur earlier than changes in tumor size. Radiological evaluation of the efficacy of immune checkpoint inhibitors during therapy is therefore limited, and delayed clinical responses have been observed in some patients (18, 19). These results suggest that conventional radiological imaging may not

Institution	Research description
Peking University People's Hospital	Comparison of ctDNA and tumor tissue DNA by targeted sequencing in non-small cell lung cancer (NSCLC)
	Urinary ctDNA detection in NSCLC: a prospective study
	ctDNA detection in surveillance of surgical lung cancer patients
	Dynamic changes of ctDNA in surgical lung cancer patients
Sun Yat-sen University	ctDNA dynamic monitoring and its prognostic role in stage 1 NSCLC by NGS
	A prospective, observational trial of the diagnostic and prognostic uses of lung neoplasms and meningeal carcinomatosis
	ctDNA for the prediction of relapse in gastric cancer
	Detecting cell-free DNA in lung cancer patients
Peking University	Liquid biopsy in monitoring the therapeutic efficacy of targeted therapy in advanced/metastatic gastric cancer
	PD-1 knockout engineered T cells for castration-resistant prostate cancer
	PD-1 knockout engineered T cells for muscle-invasive bladder cancer
Peking Union Medical College Hospital	Clinical application of ctDNA in operable breast cancer patients
	Effectiveness of circulating DNA for predicting the relapse and overall survival in NHL patients
	Circulating cell-free DNA as a predictive biomarker for hepatocellular carcinoma
Beijing Cancer Hospital	Blood detection of EGFR mutation for Iressa (gefitinib) treatment
	Study of small doses of etoposide as maintenance treatment in small cell lung cancer (SCLC)

**TABLE 2. Leading research institutions doing clinical studies of circulating tumor DNA (ctDNA) on the Chinese mainland.**  
Data from ClinicalTrials.gov (<https://www.clinicaltrials.gov>). NHL, non-Hodgkin's lymphoma; EGFR, epidermal growth factor receptor; NGS, next-generation sequencing.

for monitoring efficacy of T-cell immunotherapy. A recent study showed that malignant melanoma patients who showed an early peak in ctDNA, but then cleared their serum of BRAF V600E ctDNA, were highly likely to show a complete treatment response over the next one to two years. This result may be instructive for other kinds of T-cell transfer immunotherapy approaches (46).

In China, dendritic cell/cytokine-induced killer-cell therapy is still the most common T-cell immunotherapy, used mainly for the treatment of blood malignancies (47, 48), but its efficacy for treatment of solid tumors has yet to be demonstrated. There is an urgent need for appropriate biomarkers to predict efficacy and to screen patients to find those potentially responsive

to T-cell immunotherapy. cfDNA- and ctDNA-related biomarkers might be good options to explore.

**Clinical trials of cfDNA in China**

Not surprisingly, clinical studies of ctDNA in precision cancer research are growing in number and importance both in China and worldwide. As shown in Table 1, there are currently 59 clinical trials involving ctDNA in China, including the Chinese mainland (41 studies), Hong Kong (1 study), and Taiwan (17 studies). Only 6 have been completed, while most are recruiting or preparing to recruit patients. Twenty-nine clinical trials involve lung cancer patients (particularly NSCLC), accounting for half of all studies underway, no doubt because

lung cancer causes more deaths worldwide than any other carcinoma. Certain cancers such as hepatocellular carcinoma and nasopharyngeal cancer get more attention in China than the United States or Europe, because of higher domestic incidence. Most sponsors of Chinese clinical trials involving ctDNA are domestic universities, hospitals, or companies, with the exception of the multinational pharmaceutical company AstraZeneca. Leading research institutions conducting clinical studies of ctDNA on the Chinese mainland are shown in Table 2.

### Future directions: cfDNA studies in the real world

The application of cfDNA in the management of cancer diagnosis and treatment in China faces some challenges. First, it is important that cfDNA-related clinical data be gathered from Chinese patients, as data from other populations might not translate well to the local population. Studies should cover the entire process, from diagnosis through treatment, including early diagnosis; identification of tumor tissue sources; choice of therapy, whether chemotherapy, radiotherapy, targeted therapy, or immunotherapy following surgery; and decisions about optimization of treatment options. Complete clinical and follow-up data should be collected within a researcher-centered, normative multicenter clinical study, following international guidelines for NGS. A guide to the use of NGS with cfDNA samples must also be established so that results from these studies are comparable, accurate, reproducible, and can be easily applied in the clinic. Finally, government and industry should invest substantially to carry out this research and help to establish a model for collecting large amounts of data, including clinical data on cfDNA use, early cancer screening, treatment options, guidance on personalized treatments, prognosis, monitoring of drug resistance, real-time treatment adjustment, and monitoring of small lesions that may develop during treatment.

### References

1. W. Chen *et al.*, *CA Cancer J. Clin.* **66**, 115-132 (2016).
2. P. Mandel *et al.*, *C.R. Seances Soc. Biol. Fil.* **142**, 241-243 (1948).
3. R. H. Xu *et al.*, *Nat. Mater.* **16**, 1155-1161 (2017).
4. G. Siravegna *et al.*, *Nat. Med.* **21**, 795-801 (2015).
5. C. Abbosh *et al.*, *Nature* **545**, 446-451 (2017).
6. E. Gormally *et al.*, *Cancer Res.* **66**, 6871-6876 (2006).
7. C. Bettegowda *et al.*, *Sci. Transl. Med.* **6**, 224ra24 (2014).
8. A. M. Newman *et al.*, *Nat. Med.* **20**, 548-554 (2014).
9. T. B. Hao *et al.*, *Br. J. Cancer.* **111**, 1482-1489 (2014).
10. K. L. Spindler *et al.*, *Int. J. Cancer* **135**, 2984-2991 (2014).
11. S. Guo *et al.*, *Nat. Genet.* **49**, 635-642 (2017).
12. K. Sun *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **112**, E5503-E5512 (2015).
13. R. Wan *et al.*, *J. Thorac. Oncol.* **12**, 1376-1387 (2017).
14. A. R. Thierry *et al.*, *Nat. Med.* **20**, 430-435 (2014).
15. J. M. Xu *et al.*, *Clin. Cancer Res.* **23**, 4602-4616 (2017).
16. T. Jiang, *et al.*, *Theranostics* **7**, 4753-4762 (2017).
17. A. B. Benson, 3rd *et al.*, *J. Natl. Compr. Canc. Netw.* **15**, 370-398 (2017).
18. J. Brahmer *et al.*, *N. Engl. J. Med.* **373**, 123-135 (2015).
19. J. Larkin *et al.*, *N. Engl. J. Med.* **373**, 23-34 (2015).
20. D. T. Le *et al.*, *N. Engl. J. Med.* **372**, 2509-2520 (2015).
21. R. J. Motzer *et al.*, *N. Engl. J. Med.* **373**, 1803-1813 (2015).
22. P. T. Nghiem *et al.*, *N. Engl. J. Med.* **374**, 2542-2552 (2016).
23. A. Rittmeyer *et al.*, *Lancet* **389**, 255-265 (2017).
24. H. Borghaei *et al.*, *N. Engl. J. Med.* **373**, 1627-1639 (2015).
25. M. Reck *et al.*, *N. Engl. J. Med.* **375**, 1823-1833 (2016).
26. D. T. Le *et al.*, *Science* **357**, 409-413, (2017).
27. C. B. Bifulco *et al.*, *N. Engl. J. Med.* **375**, 888-889 (2016).
28. J. M. Zaretsky *et al.*, *N. Engl. J. Med.* **375**, 819-829 (2016).
29. P. C. Tumeh *et al.*, *Nature* **515**, 568-571 (2014).
30. L. Robert *et al.*, *Clin. Cancer Res.* **20**, 2424-2432 (2014).
31. N. A. Rizvi *et al.*, *Science* **348**, 124-128 (2015).
32. T. N. Schumacher *et al.*, *Science* **348**, 69-74 (2015).
33. J. M. Kim *et al.*, *Ann. Oncol.* **27**, 1492-1504 (2016).
34. M. Liontos *et al.*, *Ann. Transl. Med.* **4**, 264 (2016).
35. P. Sharma *et al.*, *Science* **348**, 56-61 (2015).
36. M. Giannakis *et al.*, *Cell Rep.* **17**, 1206 (2016).
37. M. Kleppe *et al.*, *Nat. Med.* **20**, 342-344 (2014).
38. H. Schwarzenbach *et al.*, *Nat. Rev. Cancer* **11**, 426-437 (2011).
39. A. A. Davis *et al.*, *J. Clin. Oncol.* **35**, 584s (2017).
40. H. E. Ghoneim *et al.*, *Cell* **170**, 142-157 (2017).
41. L. Cabel *et al.*, *Ann. Oncol.* **28**, 1996-2001 (2017).
42. G. R. Oxnard *et al.*, *Clin. Cancer Res.* **20**, 1698-1705 (2014).
43. F. Diehl *et al.*, *Nat. Med.* **14**, 985-990 (2008).
44. T. Forshew *et al.*, *Sci. Transl. Med.* **4**, 136ra68 (2012).
45. S. J. Dawson *et al.*, *N. Engl. J. Med.* **368**, 1199-1209 (2013).
46. L. Xi *et al.*, *Clin. Cancer Res.* **22**, 5480-5486 (2016).
47. H. Bao *et al.*, *Curr. Stem Cell Res. Ther.* **12**, 188-196 (2017).
48. Y. Zhang *et al.*, *Expert Opin. Biol. Ther.* **15**, 481-493 (2015).

# Next-generation sequencing–based testing for cancer precision medicine in China: A review of technologies and validation procedures

Weifeng Wang, Weiwei Shi, Ming Yao, and Kai Wang\*

With the advent of next-generation sequencing (NGS), which is much faster than traditional sequencing methods, our understanding of cancer genomics has grown rapidly and is revolutionizing our ability to treat and manage cancers with precision medicine. NGS has proven to be a reliable and precise diagnostic technology in developed countries. But in China it still faces challenges in the form of quality control and assurance problems, bioinformatics analysis pipelines, clinical annotation, and the broad selection of competing NGS platforms. Here, we review NGS-based testing in China, looking at the challenges it faces, and its clinical utility in precision medicine for Chinese cancer patients now and in the future.

