The elusive dream of regenerating the heart *p. 252*

Summer selections for your reading pleasure *p. 258*

Mesoscale eddies help drive ocean circulation *p. 322*

Science

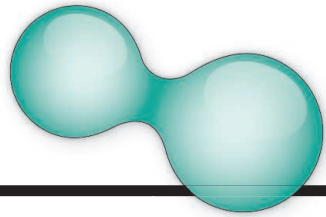
$10
18 JULY 2014
sciencemag.org

AAAS

Slicing the wheat genome

Sequencing illuminates bread wheat evolution and breeding potential *p. 285*

CONTENTS



267 & 288

Tracking electrons as iodomethane dissociates

18 JULY 2014 • VOLUME 345 • ISSUE 6194

NEWS

IN BRIEF

244 Roundup of the week's news

IN DEPTH

247 ALARM OVER BIOSAFETY BLUNDERS
Lapses stoke debate about flu virus research *By J. Cohen*

248 A RADICAL CHANGE IN PEER REVIEW
Each NSF applicant evaluated seven competing proposals *By J. Mervis*

249 PROTEINS AND A PREGNANCY WOE
Surprising insight into preeclampsia may offer early warning for at-risk women *By N. Whitehead*
▸ *SCIENCE TRANSLATIONAL MEDICINE* RESEARCH ARTICLE BY I. A. BUHIMSCHI *ET AL.*

250 RESEARCHERS AIM FOR AN ELECTRICAL MEMORY PROSTHESIS
DARPA-funded effort to develop implants for traumatic brain injuries draws skepticism *By E. Underwood*

251 HARVEST OF GENOME DATA FOR WHEAT GROWERS
DNA survey puts wheat genes in order, in a step toward a full sequence *By E. Pennisi*
▸ SLICING THE WHEAT GENOME SECTION P. 285

FEATURES

252 THE ELUSIVE HEART FIX
Researchers still don't agree whether cell infusions rejuvenate the heart *By J. Couzin-Frankel*

254 Top heart lab comes under fire
By K. Servick

INSIGHTS

BOOKS *ET AL.*

258 SUMMER READING
Eight book reviews from postdocs and grad students

PERSPECTIVES

264 THE FUTURE LIES IN UNCERTAINTY
Sophisticated statistical insights are crucial for gaining knowledge from ever-larger data sets *By D. J. Spiegelhalter*

265 MIND THE MIDZONE
Is the separation of chromosomes during cell division monitored by a checkpoint? *By M. A. Hadders and S. M. A. Lens*
▸ REPORT P. 332

267 CHARGE TRANSFER GOES THE DISTANCE
Electronic relaxation following x-ray excitation illuminates steps in molecular dissociation *By S. T. Pratt*
▸ REPORT P. 288

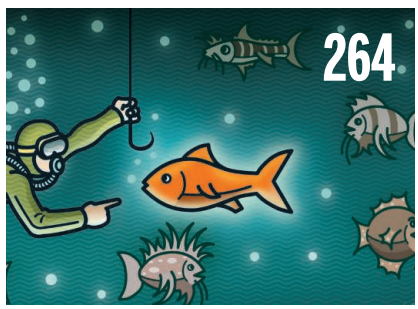
268 IMPROVING CARDIAC RHYTHM WITH A BIOLOGICAL PACEMAKER
Heart muscle cells in a large animal model are reprogrammed to restore heart rate and function *By N. V. Munshi and E. N. Olson*
▸ *SCIENCE TRANSLATIONAL MEDICINE* RESEARCH ARTICLE BY Y.-F. HU *ET AL.*

270 CREATING ANTIOXIDANTS BY OXIDATION CATALYSIS
A key component of vitamin E can be synthesized without use of expensive transition metal catalysts *By B. J. Nachtsheim*
▸ REPORT P. 291

271 ONE, TWO, THREE, CYTOPLASMIC DYNEIN IS GO!
The dynein motor protein must be part of a three-way complex to be fully active *By V. Allan*
▸ REPORT P. 337

272 QUANTUM SYSTEMS UNDER CONTROL
Optical trapping enables the building of quantum matter one atom at a time *By J. Thompson and M. D. Lukin*
▸ REPORT P. 306

274 NIH ROADMAP/COMMON FUND AT 10 YEARS
A mechanism for funding biomedical research at NIH that transcends Institute and Center boundaries is bearing fruit *By F. S. Collins* et al.



277 WALTER GEHRING (1939–2014)
A brief reflection on the life and lab of a preeminent developmental biologist *By M. Levine*

LETTERS

278 PARTNERING WITH CUBA: WEATHER EXTREMES
By J. C. Antuña Marrero et al.

278 PARTNERING WITH CUBA: EARTHQUAKE HAZARDS
By R. S. Yeats

278 DELISTED WHALES GOOD NEWS FOR PIPELINE
By J. J. Alava and J. N. Silberg

279 OUTSIDE THE TOWER: A NIGHT AT THE MUSEUM
By A. de Morrée
▸ EDITORIAL P. 243

DEPARTMENTS

243 EDITORIAL
Science advocacy, defined
By Joanne Padrón Carney
▸ OUTSIDE THE TOWER P. 279

350 WORKING LIFE
Tell the negative committee to shut up
By Fanuel Muindi

Science Staff 242
New Products 343
Science Careers 345

Published by AAAS

CONTENTS

328
Losing a catalytic domain to change function

18 JULY 2014 • VOLUME 345 • ISSUE 6194

RESEARCH

IN BRIEF

280 From *Science* and other journals

REVIEW

283 **ORGANOID GENERATION**
Organogenesis in a dish: Modeling development and disease using organoid technologies
M. A. Lancaster and J. A. Knoblich
REVIEW SUMMARY; FOR FULL TEXT: dx.doi.org/10.1126/science.1247125

REPORTS

288 **ULTRAFAST DYNAMICS**
Imaging charge transfer in iodomethane upon x-ray photoabsorption *B. Erk* et al.
▸ PERSPECTIVE P. 267

291 **ASYMMETRIC CATALYSIS**
High-turnover hypoiodite catalysis for asymmetric synthesis of tocopherols
M. Uyanik et al.
▸ PERSPECTIVE P. 270

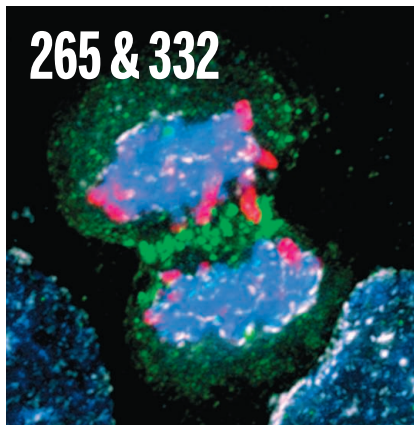
295 **PHOTOVOLTAICS**
A hole-conductor–free, fully printable mesoscopic perovskite solar cell with high stability *A. Mei* et al.

298 **APPLIED OPTICS**
Dielectric gradient metasurface optical elements *D. Lin* et al.

302 **QUANTUM COMPUTING**
Quantum computations on a topologically encoded qubit *D. Nigg* et al.

306 **QUANTUM MECHANICS**
Two-particle quantum interference in tunnel-coupled optical tweezers
A. M. Kaufman et al.
▸ PERSPECTIVE P. 272

309 **OPTICAL METROLOGY**
Electro-optical frequency division and stable microwave synthesis *J. Li* et al.



313 **ASSOCIATIVE LEARNING**
Role of synaptic phosphatidylinositol 3-kinase in a behavioral learning response in *C. elegans H. Ohno* et al.

318 **PALEOCEANOGRAPHY**
Thermohaline circulation crisis and impacts during the mid-Pleistocene transition *L. D. Pena and S. L. Goldstein*

322 **OCEANOGRAPHY**
Oceanic mass transport by mesoscale eddies *Z. Zhang* et al.

325 **FOOD SECURITY**
Leverage points for improving global food security and the environment
P. C. West et al.

328 **ALTERNATIVE SPLICING**
Human tRNA synthetase catalytic nulls with diverse functions *W.-S. Lo* et al.

332 **CELL DIVISION**
Feedback control of chromosome separation by a midzone Aurora B gradient *O. Afonso* et al.
▸ PERSPECTIVE P. 265

337 **MOLECULAR MOTORS**
Activation of cytoplasmic dynein motility by dynactin-cargo adapter complexes *R. J. McKenney* et al.
▸ PERSPECTIVE P. 271

SPECIAL SECTION

Slicing the wheat genome

INTRODUCTION

285 Slicing the wheat genome
K. Eversole et al.

RESEARCH ARTICLE ABSTRACTS

286 A chromosome-based draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome *The International Wheat Genome Sequencing Consortium (IWGSC)*

Ancient hybridizations among the ancestral genomes of bread wheat
T. Marcussen et al.

Genome interplay in the grain transcriptome of hexaploid bread wheat *M. Pfeifer* et al.

Structural and functional partitioning of bread wheat chromosome 3B
F. Choulet et al.

SEE ALSO

▸ NEWS STORY P. 251
▸ sciencemag.org/extra/wheatgenome

ON THE COVER

Red winter wheat (*Triticum aestivum* L.) growing on a farm in Idaho. As one of the largest sources of nutrition for humanity worldwide, the sequencing of its large, complex genome illuminates wheat's evolutionary history and breeding potential. See pages 251 and 285 and sciencemag.org/extra/wheatgenome. *Photo: Mark Thiessen*

PHOTO: (MIDDLE) CID LAB

SCIENCE (ISSN 0036-8075) is published weekly on Friday, except the 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 © 2014 by the American Association for the Advancement of Science. The title SCIENCE is a registered trademark of the AAAS. Domestic individual membership and subscription (51 issues): $153 ($74 allocated to subscription). Domestic institutional subscription (51 issues): $1282; Foreign postage extra: Mexico, Caribbean (surface mail) $55; other countries (air assist delivery) $85. First class, airmail, student, and emeritus rates on request. Canadian rates with GST available upon request, GST #1254 88122. 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:** $10.00 current issue, $15.00 back issue prepaid includes surface postage; bulk rates on request. **Authorization to photocopy** 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 other users registered with the Copyright Clearance Center (CCC) Transactional Reporting Service, provided that $30.00 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.

Published by AAAS

Science
AAAS
1200 New York Avenue, NW, Washington, DC 20005
Bateman House, 82-88 Hills Road, Cambridge, UK CB2 1LQ

Editor-in-Chief Marcia McNutt
Executive Editor Monica M. Bradford **News Editor** Tim Appenzeller
Managing Editor, Research Journals Katrina L. Kelner
Deputy Editors Barbara R. Jasny, Andrew M. Sugden(UK), Valda J. Vinson, Jake S. Yeston

Research and Insights

SR. EDITORS Caroline Ash(UK), Gilbert J. Chin, Lisa D. Chong, Maria Cruz(UK), Julia Fahrenkamp-Uppenbrink(UK), Pamela J. Hines, Stella M. Hurtley(UK), Paula A. Kiberstis, Marc S. Lavine(Canada), Kristen L. Mueller, Ian S. Osborne(UK), Beverly A. Purnell, L. Bryan Ray, Guy Riddihough, H. Jesse Smith, Peter Stern(UK), Phillip D. Szuromi, Brad Wible, Nicholas S. Wigginton, Laura M. Zahn **ASSOCIATE EDITORS** Brent Grocholski, Melissa R. McCartney, Margaret M. Moerchen, Jelena Stajic, Sacha Vignieri **BOOK REVIEW EDITOR** Sherman J. Suter **ASSOCIATE LETTERS EDITOR** Jennifer Sills **EDITORIAL MANAGER** Cara Tate **SR. COPY EDITORS** Jeffrey E. Cook, Chris Filiatreau, Cynthia Howe, Harry Jach, Lauren Kmec, Barbara P. Ordway, Trista Wagoner **SR. EDITORIAL COORDINATORS** Carolyn Kyle, Beverly Shields **EDITORIAL COORDINATORS** Ramatoulaye Diop, Joi S. Granger, Lisa Johnson, Anita Wynn **PUBLICATIONS ASSISTANTS** Aneera Dobbins, Jeffrey Hearn, Dona Mathieu, Le-Toya Mayne Flood, Shannon McMahon, Scott Miller, Jerry Richardson, Rachel Roberts(UK), Alice Whaley(UK), Brian White **ADMINISTRATIVE SUPPORT** Janet Clements(UK), Joan Cuthbert(UK), Maryrose Madrid, John Wood(UK)

News

NEWS MANAGING EDITOR John Travis **INTERNATIONAL EDITOR** Richard Stone **DEPUTY NEWS EDITORS** Daniel Clery(UK), Robert Coontz, Elizabeth Culotta, David Grimm, David Malakoff, Leslie Roberts **CONTRIBUING EDITORS** Martin Enserink(Europe), Mara Hvistendahl (Asia) **SR. CORRESPONDENTS** Jeffrey Mervis, Elizabeth Pennisi **NEWS WRITERS** Adrian Cho, Jon Cohen, Jennifer Couzin-Frankel, Carolyn Gramling, Eric Hand, Jocelyn Kaiser, Kelly Servick, Robert F. Service, Erik Stokstad, Emily Underwood **CONTRIBUTING INTERNS** Hassan DuRant, Xochitl Rojas-Raha, Nadia Whitehead **CORRESPONDENTS** Pallava Bagla(South Asia), Michael Balter(Paris), John Bohannon, Ann Gibbons, Sam Kean, Richard A. Kerr, Eli Kintisch, Kai Kupferschmidt(Berlin), Andrew Lawler, Christina Larson(Beijing), Mitch Leslie, Charles C. Mann, Eliot Marshall, Virginia Morell, Dennis Normile(Tokyo), Heather Pringle, Tania Rabesandratana(Brussels), Gretchen Vogel(Berlin), Lizzie Wade(Mexico City) **CAREERS** Jim Austin(Editor), Donisha Adams **COPY EDITORS** Kara Estelle, Nora Kelly, Jennifer Levin **ADMINISTRATIVE SUPPORT** Scherraine Mack

Executive Publisher Alan I. Leshner
Publisher Beth Rosner **Chief Digital Media Officer** Rob Covey

BUSINESS OPERATIONS AND ADMINISTRATION DIRECTOR Deborah Rivera-Wienhold **BUSINESS SYSTEMS AND FINANCIAL ANALYSIS DIRECTOR** Randy Yi **MANAGER OF FULFILLMENT SYSTEMS** Neal Hawkins **SYSTEMS ANALYST** Nicole Mehmedovich **ASSISTANT DIRECTOR, BUSINESS OPERATIONS** Eric Knott **MANAGER, BUSINESS OPERATIONS** Jessica Tierney **BUSINESS ANALYSTS** Cory Lipman, Cooper Tilton, Celeste Troxler **FINANCIAL ANALYST** Jeremy Clay **RIGHTS AND PERMISSIONS: ASSISTANT DIRECTOR** Emilie David **PERMISSIONS ASSOCIATE** Elizabeth Sandler **RIGHTS, CONTRACTS, AND LICENSING ASSOCIATE** Lili Kiser

MARKETING DIRECTOR Ian King **MARKETING MANAGER** Julianne Wielga **MARKETING ASSOCIATE** Elizabeth Sattler **SR. MARKETING EXECUTIVE** Jennifer Reeves **SR. ART ASSOCIATE, PROJECT MANAGER** Tzeitel Sorrosa **ART ASSOCIATE** Seil Lee **ASSISTANT COMMERCIAL EDITOR** Selby Frame **MARKETING PROJECT MANAGER** Angelissa McArthur **SR. WRITER** Bill Zimmer **PROGRAM DIRECTOR, AAAS MEMBER CENTRAL** Peggy Mihelich **FULFILLMENT SYSTEMS AND OPERATIONS** membership@aaas.org **MANAGER, MEMBER SERVICES** Pat Butler **SPECIALISTS** LaToya Casteel, Javia Flemmings, Latasha Russell **MANAGER, DATA ENTRY** Mickie Napoleoni **DATA ENTRY SPECIALISTS** JJ Regan, Jaimee Wise, Fiona Giblin

DIRECTOR, SITE LICENSING Tom Ryan **DIRECTOR, CORPORATE RELATIONS** Eileen Bernadette Moran **SR. PUBLISHER RELATIONS SPECIALIST** Kiki Forsythe **PUBLISHER RELATIONS MANAGER** Catherine Holland **PUBLISHER RELATIONS, EASTERN REGION** Keith Layson **PUBLISHER RELATIONS, WESTERN REGION** Ryan Rexroth **MANAGER, SITE LICENSE OPERATIONS** Iquo Edim **FULFILLMENT ANALYST** Lana Guz **ASSOCIATE DIRECTOR, MARKETING** Christina Schlecht **MARKETING ASSOCIATES** Thomas Landreth, Minah Kim

DIRECTOR OF WEB TECHNOLOGIES Ahmed Khadr **SR. DEVELOPER** Chris Coleman **DEVELOPER** Dan Berger **SR. PROJECT MANAGER** Trista Smith **SYSTEMS ENGINEER** Luke Johnson **PRODUCT MANAGER** Walter Jones

CREATIVE DIRECTOR, MULTIMEDIA Martyn Green **DIRECTOR OF ANALYTICS** Enrique Gonzales **SR. WEB PRODUCER** Sarah Crespi **WEB PRODUCER** Alison Crawford **VIDEO PRODUCER** Nguyen Nguyen **SOCIAL MEDIA PRODUCER** Meghna Sachdev

DIRECTOR OF OPERATIONS PRINT AND ONLINE Lizabeth Harman **PRINT PRODUCTION DIRECTOR** Wendy K. Shank **PREFLIGHT MANAGER** Marcus Spiegler **ASSISTANT MANAGER ONLINE** Lisa Stanford **ASSISTANT MANAGER PRINT** Rebecca Doshi **SR SPECIALISTS** Steve Forrester, Amy Hardcastle, Jason Hillman, Antoinette Hodal, Tara Kelly, Anthony Rosen **SPECIALISTS** Jacob Hedrick, Nichele Johnston, Yuse Lajiminmuhip, Lori Murphy, Kimberley Oster

DESIGN DIRECTOR Beth Rakouskas **ASSOCIATE ART DIRECTOR** Laura Creveling **SR. ILLUSTRATORS** Chris Bickel, Katharine Sutliff **ILLUSTRATOR** Valerie Altounian **SR. ART ASSOCIATES** Holly Bishop, Preston Huey **ART ASSOCIATES** Kay Engman, Garvin Grullón, Chrystal Smith **SR. PHOTO EDITOR** William Douthitt **PHOTO EDITOR** Leslie Blizard

DIRECTOR, GLOBAL COLLABORATION, CUSTOM PUBLICATIONS, ADVERTISING Bill Moran **EDITOR, CUSTOM PUBLISHING** Sean Sanders: 202-326-6430 **ASSISTANT EDITOR, CUSTOM PUBLISHING** Tianna Hicklin: 202-326-6463 **ADVERTISING MARKETING MANAGER** Justin Sawyers: 202-326-7061 science_advertising@aaas.org **ADVERTISING SUPPORT MANAGER** Karen Foote: 202-326-6740 **ADVERTISING PRODUCTION OPERATIONS MANAGER** Deborah Tompkins **SR. TRAFFIC ASSOCIATE** Christine Hall **SALES COORDINATOR** Shirley Young **ASSOCIATE DIRECTOR, COLLABORATION, CUSTOM PUBLICATIONS/CHINA/TAIWAN/KOREA/SINGAPORE** Ruolei Wu: +86-186 0822 9345, rwu@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** Makiko Hara: +81 (0) 3 6802 4616, FAX +81 (0) 3 6802 4615, ads@sciencemag.jp **CHINA/TAIWAN** Ruolei Wu: +86-186 0822 9345

WORLDWIDE ASSOCIATE DIRECTOR OF SCIENCE CAREERS Tracy Holmes: +44 (0) 1223 326525, FAX +44 (0) 1223 326532 tholmes@science-int.co.uk **CLASSIFIED** advertise@sciencecareers.org **U.S. SALES** Tina Burks: 202-326-6577, Nancy Toema: 202-326-6578 **SALES ADMINISTRATOR** Marci Gallun **EUROPE/ROW SALES** Axel Gesatzki, Sarah Lelarge **SALES ASSISTANT** Kelly Grace **JAPAN** Yuri Kobayashi: +81 (0)90-9110-1719, careerads@sciencemag.jp **CHINA/TAIWAN** Ruolei Wu: +86-186 0822 9345 rwu@aaas.org **MARKETING MANAGER** Allison Pritchard **MARKETING ASSOCIATE** Aimee Aponte

AAAS BOARD OF DIRECTORS RETIRING PRESIDENT, CHAIR Phillip A. Sharp **PRESIDENT** Gerald R. Fink **PRESIDENT-ELECT** Geraldine (Geri) Richmond **TREASURER** David Evans Shaw **CHIEF EXECUTIVE OFFICER** Alan I. Leshner **BOARD** Bonnie L. Bassler, May R. Berenbaum, Carlos J. Bustamante, Claire M. Fraser, Laura H. Greene, Elizabeth Loftus, Raymond Orbach, Inder M. Verma

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 202-326-6755 **REPRINTS:** Author Inquiries 800-635-7181 **COMMERCIAL INQUIRIES** 803-359-4578 **PERMISSIONS** 202-326-6765, permissions@aaas.org **AAAS Member Services** 202-326-6417 or http://membercentral.aaas.org/discounts

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 of conflicting points of view. Accordingly, all articles published in *Science*—including editorials, news and comment, and books 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 pages 680 and 681 of the 7 February 2014 issue or access www.sciencemag.org/about/authors

SENIOR EDITORIAL BOARD
A. Paul Alivisatos, *Lawrence Berkeley Nat'l. Laboratory,* **Ernst Fehr,** *U. of Zürich*
Susan M. Rosenberg, *Baylor College of Medicine,* **Michael S. Turner,** *U. of Chicago*

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

Adriano Aguzzi, *U. Hospital Zürich*
Takuzo Aida, *U. of Tokyo*
Leslie Aiello, *Wenner-Gren Foundation*
Judith Allen, *U. of Edinburgh*
Sonia Altizer, *U. of Georgia*
Virginia Armbrust, *U. of Washington*
Sebastian Amigorena, *Institut Curie*
Kathryn Anderson, *Memorial Sloan-Kettering Cancer Center*
Peter Andolfatto, *Princeton U.*
Meinrat O. Andreae, *Max-Planck Inst. Mainz*
Paola Arlotta, *Harvard U.*
Johan Auwerx, *EPFL*
David Awschalom, *U. of Chicago*
Jordi Bascompte, *Estación Biológica de Doñana CSIC*
Facundo Batista, *London Research Inst.*
Ray H. Baughman, *U. of Texas, Dallas*
David Baum, *U. of Wisconsin*
Mark Bear, *Massachusetts Inst. of Technology*
Kamran Behnia, *ESPCI-ParisTech*
Yasmine Belkaid, *NIAID, NIH*
Philip Benfey, *Duke U.*
Stephen J. Benkovic, *Penn State U.*
Gabriele Bergers, *U. of California, San Francisco*
Christophe Bernard, *Aix-Marseille U.*
Bradley Bernstein, *Massachusettes General Hospital*
Gregory C. Beroza, *Stanford U.*
Peer Bork, *EMBL*
Bernard Bourdon, *Ecole Normale Supérieure de Lyon*
Chris Bowler, *École Normale Supérieure*
Ian Boyd, *U. of St. Andrews*
Emily Brodsky, *U. of California, Santa Cruz*
Ron Brookmeyer, *U. of California, Los Angeles* **(S)**
Christian Büchel, *U. Hamburg-Eppendorf*
Joseph A. Burns, *Cornell U.*
William P. Butz, *Population Reference Bureau*
Gyorgy Buzsaki, *New York U. School of Medicine*
Blanche Capel, *Duke U.*
Mats Carlsson, *U. of Oslo*
David Clapham, *Children's Hospital Boston*
David Clary, *U. of Oxford*
Joel Cohen, *Rockefeller U., Columbia U.*
Jonathan D. Cohen, *Princeton U.*
James Collins, *Boston U.*
Robert Cook-Deegan, *Duke U.*
Alan Cowman, *Walter & Eliza Hall Inst.*
Robert H. Crabtree, *Yale U.*
Janet Currie, *Princeton U.*
Jeff L. Dangl, *U. of North Carolina*
Tom Daniel, *U. of Washington*
Frans de Waal, *Emory U.*
Stanislas Dehaene, *Collège de France*
Robert Desimone, *MIT*
Claude Desplan, *New York U.*
Ap Dijksterhuis, *Radboud U. of Nijmegen*
Dennis Discher, *U. of Pennsylvania*
Gerald W. Dorn II, *Washington U. School of Medicine*
Jennifer A. Doudna, *U. of California, Berkeley*
Bruce Dunn, *U. of California, Los Angeles*
Christopher Dye, *WHO*
Todd Ehlers, *U. of Tuebingen*
David Ehrhardt, *Carnegie Inst. of Washington*
Tim Elston, *U. of North Carolina at Chapel Hill*
Gerhard Ertl, *Fritz-Haber-Institut, Berlin*
Barry Everitt, *U. of Cambridge*
Ernst Fehr, *U. of Zurich*
Anne C. Ferguson-Smith, *U. of Cambridge*
Michael Feuer, *The George Washington U.*
Peter Fratzl, *Max-Planck Inst.*
Elaine Fuchs, *Rockefeller U.*
Daniel Geschwind, *UCLA*
Andrew Gewirth, *U. of Illinois*
Karl-Heinz Glassmeier, *TU Braunschweig*
Julia R. Greer, *Caltech*
Elizabeth Grove, *U. of Chicago*
Kip Guy, *St. Jude's Children's Research Hospital*
Taekjip Ha, *U. of Illinois at Urbana-Champaign*
Christian Haass, *Ludwig Maximilians U.*
Steven Hahn, *Fred Hutchinson Cancer Research Center*
Michael Hasselmo, *Boston U.*
Martin Heimann, *Max-Planck Inst. Jena*
Yka Helariutta, *U. of Finland*
James A. Hendler, *Rensselaer Polytechnic Inst.*
Janet G. Hering, *Swiss Fed. Inst. of Aquatic Science & Technology*
Michael E. Himmel, *National Renewable Energy Lab.*
Kai-Uwe Hinrichs, *U. of Bremen*
Kei Hirose, *Tokyo Inst. of Technology*
David Hodell, *U. of Cambridge*
David Holden, *Imperial College*
Lora Hooper, *UT Southwestern Medical Ctr. at Dallas*
Thomas Hudson, *Ontario Inst. for Cancer Research*
Raymond Huey, *U. of Washington*
Steven Jacobsen, *U. of California, Los Angeles*
Kai Johnsson, *EPFL Lausanne*
Peter Jonas, *Inst. of Science & Technology (IST) Austria*
Matt Kaeberlein, *U. of Washington*
William Kaelin Jr., *Dana-Farber Cancer Inst.*
Daniel Kahne, *Harvard U.*
Daniel Kammen, *U. of California, Berkeley*
Masashi Kawasaki, *U. of Tokyo*
Joel Kingsolver, *U. of North Carolina at Chapel Hill*
Robert Kingston, *Harvard Medical School*
Alexander Kolodkin, *Johns Hopkins U.*
Roberto Kolter, *Harvard Medical School*
Alberto R. Kornblihtt, *U. of Buenos Aires*
Leonid Kruglyak, *UCLA*
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, *Cambridge U.*
Marcia C. Linn, *U. of California, Berkeley*
Jianguo Liu, *Michigan State U.*
Luis Liz-Marzan, *CIC biomaGUNE*
Jonathan Losos, *Harvard U.*
Ke Lu, *Chinese Acad. of Sciences*
Christian Lüscher, *U. of Geneva*
Laura Machesky, *CRUK Beatson Inst. for Cancer Research*
Anne Magurran, *U. of St. Andrews*
Oscar Marin, *CSIC & U. Miguel Hernández*
Charles Marshall, *U. of California, Berkeley*
C. Robertson McClung, *Dartmouth College*
Graham Medley, *U. of Warwick*
Yasushi Miyashita, *U. of Tokyo*
Richard Morris, *U. of Edinburgh*
Alison Motsinger-Reif, *NC State U.* **(S)**
Sean Munro, *MRC Lab. of Molecular Biology*
Thomas Murray, *The Hastings Center*
James Nelson, *Stanford U. School of Med.*
Karen Nelson, *J. Craig Venter Institute*
Daniel Neumark, *U. of California, Berkeley*
Timothy W. Nilsen, *Case Western Reserve U.*
Pär Nordlund, *Karolinska Inst.*
Helga Nowotny, *European Research Advisory Board*
Ben Olken, *MIT*
Luke O'Neill, *Trinity College, Dublin*
Joe Orenstein, *U. of California Berkeley & Lawrence Berkeley National Lab*
Harry Orr, *U. of Minnesota*
Andrew Oswald, *U. of Warwick*
Steve Palumbi, *Stanford U.*
Jane Parker, *Max-Planck Inst. of Plant Breeding Research*
Giovanni Parmigiani, *Dana-Farber Cancer Inst.* **(S)**
Donald R. Paul, *U. of Texas, Austin*
John H. J. Petrini, *Memorial Sloan-Kettering Cancer Center*
Joshua Plotkin, *U. of Pennsylvania*
Philippe Poulin, *CNRS*
David Randall, *Colorado State U.*
Colin Renfrew, *U. of Cambridge*
Trevor Robbins, *U. of Cambridge*
Jim Roberts, *Fred Hutchinson Cancer Research Ctr.*
Barbara A. Romanowicz, *U. of California,Berkeley*
Jens Rostrup-Nielsen, *Haldor Topsoe*
Mike Ryan, *U. of Texas, Austin*
Mitinori Saitou, *Kyoto U.*
Shimon Sakaguchi, *Kyoto U.*
Miquel Salmeron, *Lawrence Berkeley National Lab*
Jürgen Sandkühler, *Medical U. of Vienna*
Alexander Schier, *Harvard U.*
Randy Seeley, *U. of Cincinnati*
Vladimir Shalaev, *Purdue U.*
Robert Siliciano, *Johns Hopkins School of Medicine*
Joseph Silk, *Institut d'Astrophysique de Paris*
Denis Simon, *Arizona State U.*
Alison Smith, *John Innes Centre*
Richard Smith, *U. of North Carolina* **(S)**
John Speakman, *U. of Aberdeen*
Allan C. Spradling, *Carnegie Institution of Washington*
Jonathan Sprent, *Garvan Inst. of Medical Research*
Eric Steig, *U. of Washington*
Paula Stephan, *Georgia State U. and National Bureau of Economic Research*
Molly Stevens, *Imperial College London*
V. S. Subrahmanian, *U. of Maryland*
Ira Tabas, *Columbia U.*
Sarah Teichmann, *Cambridge U.*
John Thomas, *North Carolina State U.*
Christopher Tyler-Smith, *The Wellcome Trust Sanger Inst.*
Herbert Virgin, *Washington U.*
Bert Vogelstein, *Johns Hopkins U.*
Cynthia Volkert, *U. of Göttingen*
Douglas Wallace, *Dalhousie U.*
David Wallach, *Weizmann Inst. of Science*
Ian Walmsley, *U. of Oxford*
David A. Wardle, *Swedish U. of Agric. Sciences*
David Waxman, *Fudan U.*
Jonathan Weissman, *U. of California, San Francisco*
Chris Wikle, *U. of Missouri* **(S)**
Ian A. Wilson, *The Scripps Res. Inst.* **(S)**
Timothy D. Wilson, *U. of Virginia*
Rosemary Wyse, *Johns Hopkins U.*
Jan Zaanen, *Leiden U.*
Kenneth Zaret, *U. of Pennsylvania School of Medicine*
Jonathan Zehr, *U. of California, Santa Cruz*
Len Zon, *Children's Hospital Boston*
Maria Zuber, *MIT*

BOOK REVIEW BOARD
David Bloom, *Harvard U.,* **Samuel Bowring,** *MIT,* **Angela Creager,** *Princeton U.,* **Richard Shweder,** *U. of Chicago,* **Ed Wasserman,** *DuPont*

Published by AAAS

EDITORIAL

Science advocacy, defined

When one thinks of advocacy, the first—usually cynical—idea that comes to mind is generously compensated lobbyists walking the halls of legislature and advocating on behalf of their clients over expensive dinners. While this image may make for great television drama, it is an incomplete picture, especially as regards science advocacy.

According to the Merriam-Webster dictionary, an "advocate" is "a person who works for a cause or group." This definition aptly describes a science advocate, and it was underscored in the NextGenVOICES survey on advocacy in the 4 April 2014 issue of *Science*. When young investigators were asked what they would do if they had five extra hours to advocate for science, the majority desired to engage with the public and to inspire students and teachers alike about the wonders and power of science and technology. Others wanted to bring their stories to government in the hope of inspiring support for the next generation of researchers.

"Scientists [who] know how to talk to Congress are worth their weight in gold."

The recognition that scientific research involves more than simply working alone in a lab and publishing the results demonstrates that these survey participants view their place at the intersection of science and society in a more nuanced fashion. They combine both a passion for their work and a sense of responsibility for contributing to a broader discourse. *Science*'s "Outside the Tower" (see page 279), written by scientists or citizen scientists, will highlight projects or events that engage the public.

The relationship between science policy and advocacy is also a nuanced one. Science can be a powerful tool for developing policies (e.g., public health, food safety), and at the same time policies can have a powerful impact on the conduct of science, in ways that extend far beyond public funding allocations alone.

If scientists want to advocate for science, they need to add some additional skills to an already exceptional skill set. Being outstanding in one's field does not make one an engaging communicator. Learning how to communicate science to a lay audience is as much an art as it is a science, and more institutions offer such courses.

For science policy, it is incumbent on researchers to gain a better understanding of the relationship between science and policy, the roles of the executive and legislative branches of government, and the responsibility that comes with being a reputable advocate. A communicator extraordinaire may do more harm than good if she or he does not understand or appreciate the broader context within which the debate over science policy is conducted.

Fortunately, there is a hunger and desire to learn about science policy and advocacy. The 40-year success of the AAAS Science and Technology Policy Fellowships, and the recently held Catalyzing Advocacy in Science and Engineering workshop for graduate students, demonstrate this point. The workshop (arranged by AAAS and other organizations) introduced students to the role of science in policy-making with the goal of empowering them with ways to become a voice for science throughout their career. As one participant stated, "I finally understood science policy and how it is present in different ways in a lot of aspects of our lives as citizens and scientists." Or the tweet from another student: "Scientists that know how to talk to Congress are worth their weight in gold."

To paraphrase Thomas Kuhn, the physicist and historian, science is more than just a set of discrete facts, theories, and methods. It comprises individuals across a multitude of disciplines, sectors, and nations who contribute to the "constellation" that we call science. So scientists should recognize their star quality and hone their skills to become effective advocates for science, whether at elementary schools, museums, churches, or even the halls of government.

– Joanne Padrón Carney

Joanne Padrón Carney is the director of government relations at AAAS, Washington, DC. E-mail: jcarney@aaas.org

10.1126/science.1258492

IMAGES: (RIGHT) C. SMITH/*SCIENCE*; (INSET) ELLAGRIN/ISTOCKPHOTO.COM

Downloaded from www.sciencemag.org on August 4, 2014

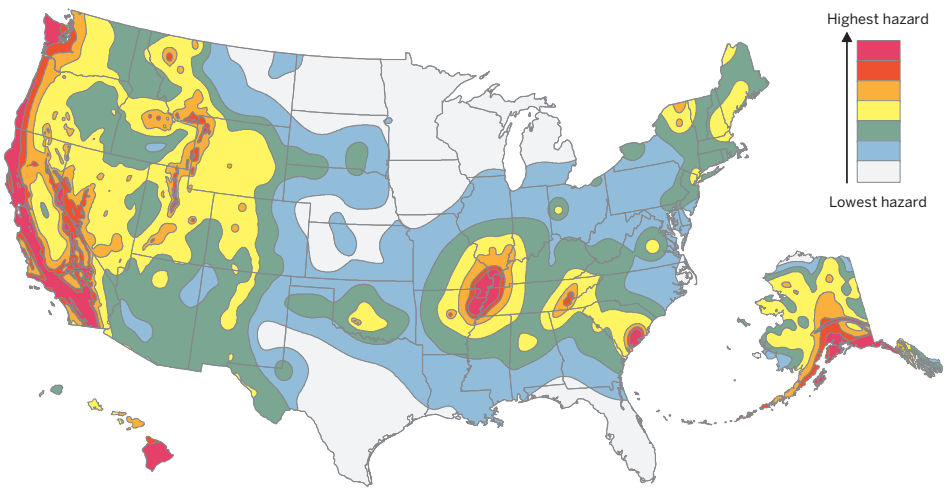
Published by AAAS

NEWS

> “It felt very much like a punch to the gut.”
>
> **Hannah Gay**, a pediatrician at the University of Mississippi Medical Center in Jackson, at a 10 July teleconference, after learning that HIV was detected in a child she treats. Infected at birth, the child appeared to be cured for more than 2 years.

IN BRIEF

Shaking out the U.S.'s earthquake risk



New hazard maps highlight seismic hot spots.

The U.S. Geological Survey's (USGS's) seismic hazard maps influence hundreds of billions of dollars in construction every year. This week, USGS released its latest earthquake hazard map for the country. On a national scale, the U.S. earthquake hazard picture hasn't changed much since the last 2008 maps, says the report's lead author, seismologist Mark Petersen of USGS in Golden, Colorado. The hazards are still high in California and the Pacific Northwest, and there's still a hot spot in the New Madrid Seismic Zone. But, Petersen notes, the devil is in the details. The maps include new ground motion models and lessons learned from powerful subduction-zone earthquakes around the world, such as the 9.0-magnitude Tohoku earthquake of 2011. The new data lead to increased hazard estimation in southern Cascadia, as well as new models of fault ruptures across California. The new maps don't include induced seismicity—earthquakes due to human activities, such as reintroducing wastewater into injection wells—because mapping out those hazards to incorporate them into building codes requires “a different logic tree,” Petersen says. USGS is planning regional workshops for future map updates to include the controversial quakes. Stay tuned. http://scim.ag/2014seishaz

AROUND THE WORLD

Building 'big data' scientists

BEIJING | In a deal brokered with China's Education Ministry, IBM pledged last week to donate $100 million in software to Chinese universities to help produce a new generation of data scientists. The agreement includes the creation of research centers focused on big data—data too large for traditional applications—at 100 institutions. The $2.3 billion big data technology and services market in China is expected to nearly quadruple in size in the next 2 years, according to CCID Consulting in Beijing. But China lacks sufficient scientists to handle that deluge of data. IBM's cash infusion could help fix that—and would give IBM valuable access to the Chinese market, notes Shi Yong, executive deputy director of the Chinese Academy of Sciences' Research Center on Fictitious Economy & Data Science in Beijing. “It's a win-win situation.”

Organic foods come up rosy

NEWCASTLE, U.K. | The largest meta-analysis of organic food studies to date has found that organic fruits and vegetables have higher levels of antioxidants compared with conventionally grown crops. The levels were up to 69% higher, which the authors say is equivalent to the amount in two servings. Pesticide levels were lower, which was expected because organic growers are not allowed to use

Vegetables at a farmers market in Seattle, Washington.

CREDITS: (TOP TO BOTTOM) USGS; (PHOTO) ISTOCKPHOTO

Downloaded from www.sciencemag.org on August 4, 2014

Published by AAAS

most kinds. Cadmium, which can be toxic in high doses, was also lower in organically grown cereals. (Both pesticides and cadmium were below the levels of concern set by regulators.) The review, which appears this week in the *British Journal of Nutrition*, examined 343 studies—enough to allow the authors to weight each study according to its quality, says the review's lead author Carlo Leifert, an agronomist at Newcastle University.

No Scripps-USC merger

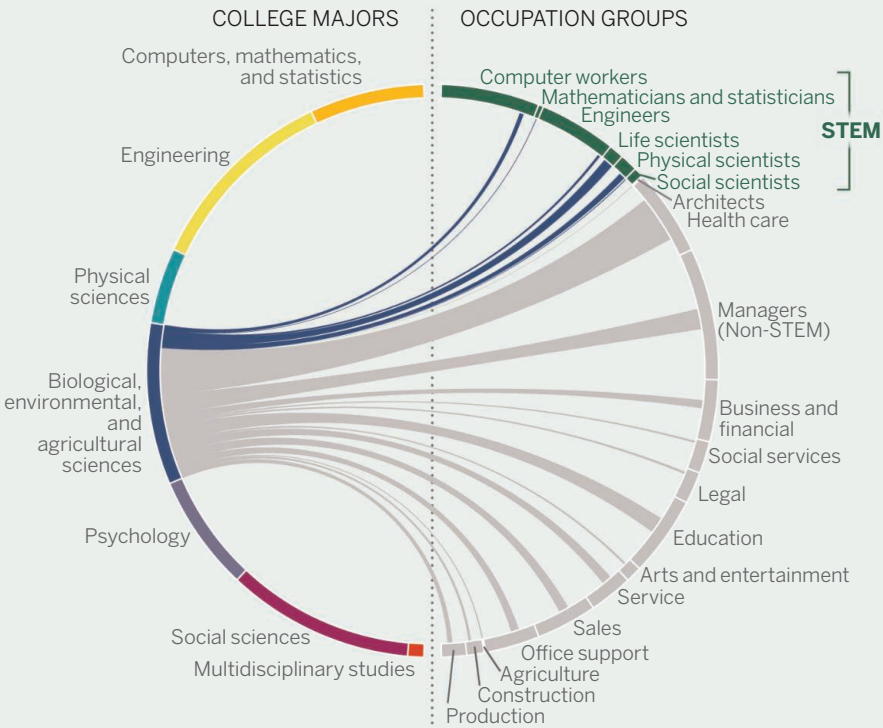
SAN DIEGO, CALIFORNIA | Officials at the Scripps Research Institute announced 9 July that they've called off discussions with the University of Southern California (USC) on a possible merger. The world's largest private biomedical research

The Scripps Florida campus.

institute, with campuses in San Diego, California, and Jupiter, Florida, Scripps has faced financial woes, in large part due to a 12% drop in research grants from the National Institutes of Health between 2007 and 2013. USC has deeper pockets and more diverse sources of income. Media reports about the proposed merger suggested USC offered $15 million over 40 years to absorb the 262-member Scripps faculty. But those researchers revolted: Ten department chairs and a dean sent a letter to Scripps's president and the chair of its board of trustees, saying the proposed terms were "not even close to what it would take to build faculty support."
http://scim.ag/noUSCmerger

First dengue vaccine limited

LYON, FRANCE | The first phase III efficacy trial of a vaccine for the tropical, mosquito-borne disease dengue has produced a glass-half-empty-or-half-full conundrum. The vaccine proved safe, but immunization reduced incidence of the disease by only 56.5%, mostly due to limited efficacy against dengue 2, one of four serotypes of the virus. Researchers reported the trial results, involving more than 10,000 children in five Southeast Asian countries, online on 11 July in *The Lancet*. Sanofi Pasteur, which developed the vaccine, plans to deliver the first doses by the end of 2015, provided the vaccine obtains market approval, says Guillaume Leroy, head of the company's dengue program. Public health officials in affected countries now face a difficult decision: whether cutting the disease burden in half is worth the cost of including the vaccine in national immunization programs.



Source: U.S. Census Bureau, 2012 American Community Survey

A fresh look at the STEM workforce

The U.S. Census Bureau has created a nifty interactive graphic showing the employment picture for college graduates with so-called STEM (science, technology, engineering, and mathematics) degrees. Both sides in the long-running debate over whether the country is producing too few—or too many—STEM workers have been handicapped by confusing data. So last fall the Census Bureau asked a team of graphic designers and statistical analysts to design a more compelling way to depict job flows: where the nation's 14.8 million STEM graduates are working and what training the 5.2 million people holding STEM jobs have received. The chord diagrams, using data from a study of the STEM workforce requested by the White House, paint a compelling picture, including nuggets such as where people with biology degrees are employed (above). The original report came out last fall (census.gov/prod/2013pubs/acs-23.pdf), but you'll probably have more fun playing with the graphic, at census.gov/dataviz/visualizations/stem/stem-html/.

U.K. science minister shuffle

LONDON | David Willetts resigned as U.K. universities and science minister on 14 July as part of a government reshuffle. U.K. scientists and policy leaders had praised Willetts as a strong advocate for research funding after the financial downturn: Under Willetts, appointed in 2010, science dodged expected cuts and was guaranteed £4.6 billion annually for the following 4 years—although, adjusted for inflation, the science budget lost £1.1 billion from October 2010 to 2015–2016, notes the U.K.

PHOTO: BRUCE R. BENNETT/ALAMY

Published by AAAS

Campaign for Science and Engineering. On 15 July, Greg Clark, a conservative minister responsible for cities policy and constitutional reform, took over Willetts's portfolio. The reshuffle is seen as an effort by Prime Minister David Cameron to bring in more women and younger ministers in his government ahead of next year's parliamentary elections. http://scim.ag/Willetts

NEWSMAKERS

New E.U. committee head

Following the E.U. elections in May, the European Parliament's research committee has a new chair: **Jerzy Buzek**, a conservative politician from Poland and a former chemical engineering professor. Buzek presided over the European Parliament from July 2009 to January 2012 and was prime minister in his country from 1997 to 2001. As chair of the Industry, Research and Energy (ITRE) Committee, Buzek will have a strong voice in the union's science and innovation policies. Now that Horizon 2020, the bloc's 7-year science funding program, is under way, energy policy—one of Buzek's pet subjects—will be higher on ITRE's agenda, says Jerzy Langer, a Polish physicist who served as deputy science minister in 2005. Langer praises Buzek as an enthusiastic "gentleman professor" and a good listener with a knack for consensus. Buzek was elected by acclamation last week, meaning that he was the only candidate for the post. The committee also decided that after 2.5 years, former E.U. budget commissioner Janusz Lewandowski will take over from Buzek, who's 74. http://scim.ag/JerzyBuzek

C. yangi's tail is the longest of any flying dinosaur.

Flying dino had long, feathery tail

Please welcome the latest member of the growing club of flying dinosaurs, *Changyuraptor yangi*, pictured here in an artist's reconstruction. This latest specimen, found in 125-million-year-old sediments in northeastern China, was about 1.2 meters long and is related to a noted group of flying dinosaurs called *Microraptor*, which has provided important insights into the evolution of powered flight. Like *Microraptor*, the new specimen had feathers on all four limbs; but its feathery tail, which takes up about 30% of its total length, is the longest known among flying dinosaurs. *Changyuraptor,* described this week in *Nature Communications*, weighed 4 kilograms, making it among the heaviest flying dinosaurs known. As for its long tail, the dino probably used it to slow itself down when descending, thus avoiding crash landings.