## Advantages of NGS-based testing technologies

NGS-based testing is an efficient and cost-effective technology that can inform cancer treatments by detecting genomic alterations in tumor tissues or liquid biopsies. The prevailing

belief in cancer biology is that genetic alterations spur cancer initiation, evolution, and progression. However, early attempts at understanding these alterations were hindered by the cost and inefficiency of available technologies, such as Sanger sequencing. NGS-based testing, in contrast, can detect all classes of genomic alterations in a single test, and at a lower per-base cost. Those alterations detected include base substitutions, long or short insertions and deletions (indels), gene copy number variations, and gene fusions/rearrangements. Such comprehensive profiling provides a detailed genetic picture of a cancer patient. For instance, researchers have been unraveling the various mutation forms of the epidermal growth factor receptor (*EGFR*) gene, known to play a role in a number of different cancer types. In the past 15 years, researchers have developed therapies that target *EGFR* mutations, including base substitutions (1, 2), short insertions (3), duplications (4), and fusions (5) (Figure 1). Detection of such complex mutation profiles in individual patients would have been impossible without NGS-based testing. The seminal discovery about 10 years ago of the correlation between *EGFR* mutations and the clinical response to drugs called *EGFR* tyrosine kinase inhibitors (TKIs) in non-small cell lung cancers (NSCLC) ignited the pursuit

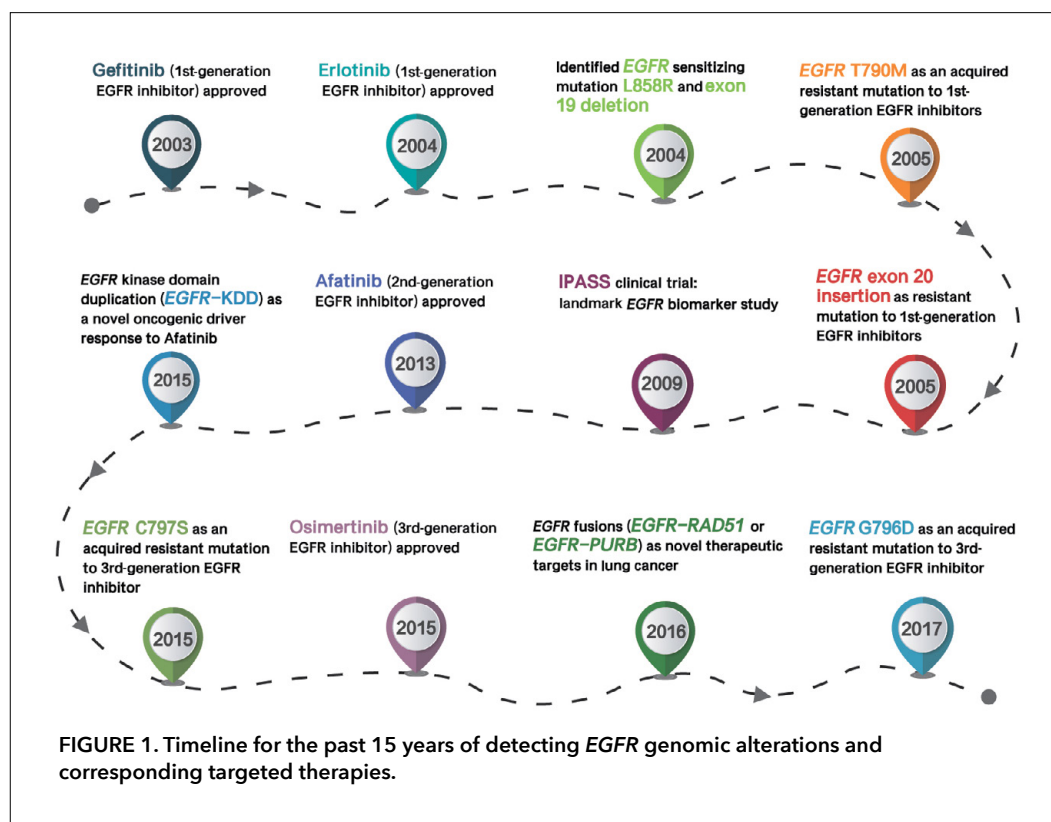


FIGURE 1. Timeline for the past 15 years of detecting *EGFR* genomic alterations and corresponding targeted therapies.

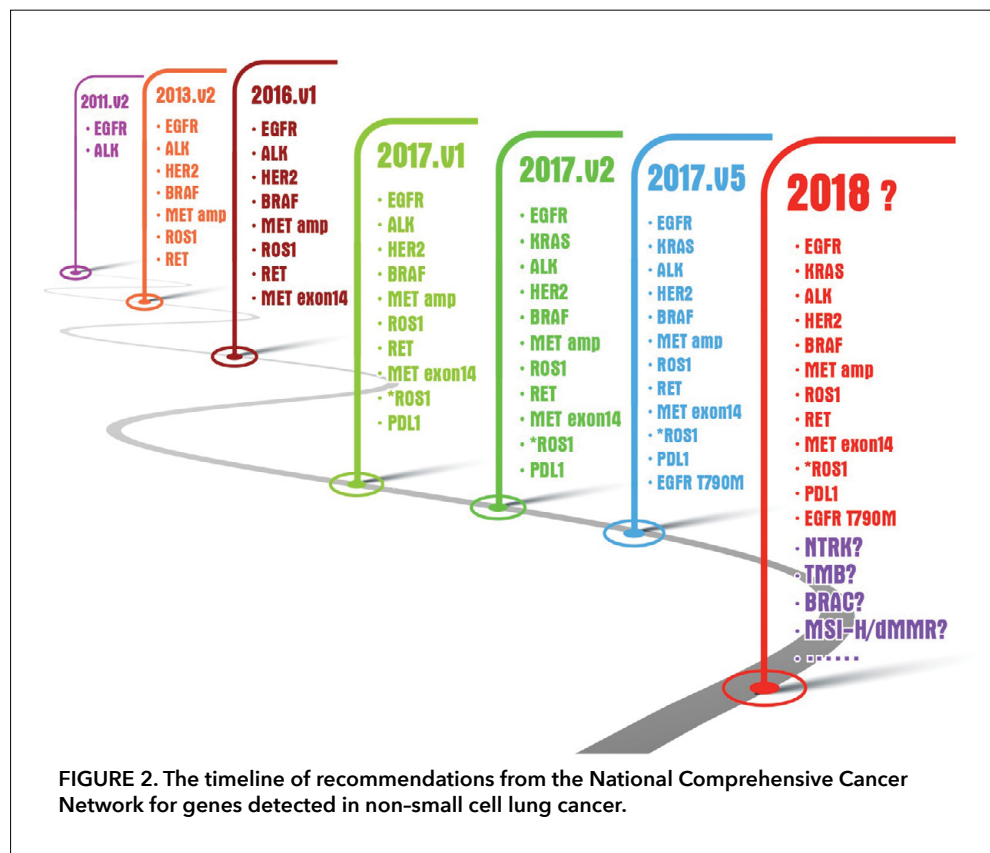


FIGURE 2. The timeline of recommendations from the National Comprehensive Cancer Network for genes detected in non-small cell lung cancer.

of precision medicine, which tailors individualized therapies according to the genetic and cellular characteristics of the patient. Thanks to its technical accuracy and clinical utility, NGS-based testing has become the primary method of informing precision cancer therapy.

In addition to its clinical utility, NGS-based testing advances the identification of those genomic changes that drive cancer, as well as the even larger number of passenger mutations that do not provide a growth advantage but simply “ride along,” significantly expanding our knowledge of the causes of cancer. With worldwide efforts from such projects as The Cancer Genome Atlas (TCGA) in 2005 and the International Cancer Genome Consortium (ICGC) in 2008, as well as vigorous bioinformatics data analyses, researchers are constructing a complete mutational atlas of all cancer types and their associated alterations. In the treatment of NSCLC, for example, analysis of the *EGFR* and anaplastic lymphoma kinase (*ALK*) genes has become part of routine testing to help doctors choose the appropriate targeted therapies in first-line treatment, substantially improving patient survival and quality of life. In addition to its proven value in targeted therapy, NGS-based diagnosis panels are finding a place in the rapidly

developing field of immunotherapy by identifying those patients who are most likely to benefit from the PD-1/PD-L1 inhibitor treatments. Tumor mutational burden (TMB)—the total number of mutations present in a tumor specimen—is becoming an important metric for predicting a patient’s response to immunotherapy (6). Tumors that harbor defects in the core DNA mismatch repair (MMR) protein complex are classified as being deficient in MMR and as microsatellite instability-high (dMMR/MSI-H) (7). In June 2017, the U.S. Food and Drug Administration (FDA) granted accelerated approval to pembrolizumab, a drug targeting adult and pediatric patients identified as MSI-H or dMMR. This was the first time the agency had approved a cancer treatment based on

a common biomarker rather than the tissue type in which the tumor originated. In light of these developments, comprehensive NGS-based gene profiling has the potential to act as a versatile tool for detecting not only gene alterations for targeted therapy, but also predictive markers associated with response rates to immunotherapy. This suggests that a well-designed combination of targeted therapy and immunotherapy based on genomic profiles could provide an impetus for new clinical trials to explore the relationship between the genomic data and the therapies.

With the anticipated approval of additional drugs, the U.S.-based National Comprehensive Cancer Network (NCCN) has been updating its list of recommendations for actionable gene testing in the treatment of NSCLC at an unprecedented pace (Figure 2). For example, a structural gene variant called “*MET* exon 14 skipping,” which is caused by a defective splicing of the gene’s RNA transcript, was added to the NCCN guideline on NSCLC shortly after the publication of papers that showed significant clinical benefits from *MET* inhibitors to patients with such mutations (8). In practice, NGS-based diagnosis can detect all of the NCCN-recommended targets, while having the flexibility to incorporate screening of

	ARMS PCR	Digital PCR	FISH	IHC	Sanger sequencing	NGS (target sequencing)
Sample usage	Inefficient	Inefficient	Inefficient	Inefficient	Inefficient	Efficient
Throughput	Low	Low	Low	Low	Low	High
Sensitivity of MAF	0.1%-1%	0.1%-0.01%	1%	1%-5%	10%	1% (depends on sequencing coverage)
Variation type	Known SNV and indel	Known SNV and indel	Known CNV and fusion	Protein expression	Known and unknown SNV, indel, fusion	Known and unknown SNV, CNV, indel, fusion
New variant discovery	No	No	No	No	Yes	Yes

**TABLE 1. Comparison of detection techniques in clinical application.**\* ARMS, amplification-refractory mutation system; FISH, fluorescence in situ hybridization; IHC, immunohistochemistry; SNV, single-nucleotide variant; indel, insertion or deletion; CNV, copy number variant; MAF, mutant allele frequency. \*Some special protocols not included. MAF sensitivity of FISH and IHC influenced by subjective interpretations.

many more mutations as they become known. Larotrectinib, the first small-molecule inhibitor specifically targeting oncogenic chimeric proteins expressed from fusion genes containing neurotrophic tyrosine receptor kinase (*NTRK*) plus other gene partners, has demonstrated consistent and durable antitumor activity in clinical trials of both adults and children with 17 different cancers. It is thus expected that the detection of *NTRK* fusions, TMB, or MSI-H using NGS sequencing panels will probably be approved as companion diagnostic tests in the near future (Figure 2). All of this will allow clinicians to identify more cancer patients with unique deleterious gene variants and help them determine therapeutic options quickly and cost-effectively.

NGS offers a single universal test that identifies predictive biomarkers for both targeted and immunotherapies. The high-throughput capacity of NGS-based testing is evident when it is compared with traditional molecular assays, such as fluorescence in situ hybridization (FISH) and polymerase chain reaction (PCR). FISH primarily detects amplifications and rearrangements of known genes, while PCR merely detects variations of a single nucleotide and known insertions or deletions (Table 1). So far, NGS is the best diagnostic technology for examining all types of genomic alterations in a single experiment, regardless of prior knowledge, and at much lower per-base cost than traditional techniques.

**Challenges facing NGS adoption**

Cancer is a major public health problem in China. According to a recent report, it is the leading cause

of death in the country, with new patients diagnosed at a rate as high as 12,000 cases each day (9). As a developing country with the world’s largest population of patients, China has a pressing need to provide precise and efficient therapies. Targeted therapy has been increasingly used in recent years in addition to traditional chemotherapy. Unfortunately, the genetic characteristics of tumors can vary widely, which means that not all tumors have the same therapeutic targets. Therefore, precise matching of treatments to each patient’s clinically actionable targets will be crucial to the success of precision cancer medicine. This is of particular importance in developing countries like China where the cost of therapies is a big concern. Many of the top hospitals and cancer centers in the United States have established NGS screening platforms for cancer patients. In China, clinical molecular diagnostic laboratories are also adopting NGS-based diagnostics. The technology is at the epicenter of a paradigm shift, and has attracted a tremendous amount of attention from doctors, scientists, and entrepreneurs, which, in turn, is catalyzing its refinement and growth.

NGS-panel-based diagnosis is still a nascent technology in China and poses many specific challenges, both in terms of techniques used and management. One issue is that not all Chinese diagnostics providers can ensure the quality of their products. When facing fierce market competition and loose regulations, many companies fail to meet basic quality control standards. In a nationwide external quality assessment carried out in 2015, only 27%

Number of genes examined	Advantages	Disadvantages	Price
1-3	Well-recognized and studied genes; traditional gene testing methods used for years	More samples needed; may miss important treatment options	\$
<10	Well-recognized genes in certain cancer types	May miss critical treatment options	\$
~40	Covers all approved drugs in China	May miss novel treatment options	\$
~100	Currently approved drugs; off-label usage	Not reliable on CNV or TMB	\$\$
~500	Covers all cancer-related genes to offer potential novel treatments, especially for relapsed patients	Relatively expensive	\$\$
Whole-exome sequencing	Covers all human genes. Offers accurate TMB determination and better CNV calling	Expensive with the risk of missing important treatment options due to missing actionable gene mutations caused by low coverage or gene fusions	\$\$\$\$

TABLE 2. Comparison of sequencing panels available in China that analyze different numbers of genes.\* CNV, copy number variant; TMB, tumor mutational burden.\*NGS on capture-based Illumina platforms only.

of participating laboratories were able to detect all known variants in standard samples provided by the National Center for Clinical Laboratories of the Chinese Ministry of Health. This prompted the Chinese Society of Clinical Oncology and the China Actionable Genome Consortium to release a consensus statement on the application of NGS technology to precision cancer medicine. Similarly, in 2017, the U.S. Association for Molecular Pathology and the College of American Pathologists published a joint recommendation on guidelines for validating NGS-based cancer sequencing screening panels (10). As expected, both sets of recommendations emphasized the importance of standard NGS operating procedures and reliable NGS equipment to guarantee high-quality genetic information from cancer samples. Stringent standard operating procedures and quality control checkpoints should be imposed at each step of the NGS-based testing pipeline. This includes sample preparation, DNA library preparation, sequencing, data analysis, annotation of results, and reporting of results. Additional attention should be paid to the criteria by which NGS panel tests are judged to be valid. These include types and number of samples, reproducibility and repeatability of variant detection, reportable range (region of the genome with acceptable quality results), reference range (the spectrum of nonpathogenic base changes observed in a population), limits of detection, interfering substances, clinical sensitivity and specificity, and if appropriate, validation of

bioinformatics pipelines and other parameters (11). As emphasized by OrigiMed, a Shanghai-based diagnostics service provider, NGS-based diagnosis in China should have meticulously designed, well-tested platforms run by trained personnel, and backed up by adequate infrastructure. Another recommendation was that annotations explaining genomic alterations and potentially actionable sites or signaling pathways should be clearly described for each gene.

The depth and breadth of sequencing coverage are two important technical factors affecting the quality and cost of NGS-based testing (Table 2). The numbers of target genes analyzed can differ substantially depending on the type of screening used, with the breadth of sequencing ranging from just a few variants to a few hundred using sequencing panels, to the entire coding gene region [whole-exome sequencing (WES)] or even to the entire genome [whole-genome sequencing (WGS)]. The selection of which genes to screen for in clinical applications more or less obeys the “Goldilocks principle,” namely that a small number of genes may generate an insufficient number of clinical clues, while a large number of genes could produce too much information. Specifically, the excessive genomic information generated by large sequencing ventures provides limited additional clinical value while posing the huge challenge of having to interpret incidental findings of variants of unknown or uncertain significance. Furthermore,

sequencing too wide a DNA region drives up expense due to the need to sequence to sufficient depth for the data to be valid, and can lead to poor sensitivity if depth of coverage is insufficient. More narrow sequencing coverage costs less, but significantly decreases detection sensitivity and may ultimately result in data that is difficult to interpret. If sequencing coverage is less than 300x, only 84% of low-frequency single-nucleotide variants (<5% mutant allele frequency) are detected, and the rate of detection for short variants is even poorer (12). Previous publications have shown that druggable mutations with low mutant allele frequency cannot be detected by low-coverage WGS/WES, or by sequencing panels, and that this may affect treatment options. On the other hand, most gene fusions, such as those found in the *ALK* gene, cannot be detected by WES alone unless the breakpoints are within or in close proximity to exons, and may require WGS.

A small NGS panel offers an affordable way to sequence fewer than 10 genes containing mutations common in certain cancer types, but risks missing several critical treatment options due to incomplete gene profiling (Table 2). The incorporation of more genes into the panel improves the quality of genomic mutation profiling and increases the number of treatment options, but it still fails to reflect the overall TMB, which is an important predictor of how effective immunotherapy will be. In order to maximize a patients' treatment options, an even larger panel consisting of several hundred relevant genes is recommended. However, this is not always available in China. The leading NGS panel-based product in the Chinese market, YuanSu from Origimed, sequences 450 well-defined cancer genes known to be important for targeted therapy, and analyzes TMB, microsatellite instability, and potential resistance to immunotherapy treatment (also incorporating RNA sequencing to detect more gene fusions and splice variants).

### **The present and future of NGS-based testing in China**

In 2015, the Chinese Ministry of Science and Technology held the first nationwide conference of experts to devise and promote a strategy for precision medicine. One of the primary accomplishments of the conference was a pledge from the government to invest 60 billion RMB (US\$9.24 billion) in the field of precision medicine by the year 2030. NGS-based technology is one of the key components of this investment.

China is preparing for the expanded use of NGS-based technology in precision medicine, especially in

the area of cancer. The American company Illumina retains the largest share of the market for sequencing hardware, with around three hundred NGS machines installed so far; other vendors provide the majority of reagents and consumables. But several domestic Chinese companies have entered the NGS market in recent years. More than 10 companies have developed their own NGS machines, some of which have been approved by the China Food and Drug Administration. These include the BGISEQ-500, developed by BGI, and the BioelectronSeq 4000, made by Capital-Bio Corporation. However, domestic companies still have a long way to go to reach the high bar set by their foreign competitors. In the sequencing services and clinical testing segment of the market, the lack of good training and supervision means that most of the approximately 160 sequencing service companies in China's eight major cities fail to meet reliability standards for their sequencing data analysis.

China is still behind the United States in the clinical use of NGS-based testing. Most hospitals in China remain heavily reliant on traditional molecular diagnostics like PCR, immunohistochemistry, and FISH as the major cancer diagnosis and prognosis tests. There are three main reasons why this situation exists. First, NGS-based testing involves a series of procedures—including sample preparation, DNA library construction, sequencing, bioinformatics analyses, and clinical annotation—that each require stringent standards of operations and quality control and assurance. As discussed above, China is not readily set up for these requirements in most domestic hospital laboratories. While the business of commercial assays has boomed in the past two years, only a few companies have the necessary analytical and clinical validation for their NGS-based assays. And the importance of validation is not yet appreciated by every doctor in China.

Second, numerous targeted cancer therapies have been approved by the U.S. FDA, and many more are being studied in clinical trials, but experience in the appropriate use of the new drugs is relatively limited in China. Doctors and patients tend to opt for traditional diagnostics like single gene-based tests, instead of comprehensive diagnostics like NGS-based assays. Finally, the cost of NGS-based testing is still too high for most Chinese patients. Generally, the price of a comprehensive NGS-based assay ranges from US\$1,500 to US\$3,000, depending on the size of the panel assay and the quality.

Although NGS-based cancer testing is still in its infancy, it is soon expected to become the major component of worldwide clinical NGS diagnostics,

including in China. It is expected that academia and industry in China will continue their efforts at building sequencing hardware and designing reagents. More importantly, China will be expanding the academic and clinical application of NGS-based assays with a focus on exploring cancer etiology and pathogenesis, and providing a genetic basis for personalized therapy with the help of cancer genome sequencing, as well as integrating various NGS-based technologies to aid ongoing research in conventional clinical medicine. The hope is that these steps will lead to significant improvements in precision medicine, including disease diagnosis, treatment, and clinical decision-making for Chinese cancer patients.

### Looking to the future

NGS-based testing is one of the most exciting developments in oncology research and clinical application in recent years. It provides a more comprehensive and efficient approach for characterizing the genomes of individual cancer patients and matching them with treatments for optimal clinical management. Most hospitals in China are not fully ready to run NGS platforms in-house. But with the concerted efforts of the government, researchers, and companies, precision cancer medicine in China is developing rapidly. More detection technologies are emerging for clinical applications, including ultrasensitive circulating tumor DNA detection and exosome and epigenome profiling. We envision a future where each cancer patient will be tested by a comprehensive gene panel at several points in time and wherever they are receiving treatment, with a range of sequencing technologies that will detect multilevel molecular abnormalities and help clinicians design a personalized, precise therapy tailored to their individual needs.

### References

1. J. G. Paez *et al.*, *Science* **304**, 1497–1500 (2004).
2. T. J. Lynch *et al.*, *N. Engl. J. Med.* **350**, 2129–2139 (2004).
3. H. Greulich *et al.*, *PLOS Med.* **2**, e313 (2005).
4. J. N. Gallant *et al.*, *Cancer Discov.* **5**, 1155–1163 (2015).
5. K. Konduri *et al.*, *Cancer Discov.* **6**, 601–611 (2016).
6. D. P. Carbone *et al.*, *N. Engl. J. Med.* **376**, 2415–2426 (2017).
7. D. T. Le *et al.*, *Science* **357**, 409–413 (2017).
8. G. M. Frampton *et al.*, *Cancer Discov.* **5**, 850–859 (2015).
9. W. Chen *et al.*, *CA Cancer J. Clin.* **66**, 115–132 (2016).
10. L. J. Jennings *et al.*, *J. Mol. Diagn.* **19**, 341–365 (2017).
11. H. L. Rehman, *Nat. Rev. Genet.* **14**, 295–300 (2013).
12. G. M. Frampton *et al.*, *Nat. Biotechnol.* **31**, 1023–1031 (2013).