CREDITS: (TOP TO BOTTOM) EUROPEAN PARLIAMENT; STEPHANIE ABRAMOWICZ/DINOSAUR INSTITUTE/NHM

Published by AAAS

IN DEPTH

LAB SAFETY

Alarm over biosafety blunders

Lapses stoke debate about flu virus research

By **Jon Cohen**

CDC microbiologist Thomas Stevens showers before leaving a BSL-4 lab.

PHOTO: CDC/SCOTT SMITH

Angry. Disappointed. Astonished. Those are the words Tom Frieden, head of the U.S. Centers for Disease Control and Prevention (CDC), chose to describe his reaction to three separate, extraordinary mishaps with potentially deadly "select agents" that occurred in U.S. government labs over the past few months. "Events like this should never happen, and that's why I will do everything in my power to make sure that nothing like this happens again," he said at an 11 July press conference at CDC headquarters in Atlanta. That includes temporarily closing two CDC labs and halting the shipment of the most dangerous pathogens.

No one was harmed in the three security breaches, two at CDC with anthrax and a particularly dangerous strain of influenza, and a third at the U.S. National Institutes of Health (NIH) in Bethesda, Maryland, where overlooked vials of smallpox that should have been destroyed decades ago were discovered. But the blunders bruised public faith in safety procedures at government labs, triggered inquiries from Congress, and have focused renewed scrutiny on controversial "gain-of-function" experiments that manipulate dangerous flu strains to better understand which factors might make them transmit more readily in humans.

"I don't know the last time I've seen a CDC director express anger and concern over a public health lab issue," says bioterrorism expert Michael Osterholm, director of the Center for Infectious Disease Research and Policy at the University of Minnesota, Twin Cities. "This is a very dark cloud over lab science in general. We've been so busy trying to keep the bad guys out of the laboratory that we forget that the good guys sometimes don't do it carefully."

In what Frieden called "the most distressing of the three incidents," a CDC lab that specializes in influenza sent what it thought was a relatively nonpathogenic avian strain to the U.S. Department of Agriculture. After researchers noticed the sample was behaving oddly, they discovered that it had been contaminated with H5N1, a bird flu virus that readily kills humans. Frieden said he was troubled that CDC's "superb laboratory" could make such a mistake and that higher-ups weren't notified for 6 weeks.

The second breach, disclosed in June, took place at CDC's Bioterrorism Rapid Response and Advanced Technology laboratory, which was testing a method to detect anthrax. As CDC described in a report released 11 July, scientists working in a biosafety level 3 (BSL-3) lab that requires them to wear spacesuitlike gear failed to properly inactivate anthrax samples before shipping them to lower security CDC labs. Sixty-seven CDC workers were potentially exposed.

Thomas Inglesby, director of the UPMC Center for Health Security in Baltimore, Maryland, says the lab errors are especially concerning because they occurred at CDC, which he says has some of the most highly qualified experts in the world. "It should make us all mindful that if we see these problems at CDC, which is quite transparent in its ways, we should presume these events are happening elsewhere."

Frieden closed both the bioterrorism and influenza labs involved and imposed a moratorium on shipment of any biological material from CDC's BSL-3 or even more secure BSL-4 labs to outside researchers. A CDC spokesman told *Science* it ships an average of 200 samples of infectious agents each month, noting that some are inactivated and that it does not know how many come from BSL-3 and -4 labs. Anthony Fauci, head of the U.S. National Institute of Allergy and Infectious Diseases in Bethesda, says the researchers who work with these dangerous pathogens "are an extremely small cohort of the group we fund." He strongly supports the shipping moratorium. "We have to call a time-out and make sure we get these things

Downloaded from www.sciencemag.org on August 4, 2014

Published by AAAS

right," Fauci says.

Daniel Perez, a researcher at the University of Maryland, College Park, who works with highly pathogenic influenza strains, says the moratorium could have an impact on his work if "CDC takes too long on deciding on additional policies on transfer of materials." Perez thinks existing policies are adequate but must be strictly followed by properly trained and supervised lab workers.

The third incident involved six vials of smallpox found on 1 July in a Food and Drug Administration refrigerated storage room at NIH. The vials, dated 10 February 1954—more than 2 decades before smallpox was eradicated—were in what amounts to a walk-in refrigerator. On 7 July, FBI and local police helped CDC staffers move the samples from Bethesda to a government plane and fly them to a CDC BSL-4 lab in Atlanta. CDC is one of only two institutions in the world authorized to hold the smallpox virus. At the press conference, Frieden announced that two of the six vials of freeze-dried virus tested so far contained live virus. After completing the testing and doing genomic sequencing, CDC will destroy the smallpox viruses and related material. Fauci said NIH did an extensive inventory of all of its freezers for select agents in 2012, but it did not inspect its cold rooms because they're typically not used to store dangerous pathogens. Now, NIH will conduct that inventory. "This sounds to me like an unusual fluke," Fauci says.

Frieden announced at the press conference that CDC would like to reduce the number of BSL-3 and -4 labs around the country and, in turn, the number of people who work with dangerous pathogens "to the absolute minimum necessary." According to a Government Accountability Office report released last year, there were 415 BSL-3 and -4 labs registered with the Federal Select Agent Program in 2004 and 1495 in 2010.

The lapses are stoking the debate about gain-of-function experiments. Inglesby is one of 17 prominent researchers who have banded together as the "Cambridge Working Group" and issued a statement on 14 July that calls for curtailing all such research until the risks have been more credibly assessed. "Discussions that have gone on in the last couple of years about laboratory safety in some cases have been so assuring that you might have the impression that lab accidents are in the realm of the impossible," Inglesby says. "That's just not correct. No matter where you are in the debates about gain-of-function research, everyone has to start with the assumption that we have human systems and they are fallible. There's no such thing as perfect systems." ■

With reporting by Jocelyn Kaiser.

NSF's George Hazelrigg with a pile of grant applications.

PHOTO: PAUL FETTERS

RESEARCH GRANTS

A radical change in peer review

Each NSF applicant evaluated seven competing proposals

By **Jeffrey Mervis**

Letting a scientist who is applying for a federal grant also serve as a reviewer of proposals from competitors has always been a no-no. But the National Science Foundation (NSF) recently crossed that line, running an experiment that not only allowed more than 100 applicants to double as reviewers, but also required them to assess seven competing proposals as the price for submitting their own. NSF officials say the pilot, aimed at easing the strain on a merit review system stretched to the breaking point, saved time and money and may even have improved the quality of reviews.

NSF used some 36,500 researchers last year as volunteer reviewers of 49,000 proposals. The number of proposals is up 53% from 2001, but because NSF's budget hasn't kept pace, success rates fell from 31% to 22% over the same period. Those trends have increased the cost of peer review and caused some scientists to gripe about wasting their time reviewing good ideas that have little chance of being funded.

Rather than simply wring his hands over the problem, longtime NSF program manager George Hazelrigg went looking for an alternative that would distribute the workload more equitably and provide reviewers with a positive incentive. He found inspiration in a 2009 paper by astronomer Michael Merrifield of the University of Nottingham in the United Kingdom and mathematician and voting system specialist Donald Saari of the University of California, Irvine. The paper proposed using applicants as reviewers to tackle the flood of applications for observing time on instruments operated by the European Southern Observatory (ESO).

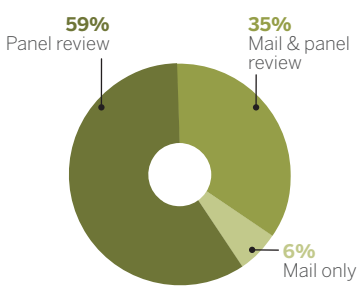
In NSF's version of what Merrifield calls "our radical alternative," applicants assessed the quality of seven proposals and then ranked the proposals from best to worst. NSF also employed a novel scoring system aimed at dissuading a reviewer from downgrading a competitor's proposal in order to boost their own; reviewers earned bonus points on their own applications if their assessments of other proposals closely matched what their colleagues thought.

The agency announced the pilot in May 2013 in a letter to those likely to seek funding from NSF's program on sensors and sensing systems within its engineering directorate. Participation was mandatory for the October cycle, it said, although scientists who didn't like the idea could wait for the next deadline, in mid-February. To Hazelrigg's surprise and delight, NSF received 131 applications, some 50 more than the norm for a fall cycle.

"I understood that I was agreeing to do

Panels rule

In 2013, 36,475 experts conducted 187,512 reviews for NSF



Source: NSF

Published by AAAS

a bunch of reviews, but that didn't affect my decision to apply," says Rolf Mueller, a bioengineer at Virginia Polytechnic Institute and State University in Blacksburg. "I had what I thought was a pretty good idea." Even so, he says "it was interesting to see some of the other proposals." (Mueller was right: He won a $360,000 grant to explore how a bat's biosonar system might help improve humanmade sensors.)

NSF says there is some evidence the experiment led to "more comprehensive reviews." Each proposal received seven reviews rather than the normal three or four, Hazelrigg notes. "And each review had, on average, 40% more words." (Only one entrant was disqualified, for failing to meet the 6-week deadline for submitting the reviews.)

The system saved time and money, says Hazelrigg, the acting director of the NSF's Civil, Mechanical and Manufacturing Innovation (CMMI) Division. It eliminated the need to spend 2 or 3 weeks winnowing a list of some 400 potential reviewers to the 16 to 20 needed for a panel. And mail reviews meant NSF didn't have to pay travel and per diem expenses to bring those reviewers to NSF headquarters in Arlington, Virginia. "Our division runs 200 panels a year," he says, "so that's a big cost savings."

The pilot has some potential downsides. One NSF program manager who asked to remain anonymous worries that a bonus system based on group consensus could discourage innovative ideas. "It rewards people for playing it safe," the program manager says. Hazelrigg plays down that possibility, however, noting that a program manager can recommend a proposal for funding even if reviewers raise doubts.

NSF officials are still evaluating whether to expand the CMMI pilot, one of seven experiments to tweak merit review that the agency conducted last year. In the meantime, one group of astronomers has decided to run its own version.

Markus Kissler-Patig was an ESO scientist when Merrifield and Saari wrote their 2009 paper. And although ESO didn't change its procedures, Kissler-Patig resurrected the idea when he became director of the Gemini Observatory, a consortium that operates twin 8-meter telescopes in Hawaii and Chile. After much discussion, the observatory will begin using applicants as reviewers next January, to allocate 10% of the viewing time on Gemini North.

"If it works, we can expand it to Gemini South," says Rachel Mason, a staff astronomer in Hawaii who is coordinating the experiment, called Fast-Turnaround. Without such changes, she says, "the existing system is simply going to break down." ■

PHOTO: 68/OCEAN/CORBIS

BIOMEDICINE

Proteins and a pregnancy woe

Surprising insight into preeclampsia may offer early warning for at-risk women

By **Nadia Whitehead**

Every obstetrician dreads it. Without warning, the blood pressure of a seemingly healthy pregnant woman surges; if the baby isn't delivered right away, she may die. This is preeclampsia, the leading cause of death in pregnant women and a medical mystery. Dubbed the disease of theories, no one knows its cause or how to prevent it. Now, a team of researchers and clinicians offers an entirely new take on this bewildering disorder: Preeclampsia, like Alzheimer's, Parkinson's, and mad cow diseases, is marked by misfolded and clumped proteins.

Doctors monitor the blood pressure of pregnant women, which can surge without warning.

Whether these bunched proteins actually cause the disorder is still unclear. But the proteins, reported this week in *Science Translational Medicine*, could open the way to a simple urine test to definitively diagnose preeclampsia before symptoms worsen. "The study is just terrific," says Daniel Kahn, a research obstetrician at the University of California, Los Angeles, who was not part of the team. "Their discovery can easily be used [as a diagnostic tool] in low-resource settings, which is just tremendous because this can be applied worldwide."

Preeclampsia affects about 5% to 10% of pregnancies and kills an estimated 76,000 women every year worldwide. Potential causes include abnormalities in immune tolerance and placental development, or even diet. The only way to prevent the mother's downward spiral into the seizures and hemorrhaging of eclampsia is to deliver the baby prematurely, risking its survival, especially in developing countries. Kahn says an early warning could allow doctors to send women developing preeclampsia to hospitals equipped to deal with preemies, saving lives.

Doctors routinely check pregnant women for symptoms of preeclampsia: high blood pressure and elevated levels of protein in their urine. But not every woman has these signs, and preexisting conditions can mask the warnings. In search of a better diagnostic tool, Irina Buhimschi, a researcher at Nationwide Children's Hospital and Ohio State University, Columbus, asked: "Could we diagnose preeclampsia by looking at the 'fingerprint' of proteins in women with this disorder?"

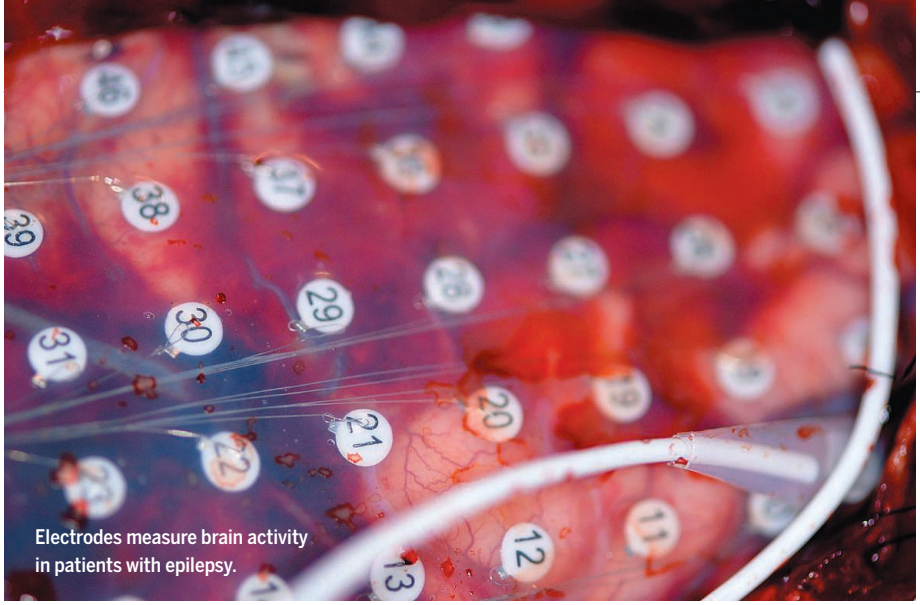
The team collected urine from more than 600 pregnant women and analyzed the samples with Congo red, a dye that binds to large clumps of protein but washes off smaller molecules. Samples stayed straw-colored in women with normal pregnancies but turned telltale red in those with suspected preeclampsia. The amount of clumped protein, as indicated by the intensity of the stain, predicted with an accuracy of 80% or higher which women would require early delivery.

Correctly folded proteins don't clump this way, Buhimschi explains, so the aggregations signal improper folding. The team found that five specific proteins repeatedly bunched together. To their surprise, one was amyloid precursor protein, notorious for its role in Alzheimer's, where it is suspected to cause dementia.

Fetal hemoglobin, known to be excreted by preeclamptic women, could be inflating the estimates of clumped proteins in some tests, cautions research obstetrician Stefan Hansson of Lund University in Sweden. He and Kahn aren't ready to label preeclampsia as a protein misfolding disease, saying the clumped proteins may be a consequence rather than a cause of the disease.

Although Buhimschi isn't sure how bunched proteins might cause the disease, she thinks they may offer new lines of treatment, perhaps by using Alzheimer's drugs that help prevent misfolding. "These diseases may look different, and have different symptoms," she says, but "the mechanism is common." ■

Published by AAAS



Electrodes measure brain activity in patients with epilepsy.

PHOTO: GORDON BALTUCH

NEUROSCIENCE

Researchers aim for an electrical memory prosthesis

DARPA-funded effort to develop implants for traumatic brain injuries draws skepticism

By **Emily Underwood**

Last fall, Geoffrey Ling, a top biotechnology research official at the Defense Advanced Research Projects Agency (DARPA), challenged neuroscientists to do something extraordinary: Develop an implantable device that can reverse memory loss in veterans with traumatic brain injuries. Dangling up to $40 million in funding, Ling said: "Here's the golden ring—who's brave enough to step up and actually grab it?"

Last week, DARPA announced that two academic teams have made the lunge. Both teams, at the University of California, Los Angeles (UCLA), and the University of Pennsylvania (Penn), aim to develop electrical prostheses that will tickle brain regions critical to memory. The U.S. Department of Energy's Lawrence Livermore National Laboratory in California and device manufacturer Medtronic will also participate, aiming to build neurostimulators at least 10 times smaller than previous devices. In spite of the scale of the efforts, however—UCLA's is funded at up to $15 million over 4 years and Penn's at up to $22.5 million over the same period—some researchers are skeptical about their prospects. Scientifically, "it's a bit early" for the approach DARPA has chosen, says neuropsychologist James Sumowski of the Kessler Foundation in West Orange, New Jersey.

Both teams will start by studying neuronal activity in people with epilepsy who are undergoing brain surgery to remove the focus of the seizures. In some patients, surgeons must blanket the brain with hundreds of electrodes and wait weeks for a seizure to occur before determining where to operate, says neuroscientist Michael Kahana, head of the Penn team. That gives researchers a chance to map "in exquisite detail" the electrophysiology of memory storage and retrieval, he says. The electrical recordings should also offer insights into memory impairment, which many epileptics experience.

By searching for electrical "biomarkers" of both normal and impaired memory retrieval and storage, Kahana hopes to build computer programs that can detect when memory goes awry and instruct a device to help repair it. Kahana's group has already detected signature patterns of brain activity associated with the proper encoding, or storage, of new memories and will soon publish evidence of the first signatures of memory retrieval, he says. "The early results are highly promising."

The UCLA group, led by neurosurgeon Itzhak Fried, will use a different method to develop a memory formation model. Fried's earlier work in people with epilepsy has shown that stimulating a brain region called the entorhinal cortex improves performance on a computer game that requires players to quickly learn and remember where to drop off taxi passengers in a virtual city. Now, Fried will use data from such studies to build simulations of how the entorhinal cortex and the hippocampus, another region crucial to memory, work together to convert daily experiences into lasting memories.

If these efforts succeed, the researchers ultimately hope to conduct the first clinical trials of devices that produce deep brain stimulation in people with traumatic brain injury. But relying too heavily on epilepsy as a model for memory loss in traumatic brain injury could prove problematic, Sumowski says. Most people with those injuries don't show the same patterns of abnormal electrical activity or areas of atrophy seen in epilepsy, he says. "They are very different disorders."

The success of the projects will also depend greatly on exactly what kind of memory loss brain injury sufferers actually have, predicts Roger Redondo, a neuroscientist at the Massachusetts Institute of Technology in Cambridge. If a memory is there, but difficult to access or retrieve, stimulation might be useful. But if the connections that form a memory never developed or were destroyed, "no implantable device is going to help," he says.

Determining which cells contain the memory and precisely tuning electrical stimulation to drive its retrieval "is going to be extremely hard," Redondo adds. "The complexity of the brain, and the hippocampus, is such that any change in voltage that a microelectrode or chip can apply, even in a tiny area, will affect multitudes of neurons in uncontrolled ways."

Kahana is encouraged by evidence from disorders such as Parkinson's disease and obsessive-compulsive disorder, which can be effectively treated with electrical stimulation even though researchers don't fully understand how zapping the brain affects circuits on a cellular level. The memory biomarkers from epileptics will serve as "beacons" for physicians attempting to coax injured brains back to normal, he says.

Such challenges define the kind of high-risk projects that DARPA's Restoring Active Memory (RAM) program is supposed to tackle, says Justin Sanchez, RAM's project manager. (RAM is part of the federal government's roughly $110 million BRAIN, or Brain Research through Advancing Innovative Neurotechnologies, Initiative.) Given the 270,000 veterans of the Iraq and Afghanistan wars who have diagnosed traumatic brain injuries, nothing less than a major scientific and technological leap is required, he says. "We're going to take an entirely new strategy compared to what is currently available out there." ■

Downloaded from www.sciencemag.org on August 4, 2014

Published by AAAS

AGRICULTURE

Harvest of genome data for wheat growers

DNA survey orders genes, edging closer to the deciphering of this complex crop

By **Elizabeth Pennisi**

Geneticists have sequenced obscure fishes, extinct beasts, and many, many microbes, not to mention crops including corn and rice. But for a long time, the genome of the world's most widely grown cereal, wheat, remained mostly uncharted, thought to be too complex for existing sequencing approaches.

In 2005, wheat farmers in Kansas, which grows 18% of the U.S. wheat crop, decided to defy the conventional wisdom, believing that a genome on par with that of rice was necessary to breed the best varieties of their crop. This week, the Kansas Wheat Commission's efforts bore fruit. On pages 286 and 287, the group it helped spawn—the International Wheat Genome Sequencing Consortium (IWGSC)—describes a coarse "survey" sequence, covering about 61% of the genome and likely more than 95% of the genes, of a strain of bread wheat, and a more detailed sequence of one chromosome. "There were people who said it couldn't be done," recalls IWGSC Executive Director Kellye Eversole.

These new data are "a resource that we've been eagerly waiting for," says James Anderson, a wheat breeder at the University of Minnesota, Twin Cities. They should allow him and his colleagues to greatly speed up the breeding of new strains of wheat, more productive than today's and more resistant to drought and pests.

The sequencing effort was indeed daunting. The "big three" cereal crops span the range of genome complexity and size. Rice, which in 2002 became one of the first plant genomes deciphered, is relatively simple. Its genome is diploid, meaning there are just two copies of each chromosome. Maize DNA—sequenced in 2009—at one point had four copies of each chromosome, but has evolved into an ostensibly diploid genome with a lot of duplication. Wheat is another matter. It essentially contains three very distinct genomes, called subgenomes A, B, and D, all very similar. (C belongs to another wheat species.) Each has seven chromosomes, for a total of 21.

That's a challenge for the now-common whole genome shotgun approach to sequencing, in which the wheat's DNA would be broken into small fragments, which would be read individually. A computer program would then piece the resulting short sequences together into the full genome. But with so many virtually identical chromosomes, riddled with repetitive DNA, assembling the different subgenomes correctly would be virtually impossible.

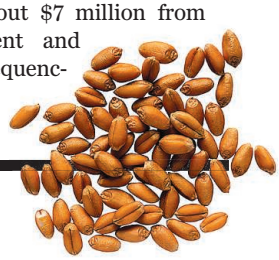
So IWGSC first separated the chromosomes and did a shotgun sequence for each one on its own. They then combined the information into the survey genome. At the same time, Catherine Feuillet, a consortium co-chair who is now a geneticist at Bayer CropScience in Morrisville, North Carolina, and her colleagues took a more detailed look at wheat's largest chromosome, 3B. They split it into 100,000-base fragments and placed each into a bacterium that proceeded to replicate, making many copies of the so-called bacterial artificial chromosomes (BACs) containing the wheat DNA. By sequencing thousands of copies of each BAC, they generated a highly accurate read of each 3B fragment despite wheat's repetitive DNA. Deciphering these BACs one by one revealed the location of genes on chromosome 3B much more precisely than the shotgun survey did for the rest of the genome.

The in-depth sequence reveals that chromosome 3B is divided into distinct regions. Other studies of the DNA showed that genes closer to each chromosome end tend to be expressed only under specific conditions, such as during drought stress or early in development, whereas genes more broadly active are closer to the center. The BAC-by-BAC approach also revealed that many genes exist in multiple copies within the chromosome, duplicates the shotgun strategy would have often missed, Feuillet says.

Already, research teams within Norway and Japan are completing two more wheat chromosomes using BACs, and an international coordinating body called the Wheat Initiative has come out in favor of that same approach for the rest of the genome. What's still missing, Eversole says, is the estimated $19 million needed to finish the job. Tackling 3B alone cost $6.8 million, but sequencing costs have recently dropped dramatically. (The total amount spent so far by IWGSC is about $68.5 million, Eversole notes.)

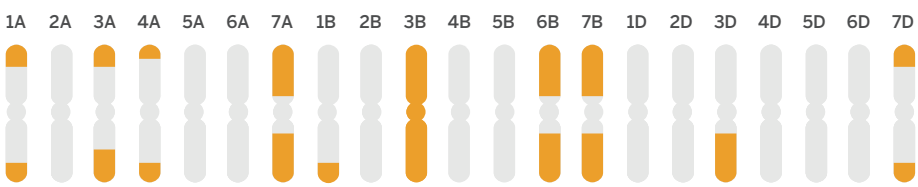
Meanwhile, an independent group that published a gene inventory of wheat 2 years ago believes it has a quicker and cheaper approach. The group, now led by Mario Caccamo, director of the Genome Analysis Centre in Norwich, U.K., and Michael Bevan, a plant geneticist at the John Innes Centre in Norwich, has about $7 million from the U.K. government and plans to use a new sequencing technology that reads very long DNA molecules. That, along with improvements in the computer programs that stitch together sequence information, might make it possible to produce a high-quality wheat genome through the shotgun method.

"The question is whether breeders need this [detailed] information," says Alan Schulman, a molecular biologist at MTT Agrifood Research Finland in Helsinki. Even the survey genome published this week will be a boon for breeders, who have become increasingly reliant on assessing DNA markers in seedlings to quickly tell if they have a desired trait, rather than waiting for the plants to grow up. The survey sequence provides many more markers and a better sense of what genes are close to those markers, enabling wheat breeders to be more precise in their efforts. But, Anderson notes, "you really need that complete sequence" for pinning down a trait's exact gene. And that could help breeders make wheat, the staff of life for a third of the world, a little sturdier. ■



Wheat chromosomes: One down, 20 more to go

3B is fully sequenced; there's progress (orange) on several others



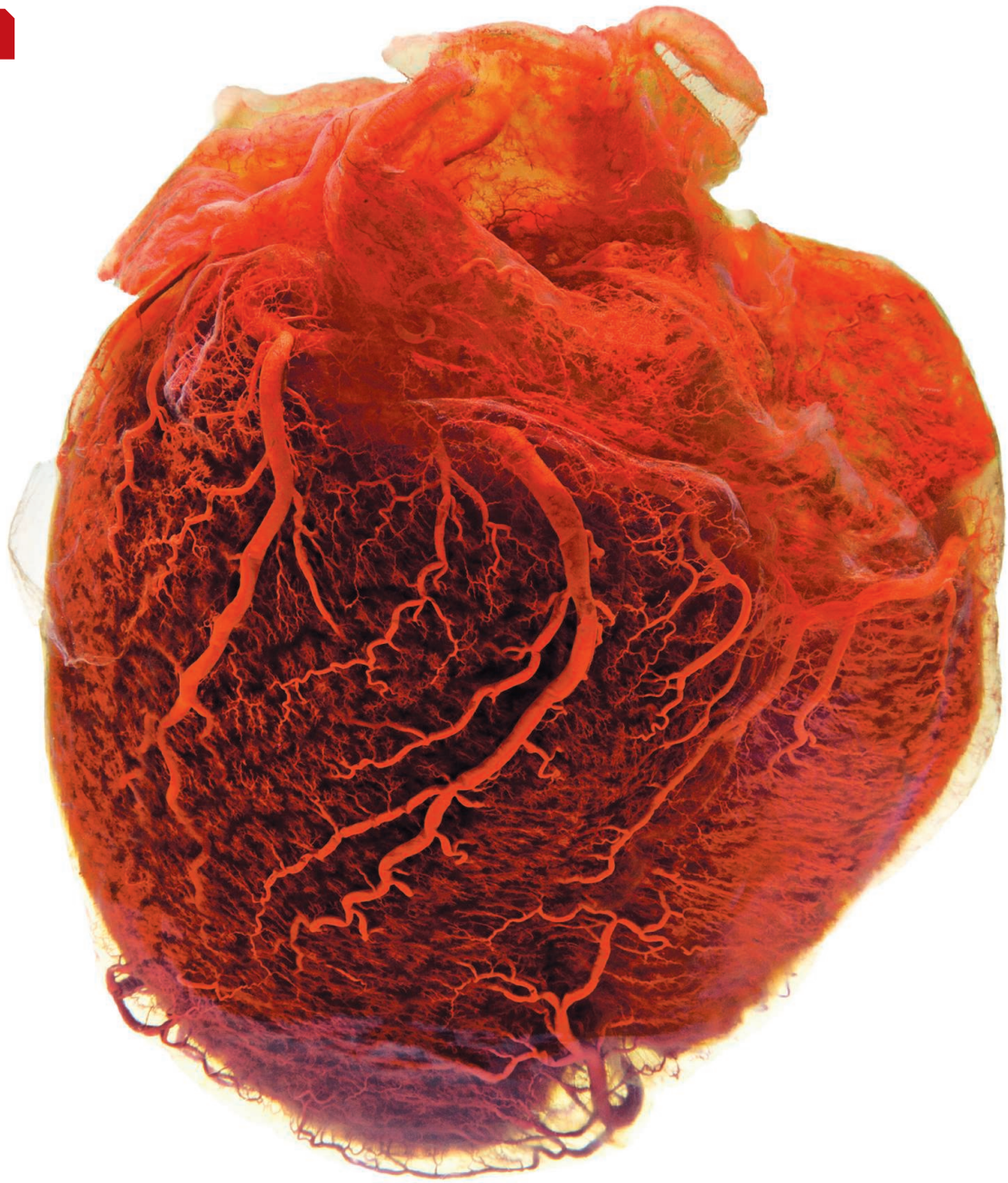
Source: IWGSC

PHOTO: (INSET) 2/SIEDE PREIS/OCEAN/CORBIS

Downloaded from www.sciencemag.org on August 4, 2014

Published by AAAS

FEATURES



The elusive heart fix

After years of hope and hype, researchers still don't agree whether cell infusions rejuvenate the human heart

By **Jennifer Couzin-Frankel**

PHOTO: ROBERT CLARK/INSTITUTE

Downloaded from www.sciencemag.org on August 4, 2014

Published by AAAS

In April, a maverick British cardiologist rocked his field with an assault on one of its cherished hopes. He suggested that a strategy for rejuvenating a damaged heart with infusions of cells, which has tantalized doctors and patients for more than a decade, does not work.

Darrel Francis treats patients with heart disease at Imperial College London, but he also has a penchant for debunking medical practices, both the accepted and the experimental. (Earlier this year, he suggested that the European recommendation to start β-blocker drugs during certain surgeries to prevent heart attacks was killing more people than it saved.) His interest in cell therapy was piqued several years ago, after a patient traveled to Germany for a cardiac cell therapy trial. Francis later came across a published study of the technique, in which cells drawn from a patient's own bone marrow are infused into the heart. The study described the treatment as a success—but what struck Francis were the differences between control patients and those getting treatment. Those differences, he thought, could easily explain why recipients of cell therapy seemed to fare better. With a nose for statistics and an appetite for controversy, Francis began digging.

He and his colleagues combed through 133 reports on 49 clinical trials of cardiac cell therapy—specifically, a type that uses a mix of a person's own bone marrow cells. What he found was enough to delight any rabble-rouser. All but five trials contained what Francis labeled "discrepancies," from minor errors in tables to dead patients listed as still taking drugs. And the more discrepancies a trial had, the more effective the therapy appeared to be. His disquieting conclusion: Knowingly or not, some cardiologists were massaging their data to lend them a veneer of efficacy.

"In clinical practice, we think we are immune to bias and wishful thinking," Francis says. "If somebody goes for a therapeutic maneuver and it works, they end up getting a Nobel Prize. … If it doesn't work, you feel that you can walk away. But in reality, you can't." With one's career built around a potentially lifesaving therapy, Francis says, acknowledging failure can be nearly impossible.

Published in *BMJ*, the study is the latest salvo directed at a promising field marked by hope and hype. Since the first heart patients received experimental cell therapy 13 years ago, controversy has come in waves, and it's now at another peak. Trials built on uncertain science have led to fights about why some can't be replicated, charges that the animal data supporting trials are too skimpy, accusations of sloppiness, and worse.

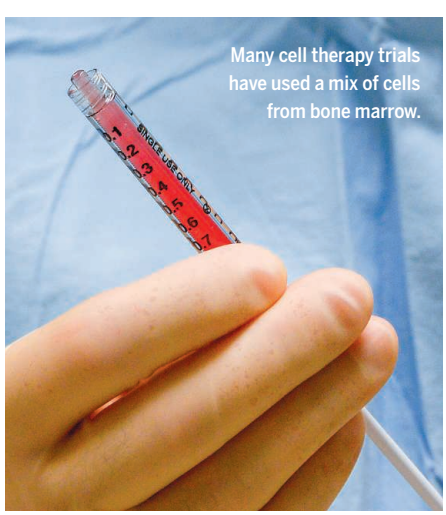
"I have a long list of things that have troubled me as I've watched the field unfold," says Jonathan Kimmelman, an ethicist at McGill University in Montreal, Canada, who has studied cell and gene therapies. "There is so much pressure to demonstrate a major treatment effect that people are not always applying careful, painstaking methodologies."

Meanwhile, thousands of people have signed up to participate in clinical trials of different forms of cardiac cell therapy. There are dozens of trials to choose from, led by clinicians eager to save patients—and to make history. Says Charles Murry, a cardiovascular pathologist at the University of Washington, Seattle: "Everybody has wanted to be the first to regenerate the heart."

> ***"Everybody has wanted to be the first to regenerate the heart."***
>
> **Charles Murry,** University of Washington, Seattle

THE SCIENTIFIC BACKSTORY BEGINS, as it often does, in mice. In 2001, a team led by Piero Anversa, now at Harvard University, and Donald Orlic of the National Institutes of Health (NIH) described injecting bone marrow cells into rodents that had suffered induced heart attacks. Within 9 days, they reported, the cells were turning into new heart muscle.

In a second study, Anversa examined hearts from people who had died of heart attacks and found that some organ muscle could regenerate. The results challenged the dogma that the heart cannot rejuvenate after losing muscle during a heart attack.



Many cell therapy trials have used a mix of cells from bone marrow.

PHOTO: TEXAS HEART INSTITUTE

"This is a breakthrough," a former president of the American Heart Association (AHA) told *The New York Times*.

Cardiac patients were in dire need of breakthroughs. Thanks to new treatments, many were surviving massive heart attacks that would have killed them several years earlier. But their hearts, temporarily deprived of oxygen, were profoundly damaged. They had heart failure, and many weren't candidates for transplants or wouldn't climb atop the waiting list in time. "They can't exercise, they can't move, they can't climb stairs," says Joshua Hare, a cardiologist at the University of Miami in Florida. "Their minute-to-minute existence is symptomatic."

Anversa's experiments, which suggested bone marrow could become heart muscle and that the heart had impressive regenerative power, seemed, to some, to hold the answer. The findings came against a backdrop of other tantalizing research in adult stem cells: reports that they could morph into lung, neurons, and other tissues. Scientists quickly moved to inject bone marrow into the human heart.

Testing began in 2001. Cardiologists initially infused an imprecise mix of cells drawn from bone marrow: mostly mature white blood cells and their progenitors, but also a very small proportion of adult stem cells. For example, about one in 10,000 cells in this mix are hematopoietic stem cells, which give rise to all the blood cells in the body. Reassuring to everyone was bone marrow's long history in medicine: Since the 1970s, doctors have been offering bone marrow transplants to cancer sufferers. Because of that, cardiologists and drug regulators were hopeful that bone marrow would be safe in heart patients.

As cell therapy shifted from mice to ailing humans, uncertainties arose. Doubts were emerging about whether adult stem cells were as versatile as thought. And in 2004, three groups reported independently that they could not replicate Anversa and Orlic's mouse findings. But by then, "the clinical trials were out of the barn," says Jonathan Epstein, a cardiologist at the University of Pennsylvania.

Murry, who published one of those three papers, says he was struck by how little effect the negative results had on a field that was gathering momentum. Trials based on now-unreplicable research quickly signed on volunteers. It was "like, mechanism, schmechanism," he says. No one seemed to care what exactly bone marrow cells were doing in the heart.

Many physicians argued that although biology mattered, so did saving people. And the very first trials, led by U.S. and European

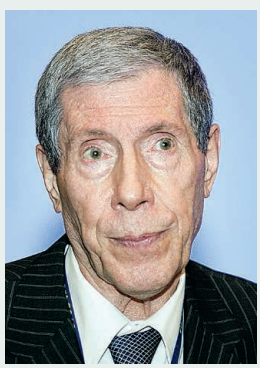
Published by AAAS

Top heart lab comes under fire

By **Kelly Servick**

In the study of cardiac regeneration, Piero Anversa is among the heavy hitters. His research into the heart's repair mechanisms helped kick-start the field of cardiac cell therapy (see main story, p. 252). After more than 4 decades of research and 350 papers, he heads a lab at Harvard Medical School's Brigham and Women's Hospital (BWH) in Boston that has more than $6 million in active grant funding from the National Institutes of Health (NIH). He is also an outspoken voice in a field full of disagreement.

So when an ongoing BWH investigation of the lab came to light earlier this year, Anversa's colleagues were transfixed. "Reactions in the field run the gamut from disbelief to vindication," says Mark Sussman, a cardiovascular researcher at San Diego State University in California who has collaborated with Anversa. By Sussman's account, Anversa's reputation for "pushing the envelope" and "challenging existing dogma" has generated some criticism. Others, however, say that the disputes run deeper—to doubts about a cell therapy his lab has developed and about the group's scientific integrity. Anversa told *Science* he was unable to comment during the investigation.



Piero Anversa

"People are talking about this all the time—at every scientific meeting I go to," says Charles Murry, a cardiovascular pathologist at the University of Washington, Seattle. "It's of grave concern to people in the field, but it's been frustrating," because no information is available about BWH's investigation. BWH would not comment for this article, other than to say that it addresses concerns about its researchers confidentially.

In April, however, the journal *Circulation* agreed to Harvard's request to retract a 2012 paper on which Anversa is a corresponding author, citing "compromised" data. *The Lancet* also issued an "Expression of Concern" about a 2011 paper reporting results from a clinical trial, known as SCIPIO, on which Anversa collaborated. According to a notice from the journal, two supplemental figures are at issue.

For some, Anversa's status has earned him the benefit of the doubt. "Obviously, this is very disconcerting," says Timothy Kamp, a cardiologist at the University of Wisconsin, Madison, but "I would be surprised if it was an implication of a whole career of research."

Throughout that career, Anversa has argued that the heart is a prolific, lifelong factory for new muscle cells. Most now accept the view that the adult heart can regenerate muscle, but many have sparred with Anversa over his high estimates for the rate of this turnover, which he maintained in the retracted *Circulation* paper.

Anversa's group also pioneered a method of separating cells with potential regenerative abilities from other cardiac tissue based on the presence of a protein called c-kit. After publishing evidence that these cardiac c-kit$^+$ cells spur new muscle growth in rodent hearts, the group collaborated in the SCIPIO trial to inject them into patients with heart failure. In *The Lancet*, the scientists reported that the therapy was safe and showed modest ability to strengthen the heart—evidence that many found intriguing and provocative. Roberto Bolli, the cardiologist whose group at the University of Louisville in Kentucky ran the SCIPIO trial, plans to test c-kit$^+$ cells in further clinical trials as part of the NIH-funded Cardiovascular Cell Therapy Research Network.

But others have been unable to reproduce the dramatic effects Anversa saw in animals, and some have questioned whether these cells really have stem cell–like properties. In May, a group led by Jeffery Molkentin, a molecular biologist at Cincinnati Children's Hospital Medical Center in Ohio, published a paper in *Nature* tracing the genetic lineage of c-kit$^+$ cells that reside in the heart. He concluded that although they did make new muscle cells, the number is "astonishingly low" and likely not enough to contribute to the repair of damaged hearts. Still, Molkentin says that he "believe[s] in their therapeutic potential" and that he and Anversa have discussed collaborating.

Now, an anonymous blogger claims that problems in the Anversa lab go beyond controversial findings. In a letter published on the blog Retraction Watch on 30 May, a former research fellow in the Anversa lab described a lab culture focused on protecting the c-kit$^+$ cell hypothesis: "[A]ll data that did not point to the 'truth' of the hypothesis were considered wrong," the person wrote. But another former lab member offers a different perspective. "I had a great experience," says Federica Limana, a cardiovascular disease researcher at IRCCS San Raffaele Pisana in Rome who spent 2 years of her Ph.D. work with the group in 1999 and 2000, as it was beginning to investigate c-kit$^+$ cells. "In that period, there was no such pressure" to produce any particular result, she says.

Accusations about the lab's integrity, combined with continued silence from BWH, are deeply troubling for scientists who have staked their research on theories that Anversa helped pioneer. Some have criticized BWH for requesting retractions in the midst of an investigation. "Scientific reputations and careers hang in the balance," Sussman says, "so everyone should wait until all facts are clearly and fully disclosed." ■

researchers, suggested that injecting a patient's bone marrow into his or her heart might somehow enhance its ability to pump, even if the cells did not proliferate. Why that appeared to be happening could be sorted out later.

One of the first trials, on 21 people with severe heart failure in Brazil, was led by researchers at the Texas Heart Institute. In the 14 patients treated (seven others were controls), they described an improvement of left-ventricular ejection fraction—the heart's ability to pump blood—from 20% to 29%. A healthy ejection fraction is 55% to 70%. Still, the improvement was meaningful and generated enthusiasm for further testing.

A year later, the Texas Heart Institute announced with much fanfare that the treatment was coming to the United States. "A procedure that for more than a year has been bringing South American heart failure patients back from near-death has been approved for testing in the Texas Medical Center," a press release declared. The release quoted one of the trial leaders effusing about the treatment: "These people in Brazil were reborn," the cardiologist said.

PHOTO: MARIUS GEANTA/*HIPOCRATE JOURNAL ROMANIA*

Published by AAAS

At research centers worldwide such as the Texas Heart Institute, physicians have been testing cardiac cell therapy for years.

PHOTO: TEXAS HEART INSTITUTE

Cardiologists acknowledge that cell therapy was marked by outsize expectations. "We all hype our work," says Steven Houser, a cardiac muscle biologist at Temple University in Philadelphia, Pennsylvania. "We want to tell people our work is important. These patients, many of them coming to enroll in these trials, they have no other hope." And, Houser suggests, in this case "the hype got ahead of the science."

Across the Atlantic there was similar exuberance. German cardiologist Bodo-Eckehard Strauer treated 10 patients who had had a heart attack days earlier. That same mix of cells from bone marrow—called mononuclear bone marrow cells—was funneled into their heart. Strauer reported in 2002 that in the group as a whole the weakened area in the heart shrunk markedly.

"We were fascinated by the new knowledge, and the possibility of building up destroyed" heart muscle, says Harald Arnesen, a cardiologist at Oslo University Hospital. The Strauer trial showed the treatment was safe—one important detail that subsequent studies have borne out.

So Arnesen launched his own trial of mononuclear bone marrow cells, containing their small subset of stem cells, with 100 heart attack patients. But the results of this effort were largely negative. It's a pattern that plagues new therapies across medicine: Often, their impressiveness fades as they are tested in ever-larger cohorts.

"In mouse studies there's always dramatic improvement," says Joseph Wu, a cardiologist studying stem cells at Stanford University in Palo Alto, California. "Once you go to a large animal study, it's moderate improvement, once you go to a phase I trial, it's decent improvement, and once you go to phase II, phase III, there's no improvement. This happens again and again and again. … It's the entire field of biological research."

There are many reasons for this. Mice in lab studies are often young and healthy, so they don't reflect the reality of patients. "It's like taking a bunch of 18-year-old humans" and giving them heart attacks, Murry says. "And then we expect it to predict the results of infarcts in 70-year-old smokers with diabetes." Early clinical trials usually rely on surrogate endpoints—biological measures like the heart's pumping capacity—that don't always translate to something more meaningful such as survival or hospitalization rates. Furthermore, scientists publish positive results far more often than negative ones, and that may imply that a novel treatment is more promising than it really is.

The cardiac cell therapy field also has peculiarities of its own, which could explain some of the conflicting outcomes. One is the natural variability between one person's bone marrow and another's. A particular patient may happen to carry more cells that can induce tissue repair than someone else. "It's not like a study where you give 60 milligrams" of a drug "and everyone gets the same thing," says Timothy Henry, a cardiologist at Cedars-Sinai Medical Center in Los Angeles, California. "Everyone does not get the same thing."

Cell therapy is unique, too, because it's not controlled by a single company. This has its advantages, but it also means that securing funds for large, definitive studies can be daunting. Instead, there have been dozens of small- and medium-sized trials, with different designs, different endpoints, different patient populations, and different methods of processing the cells. "Ideally for science and for our patients it would be better to have a concerted effort," Wu says. "Realistically, it's very difficult to do that."

AS THE TRIALS HAVE GROWN MORE AMBITIOUS, the contradictions are increasingly stark. A 204-person study led by Andreas Zeiher at Goethe University in Frankfurt, Germany, a pathbreaker in the field, reported in 2006 that patients receiving cells drawn from their own bone marrow fared better than those who didn't. But two U.S. trials, by an NIH-funded cardiac cell therapy network, and a Swiss trial, with a total of about 400 participants combined, found no benefit from the same therapy. "We are very confident we see an effect" in patients, says Stefanie Dimmeler, a cardiovascular biologist at Goethe University who has collaborated with Zeiher for years. "I still don't understand why other people don't."

Published by AAAS

One possibility, she says, is variations in the patients enrolled and the study protocols.

Cell therapists have also been rocked by two recent scandals. In April, a paper published by Anversa in 2012 on the regenerative power of the human heart was retracted, and another was slapped with an "Expression of Concern." Harvard is investigating the famed cardiologist (see sidebar, p. 254). Another giant in the field, Strauer, was investigated by the University of Düsseldorf, which announced in February that it is launching disciplinary action. Neither university has said whether the scientists' early work is in question.

Meanwhile, the science is in a state of flux. The idea that bone marrow cells morph into heart muscle has been largely abandoned. Believers in the bone marrow mix that's been widely tested, and some other cell types, favor the theory of "paracrine effects," which argues that cells secrete molecules that have healing properties, perhaps enhancing blood vessel growth or building up damaged tissue.

Mouse studies offer some support: hints that a mix of bone marrow cells could be inducing new blood vessels or new heart muscle to grow. But evidence for this in humans is thin. A study presented at the AHA meeting last November, by University of Washington, Seattle, cardiologist April Stempien-Otero, found no new blood vessel growth in the hearts of patients who had received bone marrow cells and subsequently had the organ removed for a heart transplant. In 2011, a study by Matthew Springer at the University of California, San Francisco, suggested an explanation for conflicting mouse and human results: He found that mice that got bone marrow from young donors did much better than those receiving bone marrow from animals that had had an induced heart attack. Springer speculated that a heart attack might somehow change the composition of the victim's bone marrow, the usual source of cell infusions in patients.

One goal now is to identify the specific cells that are the best candidates to rescue an ailing heart. Their identity is the subject of intense debate. The University of Miami's Hare is a believer in mesenchymal stem cells, which differentiate into connective tissue, including muscle; he has tested them in early clinical trials and reported that the cells can repair scar tissue and help new blood vessels and muscle grow. Various other cells found in bone marrow have their champions, and several small trials are now testing bone marrow cells transplanted from young, healthy donors. Wu and others say it is also worth looking to embryonic stem cells and induced pluripotent stem cells, which haven't entered human trials yet.

Then there is a class of cells in the heart itself that may or may not have wide-ranging regenerative powers. One clinical trial tested a stem cell type called c-kit$^+$ cells, pioneered by Anversa, now at the center of a firestorm of his own. Anversa had reported that c-kit$^+$ cells give rise to all different heart tissues. But in May, a study in *Nature* questioned whether the cells actually create much heart muscle at all.

All of this uncertainty, and the push to test so many different cells in people, creates a quandary, Murry says: "Without an understanding of mechanism it's impossible to rationally improve therapy." Even the doubters are optimistic that one day, cardiac cell therapy will prove a viable treatment for the masses. But, they say, we still have a lot to learn before we get there.

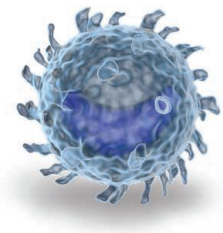
For many patients and doctors, however, the key question is practical: Does cell therapy buy time for a person with a failing heart? No clinical trial, for any cell type, has been designed to look specifically at survival, but one is about to: a 3000-person, $8.1 million study now recruiting across Europe. It is funded by the European Union and aims to resolve, finally, the bitter dispute over whether a mix of cells from a patient's own bone marrow can help. Yet it is far from universally welcomed in the fractious field. The trial, called BAMI, is "unethical," Arnesen charges, because, he believes, the therapy it's testing has already been shown not to help.

Even so, BAMI launched with an elation that echoes the earliest trials a decade ago. A 54-year-old man named Neal Grainger, the first British patient treated, told BBC that "it's fantastic to be part of this." A London hospital signing up volunteers put out a press release in February highlighting that it is "recognised for its pioneering research into stem cell therapy for heart patients." The BAMI website notes that "clinical data now exists supporting the concept that autologous bone marrow derived cells can restore cardiac function" after a heart attack. The trial, says the site, could show that the cells "will reduce the mortality rate by 25%."

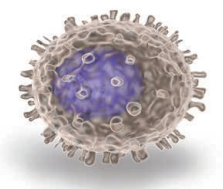
Left unmentioned: the many studies of the same therapy, in patients with the same illness, that have failed to make a difference. ■

Cast of characters

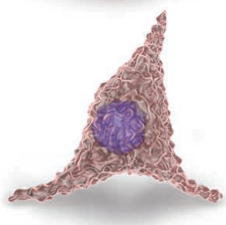
Various cell types are contenders for repairing the heart



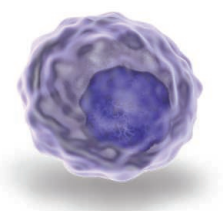
Mononuclear bone marrow cells. This mix contains mostly white blood cells (like the one shown here), and also a small proportion of adult stem cells. Widely used in clinical trials, they are now the subject of a 3000-patient study in Europe.



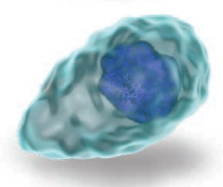
Embryonic stem (ES) cells. ES cells can transform into any cell type, including different parts of the heart. They haven't been tested in humans yet.



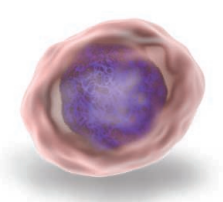
Mesenchymal stem cells. Found in the bone marrow, these cells can turn into connective tissue including muscle. Drawn from patients or healthy donors, the cells have been tested in some small heart trials.



Induced pluripotent stem cells. As versatile as ES cells, they can be turned into heart cells in the lab. They haven't been tried in humans but are generating much interest.



C-kit$^+$ cells. Some researchers argue these are cardiac stem cells that can form all of the heart's components, but a recent study challenged how often they turn into heart muscle.



Cardiosphere-derived cells. These are produced by culturing cardiac cells under special conditions. A small study showed reductions in scar mass but no improvement in cardiac function.

ILLUSTRATIONS: C. BICKEL/*SCIENCE*

Published by AAAS

Downloaded from www.sciencemag.org on August 4, 2014

BOOKS *et al.*

Summer reading

This year we turned to recent AAAS Mass Media Fellows for suggestions and reviews of some thoughtful and enjoyable science titles

DESIGN

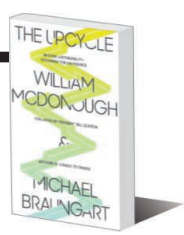
Positive steps to living green

By **Nadja C. Kunz**

From intergovernmental panels to corporate boardrooms, sustainability increasingly appears on the agenda. Accumulating evidence indicates humans now perturb global processes that underpin Earth's natural systems (*1*). The impetus to develop sustainable production and consumption systems is thus stronger today than ever before.

In *The Upcycle*, William McDonough and Michael Braungart focus on the importance of design in creating a world in which "things get better for generations in the future, not worse." The book builds on the authors' earlier work, *Cradle to Cradle* (*2*). There they argued for design that seeks to transform the technosphere (encompassing "metals, plastics, and other materials not continuously created by the biosphere") to mimic biological nutrient cycles: used products become inputs to other processes or food for living organisms. For example, rather than chopping down trees to use the wood directly for fuel, McDonough and Braungart propose a longer life cycle. Wood could instead be used first to build a table, which could subsequently be recycled to produce particleboard, paper, or insulation pulp. Finally, after exhausting alternative uses, the material could be burned to produce the same energy that would have been provided by the original wood. Reuse cycles should avoid using toxic chemicals that would preclude subsequent processing.

After 11 years of implementing cradle-to-cradle concepts, McDonough and Braungart aim to inform readers about what they have learned. They also introduce the concept of upcycling: products should not only be designed to avoid negative impacts but should create positive impacts. Changing the prevailing design philosophy from "planned obsolescence" toward upcycling will require a fundamental shift, not only by designers but within society more broadly. The authors propose several strategies for triggering that shift. These include an appeal to always keep in mind the question "what's next?"—what will happen to the product after use? The authors walk through countless examples of what they regard as insufficient attempts at



The Upcycle
Beyond Sustainability—Designing for Abundance
William McDonough and Michael Braungart
North Point Press, 2013. 254 pp.

The reviewer is at the Swiss Federal Institute of Aquatic Science and Technology, CH-8600 Dübendorf, Switzerland. E-mail: nadja.kunz@eawag.ch

ILLUSTRATION: GORDON STUDER/WWW.GORDONSTUDER.COM

Published by AAAS

upcycling, and they describe common traps in thinking that have impeded progress. They also highlight some successful efforts, such as the trend from sewage treatment plants to nutrient management systems that recover phosphate for soil fertilizer.

In the book's foreword, Bill Clinton praises the authors for their optimism, but at times I found that frustrating. In the authors' telling, sometimes the transition toward upcycling sounds too easy and the solutions obvious. However, organizational and policy changes—especially in the realm of sustainability—are rarely straightforward, and the acceptable trade-offs among competing environmental, social, and economic objectives are ultimately a matter of societal judgment. The case studies could be improved by elaborating the concrete steps that were followed to implement the upcycling philosophy from concept into practice—from the roles of individual champions to changes in organizational structures, policies, and processes. Readers would also benefit from learning about practical tools and frameworks that have facilitated consideration of competing objectives.

Quibbles aside, McDonough and Braungart's approachable language will appeal to a wide audience. *The Upcycle* should prove a rich lode of ideas to inspire innovations in sustainable design across diverse applications.

REFERENCES

1. J. Rockstrom *et al.*, *Nature* 461, 472 (2009).
2. W. McDonough, M. Braungart, *Cradle to Cradle: Remaking the Way We Make Things* (North Point, New York, 2002).

10.1126/science.1256000

PHOTO: SCIENCE PHOTO LIBRARY/CORBIS

PALEOANTHROPOLOGY

Secrets in the Bones

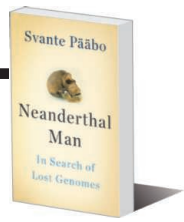
By **Jeremy Chase Crawford**

The Neanderthal is not entirely extinct. In fact, some part of this otherwise ancient hominin remains among us today, manifest in about 2% of the DNA of many modern humans as a result of ancient interbreeding between our species. One strand of Svante Pääbo's *Neanderthal Man* explains the nearly 20 years of research that eventually led to this finding. Pääbo provides a riveting, personal account of the development of paleogenetics and the technical revolution that made the field possible. The memoir also traces the sometimes unconventional path that led to his becoming a globe-trotting professor.

The author's success has stemmed from his determination to pursue what many assumed were unknowable answers to fundamental questions about human origins. In a career replete with unexpected twists, Pääbo conducted secret experiments on Egyptian mummies, researched modern kangaroo rats as well as ancient cave bears, and even used genetic analyses of his own students' sperm to advance his research. His journey was certainly not without its difficulties, and he writes about complications with merciless candor. For instance, he leaves out few raw details when discussing his own insecurities, shortcomings of his and others' research, political and cultural obstacles that often impede scientific progress, and the affairs (romantic as well as scientific) that at times threatened to turn research collaborators into competitors.

Neanderthal type material

Neanderthal Man
In Search of Lost Genomes
Svante Pääbo
Basic Books, 2014. 285 pp.



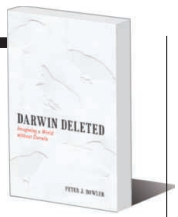
The author effectively conveys both the broad implications of his work and his concern for scientific details. While reading, you may find yourself with the sudden urge to wash your hands when Pääbo expounds on the ubiquitous nature of DNA contamination and the absolute necessity for a completely sterile workspace. These fears were particularly salient in the earlier years, when he required that all data of any importance be independently replicated by scientists in other laboratories. For perhaps his most celebrated finding—that there exists a "clearly discernible genetic contribution from Neanderthals to people outside Africa"—Pääbo would settle for no fewer than three independent lines of statistical evidence before publishing. To nonscientists, this self-professed meticulousness may at times seem to border on neurotic. But in a time of frequent headlines about scientific fraud, his portrayal of the internal and external struggles that accompany a career in academia refreshingly illustrates the high standards required for a rigorous research program.

At its core, the book is about exposing secrets: from the nucleotide sequences laboriously pried from 40,000-year-old bones to the inner workings of a prominent researcher's mind. Whether unacquainted with Pääbo's work or regular followers of his publications, readers will find that *Neanderthal Man* provides a nonpareil account of the development of the field of ancient DNA.

10.1126/science.1255859

The reviewer is at the Museum of Vertebrate Zoology and Department of Integrative Biology, University of California, Berkeley, CA 94720, USA. E-mail: jeremy.crawford@alumni.duke.edu

Published by AAAS



Darwin Deleted
Imagining a World Without Darwin
Peter J. Bowler
University of Chicago Press, 2013. 328 pp.

HISTORY OF SCIENCE

Darwin who?

By **Ana Maria R. Almeida**

Since the 1920s, the idea of evolution has generally been equated with what has been termed Darwinism, a particular evolutionary theory that explains the origin of biological diversity by means of natural selection. Largely as a result of the dominance of that theory, most scientists today would find the thought of a history without Darwin unimaginable. In *Darwin Deleted*, Peter Bowler invites readers to imagine a world in which Darwin never existed. Using counterfactual history and carefully dissecting the history of evolutionary thought, Bowler looks into the past to illuminate prominent debates we face today.

Bowler starts by refuting the "in the air" thesis: the idea that without Darwin, someone else would have come up with the same or similar ideas and history would have unfolded about as it did. Drawing on the historical record, he demonstrates that although the idea of evolution was becoming widely accepted by the time Darwin published *On the Origin of Species* (1859), natural selection was by no means part of mid-19th-century thought. Bowler argues that only Darwin, with his unique combination of diverse interests, was able to piece together all of its key components. Thus, it is very plausible that in Darwin's absence other theories would have come to play more important roles in our understanding of evolution. In fact, into the 1920s, non-Darwinian theories were the dominant explanation for evolutionary changes—which substantiates the viability of Bowler's counterfactual world.

As Bowler writes, it is unquestionable that "Darwin presented his contemporaries with the harshest possible version of nature." That contributed to his becoming the figurehead of what was perceived as an attack on traditional values. Bowler's analysis makes it clear that without Darwin's revolutionary input, evolutionism would have developed in a less confrontational manner. Darwin-like ideas would not otherwise have gained currency for another 30 or 40 years, by which time the general idea of evolution would not have posed a threat to most religious thinkers. Thus, Bowler argues, the antagonism between evolutionism and religion might well be a "product of particular historical events rather than an inevitable conflict of irreconcilable positions."

In addition, Bowler's mental experiment leads us to realize that many of the alleged consequences of what has been called social Darwinism would likely have taken place in a world without Darwin. In fact, "most of the effects … labeled as 'social Darwinism' could have emerged in a world that had no inkling of the theory of natural selection" and "some of those effects … might well have been even more strident in the absence of the Darwinian theory." Far from being a consequence of Darwinism, the idea of progress and the allied theories of directed evolution were grounded in wider social and cultural forces. It is undeniable that Darwinism is a product of its time, with the apparent materialism of a theory based on random variation and struggle. But the simplistic identification of Darwinism with harsh social policies is mistaken, argues Bowler, as most of what is called "'social Darwinism' could be justified equally well through rival theories of evolution."

Darwin Deleted offers a journey into the history of evolutionism well worth taking. Through his scenario in which the *Origin* never appeared, Bowler improves our ability to think about the assumptions underlying contemporary debates.

10.1126/science.1255865

The reviewer is at the Universidade Federal da Bahia, Instituto de Biologia, Rua Barão de Geremoabo, 147, Campus de Ondina, Salvador, Bahia 40170-290, Brazil. E-mail: anagaivota@gmail.com

Monarch butterflies massing in the Sierra Pellon, Michoacan, Mexico.

PHOTO: © RICHARD ELLIS/ALAMY

ECOLOGY

Homeward bound

By **Alexa Fritzsche**

For many people, "home" evokes a sense of comfort and safety, joy and relief. Given these quintessentially human emotions, we may assume that we are the only organisms that build and cherish such places. But *The Homing Instinct* illuminates the vast diversity of animals that locate, construct, and maintain homes. Naturalist and physiologist Bernd Heinrich explores the biology of navigation in migratory animals, describes the diversity of home-making strategies across the animal kingdom, and lastly reflects on his personal sense of home.

Despite our large brains, our ability to navigate pales in comparison to the capabilities of many other species, from monarch butterflies to sea turtles. If dropped into an unknown land, we would need a map, a landmark, and a compass to have any hope of finding home. In striking contrast, birds such as the Manx shearwater spend their lives coursing seas seemingly without landmarks, and individual sandhill cranes return year after year to their own natal pond in the Alaskan tundra after wintering in Texas. The book's opening section explores the cognitive mechanisms that long-distance migrants use in the mystifying relocation of their homes. Historically, cues from the Sun and stars were considered the primary mechanisms of navigation. But in recent decades, evidence has mounted that some species "may literally see ghost images of the Earth's magnetic field superimposed on … objects" and use these magnetic cues to guide their routes. Scientists and other readers will enjoy Heinrich's stories of the creative experiments that unveiled these discoveries, from outfitting birds with blurred goggles to raising them in a homemade planetarium with interchangeable star charts.

The author also explores the evolution, natural history, and implications of home building and maintenance in the animal world. What materials and behaviors are used to build homes and protect them from parasitic and predatory invaders? Why do some choose solitary versus communal living? I found Heinrich's comparison of human home-making to that of other animals particularly interesting. Whereas other primates have precocial newborns and may

The reviewer is at the Odum School of Ecology, University of Georgia, 140 East Green Street, Athens, GA 30602, USA. E-mail: afritzsche@gmail.com

Published by AAAS

The Homing Instinct
Meaning and Mystery in Animal Migration
Bernd Heinrich
Houghton Mifflin Harcourt, 2014. 368 pp.

use rudimentary shelters, modern humans have dependent young and are extremely reliant on homes. Could these two traits be linked? As early humans evolved bipedality and became highly mobile hunter-gatherers, carrying offspring on their backs became prohibitive. Heinrich asserts that males may have maximized reproductive success by leaving the mother and helpless children in a well-protected shelter. In these homes, "social tolerance, division of labor, cooperation, and subservience to rules or individuals ... then became necessary and hence valued survival traits." Indeed, the construction of homes may have enabled us to colonize the far reaches of the planet.

The Homing Instinct generally walks a pleasant line between fact-laden science writing and the author's reflective personal anecdotes. At the expense of deviating from the central scientific themes, the meandering final section serves as a venue for the author to nostalgically reflect on his own home, memories of hunting trips with family, and efforts to start a chestnut tree grove on his property. As a whole, the book prompts readers to appreciate homing not just as a means toward comfort and joy but also as a complex and mysterious biological trait shaped by natural selection and shared with a multitude of organisms.

10.1126/science.1255864

PARTICLE PHYSICS

Picking up grand vibrations

By **Megan Engel**

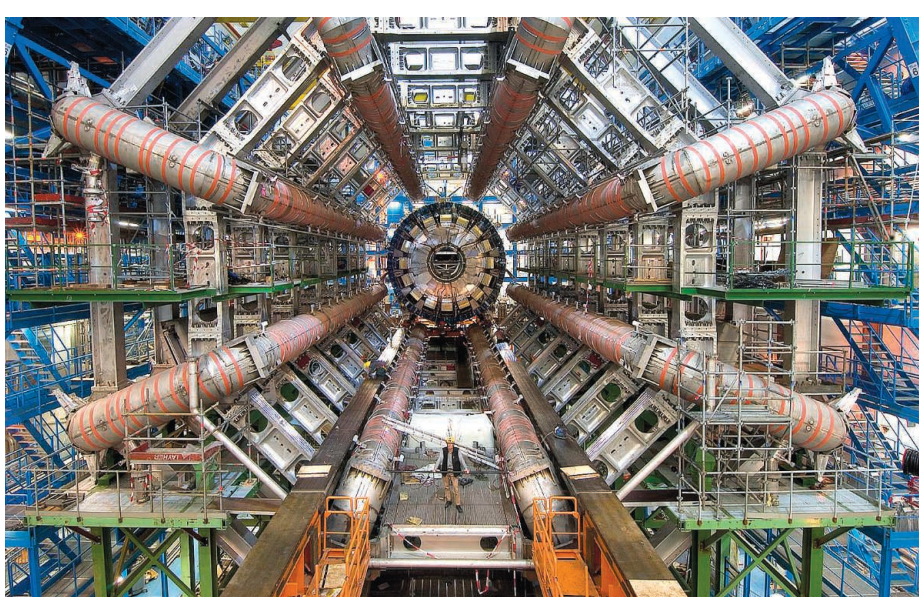
You are standing in the dark immensity of a tunnel stretching for miles before and behind you. It is colder than outer space, and vast superconducting magnets yawn above and flank you, sustained by 8000 kilometers' worth of wound cable. Every fragment of this colossal machine was crafted meticulously, in labs on every corner of the globe. At the flip of a switch, beams containing the energy of 175 metric tons of TNT will zoom around the tunnel, set to collide and explode in violent bursts of energy and light. Amazingly, all fuel for this endeavor comes from a single canister of hydrogen no larger than a fire extinguisher. Where are you? The Large Hadron Collider (LHC)—in the words of Sean Carroll, "the largest, most complex machine ever built by human beings." In *The Particle at the End of the Universe*, Carroll embarks on an ambitious voyage through the quest to find the Higgs boson (more colorfully known as the "god particle") from the development of modern physics, to the construction of the LHC, to the celebrated discovery of July 2012—and beyond.

Carroll shines in his use of digestible language to communicate esoteric physics. Although keeping up fully might prove challenging for readers lacking any physics background, much insight can be gained from Carroll's analogies. He explains, for instance, that all "particles" are actually vibrations traveling in fundamental, invisible fields that permeate space, like sound waves propagating through the air. The collision of two particles to produce the Higgs boson is like two vibrating piano strings transferring their vibrations to a third. From the ancient Greeks, to gravity, to the billions of dark matter particles passing through our bodies each second, to the extra dimensions of string theory, the book is a pedagogical tour de force, complete with ample diagrams and appendices.

The author establishes the human element of the "high stakes drama" surrounding the Higgs boson discovery and announcement in tandem with the scientific. Throughout his narrative, Carroll weaves a sense of the consuming passion that drives physicists on the "emotional roller-coaster ride" of fundamental research. He paints a rich mosaic—peppered with personal details, famous quips, and anecdotes—of the major players in physics over the centuries. The book includes poignant details such as the fact that Peter Higgs's seminal paper was initially rejected for publication and the emotion visible on Higgs's

The reviewer is at Clarendon Laboratory, Oxford University, Parks Road, Oxford OX1 3PU, UK. E-mail: megan.engel@physics.ox.ac.uk

The Particle at the End of the Universe
How the Hunt for the Higgs Boson Leads Us to the Edge of a New World
Sean Carroll
Penguin, 2012. 380 pp.



Installing the ATLAS calorimeter.

PHOTO: ATLAS EXPERIMENT © 2014 CERN

Published by AAAS

face during the standing ovation that followed the announcement that the LHC had found his eponymous boson.

Perhaps the book's strongest facet is its triumphant championing of basic scientific research. Carroll's aim is broad and sweeping: to establish the central importance of science in humanity's wrestle with the question of existence. He terms the finding of the Higgs boson a "success for the human race" in answering our "restless desire to understand our world." *The Particle at the End of the Universe* illustrates that finding it "is its own reward," emphasizes the "universality of the scientific impulse," and waxes poetic in juxtaposing art and science. The vast behemoth of modern physics, the LHC, and the scintillating discovery of the Higgs boson are not only important because they illumine the "secrets of the Universe"—the entire effort tells us something profound about ourselves.

10.1126/science.1255862

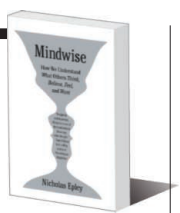
NEUROSCIENCE

Where I end, where you begin

By **Séamus A. Power**

Nicolas Epley begins *Mindwise* outside a courthouse in Addis Ababa, Ethiopia. Epley (a social psychologist at the University of Chicago) and his wife have just legally adopted two Ethiopian children. Epley is both perplexed and anxious before first meeting their biological father, who has agreed to have them adopted. What did this man think, believe, feel, and want for his children? This opening vignette is the first of many examples taken from the author's personal life, news headlines, and empirical science that drive this fast-moving and highly readable book. Epley's central concern is articulating the ways in which we both understand and, very often, misunderstand other people, even those close to us.

Epley delivers the good news that we all have a sixth sense, an ability to read minds. The bad news is that we are not very good at it. However, there is hope. In the book, Epley draws on a wealth of empirical social psychological research to help make sense of how humans understand and misunderstand one another. Moreover, his experiments show how people, even couples, overestimate how well they think they know others. Not only do the participants in his studies misunderstand one another, they are also largely unaware of their mistakes.



Mindwise
How We Understand What Others Think, Believe, Feel, and Want
Nicholas Epley
Alfred A. Knopf, 2014. 265 pp.

Some of the shortcomings of our mind-reading ability stem from how we engage and disengage it at inappropriate times. Epley reveals how we dehumanize other people. That is, how we fail to attribute minds to other people, such as the homeless or terrorists, when it is advantageous or even necessary to do so. Conversely, humans also anthropomorphize; we attribute minds to nonhumans such as gods, family pets, and technological objects like robots and even alarm clocks that force us out of bed in the morning. According to the author, these evolutionarily ingrained intuitions are only half of the reason why people converse with gods in the sky while ignoring homeless people who ask for spare change.

The book covers a variety of ways by which we try to read others' minds: We often project our own thoughts, wants, feelings, or beliefs onto the minds of others—assuming that they think what we think. We stereotype groups and consequently fail to see how individuals might be like us. We often incorrectly make inferences about others' mindsets based on our observations of their behaviors. (Most people who stayed in New Orleans despite the warnings about Hurricane Katrina were not irresponsible or stupid. They simply had no means to leave or nowhere else to go.)

These mind-reading tools allow us to make sense of complex social interactions to a degree better than chance. However, their faults indicate they could be improved, and Epley ends *Mindwise* by suggesting ways to do so. The secret lies in perspective getting rather than perspective taking. According to Epley, the latter, egocentric strategy at best reinforces—and at worst amplifies—stereotypical and incomplete understandings of others. He espouses the more direct strategy of opening the lines of communication and getting someone else's perspective as a potential solution to understanding the mind of the other. It worked for Kennedy and Khrushchev in resolving the Cuban Missile Crisis. It can also work for us in our relationships with colleagues, friends, family, and significant others.

10.1126/science.1255866

HISTORY OF CHEMISTRY

Elemental discoveries

By **James McDonagh**

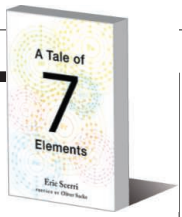
The periodic table has become a symbol of scientific understanding and the power of the scientific method. It is used by the popular media to signify scientific excellence, a logical progression to better understanding of the natural world. But how many people really appreciate the subtle complexities, trials, and tribulations that underlie this edi-

The reviewer is at the Department of Comparative Human Development, University of Chicago, 5736 South Woodlawn Avenue, Chicago, IL 60637, USA. E-mail: seamusapower@uchicago.edu

The reviewer is at the School of Chemistry, University of St. Andrews, North Haugh, St. Andrews, Fife KY16 9ST, UK. E-mail: jm222@st-andrews.ac.uk

PHOTO: PASIEKA/SCIENCE PHOTO LIBRARY/CORBIS

Published by AAAS



A Tale of Seven Elements
Eric Scerri
Oxford University Press, 2013. 304 pp.

fice of science? In *A Tale of Seven Elements*, philosopher of science Eric Scerri provides an expert perspective on the history of the periodic table, focusing principally on the last seven naturally occurring elements to be discovered.

Scerri opens the book by setting the discoveries in their historical context, the era of active science between the first and second world wars. While noting the difficulties faced by women entering the physical sciences, he points out that several of the seven elements (protactinium, hafnium, rhenium, technetium, francium, astatine, and promethium) were discovered by women, whose key contributions had often been overlooked.

The author's concise and meticulous account blends philosophy, history, and science, presenting the deep, although not immediately obvious, connections among them. He provides an informative history of the classification of chemical elements, from Dalton's atomic weights to Mendeleev's periodic table. In addition to introducing concepts such as radioactivity and isotopes, Scerri expounds on the far-reaching consequences of quantum theory for chemistry (including Bohr's deduction of atomic structure).

As Scerri notes, his seven focal elements are all "rather exotic." Many are radioactive, and they are generally of low abundance. The book's overarching theme highlights how the discoveries shaped history and how historical and social pressures shaped them. This is exemplified by the priority dispute over hafnium, in which wartime alliances stifled the acceptance of a correct claim. Scerri quotes a British journal editor's response to an article submitted in the early 1920s: "We adhere to the original word celtium given to it by Urbain as a representative of the great French nation which was loyal to us throughout the war." Scerri also notes the impact of competition for personal prestige and resources, aspects that are often overlooked in textbook accounts of the discoveries of elements.

The holistic accounts largely succeed at guiding the reader through the complex connections and numerous priority claimants, although in places the narrative can be difficult to follow. Scerri intended the book for "readers interested in digging a little deeper into the science of the elements and the periodic table." *A Tale of Seven Elements* proves his aim true.

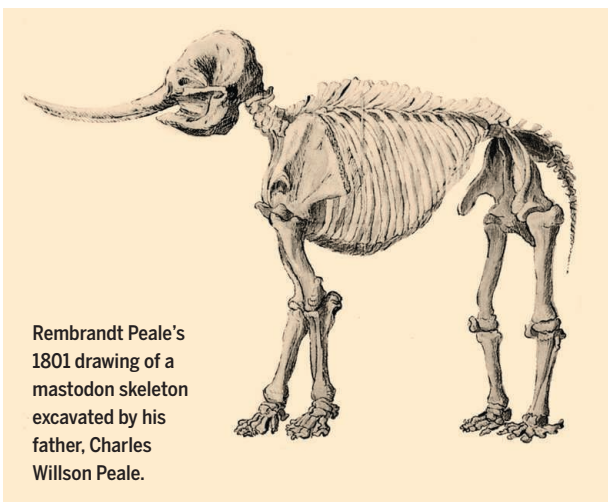
10.1126/science.1255968

EXTINCTION

Our enduring legacy?

By **Kendra Smyth**

A thought-provoking work of scientific journalism, Elizabeth Kolbert's *The Sixth Extinction* tackles the sobering reality of the large-scale loss of biodiversity happening before our eyes. Although the fossil record documents five previous great extinctions, this time the driving catalyst is neither asteroid nor volcanic eruption but a single inhabitant of Earth: us. For millennia, humans have succeeded extravagantly to the detriment of other species. "Though it might be nice to imagine there once was a time when man lived in harmony with nature," Kolbert notes, "it's not clear that he ever really did."



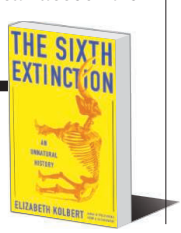
Rembrandt Peale's 1801 drawing of a mastodon skeleton excavated by his father, Charles Willson Peale.

Through masterful field reporting, Kolbert transforms a collection of daunting facts into an entertaining and enlightening journey of scientific discovery, accessible to scientists and laymen alike. She takes readers along as she visits over a dozen scientists at field stations and laboratories around the world: herpetologists in Panama, marine biologists in Australia, botanists in the Andes, paleontologists in New Jersey, and anthropologists in Germany. Intermixed with these scientific reports and personal narratives is an intriguing historical account of the origin of extinction as a concept. In the mid-18th century, the prevailing view was that species were immutable, decreed by God. When mysterious bones were unearthed in present-day Kentucky, scientists rationalized that they were remnants of elephants or hippos that had wandered astray. As Thomas Jefferson put it, "Such is the economy of nature that no instance can be produced of her having ... formed any link in her great work so weak as to be broken." Then in 1796 the French naturalist Georges Cuvier argued that life had a history, one "marked by loss and punctuated by events too terrible for human imagining." The concept of extinction was born, although not readily accepted for decades.

Today, we struggle to grasp a different concept of extinction: humans are responsible for transforming the biosphere. Despite desperate searches for alternative explanations, the evidence Kolbert presents is undeniable: at warp speed, humans have altered the composition of the atmosphere, acidified the oceans, hunted and fished large species to their brink, and destroyed ecosystems. "Right now," she writes, "... we are deciding, without quite meaning to, which evolutionary pathways will remain open and which will forever be closed. No other creature has ever managed this, and it will, unfortunately, be our most enduring legacy." The implications are haunting.

In pushing other species to extinction, what happens to us? Ecologist Paul Ehrlich warns "humanity is busy sawing off the limb on which it perches." Kolbert, halfheartedly optimistic, observes that we should not underestimate human ingenuity. Nonetheless, she gives scant reference to possible solutions, such as scattering sulfates or water droplets into the atmosphere to reflect sunlight or brighten clouds. Although her attempt to conclude on an upbeat note falls flat, it is not wholly misguided. Instead, we are left with an appreciation of the diversity and complexity of life, a desire to transition to a more sustainable world, and a reminder that every ending is a chance for a new beginning. The sixth extinction might be the end of the world we know, but it will not be the end of life.

10.1126/science.1255867

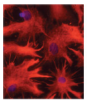


The Sixth Extinction
An Unnatural History
Elizabeth Kolbert
Henry Holt, 2014. 335 pp.

The reviewer is at the Nicholas School of the Environment, Duke University, Durham, NC 27708, USA. E-mail: kendra.smyth@duke.edu

CREDIT: NYPL/SCIENCE SOURCE/GETTY IMAGES

Published by AAAS

EXTENDED PDF FORMAT
SPONSORED BY

Travel Grants Available
Neuroscience 2014
Application Deadline: July 31st
>Apply Now
R&D systems
a bio-techne brand
www.rndsystems.com

The future lies in uncertainty

D. J. Spiegelhalter

Science **345**, 264 (2014);
DOI: 10.1126/science.1251122

This copy is for your personal, non-commercial use only.

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by clicking here.

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines here.

The following resources related to this article are available online at www.sciencemag.org (this information is current as of August 4, 2014):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:
http://www.sciencemag.org/content/345/6194/264.full.html

This article **cites 11 articles**, 3 of which can be accessed free:
http://www.sciencemag.org/content/345/6194/264.full.html#ref-list-1

This article appears in the following **subject collections:**
Computers, Mathematics
http://www.sciencemag.org/cgi/collection/comp_math

Downloaded from www.sciencemag.org on August 4, 2014

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published weekly, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. Copyright 2014 by the American Association for the Advancement of Science; all rights reserved. The title *Science* is a registered trademark of AAAS.

PERSPECTIVES

STATISTICS

The future lies in uncertainty

Sophisticated statistical insights are crucial for gaining knowledge from ever-larger data sets

By **D. J. Spiegelhalter**

Statisticians have celebrated a lot recently. 2013 marked the 300th anniversary of Jacob Bernoulli's *Ars Conjectandi,* which used probability theory to explore the properties of statistics as more observations were taken. It was also the 250th anniversary of Thomas Bayes' essay on how humans can sequentially learn from experience, steadily updating their beliefs as more data become available (*1*). And it was the International Year of Statistics (*2*). Now that the bunting has been taken down, it is a good time to take stock of recent developments in statistical science and examine its role in the age of Big Data.

Much enthusiasm for statistics hangs on the ever-increasing availability of large data sets, particularly when something has to be ranked or classified. These situations arise, for example, when deciding which book to recommend, working out where your arm is when practicing golf swings in front of a games console, or (if you're a security agency) deciding whose private e-mail to read first. Purely data-based approaches, under the title of machine-learning, have been highly successful in speech recognition, real-time interpretation of moving images, and online translation.

Statistical science does produce some excellent machine-learning tools, but it is concerned with more than just classification or ranking: It explicitly tries to deal with the uncertainty about what can be concluded from data, be it a prediction or scientific inference. The revolutionary ideas of Bernoulli, Bayes, and others form the historical basis for such learning from large data sets (*3*).

To assess the uncertainty about unknown or future quantities, statisticians tend to build probability models of underlying processes. For example, Microsoft's TrueSKILL ranking system produces a probability distribution for the unknown skill of an online gamer. Similarly, trading models can assess both the expectation and volatility of future commodity prices. Detailed risk assessment is also vital to the insurance industry, requiring complex statistical models. Sophisticated simulation-based statistical methods, such as particle filters, are increasingly used in areas such as signal processing, dynamic economic models, and systems biology (*4*).

Traditional statistical problems could be termed "large n, small p": There were many observations (n), such as participants in a clinical trial, but few parameters were measured (p), and just a handful of hypotheses tested. More recently attention has turned to "small n, large p" problems, such as a few brain scans but with millions of voxels in each, or the expression of tens of thousands of genes in a limited number of tissue samples. To deal with these problems, statisticians developed models for complex interactions—for example, when learning the structure of a network describing gene relationships (*5*), and handling observations that are not just single data points but may comprise shapes, functions, images, or phylogenetic trees (*6*).

Many such "small n, large p" problems require screening of vast numbers of hypotheses, for which the naïve use of statistical significance is inappropriate; the standard "$P < 0.05$" criterion means that 1 in 20 nonexistent relationships will be declared significant, so that if you do enough tests, some apparent discoveries will always pop up. Procedures have been developed to control the false discovery rate (FDR); that is, the proportion of apparent discoveries that turn out to be wrong (*7*). For example, the confidence required before announcing the discovery of the Higgs boson was couched in statistical terms as 5 sigma, and this assessment included an adjustment for how many hypotheses were examined (the "look elsewhere effect").

As an example of the modern use of statistics, the June 2014 issue of *Bioinformatics* features a huge range of statistical and machine-learning techniques, many using the *Bioconductor* suite of software packages. For example, Berry *et al.* use "scan" statistics to identify regions of the human genome targeted by retroviruses, assessed with FDRs that allow for the fact that the whole genome has been searched (*8*). We recently used similar methods to show that a cluster of six London cyclist deaths in a fortnight had only around a 1 in 40 chance of occurring in any fortnight over 8 years (*9*).

Statistical principles are also routinely applied in A/B experiments on Web sites, in which small tweaks in layout and design are tested by using alternative versions at random and seeing whether the innovation has, for example, increased click-through rates. More sophisticated adaptive experimental designs are making drug testing more efficient, because treatment regimes that show little promise can be rapidly dropped without compromising the overall chance of a false discovery (*10*).

As medical databases grow and large randomized trials become more expensive and difficult to conduct, there is increasing demand for trying to make causal inferences from observational data. This is a minefield: Even with masses of data, there is no automatic technique for turning correlation into causation. But statistical science has developed frameworks for clarifying the careful analysis necessary, for example, when taking into account changing interventions

ILLUSTRATION: PETER AND MARIA HOEY/WWW.PETERHOEY.COM

Statistical Laboratory, Centre for Mathematical Sciences, University of Cambridge, Cambridge, CB3 0WB, UK. E-mail: david@statslab.cam.ac.uk

Published by AAAS

over time and indirect effects of therapy through mediating factors, such as viral load in HIV (*11*).

Big Data means that we can get more precise answers; this is what Bernoulli proved when he showed how the variability in an estimate goes down as the sample size increases. But this apparent precision will delude us if issues such as selection bias, regression to the mean, multiple testing, and overinterpretation of associations as causation are not properly taken into account. As data sets get larger, these problems get worse, because the complexity and number of potential false findings grow exponentially. Serious statistical skill is required to avoid being misled.

As measurement becomes ever faster and cheaper, the trend will be toward "large n, really large p" problems as, for example, images, genomes, and electronic health records become linked together. Other challenging new areas for statistical science include those traditionally handled with deterministic models, such as weather, climate, extreme natural hazards, and epidemics. In each of these, a stochastic element can be added, such as combining weather projections from randomly perturbed starting points to provide an ensemble forecast, although the appropriate role for stochastics in climate models is still contested (*12*). In these tricky areas, statisticians can show that they are prepared to deal with messy data, but still think in terms of generalizable principles (*13*). They can enable others to produce knowledge from data by promoting their own particular skills and insights, particularly by understanding both the strengths and limitations of models.

The title of this article is deliberately ambiguous: Not only is the future uncertain, but also it will be vital to understand and promote uncertainty through the appropriate use of statistical methods rooted in probability theory. Careful application of statistical science will be essential. ■

REFERENCES

1. B. Efron, *Science* **340**, 1177 (2013).
2. M. Davidian, *J. Am. Stat. Assoc.* **108**, 1141 (2013).
3. National Academy of Sciences, *Frontiers in Massive Data Analysis* (2013); see www.nap.edu/catalog.php?record_id=18374.
4. C. Andrieu, A. Doucet, R. Holenstein, *J. R. Stat. Soc. Series B Stat. Methodol.* **72**, 269 (2010).
5. B. Li, H. Chun, H. Zhao, *J. Am. Stat. Assoc.* **107**, 152 (2012).
6. B. Aydın *et al.*, *Ann. Appl. Stat.* **3**, 1597 (2009).
7. J. Fan, X. Han, W. Gu, *J. Am. Stat. Assoc.* **107**, 1019 (2012).
8. C. C. Berry *et al. Bioinformatics* **30**, 1493 (2014).
9. J. Aberdein, D. Spiegelhalter, *Significance* **10**, 46 (2013).
10. B. Gaydos *et al.*, *Drug Inf. J.* **43**, 539 (2009).
11. O. O. Aalen, K. Røysland, J. M. Gran, B. Ledergerber, *J. R. Stat. Soc. Ser. A Stat. Soc.* **175**, 831 (2012).
12. C. Macilwain, *Science* **344**, 1221 (2014).
13. http://simplystatistics.org/2013/05/29/what-statistics-should-do-about-big-data-problem-forward-not-solution-backward/

10.1126/science.1251122

PHOTO: STEVE GSCHMEISSNER/SCIENCE SOURCE

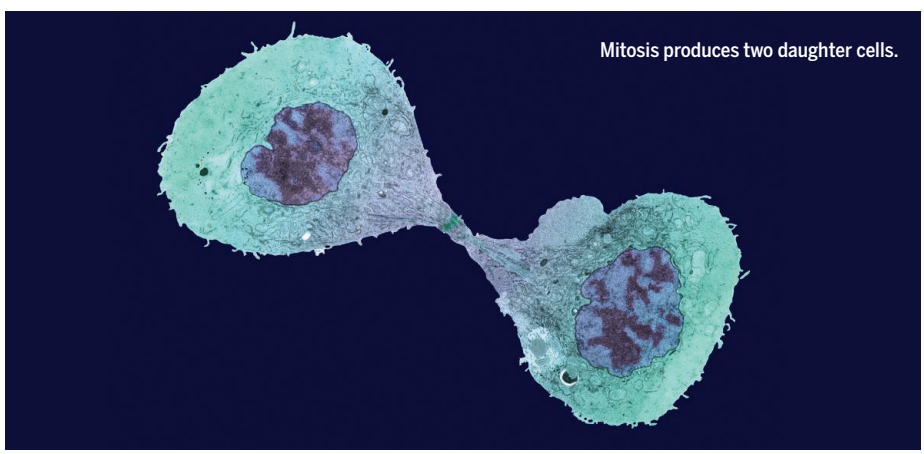
CELL BIOLOGY

Mind the midzone

Is the separation of chromosomes during cell division monitored by a checkpoint?

By **Michael A. Hadders** *and* **Susanne M.A. Lens**

When a eukaryotic cell divides, it must partition its duplicated genome into two daughter cells with a minimum of errors and DNA damage. This requires the ordering and coordination of various events during nuclear division (mitosis) and cytoplasmic division (cytokinesis). How these events are temporally coordinated remains poorly understood. On page 332 of this issue, Afonso *et al.* (*1*) propose the exciting idea of a "chromosome separation checkpoint" that controls the near final step of mitosis—the anaphase to telophase transition (see the figure).



Mitosis produces two daughter cells.

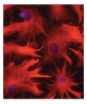
Once the duplicated chromosomes are bi-oriented on the mitotic spindle in metaphase, anaphase begins, in which the sister chromosomes start to separate. Chromosomes must then be well separated before the beginning of cytokinesis and telophase events, which include chromosome decondensation, nuclear envelope reformation (NER), and mitotic spindle disassembly. Elongation of the mitotic spindle facilitates this separation process. Anaphase onset before completion of metaphase gives rise to segregation errors, whereas the start of cytokinesis before clearance of chromosomes from the cleavage plane gives rise to DNA damage (*2*, *3*).

Afonso *et al.* found that treatments that impaired spindle elongation delayed the onset of chromosome decondensation and NER. This delay was dependent on a gradient of aurora B kinase–dependent phosphorylation that originated from the central spindle midzone (*4*). In addition, timely chromatin decondensation and NER required the activity of protein phosphatase 2A (PP2A) and protein phosphatase 1 (PP1). The findings suggest that when the spindle elongates, the separating packs of sister chromosomes are gradually pulled out of the midzone "sphere of influence" of aurora B. This would progressively tip the kinase-phosphatase balance in the region surrounding the chromosomes in favor of the phosphatases. The subsequent dephosphorylation of substrates results in chromatin decondensation and NER. In case of impaired spindle elongation, the separating chromosomes remain within the aurora B activity gradient for a longer period, thus delaying the onset of telophase.