## ctDNA-NGS: The key to unlocking a molecular diagnostic revolution in the heart of Asia

Ying Hou and Kang Ying

A wealth of information about a cancer can be obtained using liquid biopsy, a noninvasive blood test. The diagnostic potential of liquid biopsy is best illustrated by the development of the companion diagnostic for osimertinib (a third-generation targeted therapy for patients with non-small cell lung cancer). Osimertinib was approved by the Chinese Food and Drug Administration in just seven months, a landmark achievement in lung cancer therapeutics. A major contributor to its success was the codevelopment of a companion assay for detecting the *EGFR* T790M mutation (1–3), which commonly emerges in lung tumors after treatment with first-generation therapies. When tissue samples were unavailable, mutation detection could be done using liquid biopsy technology, which required only a small sample of patient blood. Liquid biopsy testing has been so effective in tracking *EGFR* T790M that the practice is now strongly recommended by the European Society for Medical Oncology (4) and in the National Comprehensive Cancer Network (5) guidelines. In addition, the Chinese Lung Cancer Summit Committee formally endorsed the use of liquid biopsy for the detection of *EGFR* T790M in 2016 (6). Although the emergence of a tissue-based molecular diagnostic market in Asia has been driven largely by the high occurrence of epidermal growth factor receptor (*EGFR*) mutations in the region, *EGFR* mutation detection is just the beginning. With a concerted effort, liquid biopsy technology has the potential to redefine the field of clinical molecular diagnostics not just in Asia, but worldwide.

### Challenges for Chinese ctDNA-NGS clinical diagnostics

The tremendous potential of liquid biopsy is offset by the technological challenges it presents. Any practical liquid biopsy diagnostic assay based on circulating tumor DNA (ctDNA) must

Metric \ Technology	Barcode-based error-correction sequencing		Concatemer-based error-correction sequencing	
	AVENIO ctDNA Expanded Kit	Oncomine cfDNA Lung Cancer Panel	Nebula: Accu-Act	Comet: Accu-Kit
Example				
Panel coverage	198kb	~2kb	100kb	~2kb
SNV/indel sensitivity	~10 copies <sup>a</sup>	<10 copies <sup>b</sup>	<10 copies <sup>c</sup>	
SNV/indel error rate	<0.6% <sup>a</sup>	0.2-2% <sup>b</sup>	0.03% <sup>c</sup>	
Workflow ease-of-use	++	+++	+	+++
Design flexibility	+++	+	+++	++

**TABLE 1. Comparison of liquid-biopsy ctDNA sequencing technologies.** SNV, single nucleotide variant; indel, insertion or deletion; kb, kilobases. <sup>a</sup>30-ng input, 0.1% allele frequency (AF) (27). <sup>b</sup>20-ng input, 0.1% AF (28). <sup>c</sup>AccuraGen analytical data for Firefly technologies. No false positive in 50 noncancerous individual plasma samples and 24 wild-type cfDNA standards. Average of 96% detection rate for 20-ng input at AF of 0.1%.

have an extremely high and readily reproducible performance sensitivity and specificity (7, 8). CtDNA—fragmented DNA from a tumor that is free-floating in the peripheral blood—is present in minuscule quantities (on the order of a few nanograms per milliliter of plasma). As a result, isolating ctDNA signals amid the noise generated by noncancer DNA fragments and sequencing errors is difficult. Droplet digital polymerase chain reaction (ddPCR) and next-generation DNA sequencing (NGS) are among the only quantitative technologies able to meet the sensitivity demands of low-frequency mutation detection in ctDNA (9, 10). Between the two, NGS offers several advantages. It has greater multiplexing capacity (where multiple genes can be analyzed in parallel from the same input) and is better at detecting de novo mutations. The performance sensitivity and specificity of NGS technologies, however, varies significantly (11–13).

CtDNA testing is also not without regulatory constraints. Unlike in the United States, medical facilities in China perform most clinical testing in-house. While this is an effective cost-saving practice for established testing protocols, the ctDNA-NGS diagnostic assay is difficult to deploy on-site without highly specialized technicians. As a result, the use of ctDNA-NGS testing is carefully monitored by the Chinese National Health and Family Planning Commission (NHFPC) (14). To date, 26 facilities have received government authorization to offer NGS testing for the purposes of oncological treatment. While the availability of these services is a promising first step, a formal assay performance evaluation conducted in 2015 by NHFPC revealed an alarming variability in results. Of the 64 sites (including the 26

facilities authorized to perform ctDNA-NGS assays), 33 passed the evaluation and among these, only 17 received perfect marks (15). The lack of technical development and operational protocols at these pilot sites has resulted in the use of off-site, third-party testing services, reducing adoption. Thus, the rapid adoption of ctDNA testing in China is contingent upon the standardization of testing protocols.

**One solution: hospital-owned ctDNA-NGS clinical labs**

One way to promote the use of ctDNA-NGS is by providing widely accessible testing guidelines or kits that would enable hospitals to confidently perform in-house NGS testing. This “kit-able” solution would include such essential features as a “low-touch workflow” requiring little or no NGS expertise and a self-contained data analysis pipeline. Such a kit could then be easily translated into an automated technology for informing personalized therapeutic treatment at scale.

Although a few kit options are available for tumor tissue sequencing, kit options for NGS-based ctDNA sequencing are far scarcer. Historically, performance comparisons of NGS-based ctDNA kits have highlighted their relative trade-offs. Capture-based assays have demonstrated high specificity and multiplexing capacity but limited sensitivity and a complex workflow. A single AVENIO capture-based assay from Roche (11, 16), for example, can cover >100 kilobases and detect single-nucleotide variants, insertion-deletions, and fusion alterations at a rate of 96%–99%, at allele frequencies of 0.5%–1%. In contrast, amplicon-based assays (17) offer high-performance sensitivity and ease of use (the Thermo Fisher

Oncomine kit has a turnaround time of just two days), but provide limited coverage and few opportunities for assay customization. Amplicon-based assays also tend to produce more false positives. We believe that AccuraGen's technology, Firefly, brings together the best of amplicon- and capture-based assays (18). Firefly can detect a breadth of genetic alterations with a sensitivity approaching the physical limit and a specificity 10-fold better than benchmark, all while maintaining the added benefits of flexible panel design and workflow integration. It has a consistent error rate of 1 in 1 million at the molecular level and can uncover 0.1% of variants in 20-ng samples of cfDNA at a detection rate greater than 90%.

The backbone of the Firefly technology is its use of rolling-circle amplification to empower consensus-based concatemer error correction. Single-stranded ligation without the usage of any barcode results in a high conversion rate, and rolling-circle template amplification minimizes the proliferation of enzymatic errors common in PCR. Finally, tandem repeats in concatemers make mutation detection both accurate and economical.

### Emerging challenges facing NGS-based ctDNA mutation assays

Although the clinical utility of ctDNA sequencing is currently oriented toward targeted therapy selection, mounting evidence suggests it may also be used to track the progression of cancer (19). Two studies monitoring postsurgical residual disease in breast (20) and colorectal (21) cancer patients demonstrated the predictive value of ctDNA in determining the likelihood of cancer relapse was far greater than estimates made using traditional clinical imaging evidence. While these studies foreshadow exciting future applications of ctDNA liquid biopsy assays, their limited genetic scope and small patient sample sizes impede their clinical actionability. Fortunately, new tools with increased panel sizes and product accessibility have since emerged. Recently, UK-based TRACERx [TRACKing Cancer Evolution through therapy (Rx)] reported the superiority of a 30-plex panel over a single plex or lowplex panel (up to 4 amplicons) in detecting cancer relapses and tumor progression (22, 23). Empowered by Firefly technology, AccuraGen currently offers Accu-Act, a 61-gene capture-based panel for residual disease monitoring and strategic cancer management. AccuraGen's amplicon-based ctDNA kit, Accu-Kit, which detects only genes included in NCCN guidelines, is easy to use, customizable, and affordable.

Increasing panel size, however, creates new challenges for developing a viable assay. The underlying

genetics of tumor cells vary tremendously both within a single tumor and from one tumor to another. How will intra- and intertumor heterogeneity be addressed? How can assay results be both highly sensitive and easily reproduced so they can be clinically actionable? How will diversified subclone mutation profiles be integrated to generate a holistic overview of the disease? Although these problems are complex, they are not insurmountable.

A combination of ctDNA-NGS kits and laboratory-developed tests will likely address the majority of known clinical needs for ctDNA. Though China's NHFPC has yet to publish official guidelines, several clinical committee consortiums have released their own recommendations to guide independent clinical testing (24–26). AccuraGen is uniquely positioned to empower NHFPC-certified facilities to act on these recommendations by equipping them with clinical kits for on-site testing. By making liquid biopsy technology more accessible and lowering barriers to workflow integration, we believe that AccuraGen can help bring cutting-edge research to the bedside.

Although the most effective way to encourage widespread liquid biopsy adoption would be CFDA approval of a ctDNA-NGS in vitro diagnostic kit, the likelihood of this occurring in the near future is low. In the short-term, ctDNA-NGS assays will need to be performed in-house at NHFPC-approved clinics. Those facilities will need the expertise to generate high-quality mutation profiles; they already regularly perform a wide array of diagnostic tests. At AccuraGen, we believe that with superior technology, ctDNA-NGS assays can be as easy to use as any other tools in the clinic's toolbox.

Regulatory processes, while slow, are necessary safeguards for delivering quality services. But regulation should not stonewall innovation. New products envisioned by teams of scientific experts and seasoned physicians will still be developed. AccuraGen is committed to advancing the frontiers of molecular diagnostic testing using ctDNA-based NGS technology, and we believe that China is in the vanguard of a revolution in this area.

### References

1. European Medicines Agency, "Summary of Product Characteristics," TAGRISSO (osimertinib) (2017); available at [http://www.ema.europa.eu/docs/en\\_GB/document\\_library/EPAR\\_-\\_Product\\_Information/human/004124/WC500202022.pdf](http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Product_Information/human/004124/WC500202022.pdf).
2. S. Jenkins *et al.*, *J. Thorac. Oncol.* **11**, S153–S154 (2016).
3. T. S. Mok *et al.*, *New Engl. J. Med.* **376**, 629–640 (2017).
4. S. Novello *et al.*, *Ann. Oncol.* **27**, v1–v27 (2016).

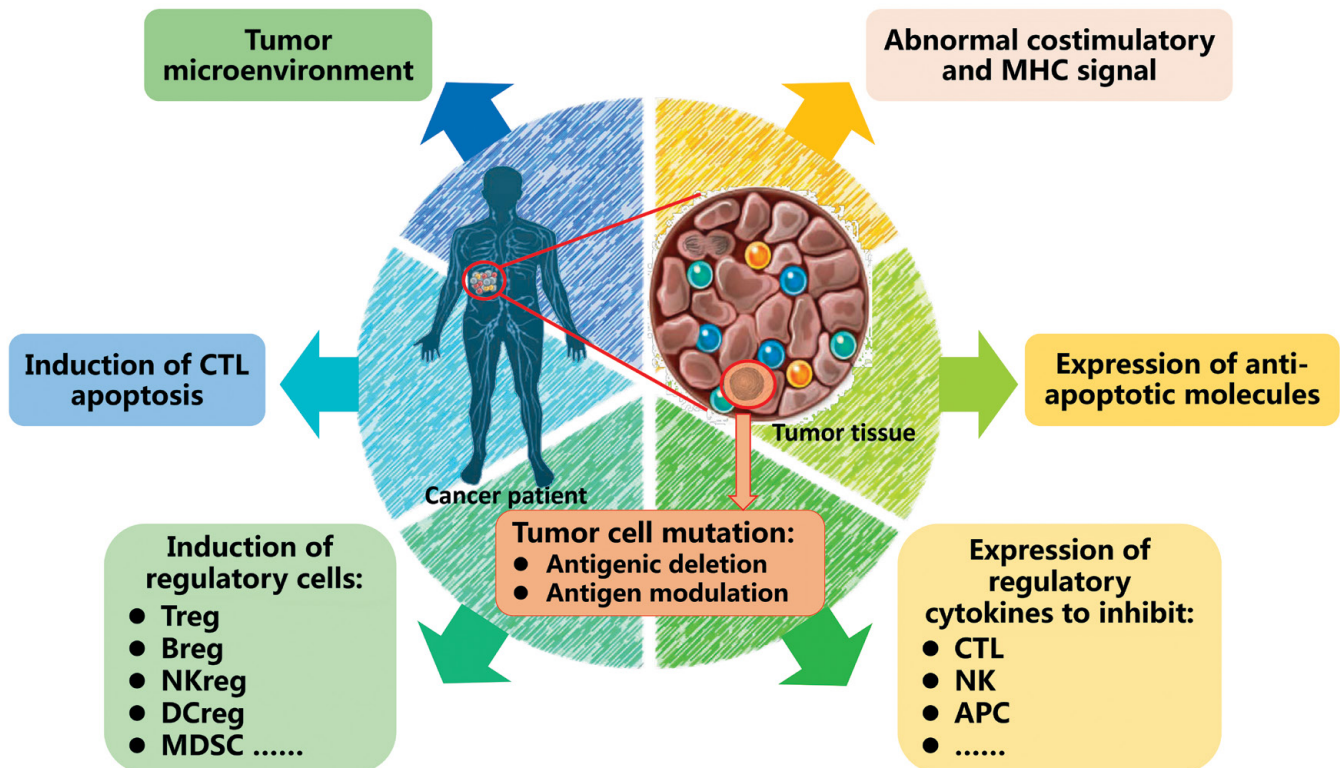
5. D. S. Ettinger *et al.*, *J. Natl. Compr. Canc. Netw.* **15**, 504–535 (2017).
6. Y. L. Wu *et al.*, *ESMO Open* **2**, e000174 (2017).
7. R. B. Lanman *et al.*, *PLOS ONE* **10**, e0140712 (2015).
8. L. A. Diaz, Jr., A. Bardelli, *J. Clin. Oncol.* **32**, 579–586 (2014).
9. T. Xu *et al.*, *Theranostics* **7**, 1437–1446 (2017).
10. K. S. Thress *et al.*, *Lung Cancer* **90**, 509–515 (2015).
11. A. M. Newman *et al.*, *Nat. Med.* **20**, 548–554 (2014).
12. I. Kinde, J. Wu, N. Papadopoulos, K. W. Kinzler, B. Vogelstein, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 9530–9535 (2011).
13. D. Gale *et al.*, “Analytical Performance and Validation of an Enhanced TAM-Seq™ Circulating Tumor DNA Sequencing Assay,” poster, inivata.com (2016); available at <http://inivata.com/wp-content/uploads/2016/06/AACRposter2016.pdf>.
14. R. Zhang, J. Li, “China’s Policies Regarding Next-Generation Sequencing Diagnostic Tests,” in *Precision Medicine in China* (Science/AAAS, Washington, DC, 2016), pp. 9–11.
15. R. Zhang *et al.*, *Oncotarget* **7**, 58500–58515 (2016).
16. A. M. Newman *et al.*, *Nat. Biotechnol.* **34**, 547–555 (2016).
17. J. Gu *et al.*, *Cancer Res.* **76** (Suppl.), Abstract 3622 (2016), doi: 10.1158/1538-7445.AM2016-3622.
18. L. Weng *et al.*, *J. Clin. Oncol.* **35** (Suppl.), Abstract e23062 (2017), doi: 10.1200/JCO.2017.35.15\_suppl.e23062.
19. J. C. Wan *et al.*, *Nat. Rev. Cancer* **17**, 223–238 (2017).
20. I. Garcia-Murillas *et al.*, *Sci. Trans. Med.* **7**, 302ra133 (2015).
21. J. Tie *et al.*, *Sci. Trans. Med.* **8**, 346ra392 (2016).
22. C. Abbosh *et al.*, *Nature* **545**, 446–451 (2017).
23. M. Jamal-Hanjani *et al.*, *New Engl. J. Med.* **376**, 2109–2121 (2017).
24. X. Yang *et al.*, *J. Mol. Diagn.* **19**, 525–536 (2017).
25. R. Zhang, J. Li, *Chin. J. Lab. Med.* **40**, 224–226 (2017), doi: 10.3760/cma.j.issn.1009-9158.2017.04.002 (in Chinese).
26. Chinese Medical Doctor Association Test Physician Branch of Molecular Diagnosis Expert Committee, *Chin. J. Lab. Med.* **39**, 897–900 (2016), doi: 10.3760/cma.j.issn.1009-9158.2016.12.007 (in Chinese).
27. Roche, “In-Depth Assessment Reveals Powerful Performance and Flexibility of the AVENIO CtDNA analysis kits”; available at [http://sequencing.roche.com/content/dam/rochesequence/worldwide/resources/SEQ100108\\_AVENIO%20ctDNA\\_Performance\\_White\\_Paper.pdf](http://sequencing.roche.com/content/dam/rochesequence/worldwide/resources/SEQ100108_AVENIO%20ctDNA_Performance_White_Paper.pdf).
28. Thermo Fisher Scientific, “Analytical Verification Methods for the Oncomine Lung cfDNA Assay Using the Ion S5 XL System,” white paper; available at <https://www.thermofisher.com/content/dam/LifeTech/Documents/PDFs/cfDNA-White-Paper-Lung-cfDNA-Assay-and-Ion-S5-XL.pdf>.

## Adoptive cell transfer therapy: A strategic rethinking of combination cancer therapy

Minghui Zhang

In recent decades, significant progress has been made in immunotherapy treatment, which is regarded by some as the latest battlefield in a long war against cancer. In 1890, the father of immunotherapy, American bone surgeon William Coley, used a mixture of two dead infectious bacteria strains to treat patients with inoperable cancers and achieved a complete remission rate of 10%. American biologist Paul Ehrlich speculated about the possible existence of immune tumor surveillance—the process by which the body’s immune system is constantly seeking out and destroying tumor cells. As more tumor antigens are identified and characterized, we gain a deeper understanding of how the immune system fights tumors. Recently, several “immune checkpoint blockers,” which promote the ability of certain cells to attack cancer, have had clinical success. Adoptive chimeric antigen receptor (CAR) T-cell transfer, a therapy in which genetically modified T cells are transferred into a patient [and a type of adoptive cell transfer (ACT)], has been approved by the U.S. Food and Drug Administration. But the efficacy of these treatments needs improvement, and researchers are still trying to mitigate their adverse side effects.

Scientists are also looking for ways to develop effective therapies that combine immune treatment with conventional cancer therapies to increase therapeutic benefit, minimize harm, and provide a cure or long-term remission. Tumors use multiple means to evade detection by the immune system. A better understanding of the mechanisms of tumor development and the reasons for success or failure of current therapeutic strategies is needed for the development of more effective treatments. ACT, which in essence provides more effective immune cells to fight against tumor cells, is playing a key role in cancer treatment today. However, there are sever-



**FIGURE 1.** Factors that allow for tumor development. MHC, major histocompatibility complex; CTL, cytotoxic T-lymphocyte; Treg, regulatory T cell; Breg, regulatory B cell; NKreg, regulatory natural killer (NK) cell; DCreg, regulatory dendritic cell; MDSC, myeloid-derived suppressor cell; APC, antigen-presenting cell.

al different forms of ACT. Here, we will discuss ways that ACT might be exploited as a broad-spectrum therapy and be combined with traditional treatments to increase its efficacy.

### Tumor development

It is well accepted that cancer is a disease that develops and progresses because of the accumulation of genetic and epigenetic changes that affect the way genes are expressed. Genetic drivers of tumor formation and growth are confirmed by gene sequencing profiles and the study of proteins expressed in and on the tumor cells. However, deeper understanding of the effects of the immune response on tumors is leading to the realization that the immune system is the most important extrinsic influence on the fate of mutant cells: If immunosurveillance works correctly, mutant cells will be recognized and eradicated (Figure 1), explaining why certain people don't get cancer. If the immunosurveillance system cannot recognize or eradicate the mutant

cells, these cells will have the opportunity to develop into a tumor. This principle is also expressed as the "three Es" of cancer immunoediting: elimination (the immune system finds and destroys cancer cells), equilibrium (cancer cells that are not destroyed right away may exist in a delicate balance between growth and control by the immune system), and escape (the tumor cells escape immune system surveillance and begin growing in an environment in which the immune response is suppressed).

Substantial genetic heterogeneity has been detected both in different samples from the same kind of primary tumors and metastases, and in individual cancers from the same patient. Precision cancer medicine in its current form is heavily dependent on genetic testing, and promotes the molecular characterization of tumors in order to develop more effective and beneficial drugs for patients carrying certain mutations. However, the benefits often don't persist, as the tumor cells can develop resistance to drugs by

the upregulation of a partially inhibited pathway, a mutation of the target gene, or an activation of alternative pathways, for example. Many patients will experience drug resistance, cancer progression, and ensuing metastasis.

The immune system can recognize mutant cells and eliminate many early malignancies. However, tumors have a variety of means to evade immune attack. They can decrease the expression of major histocompatibility complex class I molecules to escape natural killer (NK)-cell cytotoxicity, and increase the expression of coinhibitory molecules to inhibit T-cell activation and function. A tumor may also establish an immunosuppressive state in the tumor microenvironment by producing regulatory cytokines, including indoleamine 2,3-dioxygenase, vascular endothelial growth factor, and TGF- $\beta$ , and by recruiting myeloid-derived suppressor cells and regulatory T cells. Overcoming these obstacles and strengthening the immune response are the keys to successful immunotherapy.

### **The current state of ACT**

Immunotherapy can be classified into two broad categories. The first is nonspecific immune stimulation, and includes cytokine immunotherapy (which acts to stimulate the immune system to fight the cancer), oncolytic virus immunotherapy (which selectively kills tumor cells while also stimulating antitumor immunity), Toll-like receptor agonist therapy (which activates both innate and adaptive immune responses), and some forms of ACT (such as the transfer of natural killer cells into the patient), among others. The second category is specific immunotherapy, including dendritic cell (DC) vaccines, other kinds of ACT (for example, the transfer of T cells), and immune checkpoint inhibitors.

Among these approaches, ACT may be the most complicated. As eluded to above, it encompasses both nonspecific and specific immune stimulation. The cells are produced in one of two ways. In the first strategy, the doctor may isolate specific immune cells from a patient's tumor (such as tumor-infiltrating lymphocytes) or from surrounding blood, culture them in the lab to expand their numbers, and then administer them back into the patient. In the second strategy, T cells or NK cells collected from the patient are engineered with new receptors to recognize specific antigens on the surface of cancer cells, generating so-called CAR-T cells or CAR-NK cells. The recombinant cells are then infused back into the patient fol-

lowing expansion in culture. ACT therefore uses immune cells to either promote the immune response or to destroy the cancer cells that carry specific antigens.