A possible explanation for how aurora B may prevent chromatin decondensation and NER is that it directly phosphorylates substrates involved in chromatin condensation and in nuclear envelope disassembly. The condensin I complex is an interesting candidate because it is a direct substrate of aurora B and promotes chromosome compaction, and its recruitment onto chromosomes requires aurora B kinase activity (*5*–*8*). Indeed, Afonso *et al.* found that Barren—the homolog of condensin I in *Dro-*

Department of Molecular Cancer Research, Center for Molecular Medicine, University Medical Center Utrecht, Universiteitsweg 100, 3584 CG, Utrecht, Netherlands. E-mail: s.m.a.lens@umcutrecht.nl

Published by AAAS

EXTENDED PDF FORMAT
SPONSORED BY

Travel Grants Available
Neuroscience 2014
Application Deadline: July 31st
>Apply Now
R&D systems
a bietechne brand
www.rndsystems.com

Mind the midzone

Michael A. Hadders and Susanne M.A. Lens

Science **345**, 265 (2014);
DOI: 10.1126/science.1257267

This copy is for your personal, non-commercial use only.

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by clicking here.

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines here.

The following resources related to this article are available online at www.sciencemag.org (this information is current as of August 4, 2014):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:
http://www.sciencemag.org/content/345/6194/265.full.html

A list of selected additional articles on the Science Web sites **related to this article** can be found at:
http://www.sciencemag.org/content/345/6194/265.full.html#related

This article **cites 14 articles**, 7 of which can be accessed free:
http://www.sciencemag.org/content/345/6194/265.full.html#ref-list-1

This article appears in the following **subject collections:**
Cell Biology
http://www.sciencemag.org/cgi/collection/cell_biol

Downloaded from www.sciencemag.org on August 4, 2014

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published weekly, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. Copyright 2014 by the American Association for the Advancement of Science; all rights reserved. The title *Science* is a registered trademark of AAAS.

over time and indirect effects of therapy through mediating factors, such as viral load in HIV (*11*).

Big Data means that we can get more precise answers; this is what Bernoulli proved when he showed how the variability in an estimate goes down as the sample size increases. But this apparent precision will delude us if issues such as selection bias, regression to the mean, multiple testing, and overinterpretation of associations as causation are not properly taken into account. As data sets get larger, these problems get worse, because the complexity and number of potential false findings grow exponentially. Serious statistical skill is required to avoid being misled.

As measurement becomes ever faster and cheaper, the trend will be toward "large n, really large p" problems as, for example, images, genomes, and electronic health records become linked together. Other challenging new areas for statistical science include those traditionally handled with deterministic models, such as weather, climate, extreme natural hazards, and epidemics. In each of these, a stochastic element can be added, such as combining weather projections from randomly perturbed starting points to provide an ensemble forecast, although the appropriate role for stochastics in climate models is still contested (*12*). In these tricky areas, statisticians can show that they are prepared to deal with messy data, but still think in terms of generalizable principles (*13*). They can enable others to produce knowledge from data by promoting their own particular skills and insights, particularly by understanding both the strengths and limitations of models.

The title of this article is deliberately ambiguous: Not only is the future uncertain, but also it will be vital to understand and promote uncertainty through the appropriate use of statistical methods rooted in probability theory. Careful application of statistical science will be essential. ■

REFERENCES

1. B. Efron, *Science* **340**, 1177 (2013).
2. M. Davidian, *J. Am. Stat. Assoc.* **108**, 1141 (2013).
3. National Academy of Sciences, *Frontiers in Massive Data Analysis* (2013); see www.nap.edu/catalog.php?record_id=18374.
4. C. Andrieu, A. Doucet, R. Holenstein, *J. R. Stat. Soc. Series B Stat. Methodol.* **72**, 269 (2010).
5. B. Li, H. Chun, H. Zhao, *J. Am. Stat. Assoc.* **107**, 152 (2012).
6. B. Aydın *et al.*, *Ann. Appl. Stat.* **3**, 1597 (2009).
7. J. Fan, X. Han, W. Gu, *J. Am. Stat. Assoc.* **107**, 1019 (2012).
8. C. C. Berry *et al. Bioinformatics* **30**, 1493 (2014).
9. J. Aberdein, D. Spiegelhalter, *Significance* **10**, 46 (2013).
10. B. Gaydos *et al.*, *Drug Inf. J.* **43**, 539 (2009).
11. O. O. Aalen, K. Røysland, J. M. Gran, B. Ledergerber, *J. R. Stat. Soc. Ser. A Stat. Soc.* **175**, 831 (2012).
12. C. Macilwain, *Science* **344**, 1221 (2014).
13. http://simplystatistics.org/2013/05/29/what-statistics-should-do-about-big-data-problem-forward-not-solution-backward/

10.1126/science.1251122

PHOTO: STEVE GSCHMEISSNER/SCIENCE SOURCE

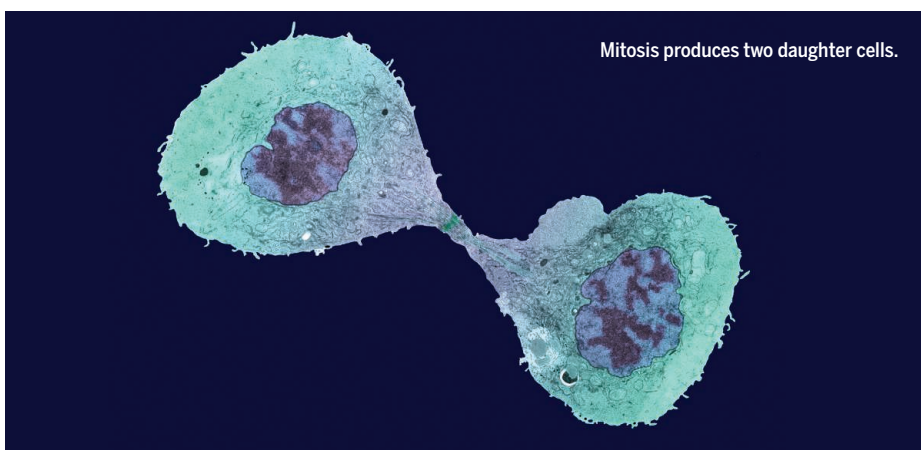
CELL BIOLOGY

Mind the midzone

Is the separation of chromosomes during cell division monitored by a checkpoint?

By **Michael A. Hadders** *and* **Susanne M.A. Lens**

When a eukaryotic cell divides, it must partition its duplicated genome into two daughter cells with a minimum of errors and DNA damage. This requires the ordering and coordination of various events during nuclear division (mitosis) and cytoplasmic division (cytokinesis). How these events are temporally coordinated remains poorly understood. On page 332 of this issue, Afonso *et al.* (*1*) propose the exciting idea of a "chromosome separation checkpoint" that controls the near final step of mitosis—the anaphase to telophase transition (see the figure).



Mitosis produces two daughter cells.

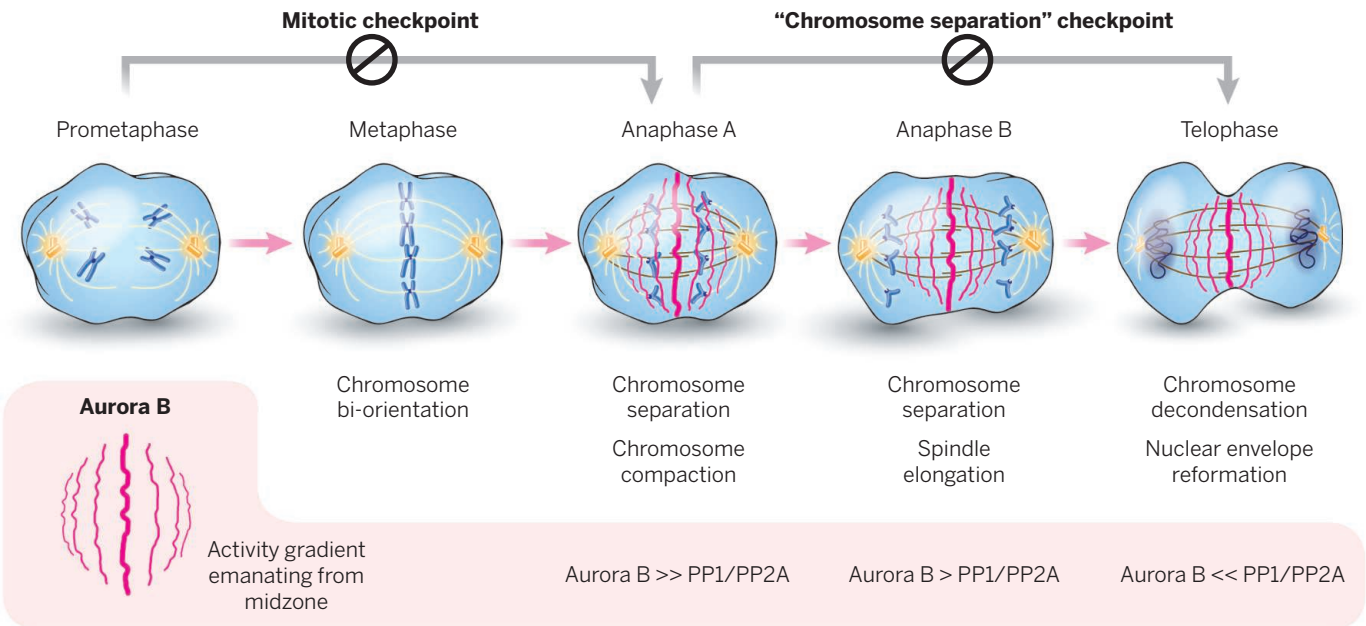
Once the duplicated chromosomes are bi-oriented on the mitotic spindle in metaphase, anaphase begins, in which the sister chromosomes start to separate. Chromosomes must then be well separated before the beginning of cytokinesis and telophase events, which include chromosome decondensation, nuclear envelope reformation (NER), and mitotic spindle disassembly. Elongation of the mitotic spindle facilitates this separation process. Anaphase onset before completion of metaphase gives rise to segregation errors, whereas the start of cytokinesis before clearance of chromosomes from the cleavage plane gives rise to DNA damage (*2*, *3*).

Afonso *et al.* found that treatments that impaired spindle elongation delayed the onset of chromosome decondensation and NER. This delay was dependent on a gradient of aurora B kinase–dependent phosphorylation that originated from the central spindle midzone (*4*). In addition, timely chromatin decondensation and NER required the activity of protein phosphatase 2A (PP2A) and protein phosphatase 1 (PP1). The findings suggest that when the spindle elongates, the separating packs of sister chromosomes are gradually pulled out of the midzone "sphere of influence" of aurora B. This would progressively tip the kinase-phosphatase balance in the region surrounding the chromosomes in favor of the phosphatases. The subsequent dephosphorylation of substrates results in chromatin decondensation and NER. In case of impaired spindle elongation, the separating chromosomes remain within the aurora B activity gradient for a longer period, thus delaying the onset of telophase.

A possible explanation for how aurora B may prevent chromatin decondensation and NER is that it directly phosphorylates substrates involved in chromatin condensation and in nuclear envelope disassembly. The condensin I complex is an interesting candidate because it is a direct substrate of aurora B and promotes chromosome compaction, and its recruitment onto chromosomes requires aurora B kinase activity (*5*–*8*). Indeed, Afonso *et al.* found that Barren—the homolog of condensin I in *Dro-*

Department of Molecular Cancer Research, Center for Molecular Medicine, University Medical Center Utrecht, Universiteitsweg 100, 3584 CG, Utrecht, Netherlands. E-mail: s.m.a.lens@umcutrecht.nl

Published by AAAS



Order in mitosis. The mitotic checkpoint prevents anaphase onset until all chromosomes are properly connected to the microtubules of the mitotic spindle. The "chromosome separation checkpoint" proposed by Afonso *et al*. delays telophase onset until the separating chromosome packs have been drawn out of the aurora B activity gradient emanating from the midzone. A change in the ratio of phosphatases (PP1 and PP2A) to aurora B kinase in the region surrounding the separating chromosomes may control the order of late mitotic events.

sophila melanogaster—dissociated from the separating packs of sister chromatids in late anaphase, but was retained on lagging chromosomes that lingered in proximity of midzone-localized aurora B. Importantly, depletion of Barren impaired spindle elongation but failed to delay NER, similar to what was seen when aurora B was inhibited. It is unknown whether components of the nuclear envelope are direct substrates of aurora B that must be dephosphorylated to allow NER. However, many of these components are substrates of cyclin B–cyclin dependent kinase 1 (Cdk1) (*9*, *10*), a protein complex considered to be the master regulator of mitosis and involved in nuclear envelope disassembly (*10*). Afonso *et al*. observed that blocking Cdk1 activity at anaphase onset accelerated NER, indicating that Cdk1-dependent phosphorylation in anaphase can also delay telophase events. Interestingly, in the Cdk1-inhibited cells, NER was still delayed on the chromosomes in proximity to the midzone. Perhaps aurora B maintains the phosphorylation status of Cdk1 substrates by regulating the phosphatases of these substrates. It would be interesting to see whether, for example, the recruitment of PP1 and PP2A to anaphase chromosomes is reduced on chromosomes that reside in the aurora B activity cloud.

Afonso *et al*. nicely establish that the midzone gradient of aurora B activity regulates chromosome decondensation and NER. But is this really a checkpoint in mitosis? The classic definition of a checkpoint (*11*) is a control mechanism that enforces the dependence of a late event in the cell division cycle on an earlier event. Moreover, there should be ways to override this dependence, thus permitting a late event to occur even when an early, normally prerequisite event is prevented. The "chromosome separation checkpoint" suggested by Afonso *et al*. seems to fulfill these criteria: It delays the onset of telophase until the chromosomes are well separated, and inhibition of aurora B relieves this dependence. But what are the consequences for genomic integrity in case this checkpoint is dysfunctional? Answering this question poses a serious challenge as it involves inhibiting the "chromosome separation checkpoint" function of aurora B without perturbing its other functions during (late) anaphase (*12*). This requires identification of the specific substrates involved.

If the "chromosome separation checkpoint" is a surveillance mechanism, then what is monitored? The most likely answer is that it senses the presence of chromatin in the midzone. If so, even one or a few chromosomes that lag behind in the midzone would delay telophase onset. However, although Afonso *et al*. note that NER was prevented on lagging chromosomes, these laggards did not prevent NER or decondensation of the packs of separating chromosomes. This could mean that rather than a checkpoint, the data support a model in which a gradual change in the ratio of phosphatases (PP1 and PP2A) to kinases (aurora B and Cdk1) results in sequential dephosphorylation of specific substrates, and thus, the ordering of events. This model is similar to that proposed for mitotic exit events in budding yeast (*13*).

The formation of two new nuclei after chromosome segregation necessitates that the nuclear envelope reforms around the two separated and compact clusters of chromosomes (*10*). The aurora B midzone gradient is in an excellent position to coordinate and order these events because it promotes compaction of individual chromosomes (*8*, *14*) and, as Afonso *et al*. report, it prevents NER before the chromosome packs are well separated. As such, the aurora B gradient helps a dividing cell face its ultimate challenge—ensuring that its two daughter cells end up with nuclei with identical genomes. ■

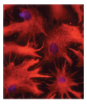
REFERENCES

1. O. Afonso *et al*., *Science* **345**, 332 (2014).
2. E. A. Foley, T. M. Kapoor, *Nat. Rev. Mol. Cell Biol.* **14**, 25 (2013).
3. A. Janssen, M. van der Burg, K. Szuhai, G. J. P. L. Kops, R. H. Medema, *Science* **333**, 1895 (2011).
4. B. G. Fuller *et al*., *Nature* **453**, 1132 (2008).
5. J. J. Lipp, T. Hirota, I. Poser, J.-M. Peters, *J. Cell Sci.* **120**, 1245 (2007).
6. R. Giet, D. M. Glover, *J. Cell Biol.* **152**, 669 (2001).
7. K. Tada, H. Susumu, T. Sakuno, Y. Watanabe, *Nature* **474**, 477 (2011).
8. F. Mora-Bermúdez, D. Gerlich, J. Ellenberg, *Nat. Cell Biol.* **9**, 822 (2007).
9. J. D. Blethrow, J. S. Glavy, D. O. Morgan, K. M. Shokat, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 1442 (2008).
10. S. Güttinger, E. Laurell, U. Kutay, *Nat. Rev. Mol. Cell Biol.* **10**, 178 (2009).
11. L. H. Hartwell, T. A. Weinert, *Science* **246**, 629 (1989).
12. M. S. van der Waal, R. C. Hengeveld, A. van der Horst, S. M. Lens, *Exp. Cell Res.* **318**, 1407 (2012).
13. C. Bouchoux, F. Uhlmann, *Cell* **147**, 8032011 (2011).
14. G. Neurohr *et al*., *Science* **332**, 465 (2011).

10.1126/science.1257267

ILLUSTRATION: C. BICKEL/*SCIENCE*

Published by AAAS

EXTENDED PDF FORMAT
SPONSORED BY

Travel Grants Available
Neuroscience 2014
Application Deadline: July 31st
>Apply Now
R&D systems
a bietechne brand
www.rndsystems.com

Charge transfer goes the distance

Stephen T. Pratt

Science **345**, 267 (2014);
DOI: 10.1126/science.1255943

This copy is for your personal, non-commercial use only.

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by clicking here.

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines here.

The following resources related to this article are available online at www.sciencemag.org (this information is current as of August 4, 2014):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:
http://www.sciencemag.org/content/345/6194/267.full.html

A list of selected additional articles on the Science Web sites **related to this article** can be found at:
http://www.sciencemag.org/content/345/6194/267.full.html#related

This article **cites 11 articles**, 3 of which can be accessed free:
http://www.sciencemag.org/content/345/6194/267.full.html#ref-list-1

This article appears in the following **subject collections:**
Chemistry
http://www.sciencemag.org/cgi/collection/chemistry

Downloaded from www.sciencemag.org on August 4, 2014

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published weekly, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. Copyright 2014 by the American Association for the Advancement of Science; all rights reserved. The title *Science* is a registered trademark of AAAS.

ULTRAFAST DYNAMICS

Charge transfer goes the distance

Electronic relaxation following x-ray excitation illuminates steps in molecular dissociation

By **Stephen T. Pratt**

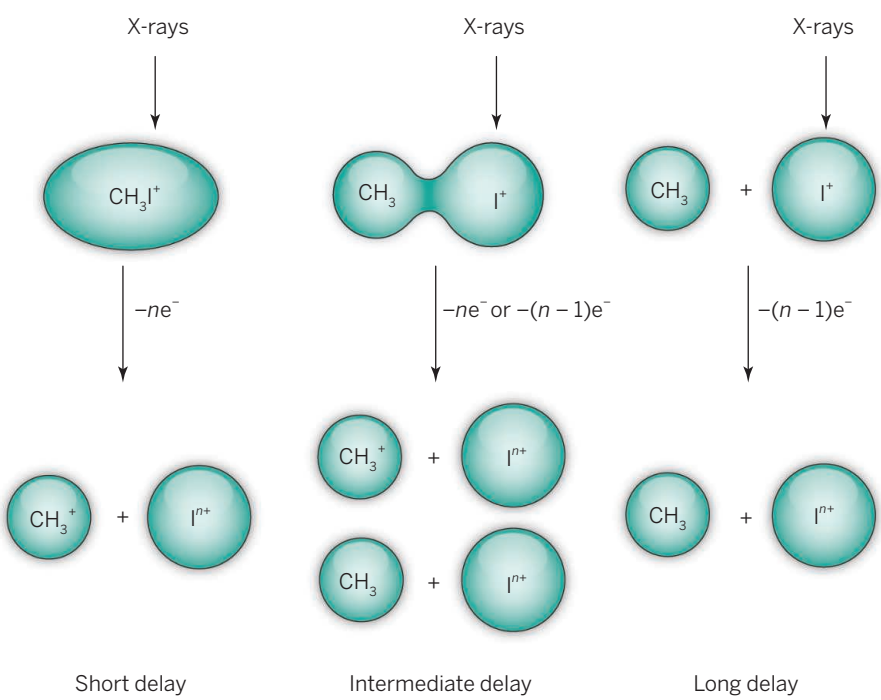
The strong covalent and ionic bonds that hold molecules together form through the sharing or transfer of electrons among the constituent atoms. The dynamics of the making and breaking of chemical bonds on the natural time scale of molecular vibrations (femtoseconds) can be followed by exciting the valence electrons in molecules with ultrafast pulses from lasers operating in the near-infrared, visible, and near-ultraviolet (*1*–*3*). However, valence electrons tend to be delocalized within molecules, so it is difficult to obtain site-specific information about bonding in this way. Ultrafast extreme ultraviolet and x-ray sources (*4*, *5*), as well as complementary ultrafast sources of electrons, now provide access to the more localized inner-shell electrons of selected atoms (*6*–*8*). As reported on page 288 of this issue, Erk *et al.* (*9*) have determined the distance at which electron or charge transfer between the fragments of a dissociating molecule becomes negligible by using a combination of ultrafast x-ray pulses, an ultrafast near-infrared laser, and a high-resolution ion-imaging spectrometer.

Erk *et al.* used an intense near-infrared laser to initiate the dissociation of methyl iodide (CH_3I). This excitation method leads primarily to three different sets of dissociation products: $CH_3^+ + I^+$, $CH_3^+ + I$, or $CH_3 + I^+$. The three processes creating these products can be readily distinguished by measuring the kinetic energies of the resulting fragments. With the near-infrared pulse defining time zero, an x-ray laser pulse was used to eject an electron from an inner-shell orbital of the neutral I atom or I^+ cation after a variable time delay. Knowledge of the potential energy curves and the measured kinetic energies allows the measured time delays to be correlated with the CH_3–I distance; thus, the decays of the inner-shell hole could be studied as a function of this distance.

The inner-shell orbital is not associated with bonding in CH_3I, and is well localized on the I atom, as is the inner-shell hole (electron vacancy) produced by the x-rays. This hole is filled by a cascade of Auger processes (*10*), in which an electron from a higher-lying (more weakly bound) orbital fills the hole, and a second electron from the same or another orbital is ejected to create a more positively charged molecule. Although some of these higher-lying orbitals are also localized on the initially excited atom, eventually the delocalized valence orbitals enter the cascade. The resulting highly charged molecule is unstable, and it dissociates. The bond is broken when electrons are no longer shared between the fragments.

For the dissociation path producing CH_3 and I^+, the holes produced in the valence orbitals by the electron cascade lead to some probability for charge to end up on the CH_3, and for the molecule to dissociate to $CH_3^+ + I^{n+}$, where n^+ is the charge state of the multiply charged ion (typically $n \geq 4$; see the figure). The repulsion created by positive charge on both species leads to a "Coulomb explosion" and causes the fragments to gain considerably more kinetic energy than the case where the CH_3 is neutral. At the shortest delays, CH_3I^+ has not yet begun to dissociate when the core hole is produced. When the valence shell holes appear after rapid Auger decay, there is substantial probability for producing $CH_3^+ + I^{n+}$. At very long delays, the molecule is fully dissociated to $CH_3 + I^+$ before the x-ray pulse arrives, so the valence orbitals are fully separated into those of the two fragments, and no charge appears on the CH_3. At intermediate times, the overlap of the fragment orbitals decreases with increasing separation, resulting in the vanishing formation of CH_3^+ that signals the "breaking" of the bonding between fragments.

The critical distance between the CH_3 and highly charged I^{n+} fragments was estimated with a simple molecular model (*11*) originally developed for ion-atom collisions that considers only the Coulomb potentials of the separating fragments and the binding energy of the most weakly bound electron. The success of this model not only supports the analysis of Erk *et al.* but also



After the breakup. In the experiment of Erk *et al.*, a near-infrared laser pulse creates CH_3I^+ in a dissociative state. As the CH_3 and I^+ fragments separate, a delayed x-ray laser pulse excites an inner-shell electron of I^+. This event sets off a cascade of processes that ultimately ejects a number of both inner-shell and valence electrons. The number of electrons will in general depend on the details of the decay process. At short delay, charge is produced on the CH_3 fragment, resulting in CH_3^+, whereas at long delay, the valence orbitals of the fragments are separated before the x-ray pulse, and the CH_3 is not ionized. At intermediate delays, the transition from CH_3^+ to CH_3 products is observed, signaling the reduction in overlap of the valence orbitals of the CH_3 and I^{n+} fragments.

ILLUSTRATION: P. HUEY/*SCIENCE*

Argonne National Laboratory, Argonne, IL 60439, USA. E-mail: stpratt@anl.gov

Published by AAAS

illuminates the close connection between the core-hole relaxation dynamics in photoexcitation and collision studies with highly charged ions. Of course, some ambiguity in the specific nature of the process remains. For example, electron transfer from the CH_3 to the I^{n+} can precede the ejection of an outer electron of the latter, or an interfragment process can occur in which an outer electron of the I^{n+} fills the hole and the energy released results in the ejection of an electron from the CH_3. Both processes depend on the overlap of the fragment wave functions, but could be distinguished in future experiments by measuring the energies of the ejected electrons.

Mapping the changing Auger processes as a function of the interfragment distance will also provide considerable insight into how the electronic structure changes along the reaction coordinate—information that is relevant for understanding processes like intermolecular Coulombic decay (*12*). Such studies should reveal additional complexity when the time scales for Auger decay and motion of the fragments are similar (*13*). In this case, the separation of the electronic versus vibrational and translational motions embodied by the Born-Oppenheimer approximation breaks down, along with the traditional view of potential energy surfaces, and alternative perspectives will be required. Future experiments will certainly address these complexities in more detail. For now, Erk *et al.* provide a tantalizing glimpse of how x-ray free-electron lasers and advanced laser and detector technologies can provide new insight into the mechanisms of molecular physics and chemistry. ■

REFERENCES AND NOTES

1. A. H. Zewail, in *Nobel Lectures, Chemistry 1996–2000*, I. Genthe, Ed. (World Scientific, Singapore, 1993), pp. 274–367.
2. A. Stolow, J. G. Underwood, *Adv. Chem. Phys.* **139**, 497 (2008).
3. O. Bixner *et al.*, *J. Chem. Phys.* **136**, 204503 (2012).
4. T. Popmintchev *et al.*, *Science* **336**, 1287 (2012).
5. R. W. Schoenlein *et al.*, *Science* **287**, 2237 (2000).
6. C. Rose-Petruck *et al.*, *Nature* **398**, 310 (1999).
7. X-ray free-electron laser (XFEL) sources for femtosecond x-ray pulses include the Linac Coherent Light Source (LCLS) in the United States, the European XFEL, and the SPRing-8 Angstrom Compact free-electron laser in Japan (*8*).
8. C. Bostedt *et al.*, *J. Phys. B* **46**, 164003 (2013).
9. B. Erk *et al.*, *Science* **345**, 288 (2014).
10. V. Jonauskas *et al.*, *J. Phys. B* **36**, 4403 (2003).
11. A. Niehaus, *J. Phys. B* **19**, 2925 (1986).
12. J. Zobeley, L. S. Cederbaum, F. Tarantelli, *J. Phys. Chem. A* **103**, 11145 (1999).
13. M. Eroms, O. Vendrell, M. Jungen, H. D. Meyer, L. S. Cederbaum, *J. Chem. Phys.* **130**, 154307 (2009).

ACKNOWLEDGMENTS

Supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, Division of Chemical Sciences, Geosciences, and Biosciences under contract DE-AC02-06CH11357.

10.1126/science.1255943

TRANSLATIONAL MEDICINE

Improving cardiac rhythm with a biological pacemaker

Heart muscle cells in a large animal model are reprogrammed to restore heart rate and function

By **Nikhil V. Munshi**[1,2,3] *and* **Eric N. Olson**[3,4]

Electronic device therapies, including implantable pacemakers and defibrillators, have revolutionized the management of cardiovascular disease (*1*). For example, patients with certain slow cardiac rhythms (bradycardia) can experience exercise intolerance, easy fatigability, or circulatory collapse. Given that currently available drugs cannot safely and sustainably elevate heart rate, the only proven treatment for symptomatic bradycardia is permanent pacemaker implantation. Contemporary electronic pacemakers have an extended battery life, contain leads that minimize inflammation and scarring, and possess advanced algorithms to contend with heart rate elevations during exercise. These features allow pacemakers to improve longevity and quality of life in patients who require them. But electronic pacemakers cannot recapitulate all aspects of the endogenous sinoatrial node, the dominant pacemaker in the uninjured heart. In this regard, a recent study by Hu *et al.* (*2*) demonstrates the feasibility of a somatic cell reprogramming strategy for creating a biological pacemaker in a large animal preclinical model, raising prospects for clinical translation.

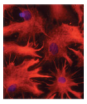
Although the efficacy of electronic pacemakers is clear, there remain specific clinical situations where a biological pacemaker might be advantageous. For example, pacemaker infections necessitate removal of all device-associated equipment (i.e., battery and leads) while the patient is treated with antibiotics (*3*). However, this is problematic for pacemaker-dependent patients during the period of treatment (typically 2 to 6 weeks). Currently, a temporary pacemaker is placed for this indication, but implantation of additional hardware is not an ideal solution given the high likelihood for re-infection. Many view this scenario as the perfect niche for a temporary biological pacemaker (*4*). A biological pacemaker should theoretically exhibit more physiological autonomic responses and accommodate for growth, which is especially important in pediatric patients. Such reasoning provides strong rationale for developing a biological pacemaker.

Previous attempts to create a biological pacemaker have focused on three general approaches: introduction of specific ion channels into heart muscle cells (cardiomyocytes) by gene transfer (transduction), ion channel expression in non-cardiomyocytes followed by cell fusion to native cardiomyocytes in situ, and the introduction of stem cells that have been previously differentiated into cardiomyocytes. In an early iteration of ion channel transduction, a dominant negative potassium channel (which antagonizes the native channel) was introduced directly into ventricular myocardium in a small animal preclinical model, resulting in a transient escape rhythm (*5*). Because such an approach required direct delivery of the gene encoding the potassium channel into cardiomyocytes, cell fusion strategies were developed to optimize the properties of heterologous "delivery" cells prior to their introduction into preclinical animal models (*6*–*8*). However, these approaches may not recapitulate additional as-yet uncharacterized features of pacemaker cells. To circumvent this limitation, embryonic stem cells were differentiated in vitro into cardiomyocytes with pacemaker activity and then introduced directly into the heart of pigs or guinea pigs (*9*, *10*). Although this strategy proved successful, the teratogenic potential and heterogeneity of differentiated embryonic stem cells have slowed clinical translation.

> ***"...the results...provide an encouraging indication that a biological pacemaker might eventually be ready for human translation."***

Recently, direct lineage reprogramming has emerged as a means for converting one cell type into another cell type, including pluripotent stem cells, neurons, and cardio-

Published by AAAS

EXTENDED PDF FORMAT
SPONSORED BY

Travel Grants Available
Neuroscience 2014
Application Deadline: July 31st
>Apply Now
R&D systems
a bio-techne brand
www.rndsystems.com

Improving cardiac rhythm with a biological pacemaker

Nikhil V. Munshi and Eric N. Olson

Science **345**, 268 (2014);
DOI: 10.1126/science.1257976

This copy is for your personal, non-commercial use only.

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by clicking here.

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines here.

The following resources related to this article are available online at www.sciencemag.org (this information is current as of August 4, 2014):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:
http://www.sciencemag.org/content/345/6194/268.full.html

A list of selected additional articles on the Science Web sites **related to this article** can be found at:
http://www.sciencemag.org/content/345/6194/268.full.html#related

This article **cites 15 articles**, 9 of which can be accessed free:
http://www.sciencemag.org/content/345/6194/268.full.html#ref-list-1

This article appears in the following **subject collections:**
Medicine, Diseases
http://www.sciencemag.org/cgi/collection/medicine

Downloaded from www.sciencemag.org on August 4, 2014

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published weekly, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. Copyright 2014 by the American Association for the Advancement of Science; all rights reserved. The title *Science* is a registered trademark of AAAS.

illuminates the close connection between the core-hole relaxation dynamics in photoexcitation and collision studies with highly charged ions. Of course, some ambiguity in the specific nature of the process remains. For example, electron transfer from the CH_3 to the I^{n+} can precede the ejection of an outer electron of the latter, or an interfragment process can occur in which an outer electron of the I^{n+} fills the hole and the energy released results in the ejection of an electron from the CH_3. Both processes depend on the overlap of the fragment wave functions, but could be distinguished in future experiments by measuring the energies of the ejected electrons.

Mapping the changing Auger processes as a function of the interfragment distance will also provide considerable insight into how the electronic structure changes along the reaction coordinate—information that is relevant for understanding processes like intermolecular Coulombic decay (*12*). Such studies should reveal additional complexity when the time scales for Auger decay and motion of the fragments are similar (*13*). In this case, the separation of the electronic versus vibrational and translational motions embodied by the Born-Oppenheimer approximation breaks down, along with the traditional view of potential energy surfaces, and alternative perspectives will be required. Future experiments will certainly address these complexities in more detail. For now, Erk *et al.* provide a tantalizing glimpse of how x-ray free-electron lasers and advanced laser and detector technologies can provide new insight into the mechanisms of molecular physics and chemistry. ■

REFERENCES AND NOTES

1. A. H. Zewail, in *Nobel Lectures, Chemistry 1996–2000*, I. Genthe, Ed. (World Scientific, Singapore, 1993), pp. 274–367.
2. A. Stolow, J. G. Underwood, *Adv. Chem. Phys.* **139**, 497 (2008).
3. O. Bixner *et al.*, *J. Chem. Phys.* **136**, 204503 (2012).
4. T. Popmintchev *et al.*, *Science* **336**, 1287 (2012).
5. R. W. Schoenlein *et al.*, *Science* **287**, 2237 (2000).
6. C. Rose-Petruck *et al.*, *Nature* **398**, 310 (1999).
7. X-ray free-electron laser (XFEL) sources for femtosecond x-ray pulses include the Linac Coherent Light Source (LCLS) in the United States, the European XFEL, and the SPRing-8 Angstrom Compact free-electron laser in Japan (*8*).
8. C. Bostedt *et al.*, *J. Phys. B* **46**, 164003 (2013).
9. B. Erk *et al.*, *Science* **345**, 288 (2014).
10. V. Jonauskas *et al.*, *J. Phys. B* **36**, 4403 (2003).
11. A. Niehaus, *J. Phys. B* **19**, 2925 (1986).
12. J. Zobeley, L. S. Cederbaum, F. Tarantelli, *J. Phys. Chem. A* **103**, 11145 (1999).
13. M. Eroms, O. Vendrell, M. Jungen, H. D. Meyer, L. S. Cederbaum, *J. Chem. Phys.* **130**, 154307 (2009).

ACKNOWLEDGMENTS

Supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, Division of Chemical Sciences, Geosciences, and Biosciences under contract DE-AC02-06CH11357.

10.1126/science.1255943

TRANSLATIONAL MEDICINE

Improving cardiac rhythm with a biological pacemaker

Heart muscle cells in a large animal model are reprogrammed to restore heart rate and function

By **Nikhil V. Munshi**[1,2,3] *and* **Eric N. Olson**[3,4]

Electronic device therapies, including implantable pacemakers and defibrillators, have revolutionized the management of cardiovascular disease (*1*). For example, patients with certain slow cardiac rhythms (bradycardia) can experience exercise intolerance, easy fatigability, or circulatory collapse. Given that currently available drugs cannot safely and sustainably elevate heart rate, the only proven treatment for symptomatic bradycardia is permanent pacemaker implantation. Contemporary electronic pacemakers have an extended battery life, contain leads that minimize inflammation and scarring, and possess advanced algorithms to contend with heart rate elevations during exercise. These features allow pacemakers to improve longevity and quality of life in patients who require them. But electronic pacemakers cannot recapitulate all aspects of the endogenous sinoatrial node, the dominant pacemaker in the uninjured heart. In this regard, a recent study by Hu *et al.* (*2*) demonstrates the feasibility of a somatic cell reprogramming strategy for creating a biological pacemaker in a large animal preclinical model, raising prospects for clinical translation.

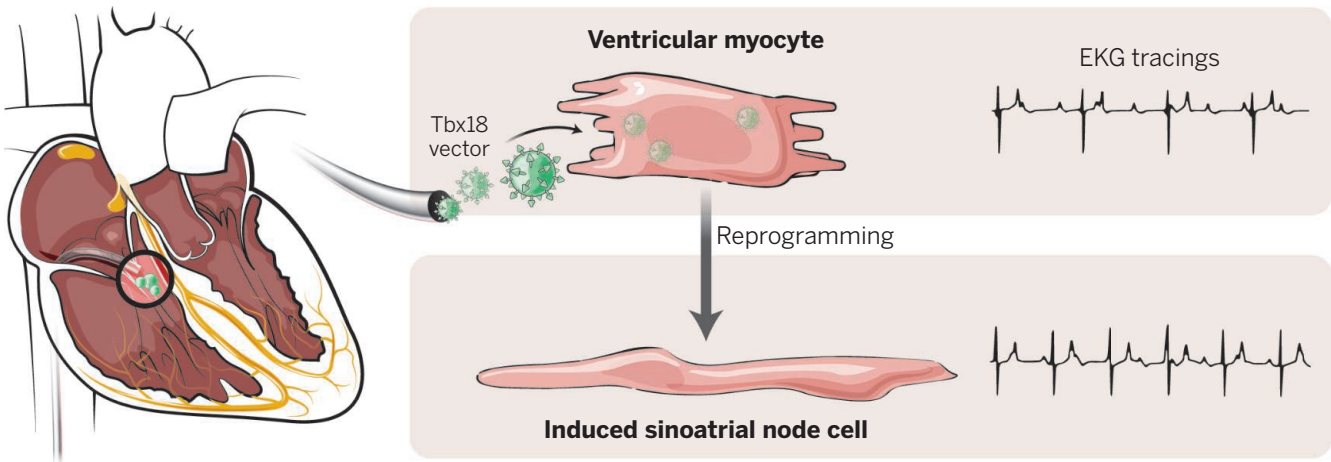
Although the efficacy of electronic pacemakers is clear, there remain specific clinical situations where a biological pacemaker might be advantageous. For example, pacemaker infections necessitate removal of all device-associated equipment (i.e., battery and leads) while the patient is treated with antibiotics (*3*). However, this is problematic for pacemaker-dependent patients during the period of treatment (typically 2 to 6 weeks). Currently, a temporary pacemaker is placed for this indication, but implantation of additional hardware is not an ideal solution given the high likelihood for re-infection. Many view this scenario as the perfect niche for a temporary biological pacemaker (*4*). A biological pacemaker should theoretically exhibit more physiological autonomic responses and accommodate for growth, which is especially important in pediatric patients. Such reasoning provides strong rationale for developing a biological pacemaker.

Previous attempts to create a biological pacemaker have focused on three general approaches: introduction of specific ion channels into heart muscle cells (cardiomyocytes) by gene transfer (transduction), ion channel expression in non-cardiomyocytes followed by cell fusion to native cardiomyocytes in situ, and the introduction of stem cells that have been previously differentiated into cardiomyocytes. In an early iteration of ion channel transduction, a dominant negative potassium channel (which antagonizes the native channel) was introduced directly into ventricular myocardium in a small animal preclinical model, resulting in a transient escape rhythm (*5*). Because such an approach required direct delivery of the gene encoding the potassium channel into cardiomyocytes, cell fusion strategies were developed to optimize the properties of heterologous "delivery" cells prior to their introduction into preclinical animal models (*6*–*8*). However, these approaches may not recapitulate additional as-yet uncharacterized features of pacemaker cells. To circumvent this limitation, embryonic stem cells were differentiated in vitro into cardiomyocytes with pacemaker activity and then introduced directly into the heart of pigs or guinea pigs (*9*, *10*). Although this strategy proved successful, the teratogenic potential and heterogeneity of differentiated embryonic stem cells have slowed clinical translation.

> ***"...the results...provide an encouraging indication that a biological pacemaker might eventually be ready for human translation."***

Recently, direct lineage reprogramming has emerged as a means for converting one cell type into another cell type, including pluripotent stem cells, neurons, and cardio-

Published by AAAS



Biological pacemaker. Tbx18-containing adenoviral particles can be delivered percutaneously in pigs that have a slow cardiac rhythm. Over a 1-week period, cardiomyocytes are reprogrammed into pacemaker cells at the site of virus injection and animals show increased heart rate, clinical improvement, and no obvious adverse sequelae.

myocytes (*11*). The conversion of fibroblasts into mouse and human cardiomyocyte-like cells has been well-documented (*12*, *13*), although it remains inefficient. However, a method for generating pacemaker cells in this way has not yet been identified. Alternatively, interconversion of distinct cardiomyocyte cell types (e.g., ventricular to pacemaker and atrial to pacemaker) has been observed in vitro and in vivo (*14*, *15*). Specifically, the human developmental transcription factor T-box 18 (Tbx18) can reprogram ventricular myocytes directly into induced sinoatrial node (iSAN) cells with many features characteristic of endogenous pacemaker cells. Moreover, this somatic reprogramming into iSAN cells has been achieved in vivo in a small animal preclinical model, resulting in durable pacemaker activity.

The study by Hu *et al.* represents the next logical step toward clinical translation of somatic reprogramming to create a biological pacemaker. The authors used an adenovirus vector to transduce the gene encoding human Tbx18 into ventricular cardiomyocytes of pigs. The procedure involved viral delivery via a specialized catheter threaded through a vein into the right ventricle. They observed meaningful pacemaker activity in the context of experimentally induced complete heart block, a clinical condition that would normally require pacemaker implantation in humans. From a functional standpoint, relative to control animals, the pigs receiving Tbx18 showed an increased heart rate with less dependence on a backup electronic pacemaker. Pacemaker activity in pigs that received Tbx18 showed diurnal variation (higher rate during the day and lower rate at night) and appropriate autonomic responses (e.g., increased rate in response to isoproterenol infusion), relative to control animals. Furthermore, the authors provide evidence that pacemaker activity emanates from the injection site, indicating that reprogrammed cardiomyocytes are indeed responsible for the observed clinical improvements. No increase in arrhythmias was observed in Tbx18-recipient animals, and all measured clinical parameters appeared similar to those of control animals.

The demonstration of stable pacemaker activity induced through percutaneous venous delivery in a clinically relevant large-animal model suggests that somatic reprogramming may be a viable strategy for creating a biological pacemaker. Prior to nonhuman primate studies and possible phase I clinical trials in humans, however, several hurdles will need to be cleared. Although Hu *et al.* suggest that the transient nature of adenoviral infection might be advantageous for a temporary biological pacemaker after removal of an infected device, it remains unclear whether the observed durability will be sufficient for prolonged antibiotic regimens (e.g., 6 weeks for pacemaker-related endocarditis). The Tbx18-induced biological pacemaker was effective for up to 14 days in the pig model. The lack of clinically adverse effects in the study is encouraging, but it remains to be seen whether viral dispersion leads to multiple pacemaker foci within the heart or whether the low-level infection in the lung and spleen that was observed might have long-term consequences. Moreover, extended studies will be needed to evaluate whether adenoviral infection leads to untoward inflammation or fibrosis at the site of injection, which may lead to dangerous ventricular arrhythmias or depressed myocardial function. Given these potential limitations, the development of alternative delivery systems capable of more durable gene expression may expand the possible indications for a biological pacemaker in the future.

Because Tbx18 is a pleiotropic developmental regulator, it could induce alternative cell fates in vivo that have escaped detection and are potentially dangerous. Nonetheless, the results of Hu *et al.* represent a critical step toward potentially filling an important clinical niche and provide an encouraging indication that a biological pacemaker might eventually be ready for human translation. ■

[1]*Department of Internal Medicine, Division of Cardiology, UT Southwestern Medical Center, Dallas, TX 75390, USA.* [2]*McDermott Center for Human Growth and Development, UT Southwestern Medical Center, Dallas, TX 75390, USA.* [3]*Department of Molecular Biology, UT Southwestern Medical Center, Dallas, TX 75390, USA.* [4]*Center for Regenerative Science and Medicine, UT Southwestern Medical Center, Dallas, TX 75390, USA. E-mail: nikhil.munshi@utsouthwestern.edu; eric.olson@utsouthwestern.edu*

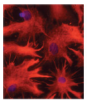
ILLUSTRATION: V. ALTOUNIAN

REFERENCES

1. R. Lampert, *Circulation* **128**, 1576 (2013).
2. Y. F. Hu, J. F. Dawkins, H. C. Cho, E. Marbán, E. Cingolani, *Sci. Transl. Med.* **6**, 245ra94 (2014).
3. L. M. Baddour, Y. M. Cha, W. R. Wilson, *N. Engl. J. Med.* **367**, 842 (2012).
4. H. C. Cho, E. Marbán, *Circ. Res.* **106**, 674 (2010).
5. J. Miake, E. Marbán, H. B. Nuss, *Nature* **419**, 132 (2002).
6. I. Potapova *et al.*, *Circ. Res.* **94**, 952 (2004).
7. A. N. Plotnikov *et al.*, *Circulation* **116**, 706 (2007).
8. H. C. Cho, Y. Kashiwakura, E. Marbán, *Circ. Res.* **100**, 1112 (2007).
9. I. Kehat *et al.*, *Nat. Biotechnol.* **22**, 1282 (2004).
10. T. Xue *et al.*, *Circulation* **111**, 11 (2005).
11. S. M. Chambers, L. Studer, *Cell* **145**, 827 (2011).
12. M. Ieda *et al.*, *Cell* **142**, 375 (2010).
13. Y. J. Nam *et al.*, *Proc. Acad. Natl. Sci. U.S.A.* **110**, 5588 (2013).
14. N. Kapoor, W. Liang, E. Marbán, H. C. Cho, *Nat. Biotechnol.* **31**, 54 (2013).
15. M. L. Bakker *et al.*, *Cardiovasc. Res.* **94**, 439 (2012).

10.1126/science.1257976

Published by AAAS

EXTENDED PDF FORMAT
SPONSORED BY

Travel Grants Available
Neuroscience 2014
Application Deadline: July 31st
>Apply Now
R&D systems
a bio-techne brand
www.rndsystems.com

Creating antioxidants by oxidation catalysis

Boris J. Nachtsheim

Science **345**, 270 (2014);
DOI: 10.1126/science.1257347

This copy is for your personal, non-commercial use only.

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by clicking here.

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines here.

The following resources related to this article are available online at www.sciencemag.org (this information is current as of August 4, 2014):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:
http://www.sciencemag.org/content/345/6194/270.full.html

A list of selected additional articles on the Science Web sites **related to this article** can be found at:
http://www.sciencemag.org/content/345/6194/270.full.html#related

This article **cites 10 articles**, 3 of which can be accessed free:
http://www.sciencemag.org/content/345/6194/270.full.html#ref-list-1

This article appears in the following **subject collections:**
Chemistry
http://www.sciencemag.org/cgi/collection/chemistry

Downloaded from www.sciencemag.org on August 4, 2014

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published weekly, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. Copyright 2014 by the American Association for the Advancement of Science; all rights reserved. The title *Science* is a registered trademark of AAAS.

CHEMISTRY

Creating antioxidants by oxidation catalysis

A key component of vitamin E can be synthesized without use of expensive transition metal catalysts

By **Boris J. Nachtsheim**

Tocopherols are naturally occurring fat-soluble antioxidants that protect living cells against highly reactive free radicals. The best-known example is *D*-α-tocopherol, the most biologically active ingredient in vitamin E. Their potent antioxidative properties result from the 6-hydroxy chroman (6-chromanol) motif, which is a hidden hydroquinone. The 6-chromanol unit readily reacts with fatty acid peroxide radicals to give stabilized chromanoxyl radicals and far less reactive hydroperoxides (*1*); both intermediates can then be degraded by the cell safely. Because of their unique antioxidative properties, tocopherols are widely used components in human and animal nutrition additives. On page 291 of this issue, Uyanik *et al.* (*2*) report a low-cost synthesis method for making tocopherols with high biological activity.

Institut für Organische Chemie, Eberhard Karls Universität Tübingen, Auf der Morgenstelle 18, 72076 Tübingen, Germany. E-mail: boris.nachtsheim@uni-tuebingen.de

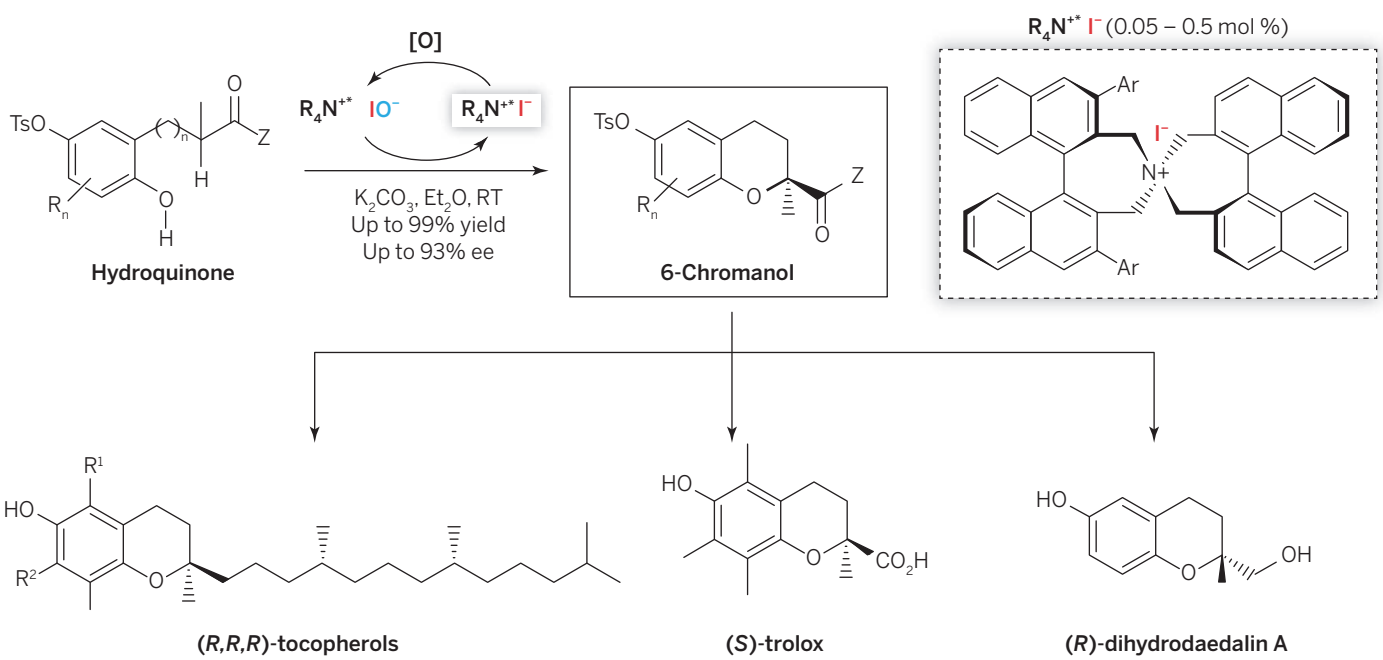
The high worldwide demand for tocopherols cannot be fulfilled by natural sources, and the vast majority of tocopherols are therefore made synthetically (*3*). Even though the naturally occurring *R,R,R*-enantiomers of tocopherols have by far the highest biological activity, current industrial processes produce their racemic forms (all rac-α-tocopherols) as a mixture of eight stereoisomers. Only four of the eight isomers show biological activity; the other four derivatives are wasted material.

Chemists around the world have developed novel methods for the enantioselective synthesis of chromans, which can be used to exclusively construct the biologically active *R,R,R* enantiomer of tocopherols (*4*). Unfortunately, most existing methods are based on cycloetherifications of allyl alcohols or alkenes, which are catalyzed by expensive and toxic transition metals such as palladium or ruthenium and require high catalyst loadings and expensive chiral ligands. Furthermore, the synthesis of the prochiral precursors can be tedious (*5*–*7*).

Uyanik *et al.* now describe a transition metal–free synthesis of 6-chromanols through an enantioselective oxidative cycloetherification of hydroquinones (see the figure). As their oxidation catalyst, the authors use chiral tetrabutylammonium iodides [$R_4N^{+*}\ I^-$] in remarkably low catalyst loadings (0.5 mol % or less), together with various organic or inorganic peroxides as stoichiometric co-oxidants. Using this combination, 6-chromanols can be synthesized efficiently in very high yields (up to 99%) and remarkable enantioselectivities (up to 93% enantiomeric excess).

Most intriguing about this approach is the fact that a 6-chromanol—the key structural element of perhaps the most abundant antioxidants in nature—can be constructed via oxidation catalysis in a medium full of oxidants. The authors construct the 6-chromanol in the natural *R* configuration in a peroxide-containing medium that is much more concentrated than those usually encountered by tocopherols in natural environments. This is like constructing a fire station in a burning forest with wooden tools.

The key to success was the introduction of a highly effective protecting group at the hydroquinone hydroxyl (OH) group. This OH group is not involved in the key oxidative C-O bond-forming step. Its derivatization with an electron-withdrawing tosyl group (tosyl = 4-toluenesulfonyl) reduces the reduction potential of the hydroquinone so much that no undesired side



Oxidative cyclization leads to versatile chroman derivatives. Uyanik *et al.* show that *N*-spirocyclic tetrabutylammonium iodides [$R_4N^{+*}\ I^-$] can be used in extremely small amounts to generate 6-chromanols. The chiral 6-chromanol core can then be used for the efficient synthesis of tocopherols and other biologically active compounds.

Published by AAAS

oxidation to the corresponding quinone can occur, neither in the starting material nor in the final 6-chromanol. This trick also makes the synthesis of the tosyl derivatives much easier than that of their silyl-protected hydroquinone analogs. Furthermore, the method provides efficient synthetic access to other chroman-derived antioxidants, such as trolox (a water-soluble analog of vitamin E) and dihydrodaedalin A (a potent kinase inhibitor) (see the figure).

With the same oxidative method, dihydrobenzofurans can also be generated with even lower catalyst loadings of 0.05%, implying catalyst turnover numbers (TONs) of 2000. This value is extremely high for transition metal–free oxidation catalysis and is competitive with the most active homogeneous transition metal catalysts. These high TONs could only be observed when potassium carbonate (K_2CO_3) was added as a base. Uyanik *et al.* proposed recently that the combination of a tetraalkylammonium iodide with an inorganic or organic peroxide leads to the formation of hypoiodites (IO^- or IOH) as the key catalytically active species in this type of oxidative C-O bond–forming reaction (*8*–*10*). Using Raman spectroscopy, they now demonstrate the presence of hypoiodite and hypoiodous acid in the reaction mixture. In addition, they observed triiodide anion (I_3^-), a stable but catalytically inactive species. The authors further show that triiodide can be forced to form a hypoiodite again by a base additive, explaining the substantial effect of base addition in this reaction.

Beyond the practical aspects of this metal-free oxidative cycloetherification, which gives fast access to various biologically active 6-chromanol derivatives, the mechanistic investigations should lead to a more profound understanding of iodide-catalyzed oxidative transformations. Finally, Uyanik *et al.*'s study shows that oxidation catalysis (in particular, hypoiodite-mediated oxidation catalysis) can be a relatively mild process, allowing even antioxidants to be generated in an efficient way in an oxidative environment. ■

REFERENCES

1. C. Schneider, *Mol. Nutr. Food Res.* **49**, 7 (2005).
2. M. Uyanik, H. Hayashi, K. Ishihara, *Science* **345**, 291 (2014).
3. T. Netscher, *Vitam. Horm.* **76**, 155 (2007).
4. H. C. Shen, *Tetrahedron* **65**, 3931 (2009).
5. L. F. Tietze, K. M. Sommer, J. Zinngrebe, F. Stecker, *Angew. Chem. Int. Ed.* **44**, 257 (2005).
6. S. Tanaka, T. Seki, M. Kitamura, *Angew. Chem. Int. Ed.* **48**, 8948 (2009).
7. B. M. Trost, H. C. Shen, L. Dong, J.-P. Surivet, C. Sylvain, *J. Am. Chem. Soc.* **126**, 11966 (2004).
8. M. Uyanik, H. Okamoto, T. Yasui, K. Ishihara, *Science* **328**, 1376 (2010).
9. A. N. French, *Science* **328**, 1365 (2010).
10. M. Uyanik, K. Ishihara, *ChemCatChem* **4**, 177 (2012).

10.1126/science.1257347

ILLUSTRATION: P. HUEY/*SCIENCE*; ADAPTED FROM (*16*)

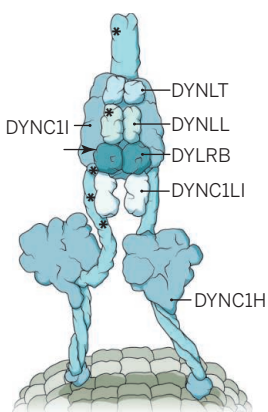
CELL BIOLOGY

One, two, three, cytoplasmic dynein is go!

The dynein motor protein must be part of a three-way complex to be fully active

By **Viki Allan**

Molecular motors are protein machines that power the movement of many different cargoes along filamentous tracks within eukaryotic cells. The dyneins move along microtubules, with the axonemal branch of the protein family working in large teams to drive the familiar beating of cilia and flagella by causing sliding of microtubules against each other. The cytoplasmic dyneins, on the other hand, move cargoes—including chromosomes, membrane-bounded organelles, mRNA particles, and protein complexes—along single microtubules (*1*). An important difference between cytoplasmic dynein 1 and axonemal dyneins is that as few as one to five molecules of cytoplasmic dynein 1 are enough to move small cargoes, whereas axonemal dyneins always work together. Another distinction is that cytoplasmic dynein 1 needs the help of an additional complex—dynactin—to function in the cell (*2*). Exactly what dynactin does, however, has been controversial. Two new papers, McKenney *et al.* on page 337 of this issue (*3*) and Schlager *et al.* (*4*), identify a missing piece of the puzzle—specific adaptor proteins—that helps explain how dynactin and cytoplasmic dynein 1 generate a single motor assembly that can move processively for many micrometers along the microtubule before falling off.



A cytoplasmic motor... The motor subunit of cytoplasmic dynein (DYNC1H) generates the force to power movement. Five nonmotor subunits form the domain that interacts with cargo. Asterisks mark the sites of mutations linked to human disease. A mouse mutation that reduces force production is arrowed.

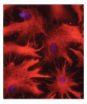
Cytoplasmic dynein 1 (called dynein hereafter) is a large (1.6 MD) protein complex that contains two copies each of the massive motor subunit (~450 kD) and five other subunits (see the first figure). The other subunits associate with the N-terminal region of the motor subunits to form a "tail" that binds to cargoes (*1*). Dynactin was purified as an activator of dynein-driven membrane movement, and it is also a large complex consisting of 11 different subunits (*2*) (see the second figure). Dynactin is needed for almost every dynein function in vivo (*1*, *2*), and although the $p150^{Glued}$ subunit of dynactin interacts with the dynein intermediate chain, it has been difficult to find evidence for stable association of the native complexes.

Several theories have been proposed to explain dynein's requirement for dynactin. First, because dynactin binds to some cargoes, it might simply drive dynein recruitment (*2*). Second, because $p150^{Glued}$ has a microtubule binding domain, it might hold on to the microtubule as dynein travels along it, so increasing dynein's processivity (*5*, *6*). However, there is evidence against both these models [reviewed in (*1*)]. Finally, dynactin might somehow directly activate dynein's motor activity, without the need for $p150^{Glued}$ to bind to microtubules (*7*).

An important step forward came with the discovery that mixing purified mammalian dynein and dynactin with Bicaudal D2 N-terminal domain (BicD2N) led to the formation of a stable complex of all three components (*8*). BicD2 in mammals helps recruit dynein to membranes in the secretory pathway and to the nuclear envelope (*8*), whereas the *Drosophila* homolog, BicD, plays a role in recruiting dynein to mRNPs (*9*). The two new studies have investigated the effects of BicD2N on dynein motility, using dynein purified from brain or cultured cells (*3*) or human dynein produced recombinantly (*4*). Both groups found that single fluorescently labeled molecules of dynein hardly moved at all, even in the presence of dynactin, and did

Faculty of Life Sciences, The Michael Smith Building, University of Manchester, Manchester, M13 9PT, UK. E-mail: viki.allan@manchester.ac.uk

Published by AAAS

EXTENDED PDF FORMAT
SPONSORED BY

Travel Grants Available
Neuroscience 2014
Application Deadline: July 31st
>Apply Now
R&D systems
a bietechne brand
www.rndsystems.com

One, two, three, cytoplasmic dynein is go!

Viki Allan

Science **345**, 271 (2014);
DOI: 10.1126/science.1257245

This copy is for your personal, non-commercial use only.

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by clicking here.

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines here.

The following resources related to this article are available online at www.sciencemag.org (this information is current as of August 4, 2014):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:
http://www.sciencemag.org/content/345/6194/271.full.html

A list of selected additional articles on the Science Web sites **related to this article** can be found at:
http://www.sciencemag.org/content/345/6194/271.full.html#related

This article **cites 16 articles**, 7 of which can be accessed free:
http://www.sciencemag.org/content/345/6194/271.full.html#ref-list-1

This article appears in the following **subject collections:**
Cell Biology
http://www.sciencemag.org/cgi/collection/cell_biol

Downloaded from www.sciencemag.org on August 4, 2014

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published weekly, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. Copyright 2014 by the American Association for the Advancement of Science; all rights reserved. The title *Science* is a registered trademark of AAAS.

oxidation to the corresponding quinone can occur, neither in the starting material nor in the final 6-chromanol. This trick also makes the synthesis of the tosyl derivatives much easier than that of their silyl-protected hydroquinone analogs. Furthermore, the method provides efficient synthetic access to other chroman-derived antioxidants, such as trolox (a water-soluble analog of vitamin E) and dihydrodaedalin A (a potent kinase inhibitor) (see the figure).

With the same oxidative method, dihydrobenzofurans can also be generated with even lower catalyst loadings of 0.05%, implying catalyst turnover numbers (TONs) of 2000. This value is extremely high for transition metal–free oxidation catalysis and is competitive with the most active homogeneous transition metal catalysts. These high TONs could only be observed when potassium carbonate (K_2CO_3) was added as a base. Uyanik *et al.* proposed recently that the combination of a tetraalkylammonium iodide with an inorganic or organic peroxide leads to the formation of hypoiodites (IO^- or IOH) as the key catalytically active species in this type of oxidative C-O bond–forming reaction (*8*–*10*). Using Raman spectroscopy, they now demonstrate the presence of hypoiodite and hypoiodous acid in the reaction mixture. In addition, they observed triiodide anion (I_3^-), a stable but catalytically inactive species. The authors further show that triiodide can be forced to form a hypoiodite again by a base additive, explaining the substantial effect of base addition in this reaction.

Beyond the practical aspects of this metal-free oxidative cycloetherification, which gives fast access to various biologically active 6-chromanol derivatives, the mechanistic investigations should lead to a more profound understanding of iodide-catalyzed oxidative transformations. Finally, Uyanik *et al.*'s study shows that oxidation catalysis (in particular, hypoiodite-mediated oxidation catalysis) can be a relatively mild process, allowing even antioxidants to be generated in an efficient way in an oxidative environment. ■

REFERENCES

1. C. Schneider, *Mol. Nutr. Food Res.* **49**, 7 (2005).
2. M. Uyanik, H. Hayashi, K. Ishihara, *Science* **345**, 291 (2014).
3. T. Netscher, *Vitam. Horm.* **76**, 155 (2007).
4. H. C. Shen, *Tetrahedron* **65**, 3931 (2009).
5. L. F. Tietze, K. M. Sommer, J. Zinngrebe, F. Stecker, *Angew. Chem. Int. Ed.* **44**, 257 (2005).
6. S. Tanaka, T. Seki, M. Kitamura, *Angew. Chem. Int. Ed.* **48**, 8948 (2009).
7. B. M. Trost, H. C. Shen, L. Dong, J.-P. Surivet, C. Sylvain, *J. Am. Chem. Soc.* **126**, 11966 (2004).
8. M. Uyanik, H. Okamoto, T. Yasui, K. Ishihara, *Science* **328**, 1376 (2010).
9. A. N. French, *Science* **328**, 1365 (2010).
10. M. Uyanik, K. Ishihara, *ChemCatChem* **4**, 177 (2012).

10.1126/science.1257347

ILLUSTRATION: P. HUEY/*SCIENCE*; ADAPTED FROM (*16*)

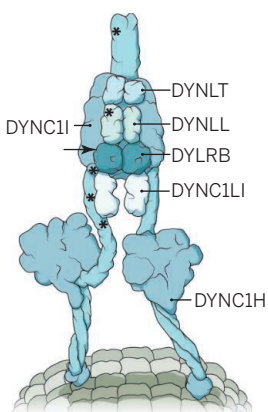
CELL BIOLOGY

One, two, three, cytoplasmic dynein is go!

The dynein motor protein must be part of a three-way complex to be fully active

By **Viki Allan**

Molecular motors are protein machines that power the movement of many different cargoes along filamentous tracks within eukaryotic cells. The dyneins move along microtubules, with the axonemal branch of the protein family working in large teams to drive the familiar beating of cilia and flagella by causing sliding of microtubules against each other. The cytoplasmic dyneins, on the other hand, move cargoes—including chromosomes, membrane-bounded organelles, mRNA particles, and protein complexes—along single microtubules (*1*). An important difference between cytoplasmic dynein 1 and axonemal dyneins is that as few as one to five molecules of cytoplasmic dynein 1 are enough to move small cargoes, whereas axonemal dyneins always work together. Another distinction is that cytoplasmic dynein 1 needs the help of an additional complex—dynactin—to function in the cell (*2*). Exactly what dynactin does, however, has been controversial. Two new papers, McKenney *et al.* on page 337 of this issue (*3*) and Schlager *et al.* (*4*), identify a missing piece of the puzzle—specific adaptor proteins—that helps explain how dynactin and cytoplasmic dynein 1 generate a single motor assembly that can move processively for many micrometers along the microtubule before falling off.



A cytoplasmic motor... The motor subunit of cytoplasmic dynein (DYNC1H) generates the force to power movement. Five nonmotor subunits form the domain that interacts with cargo. Asterisks mark the sites of mutations linked to human disease. A mouse mutation that reduces force production is arrowed.

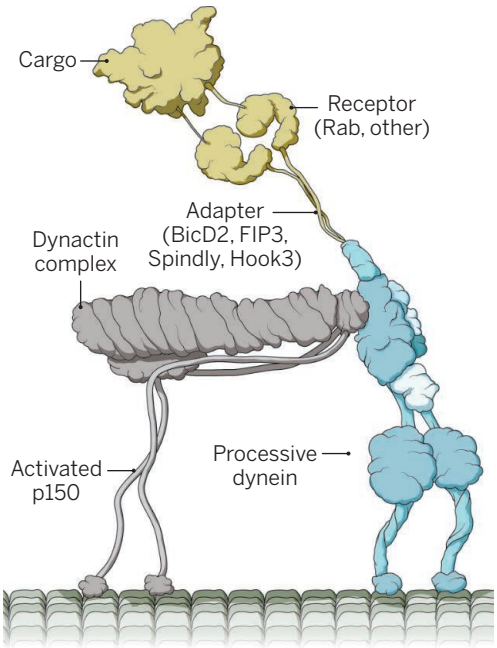
Cytoplasmic dynein 1 (called dynein hereafter) is a large (1.6 MD) protein complex that contains two copies each of the massive motor subunit (~450 kD) and five other subunits (see the first figure). The other subunits associate with the N-terminal region of the motor subunits to form a "tail" that binds to cargoes (*1*). Dynactin was purified as an activator of dynein-driven membrane movement, and it is also a large complex consisting of 11 different subunits (*2*) (see the second figure). Dynactin is needed for almost every dynein function in vivo (*1*, *2*), and although the $p150^{Glued}$ subunit of dynactin interacts with the dynein intermediate chain, it has been difficult to find evidence for stable association of the native complexes.

Several theories have been proposed to explain dynein's requirement for dynactin. First, because dynactin binds to some cargoes, it might simply drive dynein recruitment (*2*). Second, because $p150^{Glued}$ has a microtubule binding domain, it might hold on to the microtubule as dynein travels along it, so increasing dynein's processivity (*5*, *6*). However, there is evidence against both these models [reviewed in (*1*)]. Finally, dynactin might somehow directly activate dynein's motor activity, without the need for $p150^{Glued}$ to bind to microtubules (*7*).

An important step forward came with the discovery that mixing purified mammalian dynein and dynactin with Bicaudal D2 N-terminal domain (BicD2N) led to the formation of a stable complex of all three components (*8*). BicD2 in mammals helps recruit dynein to membranes in the secretory pathway and to the nuclear envelope (*8*), whereas the *Drosophila* homolog, BicD, plays a role in recruiting dynein to mRNPs (*9*). The two new studies have investigated the effects of BicD2N on dynein motility, using dynein purified from brain or cultured cells (*3*) or human dynein produced recombinantly (*4*). Both groups found that single fluorescently labeled molecules of dynein hardly moved at all, even in the presence of dynactin, and did

Faculty of Life Sciences, The Michael Smith Building, University of Manchester, Manchester, M13 9PT, UK. E-mail: viki.allan@manchester.ac.uk

Published by AAAS



...And friends. Single dynein molecules do not walk efficiently along microtubules unless they are in a tripartite complex with an adaptor molecule (e.g., BicD2) and dynactin, which also binds to microtubules. The adaptor also links the motor complex to the cargo to be transported.

not interact stably with dynactin. Furthermore, McKenney *et al.* (*3*) found that dynactin on its own did not bind to microtubules. Adding BicD2N formed a tripartite complex of one dynein, one dynactin and one BicD2N dimer. This complex showed high processivity when moving along microtubules in vitro, with run lengths and speeds approaching those seen for dynein cargoes in cells (*3*, *4*). McKenney *et al.* also showed that other dynein/dynactin "adaptors"—Rab11-FIP3, Spindly, and Hook—also formed similar highly active tripartite complexes. This suggests an attractive unified view where the interaction with cargo via adaptors not only recruits the motor but also leads to its assembly into an active complex. This would prevent dynein not bound to cargo from using up adenosine triphosphate unnecessarily. It also highlights the crucial role that adaptor regulation will play in defining where and when dynein is active.

Questions remain about how the formation of the tripartite assembly promotes dynein activity. McKenney *et al.*'s study suggests that the microtubule binding activity of $p150^{Glued}$ is needed to enhance dynein's processivity in vitro. However, these results are hard to reconcile with the observation that the $p150^{Glued}$ microtubule binding domain is not needed for dynein to move membrane cargoes in *Drosophila* S2 cells and mouse neurons in culture (*10*, *11*).

One possibility still to be tested is that the formation of the three-way complex improves the coordination between the motor heads, or perhaps increases the rigidity of the complex, which could improve processivity, enhance force production, or both. This is an intriguing idea, since a mutation in the tail region of mouse dynein reduced force production (*12*) (see the first figure). Furthermore, a number of human disease–linked mutations are found in similar positions (*13*). Negative stain electron microscopy imaging of the tripartite assembly by Schlager *et al.* (*4*) suggests that dynactin (and likely BicD2N too) lines up along dynein's tail (*4*), where it is in the right place to alter dynein's structural or mechanical properties. In contrast, McKenney *et al.* showed that there is a flexible linkage between dynein and dynactin (*3*). Of course, the organization of the complex when bound to native cargo may be different again.

There are plenty of other open questions. We do not know if this model holds true for all dynein cargoes and interactors, nor do we understand whether it matters which dynein subunit binds the adaptor, or how the adaptor is recruited to the cargo. We need to determine whether this mechanism contributes to controlling the direction of movement for cargoes that have both dynein and a plus-end–directed kinesin motor attached. Finally, we need to understand how the other known dynein regulators LIS1, Nde1, and Ndel1 fit into this scheme. It may be that the modes of action of LIS1-Nde1-Ndel1 and dynactin are fundamentally different, since LIS1 binds to dynein's motor domain (*14*) and dynein-dynactin and dynein-LIS1-Nde1 interactions are mutually exclusive (*15*). The ability to reconstitute dynein motility in vitro in a way that truly mimics its cellular function will hopefully pave the way to understanding the complexity of this intriguing motor protein. ■

REFERENCES

1. V. J. Allan, *Biochem. Soc. Trans.* **39**, 1169 (2011).
2. T. A. Schroer, *Annu. Rev. Cell Dev. Biol.* **20**, 759 (2004).
3. R. J. McKenney *et al. Science* **345**, 337 (2014).
4. M. Schlager, H. Hoang, L. Urnavicius, S. Bullock, A. Carter, *EMBO J.* 10.15252/embj.201488792 (2014).
5. T. L. Culver-Hanlon *et al.*, *Nat. Cell Biol.* **8**, 264 (2006).
6. S. J. King, T. A. Schroer, *Nat. Cell Biol.* **2**, 20 (2000).
7. J. R. Kardon, S. L. Reck-Peterson, R. D. Vale, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 5669 (2009).
8. D. Splinter *et al.*, *Mol. Biol. Cell* **23**, 4226 (2012).
9. H. C. Soundararajan, S. L. Bullock, *eLife* **3**, e01596 (2014).
10. H. Kim *et al.*, *J. Cell Biol.* **176**, 641 (2007).
11. A. J. Moughamian, E. L. F. Holzbaur, *Neuron* **74**, 331 (2012).
12. K. M. Ori-McKenney *et al. Nat. Cell Biol.* **12**, 1228 (2010).
13. G. Schiavo *et al. Trends Neurosci.* **36**, 641 (2013).
14. J. Huang *et al. Cell* **150**, 975 (2012).
15. R. J. McKenney *et al.*, *J. Biol. Chem.* **286**, 39615 (2011).
16. A. P. Carter, *J. Cell. Sci.* **126**, 705 (2013).

10.1126/science.1257245

PHYSICS

Quantum systems under control

Optical trapping enables the building of quantum matter one atom at a time

By **Jeff Thompson** *and* **Mikhail D. Lukin**

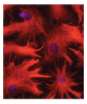
Controlled quantum systems are actively explored for efficient information processing, simulations of complex quantum systems, and secure quantum communications and metrology. In these areas, quantum devices have the potential to outperform their classical counterparts. However, achieving full control over large systems is difficult because the complex quantum mechanical correlations in large-scale systems are extremely fragile with respect to environmental noise or thermal fluctuations. One attractive approach is to use ensembles of atoms prepared in well-defined quantum states and cooled to temperatures near absolute zero. Ultracold atomic gases have emerged as one of the leading platforms in the study of many-body quantum phenomena. On page 306 of this issue, Kaufman *et al.* (*1*) demonstrate a novel twist on this approach with a "bottom-up" technique that allows complex quantum systems of ultracold neutral atoms to be put together one atom at a time.

When dealing with electrical circuits or computer code, it is often easier to develop a complex system by first testing isolated subsystems and then putting them together piecewise. This modular principle applies to quantum systems as well. Modularity is an essential property of certain physical implementations of quantum systems, such as trapped ions (*2*) and superconducting circuits (*3*), but not, so far, of ultracold atoms. Kaufman *et al.* produce fully controlled, pure quantum states of single, isolated rubidium atoms and then controllably bring them close enough together so that their quantum mechanical wave functions overlap and interfere. In combination with other recent advances, this approach opens the door to building complex quantum systems composed of atoms, molecules, and nanoscale solid-state systems.

Kaufman *et al.* start with a cloud of many atoms laser-cooled to temperatures

ILLUSTRATION: P. HUEY/*SCIENCE*

Published by AAAS

EXTENDED PDF FORMAT
SPONSORED BY

Travel Grants Available
Neuroscience 2014
Application Deadline: July 31st
>Apply Now
R&D systems
a bio-techne brand
www.rndsystems.com

Quantum systems under control

Jeff Thompson and Mikhail D. Lukin

Science **345**, 272 (2014);
DOI: 10.1126/science.1256529

This copy is for your personal, non-commercial use only.

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by clicking here.

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines here.

The following resources related to this article are available online at www.sciencemag.org (this information is current as of August 4, 2014):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:
http://www.sciencemag.org/content/345/6194/272.full.html

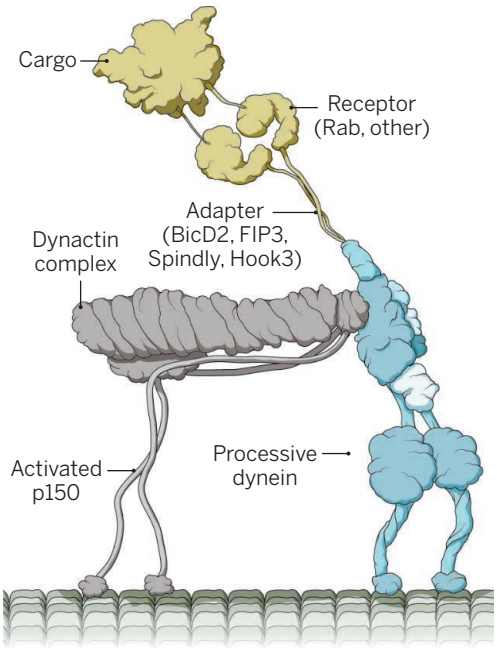
A list of selected additional articles on the Science Web sites **related to this article** can be found at:
http://www.sciencemag.org/content/345/6194/272.full.html#related

This article **cites 14 articles**, 4 of which can be accessed free:
http://www.sciencemag.org/content/345/6194/272.full.html#ref-list-1

This article appears in the following **subject collections:**
Physics
http://www.sciencemag.org/cgi/collection/physics

Downloaded from www.sciencemag.org on August 4, 2014

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published weekly, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. Copyright 2014 by the American Association for the Advancement of Science; all rights reserved. The title *Science* is a registered trademark of AAAS.



...And friends. Single dynein molecules do not walk efficiently along microtubules unless they are in a tripartite complex with an adaptor molecule (e.g., BicD2) and dynactin, which also binds to microtubules. The adaptor also links the motor complex to the cargo to be transported.

not interact stably with dynactin. Furthermore, McKenney *et al.* (*3*) found that dynactin on its own did not bind to microtubules. Adding BicD2N formed a tripartite complex of one dynein, one dynactin and one BicD2N dimer. This complex showed high processivity when moving along microtubules in vitro, with run lengths and speeds approaching those seen for dynein cargoes in cells (*3*, *4*). McKenney *et al.* also showed that other dynein/dynactin "adaptors"—Rab11-FIP3, Spindly, and Hook—also formed similar highly active tripartite complexes. This suggests an attractive unified view where the interaction with cargo via adaptors not only recruits the motor but also leads to its assembly into an active complex. This would prevent dynein not bound to cargo from using up adenosine triphosphate unnecessarily. It also highlights the crucial role that adaptor regulation will play in defining where and when dynein is active.

Questions remain about how the formation of the tripartite assembly promotes dynein activity. McKenney *et al.*'s study suggests that the microtubule binding activity of $p150^{Glued}$ is needed to enhance dynein's processivity in vitro. However, these results are hard to reconcile with the observation that the $p150^{Glued}$ microtubule binding domain is not needed for dynein to move membrane cargoes in *Drosophila* S2 cells and mouse neurons in culture (*10*, *11*).

One possibility still to be tested is that the formation of the three-way complex improves the coordination between the motor heads, or perhaps increases the rigidity of the complex, which could improve processivity, enhance force production, or both. This is an intriguing idea, since a mutation in the tail region of mouse dynein reduced force production (*12*) (see the first figure). Furthermore, a number of human disease–linked mutations are found in similar positions (*13*). Negative stain electron microscopy imaging of the tripartite assembly by Schlager *et al.* (*4*) suggests that dynactin (and likely BicD2N too) lines up along dynein's tail (*4*), where it is in the right place to alter dynein's structural or mechanical properties. In contrast, McKenney *et al.* showed that there is a flexible linkage between dynein and dynactin (*3*). Of course, the organization of the complex when bound to native cargo may be different again.

There are plenty of other open questions. We do not know if this model holds true for all dynein cargoes and interactors, nor do we understand whether it matters which dynein subunit binds the adaptor, or how the adaptor is recruited to the cargo. We need to determine whether this mechanism contributes to controlling the direction of movement for cargoes that have both dynein and a plus-end–directed kinesin motor attached. Finally, we need to understand how the other known dynein regulators LIS1, Nde1, and Ndel1 fit into this scheme. It may be that the modes of action of LIS1-Nde1-Ndel1 and dynactin are fundamentally different, since LIS1 binds to dynein's motor domain (*14*) and dynein-dynactin and dynein-LIS1-Nde1 interactions are mutually exclusive (*15*). The ability to reconstitute dynein motility in vitro in a way that truly mimics its cellular function will hopefully pave the way to understanding the complexity of this intriguing motor protein. ■

REFERENCES

1. V. J. Allan, *Biochem. Soc. Trans.* **39**, 1169 (2011).
2. T. A. Schroer, *Annu. Rev. Cell Dev. Biol.* **20**, 759 (2004).
3. R. J. McKenney *et al. Science* **345**, 337 (2014).
4. M. Schlager, H. Hoang, L. Urnavicius, S. Bullock, A. Carter, *EMBO J.* 10.15252/embj.201488792 (2014).
5. T. L. Culver-Hanlon *et al.*, *Nat. Cell Biol.* **8**, 264 (2006).
6. S. J. King, T. A. Schroer, *Nat. Cell Biol.* **2**, 20 (2000).
7. J. R. Kardon, S. L. Reck-Peterson, R. D. Vale, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 5669 (2009).
8. D. Splinter *et al.*, *Mol. Biol. Cell* **23**, 4226 (2012).
9. H. C. Soundararajan, S. L. Bullock, *eLife* **3**, e01596 (2014).
10. H. Kim *et al.*, *J. Cell Biol.* **176**, 641 (2007).
11. A. J. Moughamian, E. L. F. Holzbaur, *Neuron* **74**, 331 (2012).
12. K. M. Ori-McKenney *et al. Nat. Cell Biol.* **12**, 1228 (2010).
13. G. Schiavo *et al. Trends Neurosci.* **36**, 641 (2013).
14. J. Huang *et al. Cell* **150**, 975 (2012).
15. R. J. McKenney *et al.*, *J. Biol. Chem.* **286**, 39615 (2011).
16. A. P. Carter, *J. Cell. Sci.* **126**, 705 (2013).

10.1126/science.1257245

PHYSICS

Quantum systems under control

Optical trapping enables the building of quantum matter one atom at a time

By **Jeff Thompson** *and* **Mikhail D. Lukin**

Controlled quantum systems are actively explored for efficient information processing, simulations of complex quantum systems, and secure quantum communications and metrology. In these areas, quantum devices have the potential to outperform their classical counterparts. However, achieving full control over large systems is difficult because the complex quantum mechanical correlations in large-scale systems are extremely fragile with respect to environmental noise or thermal fluctuations. One attractive approach is to use ensembles of atoms prepared in well-defined quantum states and cooled to temperatures near absolute zero. Ultracold atomic gases have emerged as one of the leading platforms in the study of many-body quantum phenomena. On page 306 of this issue, Kaufman *et al.* (*1*) demonstrate a novel twist on this approach with a "bottom-up" technique that allows complex quantum systems of ultracold neutral atoms to be put together one atom at a time.

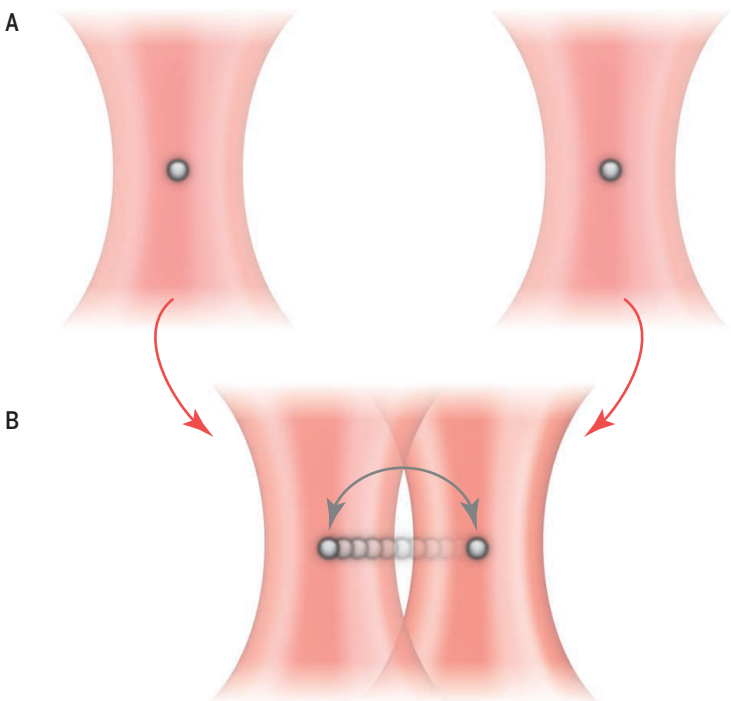
When dealing with electrical circuits or computer code, it is often easier to develop a complex system by first testing isolated subsystems and then putting them together piecewise. This modular principle applies to quantum systems as well. Modularity is an essential property of certain physical implementations of quantum systems, such as trapped ions (*2*) and superconducting circuits (*3*), but not, so far, of ultracold atoms. Kaufman *et al.* produce fully controlled, pure quantum states of single, isolated rubidium atoms and then controllably bring them close enough together so that their quantum mechanical wave functions overlap and interfere. In combination with other recent advances, this approach opens the door to building complex quantum systems composed of atoms, molecules, and nanoscale solid-state systems.

Kaufman *et al.* start with a cloud of many atoms laser-cooled to temperatures

ILLUSTRATION: P. HUEY/*SCIENCE*

Published by AAAS

within one thousandth of a degree of absolute zero. Then, a tightly focused laser beam acting as an optical tweezer forms an attractive potential well in the center of the atomic cloud. Atoms fall into the optical tweezer potential at random; however, because of the small volume of the tweezer, pairs of atoms experience an energy-boosting molecular interaction and are ejected, resulting in stable occupancies of zero or one atom only (*4*). By applying several separately focused lasers, multiple optical tweezers can be arranged to hold one atom each, with the correct occupancy verified by actually observing the atoms individually. To prepare a truly pure quantum state, the atoms are cooled further to the absolute motional ground state of their respective optical tweezers. This is accomplished using another laser-cooling technique known as Raman sideband cooling. Although this technique has long been applied to cool single trapped ions (*5*), it was only recently applied to individual neutral atoms (*6*, *7*).



Trap, then build. (**A**) Two "optical tweezer" traps are prepared with one atom each and can be rapidly repositioned with high precision. (**B**) If the tweezers are brought very close to each other, the atoms can tunnel between them, giving rise to quantum interference effects.

Once the initial state is prepared, the positions of the optical tweezers can be rapidly adjusted with high precision (see the figure). Bringing the two tweezers together until they are almost on top of each other allows the atoms to tunnel across the barrier, from one tweezer into the other. The key idea in the present experiment is that two indistinguishable atoms prepared in pure quantum mechanical states are subject to quantum interference effects that can produce correlations in their wave functions. In particular, if one atom remains in each tweezer after they have been allowed to tunnel for some time, it is impossible to determine whether neither or both of the atoms have switched sides. Therefore, the probability of finding one atom on each side must be the square of the sum of the amplitudes for both pathways. At certain times, these amplitudes can cancel each other exactly, which means that the atoms must be "bunched" on the same side. This bunching happens in the absence of any interaction between the bosonic particles, and was in fact first observed with optical photons (which are completely noninteracting) by Hong, Ou, and Mandel (*8*). To verify that it is interference between indistinguishable pathways giving rise to the bunching, Kaufmann *et al.* make the atoms distinguishable (by changing the spin state or motional state of one of them); the interference signal then disappears.

This approach of Kaufman *et al.* opens up a new avenue for building complex quantum systems in a controlled way. Although similar goals are being pursued in a variety of platforms, ranging from ultracold atomic gases in optical lattices (*9*) and trapped ions (*2*) to artificial atomlike defects in solids and superconducting quantum circuits (*3*), the approach demonstrated by Kaufman *et al.* is complementary to these techniques in a way that adds new capabilities. For example, it can be more flexible than experiments involving manipulation of entire atomic clouds, which typically begin with a quantum degenerate gas of tens of thousands of atoms loaded in a uniform periodic potential. In comparison to trapped ions, larger arrays of neutral atoms can be created in a greater variety of geometric configurations and can fully explore the effects of quantum statistics. Related techniques have recently been used to prepare and manipulate systems composed of several fermionic atoms (*10*). In such systems, the quantum interference demonstrated by Kaufman *et al.* may emerge as a powerful probe for complex quantum states of matter (*11*, *12*). Finally, the indistinguishability of atoms and ions is so far difficult to match, on a large scale, with solid-state systems.

The precise position control and tight atomic localization provided by the optical tweezer array allows atoms to be positioned in direct proximity to other objects, enabling a path toward hybrid quantum systems. For example, placing atoms near nanoscale solid-state structures with tightly confined optical, mechanical, or electrical modes allows for strong interactions, which typically scale inversely with mode volume and therefore also with the separation between the atom and the surface (*13*, *14*). A quantum interface between atoms and solid-state structures can be used to create long-range and tunable interactions between atoms or to efficiently map quantum states from atoms to photons for transmission over long distances, enabling quantum networks. Finally, this technique is a promising starting point for cold chemistry. In particular, if extended to dissimilar atomic species, it may allow for the creation of complex molecules from individual atoms. ■

Department of Physics, Harvard University, 17 Oxford Street, Cambridge, MA 02138, USA. E-mail: lukin@physics.harvard.edu

ILLUSTRATION: P. HUEY/*SCIENCE*

REFERENCES AND NOTES

1. A. M. Kaufman *et al.*, *Science* **345**, 306 (2014).
2. C. Monroe, J. Kim, *Science* **339**, 1164 (2013).
3. M. H. Devoret, R. J. Schoelkopf, *Science* **339**, 1169 (2013).
4. N. Schlosser, G. Reymond, I. Protsenko, P. Grangier, *Nature* **411**, 1024 (2001).
5. C. Monroe *et al.*, *Phys. Rev. Lett.* **75**, 4011 (1995).
6. A. M. Kaufman, B. J. Lester, C. A. Regal, *Phys. Rev. X* **2**, 041014 (2012).
7. J. D. Thompson, T. G. Tiecke, A. S. Zibrov, V. Vuletić, M. D. Lukin, *Phys. Rev. Lett.* **110**, 133001 (2013).
8. C. K. Hong, Z. Y. Ou, L. Mandel, *Phys. Rev. Lett.* **59**, 2044 (1987).
9. W. S. Bakr, J. I. Gillen, A. Peng, S. Fölling, M. Greiner, *Nature* **462**, 74 (2009).
10. F. Serwane *et al.*, *Science* **332**, 336 (2011).
11. S. Fölling *et al.*, *Nature* **434**, 481 (2005).
12. A. J. Daley, H. Pichler, J. Schachenmayer, P. Zoller, *Phys. Rev. Lett.* **109**, 020505 (2012).
13. T. G. Tiecke *et al.*, *Nature* **508**, 241 (2014).
14. A. Goban *et al.*, *Nat. Commun.* **5**, 3808 (2014).

ACKNOWLEDGMENTS

We acknowledge stimulating discussions with M. Greiner, K. K. Ni, T. G. Tiecke, and V. Vuletić.