CAR-T is currently the most promising strategy. It has demonstrated clinical benefits in some leukemias (1-4). However, because few specific tumor antigens have been identified, the effect of CAR-T on solid tumors is still marginal. In addition, the side effects of CAR-T can be dangerous or even life threatening (4-6).

Less attention is being paid to nonspecific ACT therapies (NK or DC therapy). This may be because the effect of these therapies is hard to assess using the present evaluation system. The transferred cells often trigger inflammation inside the tumor, leading to an apparently larger tumor in the short term. Another key factor is the timing of ACT. Some advanced tumor patients have tried various other treatments before they choose cellular therapy. It is therefore difficult to isolate the specific effects of the ACT treatment and evaluate what its real clinical benefits might be if used at the optimal time. Furthermore, ACT therapy must be adjusted depending on the patient's condition, the type of cell transferred, the number of effective cells, and the frequency of treatment, making it complicated to administer.

### **The future of ACT**

The complexities of tumors, including their heterogeneity and their ability to suppress an immune response, are at least partly responsible for therapeutic failures. There has been great progress made with traditional cancer treatments, including surgery, chemotherapy, radiation therapy, and targeted therapy. However, these modalities often cannot completely eradicate tumor cells. And sometimes micrometastases are occurring even as the tumor is diagnosed. Scientists face the question of how to build a new strategy that preserves the benefits of traditional treatments but overcomes their limitations and provides long-term remission or even a cure.

Combining multiple treatments offers a potentially powerful approach. The innate and adaptive parts of the immune system work together to sustain a state of equilibrium in the body. If we can understand more deeply the antitumor mechanisms of the immune system, design ACT therapies accordingly, and use them in combination with traditional cancer treatments, tumor cells could be eradicated and active immunosurveillance reinstated. Thus, in addition

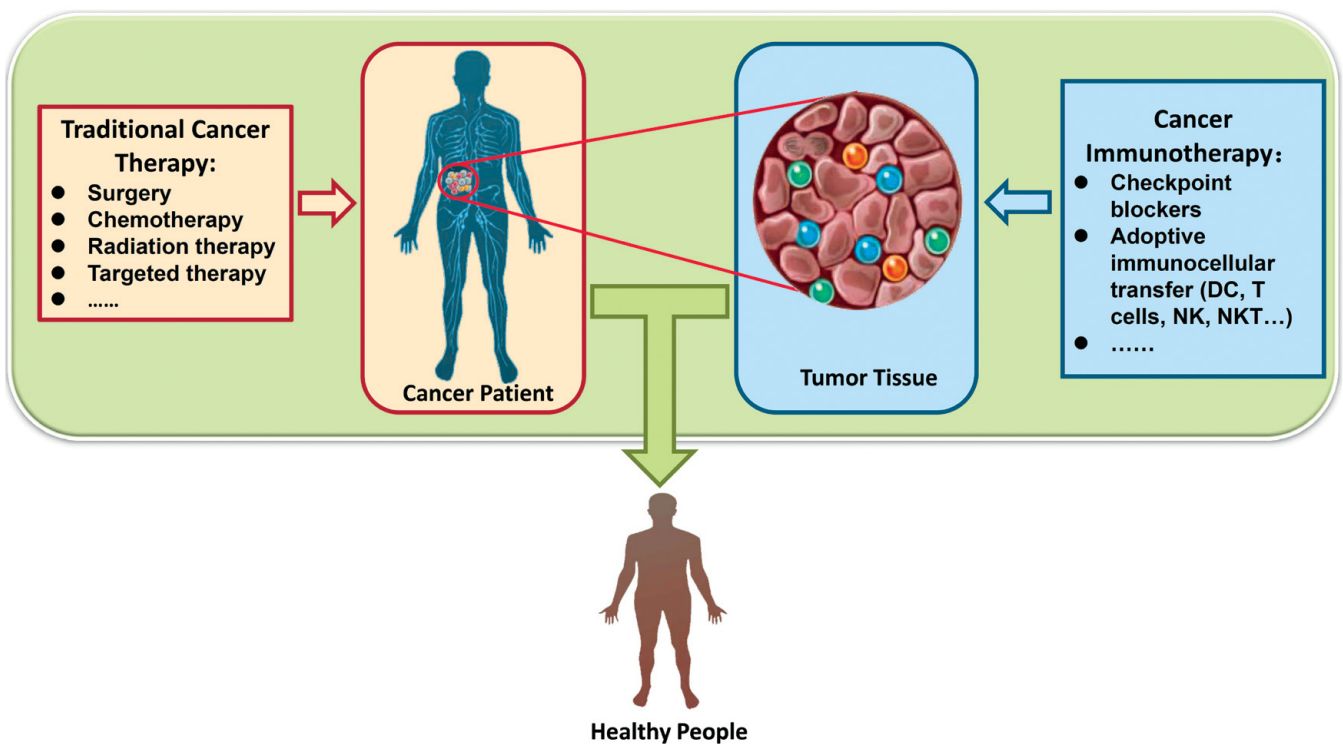


FIGURE 2. Combinatorial cancer treatment. DC, dendritic cell; NK, natural killer cell; NKT, natural killer T cell.

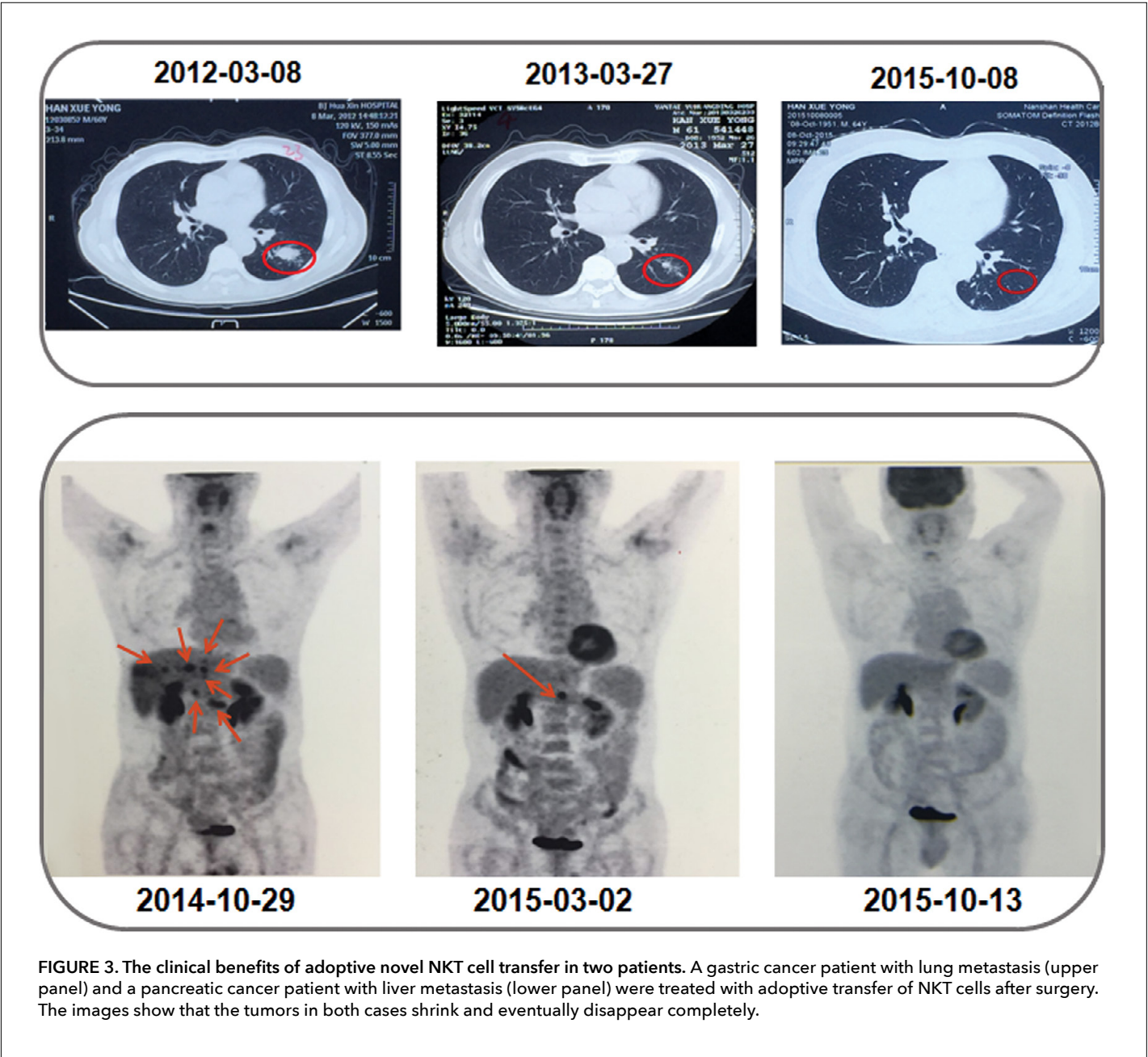
to killing tumor cells, one of the most important effects of ACT is to reverse a tumor's ability to suppress the immune response. A strengthened immune system can work as a policeman so that tumor cells that are not killed by traditional cancer treatment can be either eradicated or controlled by the immune system. We predict that this type of combination therapy will lead to the elimination of many cancers, or at least to longer survival with no worsening of the symptoms (Figure 2).

In the battle against tumors, surgery can be considered the vanguard, decreasing the tumor burden and eradicating most tumor cells. Chemotherapy, radiation therapy, and targeted therapy are the troops sent in to attack inoperable cancers. And immunotherapy, especially ACT, becomes the rear guard, killing escaped tumor cells and creating a microenvironment in which metastasis cannot occur. A "special forces" corps—for example, more recently identified cell subsets that are strongly cytotoxic to tumors—would be a major help to the body in its battle against cancer. Each patient is different, and winning this battle requires a personalized regimen in which ACT

therapy with broad-spectrum antitumor effects may be a critical factor.

Following this principal, we treated cancer patients with personalized combination treatment protocols. Based on our finding that genetically dissimilar DCs can induce the expansion of a specific subset of natural killer T (NKT) cells that are more strongly cytotoxic to tumor cells, we carried out a clinical trial of ACT with NKT cells taken from patients with solid tumors. We used NKT cells cultured in vitro to treat patients after surgery to decrease the tumor burden, or after chemotherapy or radiotherapy to kill most of the tumor cells. We observed in this study that many recipients of targeted therapy took longer to develop drug resistance when their therapy was combined with adoptive NKT transfer. The preliminary results of the study are quite promising, and some of the patients have experienced a complete remission of their cancer (Figure 3, next page) (unpublished observations). Final results are still pending.