10.1126/science.1256529

Published by AAAS

FUNDING TRANSDISCIPLINARY RESEARCH

NIH Roadmap/Common Fund at 10 years

A mechanism for funding biomedical research at NIH that transcends Institute and Center boundaries is bearing fruit

By **Francis S. Collins**[1]†, **Elizabeth L. Wilder**[1], **Elias Zerhouni**[1]*

POLICY

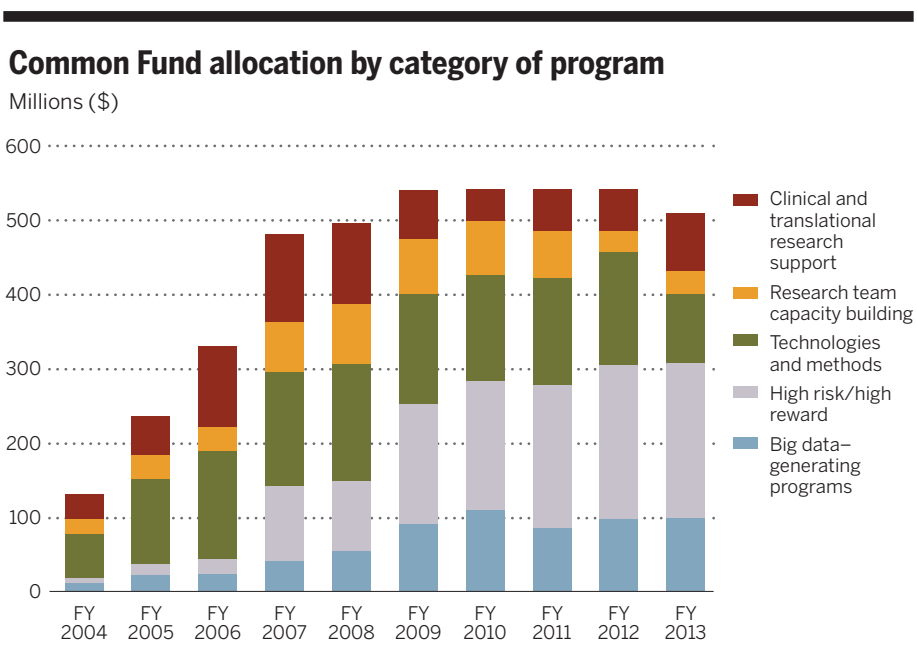
A fundamental challenge facing all institutions of science, including the National Institutes of Health (NIH), is whether their structures and disciplines, inherited from the past, continue to reflect the reality of current science and the needs of future science. Without an explicit process of adaptation to changes that often transcend established scientific structures and disciplines, the risk of missing emerging opportunities grows. This is why, 10 years ago, NIH launched an approach to the support of science that transcended all Institutes and Centers, known as the "NIH Roadmap" (*1*). The NIH Director and the Director of each of the NIH Institutes engaged in a broad priority-setting exercise, informed by extramural and intramural NIH scientists, public representatives, and leaders from other government agencies and the private sector. We asked three questions: What are today's most pressing scientific challenges? What are the roadblocks to progress and what must be done to overcome them? Which efforts are beyond the mandate of one or a few institutes, but are the responsibility of NIH as a whole? Each of the initial Roadmap programs that emerged was designed to achieve defined goals or transition to other sources of support within 10 years. As we have reached the 10th anniversary of these programs, a look back is in order.

All Institutes and Centers contributed 1% of their budgets to a common pool, and criteria were established to prioritize the many ideas that came from the consultation sessions; these criteria have changed little during the past 10 years (see the first table). With the 2006 NIH Reform Act, Congress established the NIH Common Fund within the Office of the Director, and it was authorized as a line item in the NIH budget; in fiscal year (FY) 2007, $483 million was appropriated (1.7% of the NIH budget). The Act stipulated that the Fund could not drop to a lower percentage, and anticipated a rise to 5%. As of 2014, the budget is $531 million (1.8% of NIH total appropriation) (see the first figure and second table) (Fig. 1).Fig. 1

As many of the initial programs conclude this year, final outcome assessments have not been completed, but regular external scientific panel reviews and informal assessments indicate that most of the programs have had positive outcomes. The vision-setting process for the NIH as a whole has thus delivered new technologies, research tools, experimental approaches, and large data sets that are enabling investigator-initiated research across the NIH. New ways of supporting high-risk and high-reward research have been tested. We believe that it is unlikely that these goals could have been achieved without the Common Fund.

CREATION OF INNOVATIVE TOOLS AND TECHNOLOGIES. The lack of adequate tools, methods, and technologies is a challenge that has been articulated frequently during consultations with the research community. Although some NIH R01's support tool and/or technology development, this type of research fits the transformative, catalytic requirement of Common Fund programs.

Here are a few examples: As a method to control neural activity with light, optogenetics was developed with support from Common Fund "Pioneer" and "New Innovator" Awards to Karl Deisseroth and Ed Boyden, respectively. It has revolutionized neuroscience by providing the ability to characterize neural circuits and offers potential approaches for treatment of neural disorders (*2*, *3*). Over the past decade, investigators supported by "Structural Biology of Membrane Proteins" have developed methods and tools that have contributed to an exponentially growing number of eukaryotic membrane proteins that have been structurally analyzed (*4*). "Patient-Reported Outcomes Measurement Information System" ("PROMIS") was designed to enable clinicians to obtain reproducible and quantitative feedback from patients about aspects of their health such as fatigue, anxiety, depression, social participation, and well-being. Through advances in information technology, psychometrics, and health survey research, PROMIS has generated a robust computer adaptive testing system based on item response theory that is being



Funding by categories. Roadmap/Common Fund programs were assigned to one of five categories (*12*), and funds allocated to each category are shown. Some programs have relevance to more than one category, but each program was assigned to only one category.

[1]National Institutes of Health, Bethesda, MD 20892, USA.
*Present address: Sanofi, 75008 Paris, France.
†Corresponding author. E-mail: francis.collins@nih.gov

Published by AAAS

Downloaded from www.sciencemag.org on August 4, 2014

used to assess treatment efficacy in diverse situations (*5*, *6*). The emphasis on developing transformative tools can also be seen in the "Knock-Out Mouse Phenotyping" Program, which provides broad, standardized phenotyping of a genome-wide collection of mouse knockouts generated by the International Knockout Mouse Consortium (*7*).

LARGE DATA SETS. The development of massive, publicly available data sets is perhaps the clearest example of "Common Fund-able" research: It benefits every avenue of biomedical research and would be impossible for individual investigators to achieve. The Roadmap/Common Fund has enabled large, complex data sets, such as those related to the human microbiome and the epigenome, to be provided to the entire community. End users serve as consultants to ensure that computational tools are user-friendly and meet community needs. Demonstration projects are also often included in programs that generate large data sets, (i) to provide evidence that the data are applicable to a wide spectrum of the NIH mission, (ii) to enhance the data resource, and (iii) to provide feedback to the data developers to increase the utility of the program.

ENCOURAGING RISK-TAKING. The "High Risk/High Reward" program was established in 2004 with the Pioneer Awards, which recognize and reward investigators who have demonstrated innovation in prior work and provide a mechanism for them to go in entirely new, high-impact research directions. The requirement for strong preliminary data in R01 applications was felt by many to inhibit creativity and innovation. Recently, we commissioned an independent evaluation comparing impact and innovation achieved through Pioneer, R01, and Howard Hughes Medical Institute (HHMI) funding (*8*). When controlled for funding, the Pioneer awardees, as a group, performed comparably to HHMI investigators, and their research was deemed more innovative, with higher impact than R01 investigators working in similar areas who had similar backgrounds and resources. The success of the Pioneer awards led to the expansion of the High Risk/High Reward program to include the "New Innovator," "Early Independence," and "Transformative Research" awards. The Pioneers' success also led to the distinction of the High Risk/High Reward program as the only Common Fund program without an established end date; each individual project, however, is restricted to 5 years of support. Broader implementation of similar programs is being considered across the NIH.

CHALLENGES FOR THE COMMON FUND APPROACH. For the Common Fund to take on a steady stream of new projects, it has been essential that funded programs be completed or adopted by other funding sources after a certain period of time. One issue that has become apparent is that programs that work toward broad, long-term goals, rather than aiming to achieve a specific set of goals in 5 to 10 years, do not align well with the directives of the Common Fund. Although the broad, stimulatory goals are critical to advancing science, if an end point—a deliverable—is not envisioned in advance, it becomes difficult to assess impact. Continued funding to achieve the long-term goals may also be problematic for investigators to secure.

The "Interdisciplinary Research" program included an emphasis on recognizing and rewarding team-led science, and it contributed to the development of the multiple–Principal Investigator policy at the NIH. Although an evaluation of the program indicated that the emphasis that the program placed on interdisciplinary research did have an impact on barriers to team science at research institutions (*9*) and that valuable research was conducted, the success in establishing new inter-disciplines is difficult to assess. Individual consortia and training programs that were launched ultimately faced a substantial challenge in continuation. As large, investigator-initiated projects that crossed boundaries of several NIH Institutes and Centers, they did not fit easily within the traditional funding schemes of the NIH. The research was ultimately pursued via smaller awards, but the synergies of interdisciplinary teams working together were likely compromised.

With this lesson in hand, new Common Fund programs continue to support interdisciplinary teams, but specific, short-term goals for the teams are articulated. Continuation of the science happens via uptake and use of the Common Fund deliverables rather than sustained funding for the specific projects initiated by the Common Fund.

Sustainability of infrastructure also proved to be a challenge. Although future support of some of the infrastructure was built into the design of the program (*10*), the sustainability of other programs was to be determined by requirements at the time of transition. For example, the Roadmap "Molecular Libraries" program, aimed to provide pharmaceutical-quality small-molecule screening capacity to academic investigators, enabling them to develop screening assays for diverse cellular functions and to put screening data into publicly accessible databanks with computational tools to enable wide use. The team of investigators, external consultants, and NIH staff who contributed to the management of this program established a consortium through which innovation in assay development, screening, and computational methods is combined. Although the core infrastructure continues to evolve within the National Center for Advancing Translational Sciences, the need for an integrated network of centers dissipated as additional screening facilities were developed by institutions across the country. Transition from central management and funding to decentralized partnerships between screening facilities and assay providers presented challenges. Today, we emphasize sustainability from the outset and work with awardees to establish a plan for the future.

The "National Centers for Biomedical Computing" (NCBCs) program aimed to develop software and data management tools to facilitate computational analysis of diverse biological problems. Each Center was established to tackle a core set of computational challenges but also collaborated with a series of partners. Although Centers interacted to develop training opportunities, the computational approaches developed through the program were largely developed by individual Centers. In an internal review of this program, it was recognized that, despite the many successes of the NCBC program (*11*), tighter coordination between projects—with increased expert and user

Characteristics of Common Fund programs

Transformative: Must have high potential to dramatically affect biomedical and/or behavioral research over the next decade
Catalytic: Must achieve a defined set of high-impact goals within 5 to 10 years
Synergistic: Outcomes must synergistically promote and advance individual missions of NIH Institutes and Centers
Cross-cutting: Program areas must cut across missions of multiple NIH Institutes and Centers, be relevant to multiple diseases or conditions, and be sufficiently complex to require a coordinated, trans-NIH approach
Unique: Must be something no other entity is likely or able to do

Published by AAAS

feedback obtained on a regular basis—would be necessary as the program moves forward. That experience is now being used to inform plans for the "Big Data to Knowledge" ("BD2K") initiative.

The scientific accomplishments of the Roadmap/Common Fund have required unprecedented levels of coordination and collaboration among the leadership and staff across the NIH. Although we have had anecdotal feedback that this effort has enabled the success of the programs and has worked well, we have engaged the NIH Council of Councils to evaluate the planning and management processes used for the Common Fund. Their final report is expected later this month.

LOOKING FORWARD. As mature programs come to a close, we have the opportunity to address new, emerging challenges and opportunities. We reserve ~30% of the Common Fund for the investigator-initiated, High Risk/High Reward initiatives, but otherwise, the definition and scope of new program areas are determined by the challenges and opportunities that emerge through strategic planning activities each year. Community input continues to be fundamental to the process.

We use both virtual and in-person approaches to reach out to the entire community. One lesson learned through these discussions is that, although it is easy to identify research challenges and exciting new areas, it is more difficult to identify those where recent advances—perhaps in unrelated fields—create an opportunity to overcome the challenges within 5 to 10 years and where a coordinated, goal-driven, trans-NIH approach is required. As a result, we have launched new programs that build on lessons learned from the earlier programs. New paradigms for intercellular control of gene expression and cellular function may emerge from the "Extracellular RNA" program. New methods and models for clinical research partnerships between academic centers and health care providers are being developed by the "Health Care Systems Collaboratory" program. More recently, we have established plans to launch the "4D Nucleome." A "Glycoscience" program to develop methods to simplify the functional analysis of glycans in diverse contexts is also under way.

The Common Fund also continues to develop new ways to enhance innovation and support the workforce by developing and testing novel methods for training and mentoring young scientists and for supporting early career stages. One new program is beginning with the recognition that the NIH mission requires participation of scientists from all sectors of our population, yet measures to diversify the biomedical workforce have not yielded the desired outcomes. Emerging opportunities from social science research provide the basis for the Common Fund's "Enhancing Diversity of the NIH-Funded Workforce" program: Understanding the factors that contribute to decisions by students from underrepresented groups to leave biomedical research career paths provides the opportunity to develop and test new training and mentoring methods.

CONCLUSION. Although investigator-initiated research funded through individual Institutes and Centers remains the lifeblood of the NIH, the Roadmap/Common Fund has established a tradition of bringing the biomedical research community together for coordinated, transdisciplinary collaboration when such effort is required. Although not every program has been completely successful, the synergies created by the Common Fund are inspiring, and the catalytic nature of these programs ensures that their impact will continue to grow. ■

Current Common Fund programs

- Big Data to Knowledge (BD2K)
- Enhancing the Diversity of the NIH-Funded Workforce
- Epigenomics
- Extracellular RNA Communication
- 4D Nucleome
- Genotype-Tissue Expression (GTEx)
- Global Health
- Glycoscience: Enabling Tools and Methods
- Gulf Long Term Follow-Up Study
- Health Care Systems Research Collaboratory
- Health Economics
- High-Risk/High-Reward Research:
  - Early Independence Award (EIA)
  - New Innovator Award
  - Pioneer Award
  - Transformative Research Awards
- Human Microbiome Project
- Illuminating the Druggable Genome
- Knockout Mouse Phenotyping
- Library of Integrated Network-Based Cellular Signatures (LINCS)
- Metabolomics
- Nanomedicine
- NIH Center for Regenerative Medicine
- NIH Medical Research Scholars Program
- Protein Capture Reagents
- Regulatory Science
- Science of Behavior Change
- Single Cell Analysis
- Strengthening the Biomedical Research Workforce
- Undiagnosed Diseases

REFERENCES AND NOTES

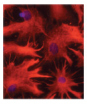
1. E. Zerhouni, *Science* **302**, 63–72 (2003).
2. E. S. Boyden, F. Zhang, E. Bamberg, G. Nagel, K. Deisseroth, *Nat. Neurosci.* **8**, 1263–1268 (2005).
3. A. Reiner, E. Y. Isacoff, *Trends Neurosci.* **36**, 557–560 (2013).
4. S. White, Membrane Proteins of Known 3D Structure (2014); http://blanco.biomol.uci.edu/mpstruc.
5. D. Cella *et al.*, *J. Clin. Epidemiol.* **63**, 1179–1194 (2010).
6. Publications linked to the Common Fund PROMIS program; http://commonfund.nih.gov/publications?pid=23.
7. S. D. M. Brown, M. W. Moore, *Mamm. Genome* **23**, 632–640 (2012).
8. B. Lal *et al.*, *An Outcome Evaluation of the National Institute of Health (NIH) Director's Pioneer Award (NDPA) Program, FY2004–2006* (Institute for Defense Analysis Science and Technology Policy Institute, Washington, DC, 2012); http://commonfund.nih.gov/sites/default/files/P-4899_Final_Redacted.pdf.
9. Interdisciplinary Research Consortium Initiative Evaluation, Executive Summary (NIH, 2011); http://commonfund.nih.gov/Interdisciplinary/evaluation.
10. E. A. Zerhouni, *N. Engl. J. Med.* **353**, 1621–1623 (2005).
11. J. M. Berg, *J. Am. Med. Inform. Assoc.* **19**, 151–152 (2012).
12. Roadmap/Common Fund programs were assigned to the categories in the figure as follows: **Big Data–Generating Programs**: National Centers for Biomedical Computing, Human Microbiome Project, Epigenomics, Genotype-Tissue Expression (GTEx), Knockout Mouse Phenotyping, Library of Integrated Network-based Cellular Signatures (LINCS), BD2K, Extracellular RNA Communications. **High Risk/High Reward program**: Pioneer, New Innovator, Transformative Research Awards, Early Independence Awards. **Technologies and Methods**: Molecular Libraries, Structural Biology of Membrane Proteins, Building Blocks/Biological Pathways/Networks, PROMIS, Protein Capture Reagents, Single-Cell Analysis. **Research Team Capacity Building:** Interdisciplinary Research, Nanomedicine, Public-Private Partnerships, Strengthening the Biomedical Workforce, Enhancing Diversity of the NIH-Funded Workforce, Global Health. **Clinical and Translational Research Support**: Clinical and Translational Science Awards and related initiatives, Clinical Research Training Program, Rapid Access to Intervention Development, Metabolomics, Health Care Systems Collaboratory, Health Economics, Gulf Long Term Follow-Up Study, NIH Center for Regenerative Medicine, Undiagnosed Diseases Network.

ACKNOWLEDGMENTS

The authors acknowledge L. A. Tabak, J. M. Anderson, and L. Skirboll for their review and comments on the manuscript.

10.1126/science.1255860

Published by AAAS

EXTENDED PDF FORMAT
SPONSORED BY

Travel Grants Available
Neuroscience 2014
Application Deadline: July 31st
>Apply Now
R&D systems
a bio-techne brand
www.rndsystems.com

Walter Gehring (1939–2014)

Michael Levine

Science **345**, 277 (2014);
DOI: 10.1126/science.1258143

This copy is for your personal, non-commercial use only.

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by clicking here.

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines here.

The following resources related to this article are available online at www.sciencemag.org (this information is current as of August 4, 2014):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:
http://www.sciencemag.org/content/345/6194/277.full.html

Downloaded from www.sciencemag.org on August 4, 2014

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published weekly, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. Copyright 2014 by the American Association for the Advancement of Science; all rights reserved. The title *Science* is a registered trademark of AAAS.

RETROSPECTIVE

Walter Gehring (1939–2014)

A brief reflection on the life and lab of a preeminent developmental biologist

By **Michael Levine**

I arrived in Basel, Switzerland, in 1982 wearing sunglasses and a Hawaiian print shirt and was nervous about my first encounter with Walter Gehring. He was stern and serious and strangely unmoved by the neurotic Jewish humor of a lowly rookie postdoc. As a prank, I used a charcoal-activated straw to drink the tap water. Walter was not amused. This antipathy turned to near rage when he found the butt of my brand of cigarette inside the eyepiece of his coveted Zeiss microscope (I think I was framed). So, it is fair to say that Walter and I were oil and water, but it is nonetheless my privilege to write this memorial on the sad occasion of his untimely death on 29 May following a traffic accident in Greece.

Walter studied zoology at the University of Zurich where he received his Ph.D. in 1965. He trained in the laboratory of the great developmental geneticist Ernst Hadorn where he worked on the development of antennae in the fruit fly *Drosophila* and distinguished himself as a brash scientific talent. The Hadorn lab studied transdetermination of *Drosophila* imaginal discs, a process whereby disc cells can undergo wholesale changes in identity to produce a different adult appendage. The lab featured a star-studded cast of future fly immortals that also included geneticists Antonio Garcia-Bellido and Gerold Schubiger (who sadly died in 2012). Hadorn himself had trained with Fritz Baltzer and Curt Stern, and thereby represented a rare fusion of classical genetics and experimental developmental biology.

Walter went on to conduct postdoctoral research at Yale University with my Ph.D. advisor, Alan Garen, one of the pioneers of bacterial molecular genetics and an early practitioner of *Drosophila* molecular biology. Between his Ph.D. and postdoctoral training, Walter obtained a broad exposure to molecular genetics and classical developmental biology. His keen sense for interesting developmental processes positioned Walter on the fast track of discovery. As a

PHOTO: BIOZENTRUM UNIVERSITY OF BASEL

postdoc, he and Lily Chan (a graduate student in the Garen lab) obtained evidence that the fate map of the adult fruit fly was already established in the early embryo. His obvious abilities led to rapid advancement to an independent faculty position at the Yale Medical School in 1969. Walter recruited future Nobelist Eric Wieschaus as his first graduate student, and hence began the long string of talented scientists who trained in the Gehring lab over a span of about 40 years.

After a few years, Walter moved from Yale to the University of Basel at the newly established Biozentrum. In the late 1970s to early 1980s, his lab was among the first to combine the emerging tools of molecular biology with classical *Drosophila* genetics to isolate and characterize genes such as *Antennapedia*, which promotes the formation of legs and suppresses the development of antennae in the thorax. During the 1980s and 1990s, the Gehring lab produced a dizzying array of discoveries: the homeobox, a 180–base pair sequence that encodes a DNA binding domain regulating a variety of developmental processes; the enhancer trap, a valuable technique in molecular biology for identifying tissue-specific regulatory sequences (enhancers) that control gene expression; and *eyeless* (encoded by *PAX6* in human), a key regulatory gene for eye development whose misexpression dramatically induces fly eyes to grow on wings, legs, and antennae.

Walter was a naturalist at heart. He was a bird watcher and amateur marine biologist. He quoted Darwin and recognized the implications of his research in the broad context of evolution. Perhaps his greatest talent was his intuitive sense for interesting research problems, combined with a remarkable ability to explain the most complex data sets as a simple story with a clear punch line. These syntheses were often accompanied by fanciful embellishments, but in the process he unlocked the mysteries of highly specialized information and rendered them plain and accessible to everyone. It is largely due to these efforts that the homeobox is widely known and taught in biology classes around the world. Walter's talents won him many awards, including the Kyoto Prize for Basic Science (2000) and the Balzan Prize for Developmental Biology (2002). He was elected a foreign member of the U.S. National Academy of Sciences and the Royal Society of London.

An amazing group of students and postdocs was attracted to the Gehring lab over the years: Eric Wieschaus (Nobelist), Christianne Nüsslein-Volhard (Nobelist), David Ish-Horowicz, Spyros Artavanis-Tsakonas, Paul Schedl, Alex Schier, Georg Halder, Hugo Bellen, and Markus Affolter, to mention just a few. I worked closely with two of my future lifelong friends and colleagues: Ernst Hafen and Bill McGinnis. The lab was an absolute blast, but a strange mix of anarchy and oppression. Walter permitted considerable independence, but was hardly laissez-faire. He could be confrontational, and did not hesitate to call us out (particularly me) when he felt we were misbehaving.

I found Walter to be a complicated character. He had the mannerisms of an authoritative Herr Doktor Professor, but was also folksy and unaffected and always ready to laugh and joke. He sometimes felt competitive with his students and postdocs, but was also highly supportive and proud of our independent careers. In short, I believe the key to Walter's success was his yin and yang embodiment of old-world scholar and modern competitive scientist. He was able to exude charm and empathy, but nothing we did seemed to be quite good enough. In other words, tough love, possibly the perfect prescription for eliciting the very best efforts from his students and postdocs. ■

Department of Molecular and Cell Biology, Center for Integrative Genomics, University of California, Berkeley, Berkeley, CA 94720–3200, USA. E-mail: mlevine@berkeley.edu

10.1126/science.1258143

Published by AAAS

LETTERS

Edited by **Jennifer Sills**

Partnering with Cuba: Weather extremes

WE SUPPORT THE RECENT Editorial on "Science diplomacy with Cuba" (G. R. Fink, A. I. Leshner, V. C. Turekian, 6 June, p. 1065). Many groups in the United States seek greater cooperation with Cuba in the earth sciences. Similar to the disease mitigation discussed in the Editorial, there is a long history of scientific advances that could save lives and reduce losses with greater cooperation.

Cuban scientists did forecast the severity of the 1900 Galveston hurricane and could have helped the nascent U.S. Weather Bureau in Galveston mitigate fatalities and losses (*1*). More recently, after the devastating 2010 M_w7.0 Haiti earthquake, the National Science Foundation supported the development of the Continuously Operating Caribbean Global Positioning System (GPS) Observational Network (COCONet) (*2*), a GPS and meteorological network in the circum-Caribbean designed to help understand tectonic and weather extremes in a complex geologic and atmospheric region. Partially facilitated by a previous AAAS delegation visit to Cuba in December 2011, a new GPS station was recently installed in Camagüey, Cuba, and is a critical component of COCONet. More broadly, COCONet involves more than 38 nations and everyone benefits from international cooperation and diplomacy that revolve around research, education, and broader impacts such as risk resiliency and enhancing the quality of life.

We encourage easing restrictions so that government science agencies may cooperate more fully on science and education. For example, the Cuban Meteorological Institute, Cuban Nuclear Agency, the National Seismological Center, and the Astronomical and Geophysical Institute could then cooperate with their U.S. counterparts on natural hazards research.

Juan Carlos Antuña Marrero,[1] M. Meghan Miller,[2] Glen Mattioli,[2]* Karl Feaux,[2] Richard Anthes,[3] John Braun,[3] Guoquan Wang,[4] Alan Robock[5]

[1]Grupo de Óptica Atmosférica de Camagüey, Centro Meteorológico de Camagüey, Instituto de Meteorología de la República de Cuba, Cuba. [2]UNAVCO, Boulder, CO 80301, USA. [3]University Corporation for Atmospheric Research, Boulder, CO 80305, USA. [4]Department of Earth and Atmospheric Sciences, University of Houston, TX 77004, USA. [5]Department of Environmental Sciences, Rutgers University, New Brunswick, NJ 08901, USA.

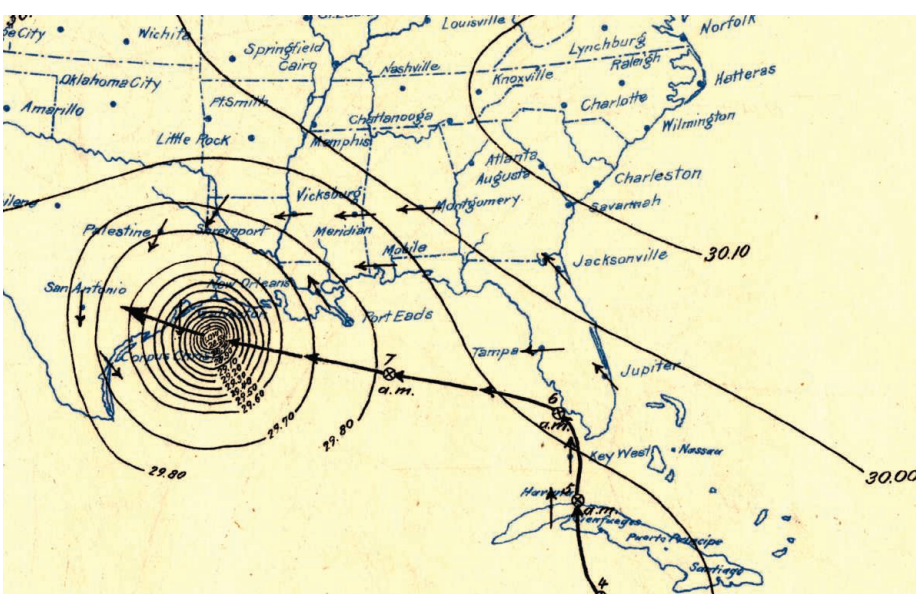
*Corresponding author. E-mail: mattioli@unavco.org

REFERENCES

1. N. C. Green, *Story of the 1900 Galveston Hurricane* (Pelican Publishing Company, Gretna, LA, 1999).
2. J. J. Braun *et al.*, *EOS Trans. Am. Geophys. Union* **93**, 9 (2012).

Partnering with Cuba: Earthquake hazards

THE 6 JUNE EDITORIAL BY G. R. Fink, A. I. Leshner, and V. C. Turekian advocates "Science diplomacy with Cuba" (p. 1065). I agree. Both countries could benefit from a joint evaluation of earthquake hazards in southeastern Cuba.



The cities of Santiago de Cuba and Guantánamo, including the U.S. base and prison, are close to the offshore, seismically active Oriente plate-boundary fault (*1*), but the earthquake danger from this fault has not been studied. The fault occupies the narrow Oriente Deep (*1*, *2*) and skims the northern boundary of Haiti into the Dominican Republic, where paleoseismic analysis of the Septentrional fault shows evidence of earthquakes and a slip rate of 6 to 12 mm/year (*3*). This is about half of the total slip rate between the North American and Caribbean plates and serves as a proxy for the slip rate estimated for the Oriente fault off the coast of Cuba. No historical earthquake has ruptured the Oriente fault, indicating that the fault may be in a seismic gap. The United States has the expertise to help Cuba respond to its earthquake hazard by upgrading their seismic and global positioning system (GPS) networks, conducting paleoseismic studies along the coast to learn the earthquake history of the Oriente fault, and upgrading buildings at risk from an Oriente fault earthquake.

Robert S. Yeats

College of Earth, Ocean, and Atmospheric Sciences, Oregon State University, Corvallis, OR 97331, USA. E-mail: yeatsr@geo.oregonstate.edu

REFERENCES

1. R. S. Yeats, *Active Faults of the World* (Cambridge Univ. Press, Cambridge, 2012).
2. E. Calais *et al.*, *Geol. Soc. Am. Spec. Pap.* **326**, 125 (1998).
3. C. S. Prentice *et al.*, *J. Geophys. Res.* **103**, 2149 (2003).

Delisted whales good news for pipeline

ON EARTH DAY (22 APRIL), the Canadian government removed the North Pacific humpback whale, which lives off British Columbia's coast, from the list of threatened species (*1*, *2*). The delisting was based on sound scientific advice by the Committee on the Status of Endangered Wildlife in Canada (COSEWIC) (*3*, *4*), but the downgrading from threatened to a species of special concern occurred just before the federal decision to approve the development of the Enbridge's Northern Gateway pipeline mega-project in the humpbacks' habitat in the summer of 2014 (*1*, *2*, *5*). The delisting from threatened status raised concerns among environmental groups (*4*, *5*), who viewed it as a fast-tracked political decision driven by economic factors (*1*, *2*). The decision has implications for the whales' critical habitat, which does not have to be protected once the species is downlisted, and questions linger about whether whale

PHOTO: NOAA.GOV

Downloaded from www.sciencemag.org on August 4, 2014

Published by AAAS

use of the habitat, an impediment for the Enbridge's project and maritime traffic of oil tankers, was suitably assessed (*5*).

After being decimated off British Columbia's coast by commercial whaling from the early 1900s to mid-1960s, the current North Pacific humpback whale population has bounced back to approximately 18,000 adult individuals (*6*–*8*). However, the local population using British Columbia's coastal waters as feeding and migrating grounds is still small (2145 whales) relative to the estimated pre-whaling levels (about 4000 whales), and the whales face increasing anthropogenic impacts, including vessel strikes, underwater noise, oil spills, and prey reduction (*3*, *7*–*9*). Because of the small population size and threats to its habitat in this region, British Columbia humpbacks should be designated as a special management unit. This management unit should be defined in the context of risk management in a given region and season, which may or may not be equivalent to a single biological stock, and it should aim to avoid local extirpations and imply that local depletion will recover to healthy levels on a management time scale (*10*).

On 17 June 2014, the Enbridge's project was approved by Canada's Prime Minister (*11*). If this project proceeds, these whales will be exposed to increased risk from tanker traffic and oil spills. The Canadian government should stand by its responsibility to preserve marine mammals and their habitats.

Juan José Alava[1,2]* and Joshua N. Silberg[1]

[1]School of Resource and Environmental Management, Faculty of Environment, Simon Fraser University, 8888 University Drive, Burnaby, BC V5A 1S6, Canada. [2]Fundación Ecuatoriana para el Estudio de Mamíferos Marinos, Guayaquil, Ecuador.

*Corresponding author. E-mail: jalavasa@sfu.ca

REFERENCES

1. E. Chung, "Humpback whale losing 'threatened' status amid Northern Gateway concerns," *CBC News* (22 April 2014); www.cbc.ca/news/technology/humpback-whale-losing-threatened-status-amid-northern-gateway-concerns-1.2617633.
2. D. Moore, "Feds to lift protections for Pacific humpback," *Global News* (22 April 2014); http://globalnews.ca/news/1283846/feds-to-lift-protections-for-pacific-humpback/.
3. COSEWIC, "COSEWIC assessment and status report on the humpback whale *Megaptera novaeangliae* North Pacific population in Canada" (Committee on the Status of Endangered Wildlife in Canada, Ottawa, Canada, 2011); http://publications.gc.ca/collections/collection_2012/ec/CW69-14-233-2011-eng.pdf.
4. A. Trites, "Attacks on whale science send shameful message," *HuffPost British Columbia* (24 June 2014); www.huffingtonpost.ca/andrewtrites/canada-humpback-whales-scientists_b_5527041.html.
5. A. J. Rosenberger, "RE: Comments on the proposed down-listing of the North Pacific humpback whale," *Raincoast Conservation Foundation* (30 April 2014); www.raincoast.org/wp-content/uploads/2014/06/humpback-de-listing-comments-RCF-to-SARA.pdf.
6. L. M. Nichol *et al.*, "British Columbia commercial whaling catch data 1908 to 1967: A detailed description of the BC historical whaling database" (Can. Tech. Rep. Fish. Aquat. Sci. 2371, Fisheries and Oceans Canada, Nanaimo, 2002).
7. Fisheries and Oceans Canada, "Recovery strategy for the North Pacific humpback whale (*Megaptera novaeangliae*) in Canada" (Species at Risk Act Recovery Strategy Series, Ottawa, Canada, 2013); www.sararegistry.gc.ca/document/default_e.cfm?documentID=1344.
8. J. Calambokidis *et al.*, "SPLASH: Structure of populations, levels of abundance and status of humpback whales in the North Pacific" (Final report for Contract AB133F-03-RP-00078, U.S. Department of Commerce, Seattle, WA, 2008).
9. J. K. B. Ford *et al.*, "An assessment of the potential for recovery of humpback whales off the Pacific Coast of Canada" (DFO Canadian Science Advisory Secretariat Research Document 2009/015, 2009); www.dfo-mpo.gc.ca/CSAS/Csas/Publications/ResDocs-DocRech/2009/2009_015_e.pdf.
10. J. A. Jackson, C. Pampoulie, "Stock definition: Terminologies revisited," Document SC/64/SD, Report of the Scientific Committee, International Whaling Commission, Panama City, Panama, 11 to 23 June 2012.
11. G. Hoekstra, P. O'Neil, D. Penner, R. Shaw, "Northern Gateway pipeline approved by Harper government," *Vancouver Sun* (18 June 2014); www.vancouversun.com/business/Northern+Gateway+pipeline+approved+Harper+government/9947588/story.html.

OUTSIDE THE TOWER

A night at the museum

At midnight I survey the scene. On the far end of the room, a levitating train makes circles by a whale exhibit, demonstrating exciting applications of superconductors. As it passes, nearby spectators in dinosaur pajamas learn about water bears. To my right, several adults giggle as they practice using hula hoops to demonstrate hypothesis-driven science. To my left, on the other end of the room, a woman in jeans holds a knife over a human brain, preparing to make a slice through the hippocampus. Throughout the room, speakers in red shirts are enthusiastically demonstrating science.

Science often requires lab work long after sunset, but tonight we've left the lab and our ivory tower behind and brought our most exciting research with us. Since 2011, the California Academy of Sciences has held Penguins and Pajamas Sleepover for grown-ups, an annual event in which the regular museum exhibits are open to the public overnight. This year, it became a public venue for science after sunset as I, along with Mary Cavanagh, invited 40 Stanford postdocs to don red shirts and bring interactive demonstrations of their current research to share (*1*, *2*).

Our visitors seemed tentative at first, looking at the myriad exhibits but reluctant to approach us. We broke the ice by starting conversations, and suddenly everyone was visibly engaged in the cutting-edge science taking place right around the corner. Some people were interested in brain function and peppered Viola Caretti with questions. Others wanted to know more about magnetism and superconductivity from Stefano Bonetti. A bar provided guests with drinks under the projection of artistic science images curated by Ermelinda Porpiglia, and late at night guests could find a unique place to sleep next to live penguins or by the light of glowing jellyfish.

Stefano Bonetti pours liquid nitrogen to cool a superconductor for a levitating train experiment.

At first light, the crowd dispersed and the postdocs packed up their materials and went home. In the months since that night, Stanford postdocs have continued to perform their science at museum events, satisfied in knowing they can make science come alive.

Antoine de Morrée

Stanford School of Medicine, Stanford University Postdoctoral Association, Stanford, CA 94305, USA.
E-mail: demorree@stanford.edu

REFERENCES

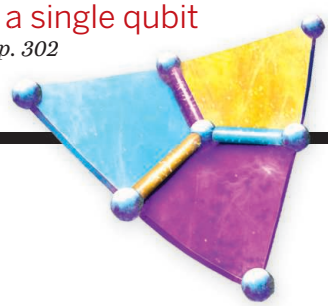
1. M. Cavanagh, Penguins, pyjamas and postdocs (www.neuwritewest.org/blog/4931).
2. E. Seigneur, NeuroTalk S2E3 Penguins & Pajamas (www.neuwritewest.org/blog/4791).

PHOTO: LUQIA HOU

Published by AAAS

RESEARCH

Trapping can reveal the quantum state of a single qubit
Nigg et al. *p. 302*



IN *SCIENCE* JOURNALS

Edited by **Stella Hurtley**

Nematode worm, *Caenorhabditis elegans*

ASSOCIATIVE LEARNING

How the worm changes its tastes

In associative learning, you link potentially unrelated things because you are exposed to them at the same time. Ohno *et al.* studied a simple associative learning task in the nematode worm *Caenorhabditis elegans*. They presented the worms with a taste substance while withholding food. After starving in the presence of the taste substance, the animals switched their behavior from being attracted to the taste to finding it aversive. A specific isoform of the insulin receptor is critical for this type of associative learning—at least in worms. — PRS

Science, this issue p. 313

QUANTUM MECHANICS

Bosons of a feather flit together

Bosons are a type of particle that likes to congregate. This property has a major effect on the behavior of identical bosons. Kaufman *et al.* demonstrated quantum interference of two bosonic Rb atoms placed in two neighboring quantum wells (see the Perspective by Thompson and Lukin). They prepared the atoms in exactly the same state so that there would be no way to tell them apart except for which well each atom was in. They then monitored the probability of the two atoms still being in separate wells. At certain times, the probability had a characteristic dip signifying that the bosons preferred to be in the same well. — JS

Science, this issue p. 306; see also p. 272

GENE THERAPY

Reprogrammed heart cells set the pace

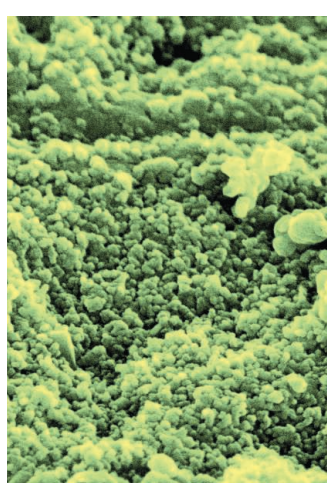
Pacemakers have revolutionized the care of patients with slow or abnormal heart rhythms, but these devices can break or become infected. With these patients in mind, Hu *et al.* created biological pacemakers to provide temporary, hardware-free support until a damaged electronic device can be replaced. They inserted a gene for a human transcription factor into heart muscle cells. This gene reprogrammed the cells to become pacemakers—cells that emit rhythmic electrical impulses to drive the beating heart. These biological pacemaker cells restored normal heart rate in pigs with complete heart block—a problem with the heart's electrical system. — MLF

Sci. Transl. Med. **6**, 245ra94 (2014).

PHOTOVOLTAICS

Improved perovskite photovoltaic performance

A recent entry in the solar cell race is perovskite cells, named for the structure adopted by salt made from metal halides and organic cations that absorb the light and generate charges. The charges generated have to be transferred to a metal oxide (typically titanium oxide), and some of these charge carriers are lost in the transfer. Mei *et al.* made this process more efficient by growing a more crystalline perovskite with fewer defects inside porous versions of titanium and zirconium oxide. They added a second organic cation that stuck to the pore walls and directed the growth of the perovskite crystals. The improved solar cells



Cross section of perovskite solar cell

PHOTOS: (COUNTERCLOCKWISE FROM TOP) HARALD RITSCH; SINCLAIR STAMMERS/SCIENCE SOURCE; MEI *ET AL.*

Downloaded from www.sciencemag.org on August 4, 2014

Published by AAAS

operated for more than 1000 hours under full sunlight. — PDS

Science, this issue p. 295

PALEOCEANOGRAPHY

Circulating between different cycles

Around a million years ago, large-scale ocean circulation changed dramatically during the switch from ~41,000- to ~100,000-year long glacial cycles. Pena and Goldstein analyzed the isotopic composition of neodymium in marine sediments from the South Atlantic. The results suggest how the contributions of deep water from northern and southern sources varied across the transition. The boundary between the two glacial states appears to have been marked by a dramatic weakening, perhaps even a shutdown, of deep-water currents. — HJS

Science, this issue p. 318

FOOD SECURITY

How to optimize global food production

Keeping societies stable and managing Earth's resources sustainably depend on doing a good, steady job producing and distributing food. West *et al*. asked what combinations of crops and regions offer the best chance of progress. Their analysis focused on reducing greenhouse gas emissions, nutrient pollution, water use, and food waste. They identify regions that are likely to yield the best balance between applying fertilizer to increase crop yields versus the resulting environmental impact. — AMS

Science, this issue p. 325

ALTERNATIVE SPLICING

Evolving from an enzyme and into a regulator

Proteins, the work-horses of the cell, are made on a messenger RNA (mRNA) template. An enzyme called aminoacyl tRNA synthetases (AARSs) attaches the correct amino acid to a transfer RNA so that mRNA is accurately translated. Over evolution, additional sequences have been added to AARSs. Lo *et al.* found a large number of AARS variants in which the domain responsible for enzyme function was deleted. Ninety-four such variants had diverse signaling activities. Thus, AARSs are used both as enzymes and alternately as regulators of signaling pathways. — VV

Science, this issue p. 328

MITOSIS

Taking a check on chromosome spacing

Animal cells divide by mitosis. Chromosomes become condensed and congregate on the mitotic spindle in the center of the cell—the midzone. The spindle then separates sister chromosomes, pulling them to opposite ends of the cell, ready to form new daughter nuclei. Afonso *et al.* now show that chromosome separation is monitored by the level of midzone-associated Aurora B kinase activity (see the Perspective by Hadders and Lens). This process ensures that daughter nuclei only reassemble after sister chromosomes have successfully separated. — SMH

Science, this issue p. 322; see also p. 265

SYNTHETIC BIOLOGY

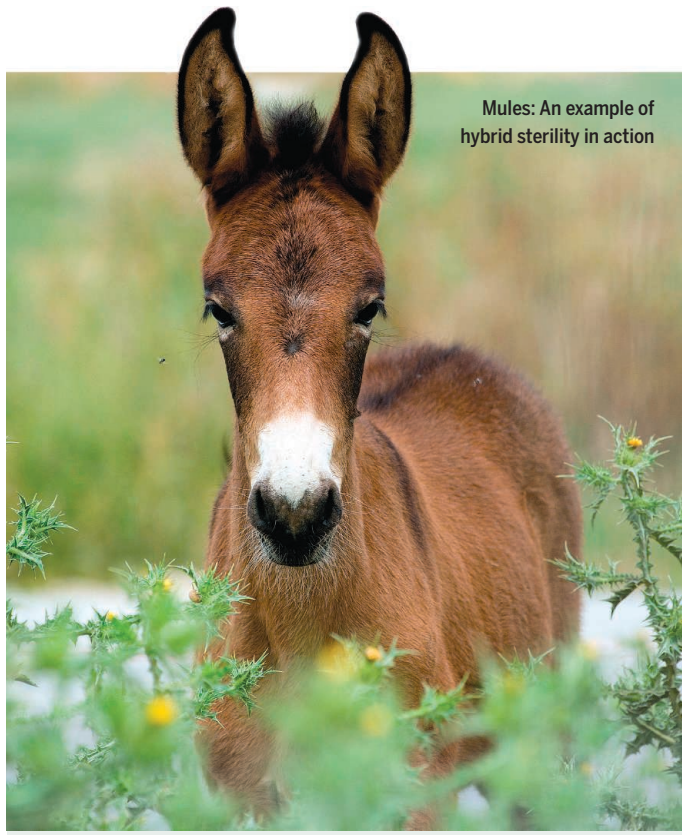
Turning cells into garbage collectors

To protect other cells or tissue from injury, professional garbage-collecting cells called phagocytes gobble up dead cells. Sometimes phagocytes become overwhelmed and could use some help. Onuma *et al*. manipulated nonphagocytic cells to express two proteins: one that recognizes a lipid that marks the surface of dying cells and another that alters the plasma membrane to make the cell phagocytic. The engineered cells bound to and consumed dying cells. Such engineered cells could potentially be used to help the body remove undesirable targets. — JFF

Sci. Signal. **7**, rs4 (2014).

IN OTHER JOURNALS

Edited by **Kristen Mueller** *and* **Jesse Smith**



Mules: An example of hybrid sterility in action

SPECIATION

Putting the genetic breaks on breeding

Several genetic barriers prevent species from interbreeding. For instance, when two different species mate, their offspring are often infertile. Identifying the genetic causes of this hybrid sterility provides insight into how species evolve genetically. Zanders *et al*. examined two closely related yeast species, which share 99.5% of their DNA sequence. They found that crosses between the species could not produce fertile offspring because of meiotic drive loci, genetic elements that are preferentially transmitted to the offspring within a species. The three DNA sites, which were present in one of the species, blocked the hybrid from developing spores that contained chromosomes from the other species. This study highlights how quickly barriers to fertility can arise between two closely related species. — LMZ

eLife 10.7554/eLife.02630.001 (2014).