Our strategy suggests that well-designed, personalized combination cancer treatments, especially those using adoptive transfer of cells



**FIGURE 3.** The clinical benefits of adoptive novel NKT cell transfer in two patients. A gastric cancer patient with lung metastasis (upper panel) and a pancreatic cancer patient with liver metastasis (lower panel) were treated with adoptive transfer of NKT cells after surgery. The images show that the tumors in both cases shrink and eventually disappear completely.

with strong cytotoxicity against tumors, might provide enhanced clinical benefit for cancer patients. Further investigation is needed to optimize the strategy, and to clarify the effects and mechanisms underlying the combination of targeted therapy with ACT. We believe we are on the right track and that patients will benefit from this strategy.

**References**

- 1. S. A. Grupp *et al.*, *N. Engl. J. Med.* **368**, 1509-1518 (2013).
- 2. S. L. Maude *et al.*, *N. Engl. J. Med.* **371**, 1507-1517 (2014).
- 3. D. L. Porter *et al.*, *Sci. Transl. Med.* **303**, 1-12 (2015).
- 4. S. S. Neelapu *et al.*, *Nat. Rev. Clin. Oncol.*, Epub before print (2017), doi: 10.1038/nrclinonc.2017.148.
- 5. C. J. Turtle *et al.*, *Sci. Transl. Med.* **355**, 1-12 (2016).
- 6. C. J. Turtle *et al.*, *J. Clin. Invest.* **126**, 2123-2138 (2016).

GiantCRO is a full-service Contract Research Organization (CRO), one of the leading CROs in China. Founded in 2001 by medical doctors and statisticians, it is growing rapidly, with its headquarters in Beijing, branch offices in Shanghai, Wuhan, Xi'An, Shenyang, Changsha, Guangzhou, and international offices in the U.S., UK, Germany, Australia, Korea and Japan.



**精诚CRO**

医药智慧 出于精诚

GiantCRO provides efficient and cost-effective services for clinical trial monitoring and management, regulatory affairs consulting, and healthcare industry research to our clients. Working closely with our clients to understand their objectives, we provide solutions tailored to address their specific issues and goals. We are committed to quality customized services and protection of client confidentiality.

GiantCRO is the most recognized Chinese CRO in the field of stroke therapy, embracing its rich medical resources and reputable clinical trial experience.

- ✓ Fully involved in the CHANCE study led by Professor Yongjun Wang (Beijing Tiantan Hospital); the research was published in *The New England Journal of Medicine* in 2013, and changed the international guidelines for stroke therapy.

GiantCRO is ranked one of the top five CROs for clinical studies, including:

- ✓ Cardiovascular
- ✓ Kidney Disease
- ✓ Ophthalmology
- ✓ Psychiatry
- ✓ Endocrinology
- ✓ Medical Cosmetology

#### Clinical Trial Monitoring and Management

- Protocol & Study Design
- CRF Design
- Site Identification and Selection
- Study Subjects Recruitment
- Medical Writing

#### Biostatistics Center

- Data Management
- Statistical Analysis

#### Quality Control Center

- QA & QC
- ICH-GCP & SOP Training

#### Regulatory Affairs

- Drug Registration
- Medical Device Registration
- Medical Device and Drug Distribution Company Registration
- Healthcare Products Registration
- Cosmetics Registration
- Hospital and Clinic Registration

#### Medical Market and Policy Research

- Competitive Information Collection & Analysis
- Brand Studies
- Healthcare Industry Research
- Market Research
- Industry Policy Analysis
- Investment Feasibility Studies

#### Intelligence Center

- Preparation of protocol manuscripts for branch disciplines
- Consultancy on protocol design for clinical studies
- Analysis of medical information



#### **GIANT Med-Pharma Service Group**

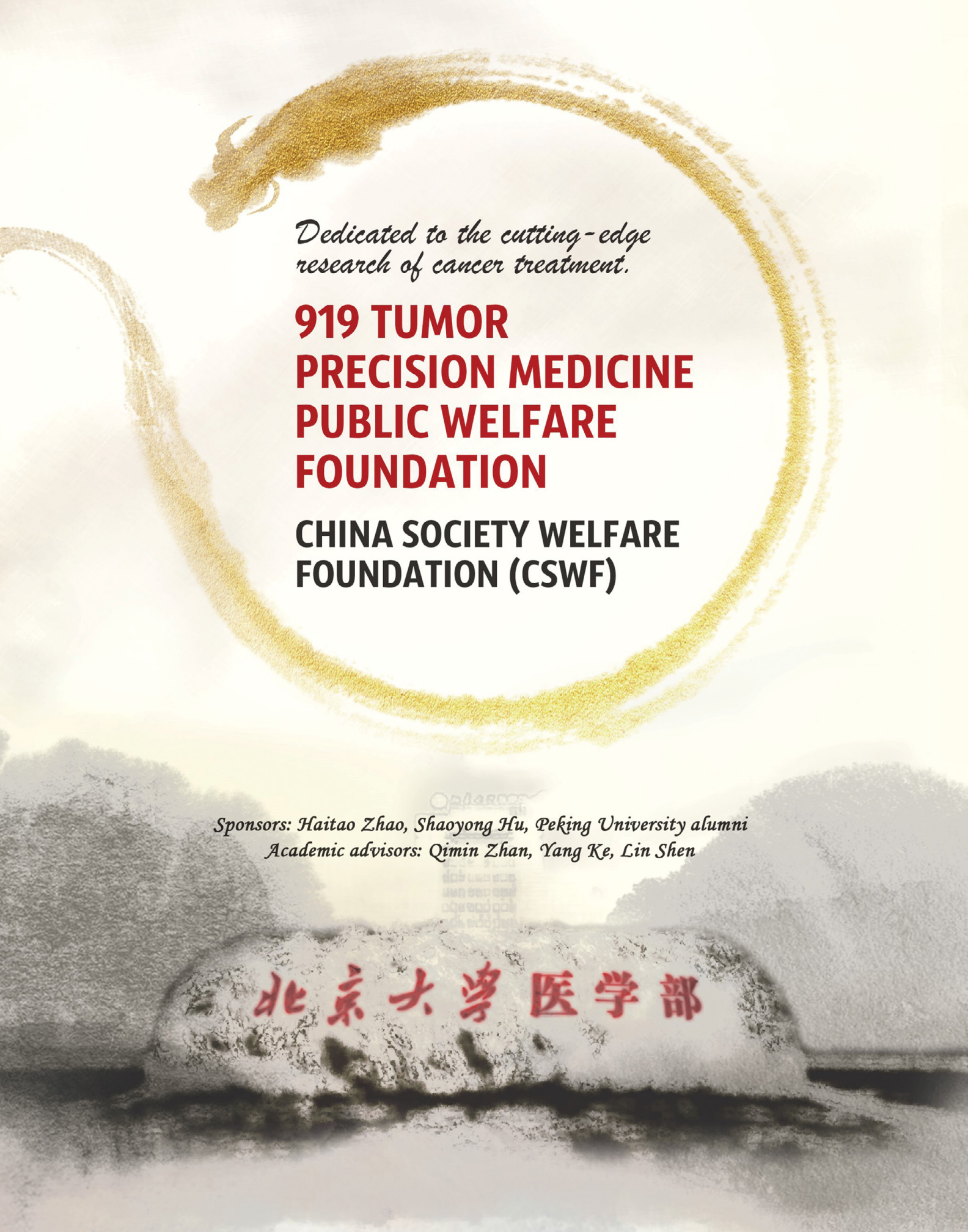
Address: Rm. 2002, Jincheng Jianguo Wuhao Plaza,  
No. 5 Jianguomen North Street, Dongcheng District,  
Beijing 100005

Tel: +86 10-51281119

Fax: +86 10-66112200

Mailbox: [giant@giantcro.com](mailto:giant@giantcro.com)

Website: [www.giantcro.com](http://www.giantcro.com)



*Dedicated to the cutting-edge  
research of cancer treatment.*

**919 TUMOR  
PRECISION MEDICINE  
PUBLIC WELFARE  
FOUNDATION**

**CHINA SOCIETY WELFARE  
FOUNDATION (CSWF)**

*Sponsors: Haitao Zhao, Shaoyong Hu, Peking University alumni*

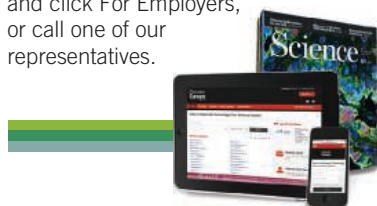
*Academic advisors: Qimin Zhan, Yang Ke, Lin Shen*

**北京大学医学部**

# Science Careers

## SCIENCE CAREERS ADVERTISING

For full advertising details, go to [ScienceCareers.org](http://ScienceCareers.org) and click For Employers, or call one of our representatives.



### AMERICAS

+1 202 326-6577  
+1 202 326-6578  
[advertise@sciencecareers.org](mailto:advertise@sciencecareers.org)

### EUROPE, INDIA, AUSTRALIA, NEW ZEALAND, REST OF WORLD

+44 (0) 1223 326527  
[advertise@sciencecareers.org](mailto:advertise@sciencecareers.org)

### CHINA, KOREA, SINGAPORE, TAIWAN, THAILAND

+86 131 4114 0012  
[advertise@sciencecareers.org](mailto:advertise@sciencecareers.org)

### JAPAN

+81 3-6459-4174  
[advertise@sciencecareers.org](mailto:advertise@sciencecareers.org)

## CUSTOMER SERVICE

### AMERICAS

+1 202 326-6577

### REST OF WORLD

+44 (0) 1223 326528

[advertise@sciencecareers.org](mailto:advertise@sciencecareers.org)

All ads submitted for publication must comply with applicable U.S. and non-U.S. laws. *Science* reserves the right to refuse any advertisement at its sole discretion for any reason, including without limitation for offensive language or inappropriate content, and all advertising is subject to publisher approval. *Science* encourages our readers to alert us to any ads that they feel may be discriminatory or offensive.

**ScienceCareers**

FROM THE JOURNAL SCIENCE AAAS

[ScienceCareers.org](http://ScienceCareers.org)

## POSITIONS OPEN



**PennState**  
College of Medicine

### FACULTY POSITION IN PROTEIN STRUCTURE

Applicants are invited for a tenured or tenure-track faculty position at the Assistant, Associate or Full Professor level in the Department of Cellular and Molecular Physiology, Penn State University College of Medicine. We are seeking Ph.D. and/or M.D. candidates with outstanding research accomplishments in structural biology, and strengths in cryo-EM technology. Further details are given at website: <https://jobs.sciencecareers.org/job/475106/assistant-associate-or-professor-cellular-and-molecular-physiology/> and applications submitted to website: <https://psu.jobs/job/76573>. Applications should include a cover letter, Curriculum Vitae, and statement of research plans, as a single PDF. *Penn State is an Equal Opportunity, Affirmative Action Employer, and is committed to providing employment opportunities to all qualified applicants without regard to race, color, religion, age, sex, sexual orientation, gender identity, national origin, disability or protected veteran status.*

## Post Your Jobs

■ **1,877,103**  
unique job seekers

■ **250,657**  
job applications in 2016

**ScienceCareers**



**Wake Forest™**  
School of Medicine

The Department of Neurobiology and Anatomy at Wake Forest Medical School announces open positions for

### Postdoctoral Training in Multisensory Processes

We seek strong candidates for postdoctoral training funded by an NIH T32 Training Grant. The training program provides a rich collaborative research environment that fosters interdisciplinary approaches to understanding how the brain integrates information from multiple senses to produce perception and adaptive behavior. Candidates with direct experience as well as those in related fields are encouraged to apply. Trainees will have access to any of 9 laboratories using human subjects and/or a variety of animal models rodents-primates with approaches spanning molecular/cellular to perceptual/behavioral. Fellowships are awarded on a competitive basis.