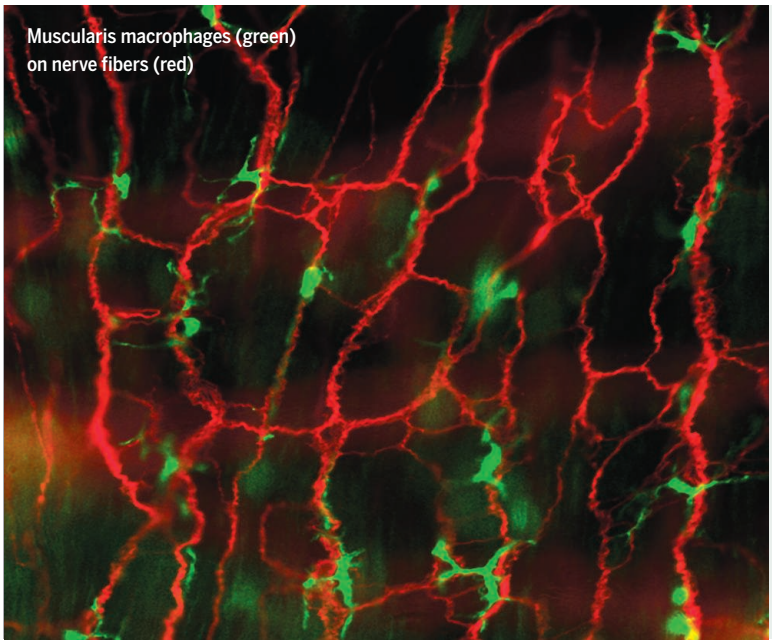
CELL NUCLEUS

Packaging DNA for a better nucleus

Nuclear pores oversee the movement of proteins, DNA, and RNA across the nuclear membrane between the nucleus and cytoplasm—an essential duty for proper cell function and survival. Two studies now report that for these pores to form, the cell requires nucleosomes, short lengths of DNA wrapped around a protein core that package chromosomes into the nucleus.

PHOTO: © IMAGEBROKER/ALAMY

Published by AAAS


Muscularis macrophages (green) on nerve fibers (red)

GUT MOTILITY

Macrophages help food move through

Food needs a complex array of cellular interactions to move through the body. Neurons, muscle cells, and interstitial cells all cooperate to ease it through the gastrointestinal (GI) tract. Now Muller *et al*. report intestinal muscularis macrophages, a type of immune cell that resides in the smooth muscles that surround the GI tract, participate, too. These macrophages secrete a substance called bone morphogenetic protein 2 (BMP2), which binds to enteric neurons and directs them to coordinate the muscle cell contractions that squeeze food through. The neurons, in turn, produce a growth factor required by the macrophages. Macrophage-neuron crosstalk is essential: When mice don't have enough of the growth factor, BMP2, or muscularis macrophages, they have defects in gut muscle contractions. — KLM

Cell 10.1016/j.cell.2014.04.050 (2014).

Inoue and Zhang and Zierhut *et al*. use either mouse zygotes or frog (*Xenopus*) egg extracts to show that nucleosomes bind to and recruit the nuclear pore protein ELYS (embryonic large molecule derived from yolk sac) to DNA, driving the initial steps of nuclear pore assembly. In this way, cells might ensure that a functional nucleus only forms around properly packaged chromosomes. — GR

Nat. Struct. Mol. Biol. 10.1038/nsmb.2839; 10.1038/nsmb.2845 (2014).

ATMOSPHERIC SCIENCE

Looking for signs of ozone recovery

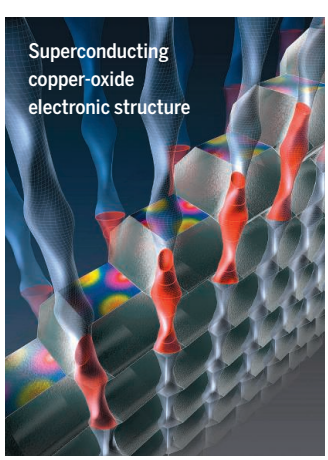
How effective is the Montreal Protocol, the international treaty aimed at protecting the stratospheric ozone layer? Since the treaty was adopted in 1989, manufacturers have gradually stopped producing numerous chemicals that deplete ozone. Scientists have predicted that ozone will recover as a result. Coldewey-Egbers *et al.* use an extensive data set to test whether recovery can be detected yet. They conclude that in mid-latitudes, at least five more years of data are needed to unequivocally detect signs of recovery, because natural variability masks the ozone recovery. At lower latitudes, where the trend in ozone concentration is smaller, even more data will be needed. — JFU

Geophys. Res. Lett. 10.1002/2014GL060212 (2014).

PHYSICS

A copper oxide's electronic structure

Physicists still do not understand why some copper-oxide compounds become superconducting at relatively high temperatures. Even more basic issues have remained controversial as well, such as the electronic structure of the normal (i.e., nonsuperconducting) state from which superconductivity emerges. Sebastian *et al.* chose $YBa_2Cu_3O_{6.56}$ as a good representative of a subclass of these compounds and used high magnetic fields to suppress its superconductivity and reach its normal state. They analyzed the wiggles of the electrical resistivity to map out the shape of the so-called Fermi surface, which separates the quantum states filled with electrons from the empty ones. They found an undulating shape that suggested that the electronic structure was organized differently from what one expects based on the crystal lattice structure alone. — JS


Superconducting copper-oxide electronic structure

Nature 10.1038/nature13326 (2014).

CIRCADIAN RHYTHMS

Tick tock, synchronizing biological clocks

Studies of coupled oscillators started in the 1600s, when the man who invented the pendulum clock set a pair of clocks side by side in a single case and noticed that they started ticking in unison. In mammalian cells, the machinery that controls the cell division cycle turns out to be similarly synchronized with the daily circadian clock, which allows cells to get on the same seasonal and day-night schedules. Feillet *et al.* imaged single live mammalian cells in culture and performed mathematical modeling. They showed that the daily circadian clock and the cell division cycle oscillate together at the same frequency. This may have clinical relevance: Circadian disruption is a risk factor for some cancers. — LBR

Proc. Natl. Acad. Sci. U.S.A. 10.1073 pnas.1320474111 (2014).

CLIMATE CHANGE

Measuring sea ice through thick and thin

Sea ice in the Arctic's Fram Straight is only half as thick as it was a decade ago, Renner and colleagues report. Measuring ice both at the surface and from the air with an instrument towed below a helicopter, they found that the ice in that region thinned by more than 50% between 2003 and 2012. The patterns of variability that they see most likely are caused by thinning occurring more broadly across the Arctic, rather than being just a regional effect. These observations are consistent with the contemporaneous reduction of the area covered by sea ice that has occurred in the Arctic and should help provide a better understanding of both the magnitude and the causes of the loss. — HJS

Geophys. Res. Lett. 10.1002/2014GL060369 (2014).

PHOTOS: (TOP TO BOTTOM) MILENA BOGUNOVIC; NICOLLE R. FULLER

Published by AAAS

ALSO IN *SCIENCE* JOURNALS

Edited by **Stella Hurtley**

ORGANOID GENERATION

The making of bodies part by part

Mention of organoids—organ-like structure growing in a Petri dish—might conjure up images of science fiction. However, the generation of organoids is very real, as is the increased understanding of organ form and function that comes from studying them. Lancaster and Knoblich review organoids as structures that include more than one cell type of an organ that exhibit structural and functional features of the natural counterpart. Knowledge of normal organ developmental pathways guides the formation of these structures. Organoids show great promise for modeling human development and disease and for biomedical research and regenerative medicine. — BAP

Science, this issue p. 283

ASYMMETRIC CATALYSIS

Iodine blooms as an oxidation catalyst

Most catalysts for organic oxidation chemistry—whether biochemical or artificial—contain a transition metal like iron or palladium. Uyanik *et al.* now show that iodine can take the place of a metal in catalyzing efficient oxidative ring closures to make chromans—hexagonal rings incorporating oxygen that are perhaps best known as a constituent of the vitamin E structure (see the Perspective by Nachtsheim). The iodine is added as a salt with a chiral cation, which directs the reaction to form just one of two possible mirror-image variants of the product. Key to the success of the system was the addition of a base, which maintained the viability of an unstable, partially oxidized iodine intermediate critical to the reaction cycle. The results bode well for more general use of iodine salts as asymmetric oxidation-reduction catalysts. — JSY

Science, this issue p. 291; see also p. 270

ULTRAFAST DYNAMICS

Tightly tracking charge migration

Electron transfer dynamics underlie many chemical and biochemical reactions. Erk *et al.* examined the charge migration between individual carbon and iodine atoms during dissociation of iodomethane (ICH_3) molecules (see the Perspective by Pratt). After initiating scission of the C-I bond with a relatively low-energy laser pulse, they introduced a higher-energy x-ray pulse to instigate ionization and charge migration. Delaying the arrival time of the x-ray pulse effectively varied the separation distance being probed as the fragments steadily drifted apart. The experimental approach should also prove useful for future studies of charge transfer dynamics in different molecular or solid-state systems. — JSY

Science, this issue p. 288; see also p. 267

APPLIED OPTICS

Extending the range of planar optics

To build miniature optical devices, scientists are using silicon to replace bulky three-dimensional devices with flat versions. A patterned surface consisting of dense arrays of nanoscale silicon strips, which act as antennae, can be designed to work as transparent optical devices for the manipulation of light. Lin *et al.* used their versatile patterning technique to create a suite of planar optical elements. By patterning a 100-nm layer of silicon into a dense arrangement of nano-antennae, they were able to fabricate gratings, lenses, and axicons—a device that can add a shape to a propagating light beam. — ISO

Science, this issue p. 298

QUANTUM COMPUTING

Fault-tolerant quantum computing

Quantum states can be delicate. Attempts to process and manipulate quantum states can destroy the encoded information. Nigg *et al.* encoded the quantum state of a single qubit (in this case, a trapped ion) over the global properties of a series of trapped ions. These so-called stabilizers protected the information against noise sources that can degrade the single qubit. The protocol provides a route to fault-tolerant quantum computing. – ISO

Science, this issue p. 302

OPTICAL METROLOGY

A more simple route to metrology standards

The time standards set by atomic clocks are usually the domain of national metrology laboratories. Li *et al.* modulated the phase of two closely spaced laser lines to generate a comb of stable and equally spaced frequencies. They then used that comb to generate stable microwaves. Being somewhat simpler than existing optical comb generation techniques and also tunable, the approach may offer a new era of metrology and precision timekeeping. — ISO

Science, this issue p. 309

OCEANOGRAPHY

Spinning up ocean circulation discretely

Ocean eddies with diameters of tens to hundreds of kilometers are an important component of water transport on a global scale. Averaged over periods of many years, ocean circulation is regular and coherent enough to have been described as a conveyor belt. However, over shorter time scales, irregular, discrete events cause significant amounts of water movement—termed mesoscale eddies. Zhang *et al.* combined data from thousands of autonomous observational devices called ARGO floats with satellite data. The mesoscale eddies moved around as much ocean water as wind or deep-water currents. — HJS

Science, this issue p. 322

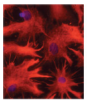
MOLECULAR MOTORS

How dynein makes the right moves

The molecular motor cytoplasmic dynein moves a wide range of different intracellular cargoes. Dynein's activity in vivo requires another protein, dynactin, but exactly why that should be has been very unclear. Although in vitro experiments have provided some evidence that dynactin increases dynein's processivity, the resulting dynein motility has never come close to matching dynein's cargo-transporting activity in living cells. Now, McKenney *et al.* show that tripartite complexes of dynein, dynactin, and an adaptor molecule are highly processive in vitro, moving the sort of distances that dynein transports cargo in vivo (see the Perspective by Allan). — SMH

Science, this issue p. 337; see also p. 271

Published by AAAS

EXTENDED PDF FORMAT
SPONSORED BY

Travel Grants Available
Neuroscience 2014
Application Deadline: July 31st
>Apply Now
R&D systems
a bio-techne brand
www.rndsystems.com

Organogenesis in a dish: Modeling development and disease using organoid technologies

Madeline A. Lancaster and Juergen A. Knoblich

Science **345**, (2014);
DOI: 10.1126/science.1247125

This copy is for your personal, non-commercial use only.

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by clicking here.

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines here.

The following resources related to this article are available online at www.sciencemag.org (this information is current as of August 4, 2014):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:
http://www.sciencemag.org/content/345/6194/1247125.full.html

This article **cites 120 articles**, 37 of which can be accessed free:
http://www.sciencemag.org/content/345/6194/1247125.full.html#ref-list-1

This article appears in the following **subject collections:**
Development
http://www.sciencemag.org/cgi/collection/development

Downloaded from www.sciencemag.org on August 4, 2014

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published weekly, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. Copyright 2014 by the American Association for the Advancement of Science; all rights reserved. The title *Science* is a registered trademark of AAAS.

REVIEW SUMMARY

ORGANOID GENERATION

Organogenesis in a dish: Modeling development and disease using organoid technologies

Madeline A. Lancaster *and* Juergen A. Knoblich*

BACKGROUND: Because of their differentiation potential, pluripotent stem cells can generate virtually any cell type and, as such, can be used to model development and disease and even hold the promise of providing cell-replacement therapies. Recently, structures resembling whole organs, termed organoids, have been generated from stem cells through the development of three-dimensional culture systems.

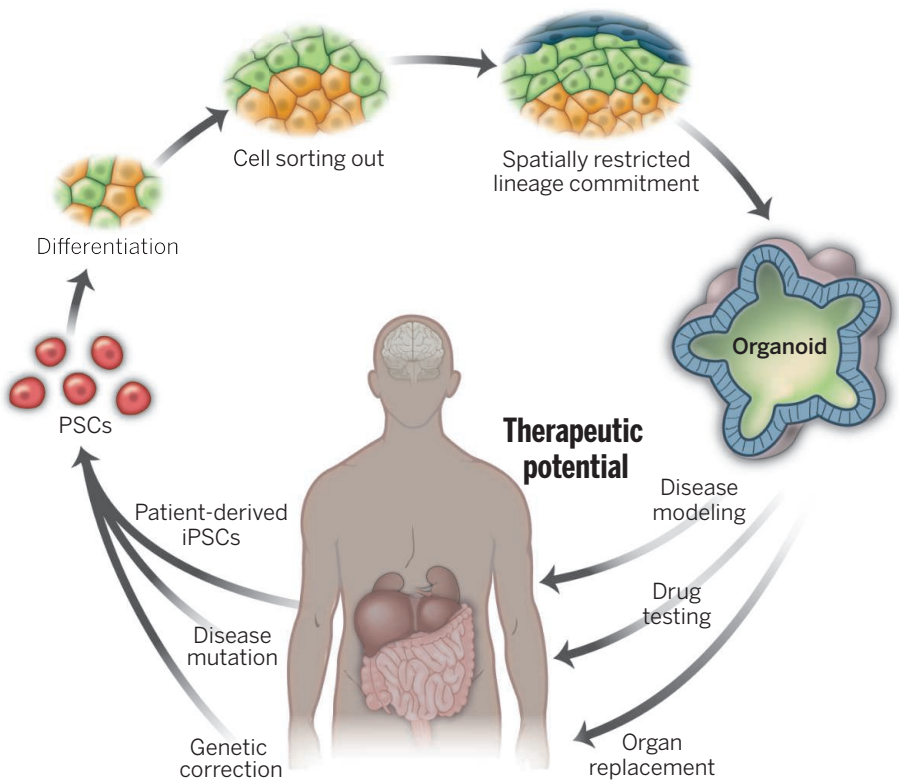
Organoids are derived from pluripotent stem cells or isolated organ progenitors that differentiate to form an organlike tissue exhibiting multiple cell types that self-organize to form a structure not unlike the organ in vivo. This technology builds upon a foundation of stem cell technologies, as well as classical developmental biology and cell-mixing experiments. These studies illustrated two key events in structural organization during organogenesis: cell sorting out and spatially restricted lineage commitment. Both of these processes are recapitulated in organoids, which self-assemble to form the cellular organization of the organ itself.

ADVANCES: Organoids have been generated for a number of organs from both mouse and human stem cells. To date, human pluripotent stem cells have been coaxed to generate intestinal, kidney, brain, and retinal organoids, as well as liver organoid-like tissues called liver buds. Derivation methods are specific to each of these systems, with a focus on recapitulation of endogenous developmental processes.

These complex structures provide a unique opportunity to model human organ development in a system remarkably similar to development in vivo. Although the full extent of similarity in many cases still remains to be determined, organoids are already being applied to human-specific biological questions. Indeed, brain and retinal organoids have both been shown to exhibit properties that recapitulate human organ development and that cannot be observed in animal models. Naturally, limitations exist, such as the lack of blood supply, but future endeavors will advance the technology and, it is hoped, fully overcome these technical hurdles.

OUTLOOK: The therapeutic promise of organoids is perhaps the area with greatest potential. These unique tissues have the potential to model developmental disease, degenerative conditions, and cancer. Genetic disorders can be modeled by making use of patient-derived induced pluripotent stem cells or by introducing disease mutations. Indeed, this type of approach has already been taken to generate organoids from patient stem cells for intestine, kidney, and brain.

Furthermore, organoids that model disease can be used as an alternative system for drug testing that may not only better recapitulate effects in human patients but could also cut down on animal studies. Liver organoids, in particular, represent a system with high expectations, particularly for drug testing, because of the unique metabolic profile of the human liver. Finally, tissues derived in vitro could be generated from patient cells to provide alternative organ replacement strategies. Unlike current organ transplant treatments, such autologous tissues would not suffer from issues of immunocompetency and rejection. ■



Organoid generation and therapeutic potential. Organoids can be derived for a number of organs from human pluripotent stem cells (PSCs). Like organogenesis in vivo, organoids self-organize through both cell sorting out and spatially restricted lineage commitment of precursor cells. Organoids can be used to model disease by introducing disease mutations or using patient-derived PSCs. Future applications could include drug testing and even tissue replacement therapy.

ILLUSTRATION: K. SUTLIFF/*SCIENCE*

ON OUR WEBSITE

Read the full article at http://dx.doi.org/10.1126/science.1247125

IMBA—Institute of Molecular Biotechnology of the Austrian Academy of Sciences, Vienna 1030, Austria.
* Corresponding author. E-mail: juergen.knoblich@imba.oeaw.ac.at

Cite this article as M. A. Lancaster and J. A. Knoblich, *Science* 345, 1247125 (2014). DOI: 10.1126/science.1247125

Published by AAAS

REVIEW

ORGANOID GENERATION

Organogenesis in a dish: Modeling development and disease using organoid technologies

Madeline A. Lancaster[1] and Juergen A. Knoblich[1]*

Classical experiments performed half a century ago demonstrated the immense self-organizing capacity of vertebrate cells. Even after complete dissociation, cells can reaggregate and reconstruct the original architecture of an organ. More recently, this outstanding feature was used to rebuild organ parts or even complete organs from tissue or embryonic stem cells. Such stem cell–derived three-dimensional cultures are called organoids. Because organoids can be grown from human stem cells and from patient-derived induced pluripotent stem cells, they have the potential to model human development and disease. Furthermore, they have potential for drug testing and even future organ replacement strategies. Here, we summarize this rapidly evolving field and outline the potential of organoid technology for future biomedical research.

Stem cell technologies hold promise for modeling development, analyzing disease mechanisms, and developing potential therapies. Until quite recently, most stem cell methods focused on pure populations of particular stem cell–derived cell types (*1*), rather than the complete set of cell types of an organ. However, this is beginning to change with the development of three-dimensional (3D) cultures of developing tissues, called organoids.

As organoid technology is on the verge of becoming an independent research field, a precise definition of the term becomes increasingly important. The term organoid, simply defined as resembling an organ, has been used quite loosely for a variety of structures, both in vitro and in vivo (*2*–*4*). The basic definition, however, implies several important features that are characteristics of organs (Box 1). First, it must contain more than one cell type of the organ it models; second, it should exhibit some function specific to that organ; and third, the cells should be organized similarly to the organ itself. This also implies similarity to the manner in which the organ establishes its characteristic organization during development. Thus, we would like to define an organoid as containing several cell types that develop from stem cells or organ progenitors and self-organize through cell sorting and spatially restricted lineage commitment, similar to the process in vivo (Box 1).

Self-organization: The foundation of organoid formation

Organoid methods build upon an extensive foundation of classic developmental biology and cell dissociation and reaggregation experiments (Fig. 1A). Two distinct approaches have been taken to understanding tissue patterning (*5*). In vivo examination of cell movements has revealed mechanisms of cell segregation into discrete domains during tissue morphogenesis (*6*). This process has been extensively examined in, for example, the *Drosophila* wing disc where the anterior-posterior boundary is established through mutually repressive interactions (*7*). A similar process occurs during vertebrate embryonic development at the midbrain-hindbrain boundary, which then acts as an organizer for subsequent tissue morphogenesis (*8*).

The second approach to understanding tissue patterning has been dissociation and reaggregation of tissues to examine relative morphogenetic movements of cells in vitro (Fig. 2A). This approach has been applied to virtually all developing vertebrate organs in a number of classic studies with embryonic chick tissues (*9*, *10*) (Fig. 1A). The results point to a general capacity of cells to reorganize and segregate in a process termed "cell sorting out" to form structures with much the same histogenic properties as those in vivo (*6*, *11*). For example, studies from the early 1960s have utilized dissociated cells from the developing chick kidney (*9*) to form reaggregates that recapitulate virtually complete renal development.

The basis of this organ self-assembly seems to arise from segregation of cells with similar adhesive properties into domains that achieve the most thermodynamically stable pattern (Fig. 2A). Known as Steinberg's differential adhesion hypothesis (*12*) (Fig. 1A), the theory is supported by a range of in vitro cell-mixing experiments (*13*). Differential adhesion is mediated by cell surface adhesion proteins, for example, in separation of vertebrate neural and epidermal ectoderm (*14*, *15*), where differential epithelial and neural and cadherin expression mediates cell sorting out.

A second mechanism that can influence tissue morphogenesis is proper spatially restricted progenitor fate decisions (Fig. 2B). An excellent example is the developing vertebrate retina, where neuroepithelial cells give rise to a complex lineage that generates the various layers of the retina in a temporally and spatially restricted manner.

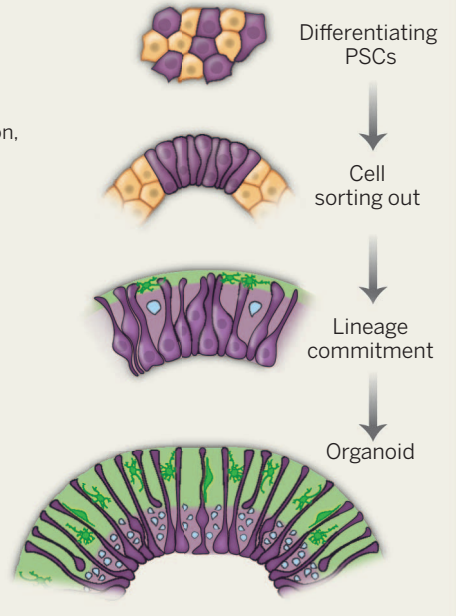
Box 1. Defining organoids.

Organoid *n.* Resembling an organ.

This implies:
1. Multiple organ-specific cell types
2. Capable of recapitulating some specific function of the organ (eg. excretion, filtration, neural activity, contraction)
3. Grouped together and spatially organized similar to an organ

Organoid formation recapitulates both major processes of self-organization during development: cell sorting out and spatially restricted lineage commitment

Definition:
A collection of organ-specific cell types that develops from stem cells or organ progenitors and self-organizes through cell sorting and spatially restricted lineage commitment in a manner similar to *in vivo*.



[1]IMBA—Institute of Molecular Biotechnology of the Austrian Academy of Science Vienna 1030, Austria.
*Corresponding author. E-mail: juergen.knoblich@imba.oeaw.ac.at

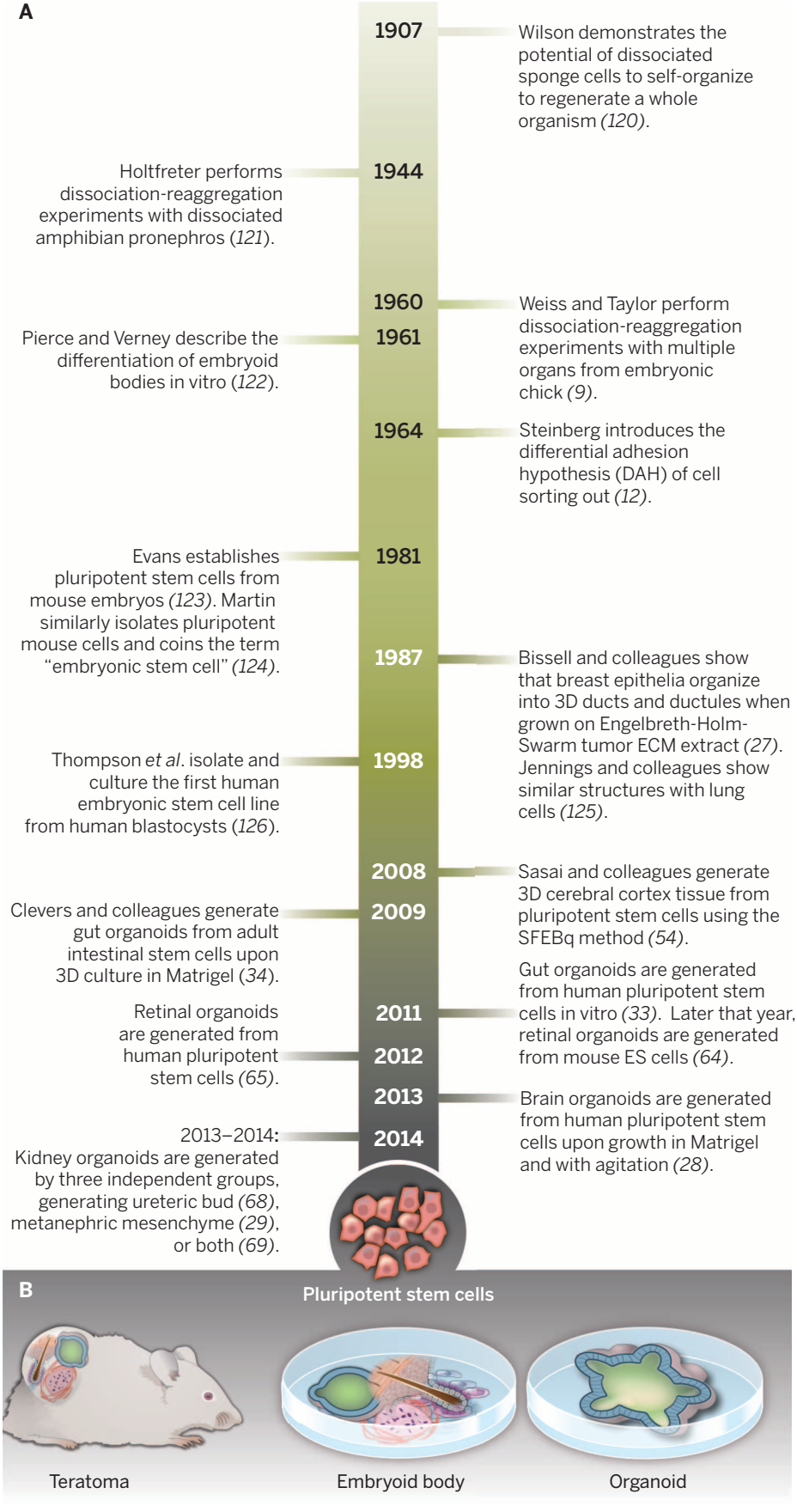


Fig. 1. History of organoid methodologies. (**A**) Key events in the history leading up to various organoid methodologies. (**B**) Comparison of paradigms of self-organization from pluripotent stem cells. Teratomas develop various tissues in vivo, either as spontaneous tumors that can arise in animals and humans or from injected PSCs in a rodent host. Embryoid bodies are 3D aggregates of stem cells that self-organize to develop tissues, similar to teratomas in many ways, but formed in vitro. Organoids are similarly 3D in vitro–derived tissues but are driven using specific conditions to generate individual, isolated tissues.

This stratification depends upon proper stem cell division orientation, the interplay of symmetric and asymmetric divisions, and migration of differentiated daughter cells to defined locations within the tissue (*16*, *17*). Remarkably, this organization can also be recapitulated upon in vitro dissociation and reaggregation (*18*) but only when retinal precursor cells are taken from a chick younger than embryonic day 6 (E6) (*19*, *20*). This suggests that retinal layering depends not only on cell sorting out but also proper execution of lineage decisions by retinal progenitors.

The combination of both sorting out and fate specification in governing self-organization is particularly evident in tumors called teratomas. Teratomas develop from pluripotent stem cells (PSCs) of the germ line and therefore display a variety of organized tissues (Fig. 1B). These include epidermis, nervous tissue, gut, and bone, as well as eyes (*21*) and limbs (*22*). The spontaneous development of these tissues from PSCs is presumably because of a recapitulation of both cell segregation and fate specification. Similarly, these two processes allow for the self-organization seen in organoids (Box 1).

In many ways, organoids represent the methodological evolution of an in vitro system called an embryoid body (EB) that is similar to an early teratoma (Fig. 1B). EBs are 3D aggregates of PSCs (*23*) that undergo initial developmental specification in much the same manner as the pregastrulating embryo (*24*). EBs can further differentiate to form various organized tissues, much like a teratoma. However, their growth in vitro allows for the application of patterning factors to drive particular identities. Not all organoid methods make use of an initial EB stage, but they all involve exogenous tissue patterning and eventually reaggregation to form a 3D self-organized tissue, an organoid.

Self-organization is possible in organoids because of a growing movement away from two-dimensional culture. This movement was triggered by the discovery that epithelial cells, such as kidney (*25*) or breast epithelia (*26*), could develop tubules and ducts when embedded in extracellular matrix hydrogels. Similarly, organoids often make use of such gels, particularly the laminin-rich extracellular matrix secreted by the Engelbreth-Holm-Swarm tumor line (*27*), also called Matrigel.

The resulting self-organizing structures exhibit typical tissue architecture, but note that they are highly heterogeneous. Thus, each organoid

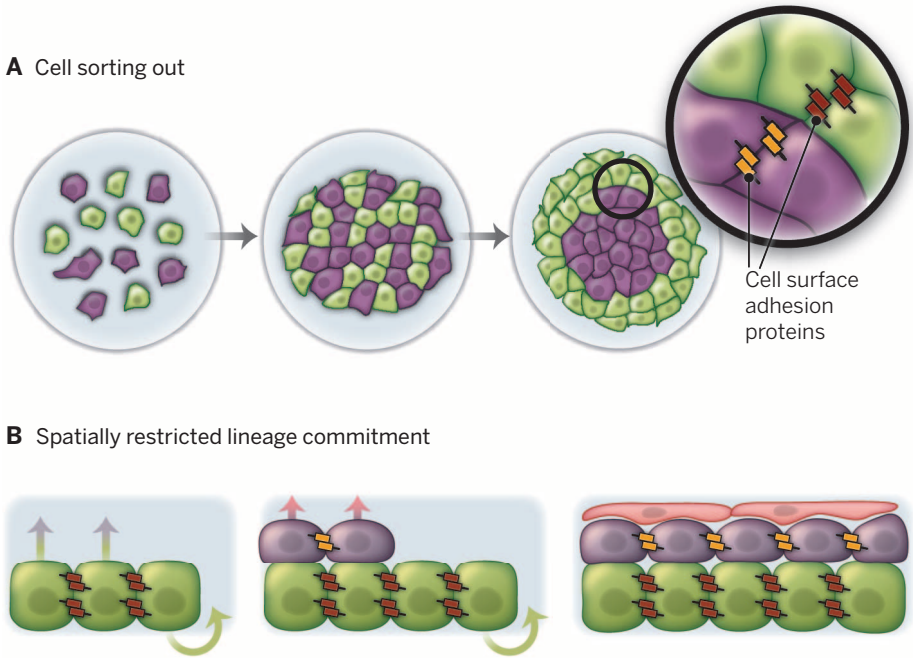


Fig. 2. Principles of self-organization. (**A**) Cell sorting out describes the movement of cells into different domains. Different cell types (purple or green) sort themselves because of different adhesive properties conferred by their differential expression of distinct cell adhesion molecules (shown as brown or orange bars). (**B**) Spatially restricted cell-fate decisions also contribute to self-organization in vivo and in organoids. Progenitors (green) give rise to more differentiated progeny (purple), which, because of spatial constraints of the tissue and/or division orientation, are forced into a more superficial position that promotes their differentiation. These cells can sometimes further divide to give rise to more differentiated progeny (pink), which are further displaced.

is unique and exhibits relative positioning of tissue regions that are often random, possibly because of a lack of embryonic axis formation. For example, brain organoids display various brain regions that individually develop quite similarly to those in vivo (*28*) but are not reliably organized relative to one another because of a lack of anterior-posterior and dorsal-ventral axes. Similarly, kidney organoids develop tubules corresponding to regions of the nephron segment (*29*), but they are randomly positioned rather than displaying medullary and cortical organization. This heterogeneity makes it difficult to generate pure populations of single cell types, but it can be a powerful tool for modeling development and disease on a whole-organ scale.

Current organoid technologies

Organoids derived from human PSCs have so far been established for gut, kidney, brain, and retina, among others (Fig. 3). Many of the organs studied had already been demonstrated to self-organize in reaggregation experiments from embryonic tissues (Table 1), which suggests that organoids could, in principle, be generated if organ progenitors could be derived from PSCs. Below, we will describe the evolution of each of these organoid approaches from this developmental foundation.

Gut organoids

The gastrointestinal (GI) tract develops primarily from the endoderm (*6*), which forms an epithelial tube that develops into three distinct portions, the foregut, midgut, and hindgut (*30*). The foregut gives rise to the oral cavity, the pharynx, the respiratory tract, the stomach, the pancreas, and the liver. The midgut gives rise to the small intestine and the ascending colon. The hindgut gives rise to the remaining portion of the colon, or large intestine, and the rectum. The separation of these three domains involves the combinatorial response to growth factors that have anteriorizing or posteriorizing effects. In particular, Wnt and fibroblast growth factor (FGF) signaling have been shown to inhibit anterior gut fate and instead promote posterior fate, which can lead to midgut and hindgut identities (*31*, *32*).

This knowledge of the posteriorizing effects of Wnts and FGFs provides the foundation on which human intestinal organoids are built (*33*) (Fig. 3). Human PSCs can be driven toward a hindgut identity by initially applying activin A, a nodal-related molecule, to drive mesendodermal identity. The subsequent addition of posteriorizing Wnt3a and Fgf4 then specifies the hindgut, the precursor to the intestine. Originally, this specification was performed in 2D, but surprisingly, the cells spontaneously formed hindgut tubes that budded off to form spheroids. This illustrates the remarkable self-organizing ability of these progenitors, a property that allows them to generate complete 3D organoids when grown in a permissive environment. The laboratory of Hans Clevers had previously shown that adult intestinal stem cells could form organoids when cultured in 3D in Matrigel (*34*). These adult-derived organoids self-organized to form 3D crypt-villus structures that mimicked the physiology and organization of the intestine and could even be transplanted into mice (*35*). Similarly, hindgut spheroids generated from human PSCs can be grown in Matrigel 3D growth conditions to further develop to mature intestinal organoids (*33*).

Intestinal organoids develop crypt-villus structures with stratified epithelium consisting of all the major cell lineages of the gut (*33*, *34*). These include columnar epithelial enterocytes with a brush border of apical microvilli. Furthermore, cell divisions occur at the base of the villus-like protrusions, and intestinal stem cells could be identified by their expression of Lgr5 in more advanced organoids. Finally, these organoids displayed intestinal functions including absorptive and secretory activity.

Although the intestine is the only gut region so far generated from PSCs, other regions of the digestive tract have been developed into organoids from adult stem cells. In particular, gastric organoids have been generated from adult pyloric stem cells (*36*) or chief cells of the stomach (*37*). Lingual organoids have been established from adult tongue epithelium (*38*). These approaches similarly use the 3D Matrigel environment, which suggests that Matrigel is a general requirement in GI tract organoid formation. Furthermore, the use of adult progenitor populations in this manner provides an additional, often more direct, route to the generation of organoids.

Liver organoids

The liver is primarily derived from endoderm, developing from an outgrowth of ventral foregut epithelium that develops into a hepatic bud structure (*39*). This hepatic bud produces the hepatoblasts that generate both hepatocytes and biliary epithelium, whereas adjacent mesoderm-derived mesenchyme contributes liver fibroblasts and stellate cells. The growth of the liver bud involves extensive vascularization and eventually it develops into the major fetal site of hematopoiesis. Thus, liver development represents a complex interplay of both endoderm- and mesoderm-derived tissues.

Early reaggregation studies had shown that dissociated chick embryonic hepatic tissue can reaggregate and organize into secretory units typical of the liver and consistent with formation of functional bile ducts (*9*). More recently, a progenitor population in adult mouse liver that is activated after injury was identified that could generate 3D liver organoids when grown in Matrigel (*40*). These adult-derived liver organoids display cells with biliary ductal identities and can be differentiated to form mature, functional hepatocytes. Finally, liver organoids can be transplanted into mice and were shown to partially rescue mortality in a mouse model of liver disease, pointing to their functionality.

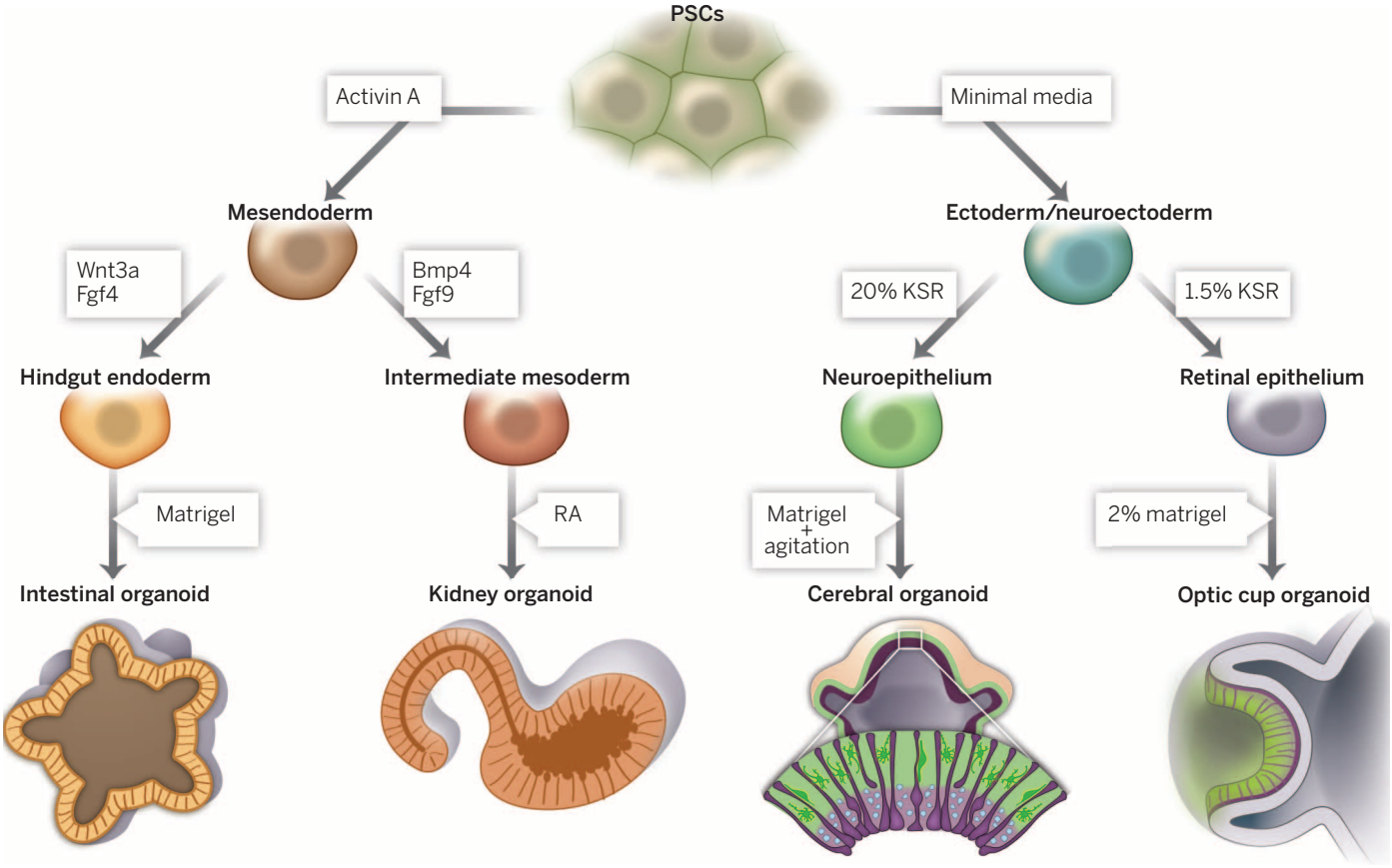


Fig. 3. Overview of organoid methodologies. Organoid differentiation strategies developed so far from human PSCs. Conditions and growth factors are indicated for the derivation of progenitor identities. For neuroectoderm, minimal medium without serum is used. KSR is knockout serum replacement, a serum-free growth-promoting alternative. Limiting its use, along with a low concentration of Matrigel dissolved in the medium, promotes retinal neuroepithelium, whereas higher KSR and embedding in pure Matrigel promotes the formation of various brain regions. Renal organoids have been generated several ways, but growth factors in common are shown.

Although similar human liver organoids have not yet been generated, a very different approach was recently established to generate tissues reminiscent of human liver buds (*41*). Beginning with differentiation of human PSCs into hepatic endodermal cells in 2D, this method mixes three cell populations: the human PSC–derived hepatic cells, human mesenchymal stem cells, and human endothelial cells. This mixed-cell population mimics the early cell lineages of the developing liver. When mixed to a high density on a layer of Matrigel, the cells spontaneously form a 3D aggregate. The liver bud–like aggregates display vascularization and can be ectopically transplanted into mice to allow blood supply. Perhaps most promising is the finding that mice with transplants of these liver bud tissues exhibit human-specific metabolites in the blood. Furthermore, survival of mice subjected to liver injury increased when liver buds were transplanted into them.

Brain organoids

The vertebrate central nervous system derives from the neural ectoderm (*6*). This tissue gives rise to the neural plate, which folds and fuses to form the neural tube, an epithelium with apical-basal polarity radially organized around a fluid-filled lumen that eventually forms the brain ventricles. Axes are established through the concerted action of morphogen gradients, such as the ventral-dorsal Shh-Wnt/Bmp axis, and the rostral-caudal axis influenced by factors such as retinoic acid and FGF (*42*). These axes allow the epithelial tube to subdivide into four major regions, the forebrain, midbrain, hindbrain, and spinal cord. The forebrain gives rise to the majority of the human brain, including the neocortex, hippocampus, and ventral telencephalic structures, such as amygdala and hypothalamus. The midbrain gives rise to the tectum, whereas the hindbrain gives rise to the cerebellum, pons, medulla, and brainstem.

Generally, neurons are generated from neural stem cells that reside next to ventricles (*43*). Neural stem cells initially expand through symmetric proliferative divisions. During neurogenesis, stem cells switch to asymmetric divisions to give rise to self-renewing progenitors and more differentiated cell types, including neurons and intermediate progenitors (*44*). These more differentiated cells migrate outward to generate stratified structures such as the three layers of the medulla, the seven layers of the optic tectum, and the six layers of the cerebral cortex.

Although the final product of neural development is a highly complex interconnected brain, earlier reaggregation studies suggest that this organ has an intrinsic self-organizing capacity (*45*). When taken at early stages of brain development, chick neural progenitors self-organized to form clusters of neuroepithelial cells organized in a radial manner surrounding a lumen, reminiscent of the neural tube. The implication of these classic experiments is that if neuroepithelium can be derived from PSCs, spontaneous self-organization is likely to occur.

Numerous previous studies have made use of in vitro–derived neural stem cells (NSCs) from PSCs to study neural differentiation (*46*). However, these homogeneous NSCs lack the characteristic apical-basal polarity and do not recapitulate the complex lineage of NSCs in vivo. As an alternative approach, neurospheres (*47*) are aggregates of NSCs that can be used to assess their self-renewing capacity. However, neurospheres are likewise not well organized and, therefore, are limited in their capacity to model many aspects of brain development.

More recently, 2D neural tube–like structures called neural rosettes were established from isolated neuroepithelium or the directed differentiation

Table 1. Current state of the art for in vitro self-organizing tissues of various organs. mESCs, mouse embryonic stem cells.