Applications including a current curriculum vitae or nominations should be sent to the Training Grant Director: Dr. Barry E. Stein, [bestein@wakehealth.edu](mailto:bestein@wakehealth.edu), or to its Co-Directors: Dr. Terrence Stanford, [stanford@wakehealth.edu](mailto:stanford@wakehealth.edu) and Dr. Dwayne Godwin, [dgodwin@wakehealth.edu](mailto:dgodwin@wakehealth.edu). A description of the faculty and the program can be accessed via the website: [http://graduate.wfu.edu/admissions/t32/training\\_tpm.html](http://graduate.wfu.edu/admissions/t32/training_tpm.html). Please also submit your application and CV to job opening 26486 at [www.wakehealth.edu/jobssearch/](http://www.wakehealth.edu/jobssearch/).

*EOE/AA: Minorities/Females/Disabled/Vets*

Advance your career  
with expert advice from  
**Science Careers.**



**Download Free Career Advice Booklets!**

[ScienceCareers.org/booklets](http://ScienceCareers.org/booklets)

### Featured Topics:

- Networking
- Industry or Academia
- Job Searching
- Non-Bench Careers
- And More



**ScienceCareers**

FROM THE JOURNAL SCIENCE AAAS



Special Job Focus:

# Chemistry

Issue date: March 2

Book ad by February 15

Ads accepted until February 23 if space allows

**129,562**

subscribers in print  
every week

**62,906**

yearly unique active job  
seekers searching for  
chemistry

**40,352**

yearly applications  
submitted for chemistry  
positions

To book your ad:  
[advertise@sciencecareers.org](mailto:advertise@sciencecareers.org)

**The Americas**

+ 202 326 6577

**Europe**

+44 (0) 1223 326527

**Japan**

+81 3 6459 4174

**China/Korea/Singapore/  
Taiwan**

+86 131 4114 0012

Produced by the *Science*/AAAS  
Custom Publishing Office.

## What makes *Science* the best choice for recruiting?

- Read and respected by 400,000 readers around the globe
- 80% of readers read *Science* more often than any other journal
- Your ad dollars support AAAS and its programs, which strengthens the global scientific community.

## Why choose this Chemistry Focus for your advertisement?

- Relevant ads lead off the career section with a special "Chemistry" banner
- Bonus distributions to:  
American Chemical Society Spring, March 18–22, New Orleans, LA  
Materials Research Society Spring Meeting, April 2–6, Phoenix, AZ

## Expand your exposure.

### Post your print ad online to benefit from:

- Link on the job board homepage directly to chemistry jobs
- Dedicated landing page for jobs in chemistry
- Additional marketing driving relevant job seekers to the job board.



**ScienceCareers**

FROM THE JOURNAL *SCIENCE* AAAS

SCIENCECAREERS.ORG

FOR RECRUITMENT IN SCIENCE, THERE'S ONLY ONE *SCIENCE*.



## Research Position at ICYS, NIMS, Japan

The International Center for Young Scientists (ICYS) of the National Institute for Materials Science (NIMS) is now seeking a few researchers. Successful applicants are expected to pursue innovative research on broad aspects of materials science using most advanced facilities in NIMS (<http://www.nims.go.jp/eng/index.html>).

In the ICYS, we offer a special environment that enables young scientists to work independently based on their own idea and initiatives. All management and scientific discussions will be conducted in English. An annual salary approximately 5.35 million yen (level of 2017) will be offered depending on qualification and experience. Additional research grant of 2 million yen per year will be supplied to each ICYS researcher. The initial contract term is two years and may be extended by one more year depending on the person's performance.

All applicants must have obtained a PhD degree within the last ten years. Applicants should submit an **application form** including a research proposal to be conducted during the ICYS tenure, **CV Header**, CV with list of publications and patents (Be sure to attach the header), list of DOI of journal publications following our instruction, reprints of three significant publications to ICYS Recruitment Desk by March 29, 2018 JST. The **application form** and **CV header** can be downloaded from **our website**. Please visit our website for more details.

ICYS Recruitment Desk,  
National Institute for Materials Science  
<http://www.nims.go.jp/icys/recruitment/index.html>

## THE UNIVERSITY of TENNESSEE HEALTH SCIENCE CENTER


### NEUROSCIENCE FACULTY POSITION

The Department of Anatomy and Neurobiology (<http://www.uthsc.edu/anatomy-neurobiology/>) at the University of Tennessee Health Science Center seeks an outstanding researcher in the area of mechanisms of neurodegenerative diseases to fill an open rank tenure track position (Assistant, Associate, Full Professor). We seek candidates to complement and extend departmental and campus-wide research in neurodegenerative diseases, especially, but not limited to, traumatic brain injury, chronic traumatic encephalopathy, stroke, Alzheimer's disease, Parkinson's disease, and amyotrophic lateral sclerosis. Research incorporating contemporary virally based circuit tracing, optogenetic and DREAD approaches are encouraged. Successful candidates are expected to maintain an independent, extramurally funded research program, and contribute to the department's teaching mission. The department and campus have state-of-the-art laboratories, core facilities, and unique populations of inbred mouse strains for elucidating the pathogenesis and treatment of neurodegenerative diseases (<https://www.uthsc.edu/citg/mouse-strain-and-pilot-projects.php>). We offer competitive salary and start-up packages, and membership in our vibrant interdepartmental Neuroscience Institute (<http://www.uthsc.edu/neuroscience/>).

Candidates must have a Ph.D. or M.D. and significant extramural funding. For best consideration applications should be submitted by **April 1, 2018**. Submit curriculum vitae, summary of research interests, and contact information for three references, in a single Word or PDF file to [hsimmers@uthsc.edu](mailto:hsimmers@uthsc.edu).

*The University of Tennessee is an EEO/AA/Title VI/Title IX, Section 504/ADA/ADEA institution in the provision of its education and employment programs and services.*

# ScienceCareers

FROM THE JOURNAL SCIENCE  AAAS

## Follow us for jobs, career advice & more!



@ScienceCareers



/ScienceCareers



Science Careers

[ScienceCareers.org](http://ScienceCareers.org)

# RUTGERS

New Jersey Medical School

### Chief, Division of Allergy, Immunology and Infectious Disease

The Department of Medicine at Rutgers Robert Wood Johnson School of Medicine in New Brunswick, New Jersey, seeks an academic-physician for the position of Chief. The desirable candidate should have a history of extramural funding in basic or translational research and demonstrable leadership qualities. She or he must also have been successful in promoting teaching, scholarship, and research, as well as developing, managing, and implementing clinical programs. Extensive collaboration opportunities are available within Rutgers University's broad range of research programs including microbiology, immunology, oncology, pharmacology and global health. A competitive salary with a generous startup package is available to support the recruitment of the Chief and additional research and clinical faculty members. Candidates must have an MD or MD/PhD degree, have achieved the position of Associate Professor or Professor, be board certified in infectious disease, and eligible for licensure in New Jersey. Applicants should submit a letter of interest and current curriculum vitae to: **Steven R. Brant, MD, Professor of Medicine, Chair of ID Chief Search Committee, via email at [steven.brant@rwjms.rutgers.edu](mailto:steven.brant@rwjms.rutgers.edu)**

*Rutgers, the State University of New Jersey, is an Equal Opportunity/Affirmative Action employer, and is compliant with the Americans with Disabilities Act (ADA). For more information, please visit <http://recruitment.rutgers.edu/TheRUCCommitment.htm>. Women and minorities are encouraged to apply.*

# In praise of slow

I huff and puff my way up the moderate slope. Even by my own abysmal standards, this is a poor run. In the past hour, I have been overtaken by both an octogenarian and a mum pushing her toddlers in a buggy. Yet I am smiling. I am a happy runner, despite my utter mediocrity at this sport. But at work, happiness had become elusive. After a relatively relaxed Ph.D. and postdoc, I had been thrilled when I landed a tenured job. But as I worked to establish myself as a group leader, I began to feel intense pressure to be more competitive and publish more. Recently, as I wondered why I felt so discontented at my job, I realized that I could apply some lessons from running to my research.

Lesson one: In the right race, your weakness can become your strength. For more than 30 years I equated running to speed and so, lacking the fast-twitch muscles of a sprinter, I chose not to run at all. But once I finally started running, I learned that not being able to sprint makes me a better endurance runner. Similarly, I often perceived my inability to focus on a single research topic as a barrier to success as an academic scientist. But I've found that my desire to branch out to different fields helps me make connections across disciplines and see my work in new ways, which has led to unexpected and exciting insights.

Lesson two: Choose the right pace for your race. Last Christmas, a silver-haired gentleman helped me beat my (lamentable) 5K personal best by whispering, "Go at your pace, not theirs!" when I got stuck behind slower runners. I now realize that this would have been excellent advice early in my career. Some Ph.D. students push themselves too hard and burn out, but I had the opposite problem. I was happy to trundle along at the slowest pace I could get away with—but it ultimately held me back. Even though I like my slow-but-steady pace, I still needed to learn to push myself rather than drag my feet.

Lesson three: An honest race is the only race worth running. I invariably finish in the bottom quartile of the local 5K run. All I would have to do to move up a couple of hundred places is take a shortcut through the fields. Yet I don't. Nobody does. Scientists are usually like that, too—but not always. Principal investigators are pressured to keep their spot in the fast lane, postdocs are chasing the elusive permanent contract, and students are keen to make their mark. Several of my publications would have had a



***"Not being able to sprint makes me a better endurance runner."***

much easier ride through the reviewing system had I been slightly less honest about our findings. The temptation to cheat to get an advantage can be great. However, and this is something that is often overlooked, an advantage is only useful if you are, in fact, engaged in a competition. This brings me to the last lesson, which is the most important of all.

Lesson four: There really is no race. For me, running isn't about being faster than other runners. Likewise, my goal in research is not to "beat" my colleagues. Mark Rowlands, a philosopher, academic, and runner, argues that running makes us happy because it is a form of play and as such has intrinsic value. I don't run just to eat more peanut butter or to

save money on psychotherapy (although these are strong motivating factors in my case). I run because doing so offers a glimpse of life's real value. I now think this is the secret to being happy in research, too. I don't do research only to get invited to conferences, see my name in print, or be promoted. Like running, research is a game with its own intrinsic value. Playing this game of discovery gives me enough joy to keep me going.

Do I recommend an academic career in the slow lane? It doesn't work for everybody. Letting go of ambition in academia is a bit like leaving your GPS watch at home when heading out for a run. Scientists are ambitious; they want to be the hares leading the race ahead. But the tortoise's secret is that there is a lot of fun to be had at the back of the pack. ■

*Irene Nobeli is a lecturer at Birkbeck, University of London. Send your career story to [SciCareerEditor@aaas.org](mailto:SciCareerEditor@aaas.org).*