	Organ	Early reaggregation experiments	Identity derivation from PSCs	3D self-organizing structure or organoid
Endoderm	Thyroid	Embryonic chick thyroid (*90*), adult rat thyroid (*91*)	Thyroid progenitors from mESCs (*92*)	Functional thyroid organoid from mESCs (*70*)
	Lung	Embryonic chick lung (*93*)	Lung progenitors from mESCs and hiPSCs (*92*, *94*)	Bronchioalveolar structures from mouse adult lung stem cells (*71*)
	Pancreas	Mouse embryonic pancreas (*95*)	Pancreatic endocrine cells from mESC (*96*) and hESCs (*97*)	Pancreatic organoids from mouse embryonic pancreatic progenitors (*72*)
	Liver	Chick embryonic liver (*9*)	Hepatocytes from mESCs and hESCs (*98*)	Liver organoids from adult stem cells (*40*); liver buds from human iPSCs (*41*)
	Stomach	Chick embryonic gizzard and proventriculus (*99*)	None	Stomach organoids from adult stem cells (*36*, *37*)
	Intestine	Rat embryonic intestine (*100*)	Intestinal cells from mESCs (*101*) and hPSCs (*33*)	Intestinal organoids from human PSCs (*33*)
Mesoderm	Heart	Chick (*102*) and rat (*103*) cardiac tissue	Spontaneous and directed differentiation of mESCs and hESCs (*104*)	Vascularized cardiac patch from hESCs (*105*)
	Skeletal muscle	Embryonic chick leg skeletal muscle (*76*)	Mesoangioblasts from human iPSCs (*106*)	Anchored contracting skeletal muscle in 3D matrix derived from myoblast progenitors (*107*)
	Bone	Skeletal bone of chick embryonic leg (*77*)	Osteoblasts from mESCs (*108*) and hESCs (*109*)	Bone spheroids from human osteogenic cells (*110*)
	Kidney	Chick embryonic kidney (*9*)	Intermediate mesoderm from mouse (*111*) and human (*112*) PSCs	Ureteric bud (*68*) and metanephric mesenchyme (*29*) renal organoids (*69*) from human and mouse PSCs
Ectoderm	Retina	Embryonic chick retina (*61*)	Retinal progenitors from mouse (*113*) and human PSCs (*114*)	Optic cup organoids from mouse (*64*) and human (*65*) PSCs
	Brain	Embryonic chick brain cells (*45*)	Neural rosettes from mouse and human PSCs (*48*, *49*)	Cerebral organoids from mouse and human PSCs (*28*, *55*)
	Pituitary	Chick anterior pituitary (*115*)	None	Adenohypophysis organoids from mouse PSCs (*73*)
	Mammary gland	Mammary gland from adult virgin mice (*75*)	None	3D breast epithelia embedded in Matrigel (*116*)
	Inner ear	Embryonic chick otocysts (*117*)	Inner ear hair cells from mESCs (*118*)	Inner ear organoids from mESCs (*74*)
	Skin	Embryonic chick skin and feather follicles (*9*)	Keratinocytes from mESCs (*119*)	Stratified epidermis from keratinocytes derived from mESCs (*119*)

of PSCs (*48*, *49*). Because neural rosettes recapitulate apical-basal polarity and exhibit spontaneous radial organization similar to that of the neural tube, they are more capable of recapitulating many aspects of brain development. These include the production of intermediate progenitor types, as well as the timed production of layer identities similar to those in vivo (*50*). However, because of the 2D nature of the method, it has many limitations in modeling the overall organization of the developing brain.

Therefore, alternative 3D culture methods with the potential to recapitulate brain tissue organization have been used extensively for investigations in the past several years. In particular, work from the lab of Yoshiki Sasai has focused on developing various isolated brain regions in 3D from mouse or human PSCs (*51*). Beginning with EB formation, particular brain region identities can be generated from neuroectoderm. Specifically, forebrain tissues are generated by plating mouse (*52*) or human (*53*) EBs in 2D and examining adherent cells. However, aggregates develop more complex structures when allowed to continue growing in 3D (*54*), eventually generating dorsal forebrain. This method has further been improved in a recent study (*55*) that also showed neuronal layering reminiscent of early cerebral cortical development.

Other regions can also be generated by mimicking endogenous patterning with growth factors. For example, Hedgehog signaling drives ventral forebrain tissue (*56*). In addition, cerebellar identities can be generated by treatment with either Bmp4 and Wnt3a to generate granule neurons (*57*) or Hedgehog inhibition to generate Purkinje neurons of the cerebellum (*58*). Conversely, minimizing exogenous bioactive factors, such as serum proteins, promotes hypothalamic identity (*59*). Thus, by stimulating neuroectoderm through an EB stage followed by the application of specific growth factors, organoids can be generated for a variety of individual brain regions.

More recently, heterogeneous neural organoids were established, termed cerebral organoids, that contain several different brain regions within individual organoids (*28*) (Fig. 3). The approach similarly begins with EBs, but growth factors are not added to drive particular brain region identities. Instead, the method is influenced by the intestinal organoid protocol, namely, by embedding the tissues in Matrigel. The extracellular matrix provided by the Matrigel promotes outgrowth of large buds of neuroepithelium, which then expand and develop into various brain regions. Cerebral organoids can reach sizes of up to a few millimeters when grown in a spinning bioreactor, which improves nutrient and oxygen exchange. This expansion allows the formation of a variety of brain regions, including retina, dorsal cortex, ventral forebrain,

midbrain-hindbrain boundary, choroid plexus, and hippocampus.

Retinal organoids

The retina is the light-receptive neural region of the eye and is derived from the neural ectoderm. Initially, the optic vesicle forms as an outgrowth of the diencephalon, the most posterior region of the forebrain (*60*). Like the rest of the CNS, this optic vesicle begins as a pseudostratified neuroepithelium with a fluid-filled lumen. However, concerted movements at two hinge regions force the vesicle to fold in on itself, forming the optic cup. Thus, early in retinal development, two adjacent epithelial layers are established: the outer retinal pigmented epithelium and the inner neural retina. This trend of stratification continues and eventually leads to a fully laminated tissue containing layers of photoreceptors and supportive cell types, such as bipolar cells and amacrine cells.

The retina has a long history of in vitro reaggregation studies and has been used as a model of retinal layer formation for decades (*18*). The first reaggregates were generated from chick retina in the early 1960s and demonstrated the robustness of retinal self-organization in vitro (*61*, *62*). Subsequent studies have used retinal reaggregates to examine the relations between different cell types and their differentiation and organization (*63*).

The evolution to PSC-derived retinal organoids, like other organoid approaches, is built upon a foundation rooted in developmental biology (Fig. 3). EBs are derived in minimal medium to generate neuroectoderm (*64*). A nominal amount of Matrigel is dissolved in the medium at an early stage to allow the formation of more rigid neuroepithelial tissues, a prerequisite of retinal pigmented epithelium formation. This promotes the formation of buds of retinal primordial tissue similar to the optic vesicle. These buds are then cut away from the rest of the neuroepithelial tissues and maintained in a medium that supports retinal tissue identity.

The resulting optic cup organoids very closely mimic early retina. They display proper markers of neural retina and retinal pigmented epithelium, they display retinal stratification with proper apical-basal polarity, and they undergo morphological tissue shape changes that mimic the stepwise evagination and invagination of the optic cup in vivo.

More recently, optic cup organoids were generated from human PSCs (*65*). These human retinal organoids show many of the characteristics common to mouse retina; however, they display a number of human specific features as well. In particular, the human retinal organoids are larger than mouse organoids, they require more time to develop, and they display certain tissue morphological differences, such as apical nuclear positioning.

Kidney organoids

The kidney arises from an early embryonic tissue called the intermediate mesoderm, a subdivision of mesodermal identity that develops from the primitive streak (*66*). In vivo, the primitive streak displays opposing gradients of Bmp4 and activin A, which combinatorially specify the endoderm or mesoderm. The intermediate mesoderm is further subspecified through the action of Fgf and Wnt signaling. This tissue then develops into two closely interacting domains, the ureteric bud and the metanephric mesenchyme, which promote each other's growth and branching to develop early renal tubules.

Like many of the tissues for which organoids have so far been developed, evidence that kidney tissue might be capable of self-organization comes from early reaggregation experiments of chick embryonic kidney (*9*). The resultant tissues displayed various segments of the nephron, including collecting duct, distal and proximal tubules, and glomeruli formed by the interaction with allantoic vessels upon transplantation on the chick allantoic membrane. Furthermore, the tissues could develop the stereotypic organization of the kidney with cortical and medullary region. These experiments suggest that, if kidney progenitors can be made from PSCs, these would, in principle, be capable of forming organized tissues if grown in a permissive environment. This principle is what has now been demonstrated by three independent studies (*67*) (Fig. 3).

Each of the recently published methods uses various combinations of growth factors to mimic endogenous signaling to drive renal differentiation. Specifically, ureteric bud identity can be generated by exposing human induced PSCs in 2D to Bmp4 and Fgf2 to drive mesodermal identities (*68*), followed by subsequent application of retinoic acid, Bmp2, and activin A. Such ureteric bud cells can be cocultured with dissociated mouse embryonic kidney to self-organize within the mouse aggregate and populate 3D ureteric bud structures.

The second major renal precursor tissue, the metanephric mesenchyme, can instead be generated beginning with an initial EB stage from mouse and human PSCs (*29*). Sequential application of activin A followed by Bmp4 and the Wnt agonist CHIR99021 then induces posterior mesoderm, the precursor to the intermediate mesoderm. Finally, application of retinoic acid followed by Fgf9 then stimulates the tissues to take on a metanephric mesenchyme identity. By coculturing with spinal cord tissue, a known nephric inducer, this tissue can produce well-organized nephric tubules and even nascent glomeruli.

Finally, both principal lineages of the kidney can be generated together (*69*) by applying activin A and Bmp4 to human embryonic stem cells (hESCs) grown in 2D to generate primitive streak identity. These cells transition to an intermediate mesoderm identity upon exposure to Fgf9 and spontaneously develop further into ureteric bud and metanephric mesenchyme in the absence of further growth factors. Although these specification events were initially performed in 2D, the cells take on 3D morphologies by either growing at low density to allow dome-like colonies to form or when cocultured with mouse kidney reaggregates. In both cases, more complex tissues arise in 3D with structures resembling ureteric epithelium and proximal tubules.

Organoids from model organisms

Despite the relatively few human-derived organoids so far described, several others have already been established from mouse PSCs or adult tissue stem cells. These include the endoderm-derived thyroid, lung, and pancreas. Thyroid organoids can be produced by overexpression of two factors important for thyroid specification, Nkx2.1 and Pax8, followed by treatment of EBs with thyroid-stimulating hormone (*70*). Lung organoids can develop from cocultured adult bronchioalveolar stem cells and lung endothelial cells (*71*) in Matrigel. And finally, pancreatic organoids can develop from simply plating pancreatic progenitor cells in Matrigel (*72*). All three systems give rise to self-organized characteristic epithelia and, in the case of thyroid organoids, even synthesis and secretion of functional thyroid hormone.

Organoids have also been derived from ectoderm-derived pituitary and inner ear. Specifically, the two identities of the developing pituitary, the neural ectodermal infundibulum and the adenohypophysis, could be generated in large EBs grown under ectoderm-promoting conditions (*73*). Remarkably, these pituitary organoids can mature and synthesize pituitary hormones, such as growth hormone, follicle-stimulating hormone, and thyroid-stimulating hormone. Additionally, sensory epithelia of the inner ear can be generated from EBs grown under ectoderm-promoting conditions with subsequent treatment with Bmp4 followed by Fgf2 to drive otic identity (*74*). The resulting otic vesicles generate functional inner ear sensory epithelia with stereocilia and kinocilia.

The future of organoid technologies

The generation of 3D organoids from human PSCs is currently in its infancy, but the field is rapidly evolving. In the near future, human organoids may be generated for those organs that have already been established in mouse or where a principle of self-organization has already been demonstrated in reaggregation studies. These include skin (*9*), mammary gland (*75*), muscle (*76*), and bone (*77*), to name a few.

Paradigms of organ development

Because organoids represent an easily accessible model system, they have the potential to open doors to developmental questions that have been difficult or impossible to answer using traditional techniques. This is particularly true for biological principles that are specific to humans. For example, human brain organoids have already been used to examine the unique division mode of human neural stem cells (*28*). Similarly, retinal organoids have been used to test differences between human and rodent tissue morphogenesis and timing (*65*). Additionally, organoids for the GI tract can be applied to the study of coordinated development of GI organs, a process

that exhibits important differences in humans compared with laboratory animals (78).

Furthermore, organoids hold the potential to model adult homeostasis as well. Indeed, intestinal organoids have already been used to examine the role of the crypt niche in stem cell self-renewal and differentiation (79, 80). This is particularly true for organoids that have been derived from adult progenitors, such as liver and stomach, that closely recapitulate regenerative events seen in the adult organ.

Therapeutic potential

Disease modeling will likely be a primary focus of future organoid studies (Fig. 4). These can range from developmental disorders, cancer, infectious disease, and degeneration. For example, gut organoids are already being used to examine infectious diseases (81, 82); tumor biology (83, 84); and genetic conditions (85, 86).

Along these lines, patient iPSCs will be a valuable tool in future disease modeling. Recently, kidney organoids were generated from iPSCs derived from a patient with polycystic kidney disease (68). Although the method did not test for a phenotype, this will likely represent an important tool in studying this and other genetic kidney disorders. Similarly, retinal organoids have the potential to model human genetic disorders that lead to blindness, such as retinitis pigmentosa. These types of disorders can be modeled by making use of patient iPSCs or, alternatively, through the introduction of patient mutations into human PSCs using modern genome-editing technologies.

Brain organoids, in particular, have huge potential in this respect. They could, in principle, be used to model various neurodevelopmental disorders that have been difficult or impossible to model in animals. Indeed, brain organoids were the first organoids to make use of patient iPSCs in this manner and to model the developmental disorder microcephaly (28). In the future, cerebral organoids may even have the potential to model disorders such as autism, schizophrenia, or epilepsy, and perhaps even adult-onset disorders like neurodegenerative diseases.

Organoids also have the potential to be used for testing efficacy and toxicity of drug compounds (Fig. 4). This could be applied to organoids that model degenerative conditions—for example, liver fibrosis or cystic kidney diseases—where one could screen for effective treatments. If successful, this approach could even cut down on the use of animal testing, reserving it for studies requiring whole-organism readouts. For this, development of human liver organoids would be of particular relevance (Fig. 4), because the human liver often metabolizes drugs in a manner distinct from animals' metabolism. Drugs can be removed at early stages of screening when they could otherwise be functional in humans, or more drastically, toxic metabolites can be produced specifically in humans but not in tested animals. Methods to screen compounds in an in vitro human liver model are therefore being investigated as an alternative in the drug discovery process (87). Human liver buds have already been shown to produce human-specific metabolites (41), which suggests that liver organoids could represent an ideal system to perform these types of studies.

Finally, organoids have the potential to provide alternative approaches to cell or even whole-organ replacement strategies in the clinic (Fig. 4). Organoids could provide a source of autologous tissue for transplantation. In this respect, renal organoids hold enormous therapeutic potential as this is the organ with the highest rate of end-stage failure leading to the highest organ demand for transplants. Already, Taguchi *et al.* succeeded in transplanting kidney organoids under the renal capsule of adult mice, which led to vascularization, a promising step toward a replacement strategy (29).

Additionally, retinal organoids could be used to treat certain types of retinal degeneration and blindness. Indeed, stem cell–therapy clinical trials are already under way to replace certain degenerating retinal cell types (88). Retinal organoids could provide an alternative approach that may better recapitulate development and, therefore, produce particular cell types of interest for transplantation. Finally, intestinal organoids could provide a treatment option for replacement of damaged colon after injury or following removal of diseased tissue. Remarkably, intestinal organoid transplantation has already been performed in mice and can contribute to colon repair after injury (35, 89). Organoid approaches could even allow for gene correction in the case of genetic defects, using modern genome-editing technologies to replace damaged organ with repaired tissue.

Although it is clear that there are many potential uses for organoids, it is important to keep in mind their current limitations. In particular, all of the organoid systems so far established remain to be thoroughly characterized with regard to the extent of recapitulation of in vivo development. For example, although retinal organoids nicely display typical laminar organization, outer segments fail to form; for example, photoreceptors fail to fully mature to become light-sensitive. Likewise, cerebral organoids recapitulate fairly early events in brain development, but later features, such as cortical plate layers, fail to fully form.

The issue of maturation seems to be a common hurdle in organoid technologies, and it remains to be seen whether this will significantly affect their therapeutic and research potential.

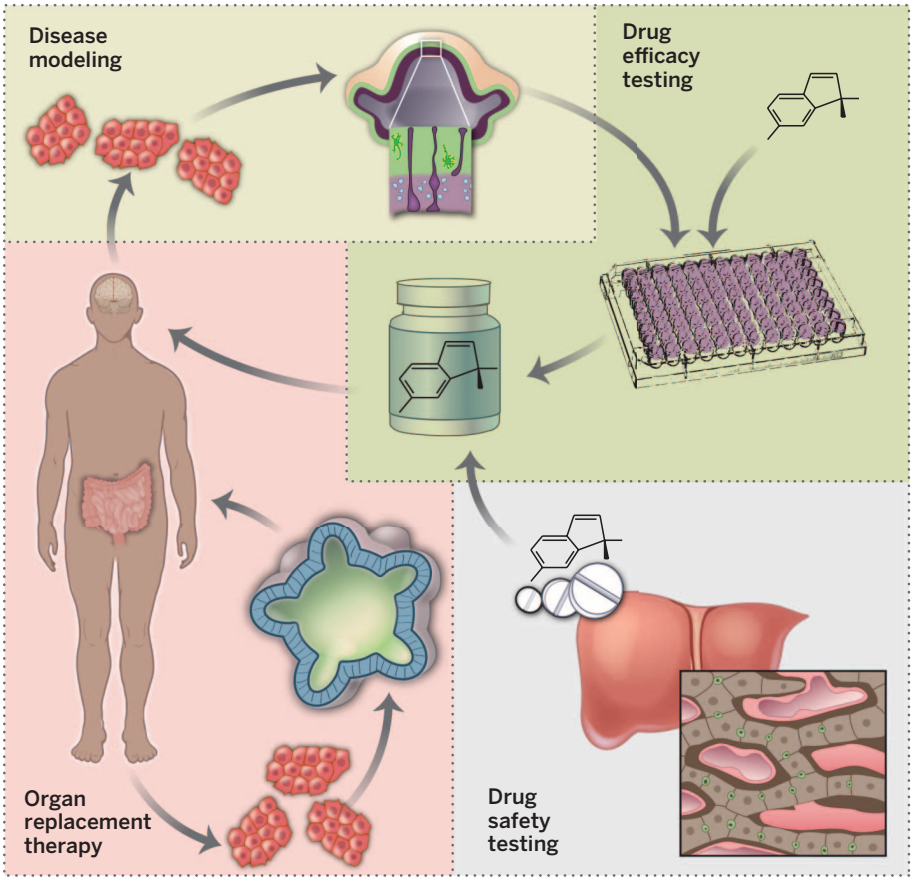


Fig. 4. Therapeutic potential of organoids. Organoids can be used to model diseases (beige box), for example modeling neurodevelopmental disorders with cerebral organoids. These types of disease models can then be used for testing drug efficacy in vitro before moving to animal models (green box). Drug compounds can be tested for toxicity and metabolic profile in liver organoids (gray box). And finally, organoids could be made from patient cells to provide autologous transplant solutions (pink box).

Human intestinal organoids have been shown to display characteristics of mature intestine, producing Lgr5$^+$ adult stem cells (*33*). Other organoids could perhaps be coaxed to fully mature once transplanted, either ectopically or for therapeutic purposes. These studies will likely be a primary focus of future organoid research.

Finally, the lack of vascularization is generally an issue with organoids in vitro. Because of limitations in nutrient supply, organoids have a limited growth potential, which can also affect their maturation. Vascularization is an issue in tissue engineering as a whole, and various approaches have been taken to address it. In the case of organoids, spinning bioreactors can provide better nutrient exchange allowing for sizes of up to a few millimeters (*28*). Alternatively, coculture with endothelial cells can generate vascular-like networks (*41*). Perhaps the most promising solution, however, is the transplantation of these tissues, as has been done for liver buds and kidney organoids, which stimulates invasion from host vasculature (*29*).

Overall, organoids have enormous potential to model development and disease, as a tool for drug testing, and as a therapeutic approach. Future efforts will no doubt bring them closer to reaching that potential.

REFERENCES AND NOTES

1. R. Lanza, J. Gearhart, B. Hogan, D. Melton, R. Pedersen, Eds., *Essentials of Stem Cell Biology* (Elsevier, San Diego, CA, ed. 2, 2009).
2. Y. Hachitanda, M. Tsuneyoshi, Neuroblastoma with a distinct organoid pattern: A clinicopathologic, immunohistochemical, and ultrastructural study. *Hum. Pathol.* **25**, 67–72 (1994). doi: 10.1016/0046-8177(94)90173-2
3. B. Zimmermann, Lung organoid culture. *Differentiation* **36**, 86–109 (1987). doi: 10.1111/j.1432-0436.1987.tb00183.x; pmid: 3446562
4. C. Xinaris *et al*., In vivo maturation of functional renal organoids formed from embryonic cell suspensions. *J. Am. Soc. Nephrol.* **23**, 1857–1868 (2012). doi: 10.1681/ASN.2012050505; pmid: 23085631
5. P. B. Armstrong, Cell sorting out: The self-assembly of tissues in vitro. *Crit. Rev. Biochem. Mol. Biol.* **24**, 119–149 (1989). doi: 10.3109/10409239909086396; pmid: 2651008
6. S. F. Gilbert, Ed., *Developmental Biology* (Sinauer Associates, Sunderland, MA, ed. 6, 2000).
7. C. Dahmann, A. C. Oates, M. Brand, Boundary formation and maintenance in tissue development. *Nat. Rev. Genet.* **12**, 43–55 (2011). doi: 10.1038/nrg2902; pmid: 21164524
8. C. Kiecker, A. Lumsden, Compartments and their boundaries in vertebrate brain development. *Nat. Rev. Neurosci.* **6**, 553–564 (2005). doi: 10.1038/nrn1702; pmid: 15959467
9. P. Weiss, A. C. Taylor, Reconstitution of complete organs from single-cell suspensions of chick embryos in advanced stages of differentiation. *Proc. Natl. Acad. Sci. U.S.A.* **46**, 1177–1185 (1960). doi: 10.1073/pnas.46.9.1177; pmid: 16590731
10. A. Moscona, Rotation-mediated histogenetic aggregation of dissociated cells. A quantifiable approach to cell interactions in vitro. *Exp. Cell Res.* **22**, 455–475 (1961). doi: 10.1016/0014-4827(61)90122-7; pmid: 13773012
11. L. Wolpert, Ed., *Principles of Development* (Oxford University Press, Oxford, ed. 3, 2007).
12. M. S. Steinberg, in *Cellular Membranes in Development*, M. Locke, Ed. (Academic Press, New York, 1964), pp. 321–366.
13. R. A. Foty, C. M. Pfleger, G. Forgacs, M. S. Steinberg, Surface tensions of embryonic tissues predict their mutual envelopment behavior. *Development* **122**, 1611–1620 (1996).pmid: 8625847
14. T. Fujimori, S. Miyatani, M. Takeichi, Ectopic expression of N-cadherin perturbs histogenesis in *Xenopus* embryos. *Development* **110**, 97–104 (1990).pmid: 2081473
15. R. J. Detrick, D. Dickey, C. R. Kintner, The effects of N-cadherin misexpression on morphogenesis in *Xenopus* embryos. *Neuron* **4**, 493–506 (1990). doi: 10.1016/0896-6273(90)90108-R; pmid: 2322458
16. M. Agathocleous, W. A. Harris, From progenitors to differentiated cells in the vertebrate retina. *Annu. Rev. Cell Dev. Biol.* **25**, 45–69 (2009). doi: 10.1146/annurev.cellbio.042308.113259; pmid: 19575661
17. M. Žigman *et al*., Mammalian inscuteable regulates spindle orientation and cell fate in the developing retina. *Neuron* **48**, 539–545 (2005). doi: 10.1016/j.neuron.2005.09.030; pmid: 16301171
18. P. G. Layer, A. Robitzki, A. Rothermel, E. Willbold, Of layers and spheres: The reaggregate approach in tissue engineering. *Trends Neurosci.* **25**, 131–134 (2002). doi: 10.1016/S0166-2236(00)02036-1; pmid: 11852139
19. G. Vollmer, P. G. Layer, Reaggregation of chick retinal and mixtures of retinal and pigment epithelial cells: The degree of laminar organization is dependent on age. *Neurosci. Lett.* **63**, 91–95 (1986). doi: 10.1016/0304-3940(86)90019-4; pmid: 2419803
20. A. Rothermel, E. Willbold, W. J. Degrip, P. G. Layer, Pigmented epithelium induces complete retinal reconstitution from dispersed embryonic chick retinae in reaggregation culture. *Proc. Biol. Sci.* **264**, 1293–1302 (1997). doi: 10.1098/rspb.1997.0179; pmid: 9332014
21. C. Sergi, V. Ehemann, B. Beedgen, O. Linderkamp, H. F. Otto, Huge fetal sacrococcygeal teratoma with a completely formed eye and intratumoral DNA ploidy heterogeneity. *Pediatr. Dev. Pathol.* **2**, 50–57 (1999). doi: 10.1007/s100249900089; pmid: 9841706
22. N. Kuno *et al*., Mature ovarian cystic teratoma with a highly differentiated homunculus: A case report. *Birth Defects Res. A Clin. Mol. Teratol.* **70**, 40–46 (2004). doi: 10.1002/bdra.10133; pmid: 14745894
23. M. Evans, Discovering pluripotency: 30 years of mouse embryonic stem cells. *Nat. Rev. Mol. Cell Biol.* **12**, 680–686 (2011). doi: 10.1038/nrm3190; pmid: 21941277
24. G. Weitzer, Embryonic stem cell-derived embryoid bodies: An in vitro model of eutherian pregastrulation development and early gastrulation. *Handbook Exp. Pharmacol.* **2006**, 21–51 (2006).
25. R. Montesano, G. Schaller, L. Orci, Induction of epithelial tubular morphogenesis in vitro by fibroblast-derived soluble factors. *Cell* **66**, 697–711 (1991). doi: 10.1016/0092-8674(91)90115-F; pmid: 1878968
26. C. Hagios, A. Lochter, M. J. Bissell, Tissue architecture: The ultimate regulator of epithelial function? *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **353**, 857–870 (1998). doi: 10.1098/rstb.1998.0250; pmid: 9684283
27. M. L. Li *et al*., Influence of a reconstituted basement membrane and its components on casein gene expression and secretion in mouse mammary epithelial cells. *Proc. Natl. Acad. Sci. U.S.A.* **84**, 136–140 (1987). doi: 10.1073/pnas.84.1.136; pmid: 3467345
28. M. A. Lancaster *et al*., Cerebral organoids model human brain development and microcephaly. *Nature* **501**, 373–379 (2013). doi: 10.1038/nature12517; pmid: 23995685
29. A. Taguchi *et al*., Redefining the in vivo origin of metanephric nephron progenitors enables generation of complex kidney structures from pluripotent stem cells. *Cell Stem Cell* **14**, 53–67 (2014). doi: 10.1016/j.stem.2013.11.010; pmid: 24332837
30. J. K. Heath, Transcriptional networks and signaling pathways that govern vertebrate intestinal development. *Curr. Top. Dev. Biol.* **90**, 159–192 (2010). doi: 10.1016/S0070-2153(10)90004-5; pmid: 20691849
31. J. Dessimoz, R. Opoka, J. J. Kordich, A. Grapin-Botton, J. M. Wells, FGF signaling is necessary for establishing gut tube domains along the anterior-posterior axis in vivo. *Mech. Dev.* **123**, 42–55 (2006). doi: 10.1016/j.mod.2005.10.001; pmid: 16326079
32. V. A. McLin, S. A. Rankin, A. M. Zorn, Repression of Wnt/beta-catenin signaling in the anterior endoderm is essential for liver and pancreas development. *Development* **134**, 2207–2217 (2007). doi: 10.1242/dev.001230; pmid: 17507400
33. J. R. Spence *et al*., Directed differentiation of human pluripotent stem cells into intestinal tissue in vitro. *Nature* **470**, 105–109 (2011). doi: 10.1038/nature09691; pmid: 21151107
34. T. Sato *et al*., Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche. *Nature* **459**, 262–265 (2009). doi: 10.1038/nature07935; pmid: 19329995
35. S. Yui *et al*., Functional engraftment of colon epithelium expanded in vitro from a single adult Lgr5$^+$ stem cell. *Nat. Med.* **18**, 618–623 (2012). doi: 10.1038/nm.2695; pmid: 22406745
36. N. Barker *et al*., Lgr5(+ve) stem cells drive self-renewal in the stomach and build long-lived gastric units in vitro. *Cell Stem Cell* **6**, 25–36 (2010). doi: 10.1016/j.stem.2009.11.013; pmid: 20085740
37. D. E. Stange *et al*., Differentiated *Troy*+ chief cells act as reserve stem cells to generate all lineages of the stomach epithelium. *Cell* **155**, 357–368 (2013). doi: 10.1016/j.cell.2013.09.008; pmid: 24120136
38. H. Hisha *et al*., Establishment of a novel lingual organoid culture system: Generation of organoids having mature keratinized epithelium from adult epithelial stem cells. *Sci Rep* **3**, 3224 (2013). doi: 10.1038/srep03224; pmid: 24232854
39. K. S. Zaret, Regulatory phases of early liver development: Paradigms of organogenesis. *Nat. Rev. Genet.* **3**, 499–512 (2002). doi: 10.1038/nrg837; pmid: 12094228
40. M. Huch *et al*., In vitro expansion of single Lgr5+ liver stem cells induced by Wnt-driven regeneration. *Nature* **494**, 247–250 (2013). doi: 10.1038/nature11826; pmid: 23354049
41. T. Takebe *et al*., Vascularized and functional human liver from an iPSC-derived organ bud transplant. *Nature* **499**, 481–484 (2013). doi: 10.1038/nature12271; pmid: 23823721
42. T. J. Petros, J. A. Tyson, S. A. Anderson, Pluripotent stem cells for the study of CNS development. *Front Mol Neurosci* **4**, 30 (2011). doi: 10.3389/fnmol.2011.00030; pmid: 22016722
43. M. K. Lehtinen, C. A. Walsh, Neurogenesis at the brain-cerebrospinal fluid interface. *Annu. Rev. Cell Dev. Biol.* **27**, 653–679 (2011). doi: 10.1146/annurev-cellbio-092910-154026; pmid: 21801012
44. M. A. Lancaster, J. A. Knoblich, Spindle orientation in mammalian cerebral cortical development. *Curr. Opin. Neurobiol.* **22**, 737–746 (2012). doi: 10.1016/j.conb.2012.04.003; pmid: 22554882
45. K. Ishii, Reconstruction of dissociated chick brain cells in rotation-mediated culture. *Cytologia (Tokyo)* **31**, 89–98 (1966). doi: 10.1508/cytologia.31.89; pmid: 5959228
46. L. Conti, E. Cattaneo, Neural stem cell systems: Physiological players or in vitro entities? *Nat. Rev. Neurosci.* **11**, 176–187 (2010). doi: 10.1038/nrn2938; pmid: 20107441
47. B. A. Reynolds, S. Weiss, Generation of neurons and astrocytes from isolated cells of the adult mammalian central nervous system. *Science* **255**, 1707–1710 (1992). doi: 10.1126/science.1553558; pmid: 1553558
48. Y. Elkabetz *et al*., Human ES cell-derived neural rosettes reveal a functionally distinct early neural stem cell stage. *Genes Dev.* **22**, 152–165 (2008). doi: 10.1101/gad.1616208; pmid: 18198334
49. S. C. Zhang, M. Wernig, I. D. Duncan, O. Brüstle, J. A. Thomson, In vitro differentiation of transplantable neural precursors from human embryonic stem cells. *Nat. Biotechnol.* **19**, 1129–1133 (2001). doi: 10.1038/nbt1201-1129; pmid: 11731781
50. N. Gaspard *et al*., An intrinsic mechanism of corticogenesis from embryonic stem cells. *Nature* **455**, 351–357 (2008). doi: 10.1038/nature07287; pmid: 18716623
51. D. Cyranoski, Tissue engineering: The brainmaker. *Nature* **488**, 444–446 (2012).doi: 10.1038/488444a; pmid: 22914148
52. K. Watanabe *et al*., Directed differentiation of telencephalic precursors from embryonic stem cells. *Nat. Neurosci.* **8**, 288–296 (2005). doi: 10.1038/nn1402; pmid: 15696161
53. K. Watanabe *et al*., A ROCK inhibitor permits survival of dissociated human embryonic stem cells. *Nat. Biotechnol.* **25**, 681–686 (2007). doi: 10.1038/nbt1310; pmid: 17529971
54. M. Eiraku *et al*., Self-organized formation of polarized cortical tissues from ESCs and its active manipulation by extrinsic signals. *Cell Stem Cell* **3**, 519–532 (2008). doi: 10.1016/j.stem.2008.09.002; pmid: 18983967
55. T. Kadoshima *et al*., Self-organization of axial polarity, inside-out layer pattern, and species-specific progenitor dynamics in human ES cell-derived neocortex. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 20284–20289 (2013). doi: 10.1073/pnas.1315710110; pmid: 24277810
56. T. Danjo *et al*., Subregional specification of embryonic stem cell-derived ventral telencephalic tissues by timed and combinatory treatment with extrinsic signals. *J. Neurosci.* **31**, 1919–1933 (2011). doi: 10.1523/JNEUROSCI.5128-10.2011; pmid: 21289201
57. H.-L. Su *et al*., Generation of cerebellar neuron precursors from embryonic stem cells. *Dev. Biol.* **290**, 287–296 (2006). doi: 10.1016/j.ydbio.2005.11.010; pmid: 16406324
58. K. Muguruma *et al*., Ontogeny-recapitulating generation and tissue integration of ES cell-derived Purkinje cells. *Nat. Neurosci.* **13**, 1171–1180 (2010). doi: 10.1038/nn.2638; pmid: 20835252
59. T. Wataya *et al*., Minimization of exogenous signals in ES cell culture induces rostral hypothalamic differentiation. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 11796–11801 (2008). doi: 10.1073/pnas.0803078105; pmid: 18697938

60. W. Heavner, L. Pevny, Eye development and retinogenesis. *Cold Spring Harb. Perspect. Biol.* **4**, a008391 (2012). doi: 10.1101/cshperspect.a008391; pmid: 23071378
61. A. Stefanelli, A. M. Zacchei, V. Ceccherini, Retinal reconstitution in vitro after disaggregation of embryonic chicken eyes. *Acta Embryol. Morphol. Exp.* **4**, 47–55 (1961).
62. A. A. Moscona, Analysis of cell recombinations in experimental synthesis of tissues in vitro. *J. Cell. Comp. Physiol.* **60** (suppl. S1), 65–80 (1962). doi: 10.1002/jcp.1030600407
63. S. Nakagawa, S. Takada, R. Takada, M. Takeichi, Identification of the laminar-inducing factor: Wnt-signal from the anterior rim induces correct laminar formation of the neural retina in vitro. *Dev. Biol.* **260**, 414–425 (2003). doi: 10.1016/S0012-1606(03)00320-8; pmid: 12921742
64. M. Eiraku *et al*., Self-organizing optic-cup morphogenesis in three-dimensional culture. *Nature* **472**, 51–56 (2011). doi: 10.1038/nature09941; pmid: 21475194
65. T. Nakano *et al*., Self-formation of optic cups and storable stratified neural retina from human ESCs. *Cell Stem Cell* **10**, 771–785 (2012). doi: 10.1016/j.stem.2012.05.009; pmid: 22704518
66. M. H. Little, A. P. McMahon, Mammalian kidney development: Principles, progress, and projections. *Cold Spring Harb. Perspect. Biol.* **4**, a008300 (2012). doi: 10.1101/cshperspect.a008300; pmid: 22550230
67. B. D. Humphreys, Kidney structures differentiated from stem cells. *Nat. Cell Biol.* **16**, 19–21 (2014). doi: 10.1038/ncb2904; pmid: 24366031
68. Y. Xia *et al*., Directed differentiation of human pluripotent cells to ureteric bud kidney progenitor-like cells. *Nat. Cell Biol.* **15**, 1507–1515 (2013). doi: 10.1038/ncb2872; pmid: 24240476
69. M. Takasato *et al*., Directing human embryonic stem cell differentiation towards a renal lineage generates a self-organizing kidney. *Nat. Cell Biol.* **16**, 118–126 (2014). doi: 10.1038/ncb2894; pmid: 24335651
70. F. Antonica *et al*., Generation of functional thyroid from embryonic stem cells. *Nature* **491**, 66–71 (2012). doi: 10.1038/nature11525; pmid: 23051751
71. J.-H. Lee *et al*., Lung stem cell differentiation in mice directed by endothelial cells via a BMP4-NFATc1-thrombospondin-1 axis. *Cell* **156**, 440–455 (2014). doi: 10.1016/j.cell.2013.12.039; pmid: 24485453
72. C. Greggio *et al*., Artificial three-dimensional niches deconstruct pancreas development in vitro. *Development* **140**, 4452–4462 (2013). doi: 10.1242/dev.096628; pmid: 24130330
73. H. Suga *et al*., Self-formation of functional adenohypophysis in three-dimensional culture. *Nature* **480**, 57–62 (2011). doi: 10.1038/nature10637; pmid: 22080957
74. K. R. Koehler, A. M. Mikosz, A. I. Molosh, D. Patel, E. Hashino, Generation of inner ear sensory epithelia from pluripotent stem cells in 3D culture. *Nature* **500**, 217–221 (2013). doi: 10.1038/nature12298; pmid: 23842490
75. F. Prop, G. J. Wiepjes, in *Tissue Culture: Methods and Applications*, P. F. Kruse Jr. and M. K. Patterson Jr., Eds. (Academic Press, New York, 1973) pp. 21–24.
76. I. R. Konigsberg, Clonal analysis of myogenesis: Its relevance to the general problem of the stability of cell-type in cultured animal cells is discussed. *Science* **140**, 1273–1284 (1963). doi: 10.1126/science.140.3573.1273; pmid: 14034592
77. A. Moscona, H. Moscona, The dissociation and aggregation of cells from organ rudiments of the early chick embryo. *J. Anat.* **86**, 287–301 (1952).pmid: 12980879
78. E. Lévy, E. Delvin, D. Ménard, J.-F. Beaulieu, Functional development of human fetal gastrointestinal tract. *Methods Mol. Biol.* **550**, 205–224 (2009). doi: 10.1007/978-1-60327-009-0_13; pmid: 19495706
79. T. Sato *et al*., Paneth cells constitute the niche for Lgr5 stem cells in intestinal crypts. *Nature* **469**, 415–418 (2011). doi: 10.1038/nature09637; pmid: 21113151
80. A. Durand *et al*., Functional intestinal stem cells after Paneth cell ablation induced by the loss of transcription factor Math1 (Atoh1). *Proc. Natl. Acad. Sci. U.S.A.* **109**, 8965–8970 (2012). doi: 10.1073/pnas.1201652109; pmid: 22586121
81. S. R. Finkbeiner *et al*., Stem cell-derived human intestinal organoids as an infection model for rotaviruses. *MBio* **3**, e00159-12 (2012). doi: 10.1128/mBio.00159-12; pmid: 22761392
82. A. Castellanos-Gonzalez, M. M. Cabada, J. Nichols, G. Gomez, A. C. White Jr., Human primary intestinal epithelial cells as an improved in vitro model for *Cryptosporidium parvum* infection. *Infect. Immun.* **81**, 1996–2001 (2013). doi: 10.1128/IAI.01131-12; pmid: 23509153
83. T. M. Yeung, S. C. Gandhi, J. L. Wilding, R. Muschel, W. F. Bodmer, Cancer stem cells from colorectal cancer-derived cell lines. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 3722–3727 (2010). doi: 10.1073/pnas.0915135107; pmid: 20133591
84. K. Onuma *et al*., Genetic reconstitution of tumorigenesis in primary intestinal cells. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 11127–11132 (2013). doi: 10.1073/pnas.1221926110; pmid: 23776211
85. A. E. Bigorgne *et al*., *TTC7A* mutations disrupt intestinal epithelial apicobasal polarity. *J. Clin. Invest.* **124**, 328–337 (2013). doi: 10.1172/JCI71471; pmid: 24292712
86. J. F. Dekkers *et al*., A functional CFTR assay using primary cystic fibrosis intestinal organoids. *Nat. Med.* **19**, 939–945 (2013). doi: 10.1038/nm.3201; pmid: 23727931
87. C. Luni, E. Serena, N. Elvassore, Human-on-chip for therapy development and fundamental science. *Curr. Opin. Biotechnol.* **25**, 45–50 (2014). doi: 10.1016/j.copbio.2013.08.015; pmid: 24484880
88. C. M. Ramsden *et al*., Stem cells in retinal regeneration: Past, present and future. *Development* **140**, 2576–2585 (2013). doi: 10.1242/dev.092270; pmid: 23715550
89. R. P. Fordham *et al*., Transplantation of expanded fetal intestinal progenitors contributes to colon regeneration after injury. *Cell Stem Cell* **13**, 734–744 (2013). doi: 10.1016/j.stem.2013.09.015; pmid: 24139758
90. S. R. Hilfer, L. B. Iszard, E. K. Hilfer, Follicle formation in the embryonic chick thyroid. II. Reorganization after dissociation. *Z. Zellforsch. Mikrosk. Anat.* **92**, 256–269 (1968). doi: 10.1007/BF00335651; pmid: 4907211
91. J. M. Mallette, A. Anthony, Growth in culture of trypsin dissociated thyroid cells from adult rats. *Exp. Cell Res.* **41**, 642–651 (1966). doi: 10.1016/S0014-4827(66)80115-5; pmid: 4952051
92. T. A. Longmire *et al*., Efficient derivation of purified lung and thyroid progenitors from embryonic stem cells. *Cell Stem Cell* **10**, 398–411 (2012). doi: 10.1016/j.stem.2012.01.019; pmid: 22482505
93. J. W. Grover, The enzymatic dissociation and reproducible reaggregation in vitro of 11-day embryonic chick lung. *Dev. Biol.* **3**, 555–568 (1961). doi: 0.1016/0012-1606(61)90032-X; pmid: 13902340
94. H. Mou *et al*., Generation of multipotent lung and airway progenitors from mouse ESCs and patient-specific cystic fibrosis iPSCs. *Cell Stem Cell* **10**, 385–397 (2012). doi: 10.1016/j.stem.2012.01.018; pmid: 22482504
95. M. R. Bernfield, P. E. Fell, Separation of the proliferating and differentiating cell populations of cultured embryonic pancreatic epithelium. *Proc. Natl. Acad. Sci. U.S.A.* **58**, 2227–2234 (1967). doi: 10.1073/pnas.58.6.2227; pmid: 5242199
96. N. Lumelsky *et al*., Differentiation of embryonic stem cells to insulin-secreting structures similar to pancreatic islets. *Science* **292**, 1389–1394 (2001). doi: 10.1126/science.1058866; pmid: 11326082
97. K. A. D'Amour *et al*., Production of pancreatic hormone-expressing endocrine cells from human embryonic stem cells. *Nat. Biotechnol.* **24**, 1392–1401 (2006). doi: 10.1038/nbt1259; pmid: 17053790
98. H. Shirahashi *et al*., Differentiation of human and mouse embryonic stem cells along a hepatocyte lineage. *Cell Transplant.* **13**, 197–211 (2004). doi: 10.3727/000000004783984016; pmid: 15191158
99. K. Urase, S. Yasugi, Induction and inhibition of epithelial differentiation by the mixed cell aggregates of the mesenchymes from the chicken embryonic digestive tract. *Dev. Growth Differ.* **35**, 33–40 (1993). doi: 10.1111/j.1440-169X.1993.00033.x
100. R. K. Montgomery, H. M. Zinman, B. T. Smith, Organotypic differentiation of trypsin-dissociated fetal rat intestine. *Dev. Biol.* **100**, 181–189 (1983). doi: 10.1016/0012-1606(83)90209-9; pmid: 6617990
101. L. Cao *et al*., Intestinal lineage commitment of embryonic stem cells. *Differentiation* **81**, 1–10 (2011). doi: 10.1016/j.diff.2010.09.182; pmid: 20934799
102. D. A. Fischman, A. A. Moscona, in *Cardiac Hypertrophy*, N. K. Alpert, Ed. (Academic Press, New York, 1971), pp. 125–139.
103. S. P. Halbert, R. Bruderer, T. M. Lin, In vitro organization of dissociated rat cardiac cells into beating three-dimensional structures. *J. Exp. Med.* **133**, 677–695 (1971). doi: 10.1084/jem.133.4.677; pmid: 4926207
104. K. R. Boheler *et al*., Differentiation of pluripotent embryonic stem cells into cardiomyocytes. *Circ. Res.* **91**, 189–201 (2002). doi: 10.1161/01.RES.0000027865.61704.32; pmid: 12169644
105. K. R. Stevens *et al*., Physiological function and transplantation of scaffold-free and vascularized human cardiac muscle tissue. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 16568–16573 (2009). doi: 10.1073/pnas.0908381106; pmid: 19805339
106. F. S. Tedesco *et al*., Transplantation of genetically corrected human iPSC-derived progenitors in mice with limb-girdle muscular dystrophy. *Sci. Transl. Med.* **4**, 140ra89 (2012). doi: 10.1126/scitranslmed.3003541; pmid: 22745439
107. D. W. J. van der Schaft *et al*., Engineering skeletal muscle tissues from murine myoblast progenitor cells and application of electrical stimulation. *J. Vis. Exp.* **2013**, e4267 (2013).pmid: 23542531
108. L. D. Buttery *et al*., Differentiation of osteoblasts and in vitro bone formation from murine embryonic stem cells. *Tissue Eng.* **7**, 89–99 (2001). doi: 10.1089/107632700300003323; pmid: 11224927
109. R. C. Bielby, A. R. Boccaccini, J. M. Polak, L. D. K. Buttery, In vitro differentiation and in vivo mineralization of osteogenic cells derived from human embryonic stem cells. *Tissue Eng.* **10**, 1518–1525 (2004). doi: 10.1089/ten.2004.10.1518; pmid: 15588411
110. S. Kale *et al*., Three-dimensional cellular development is essential for ex vivo formation of human bone. *Nat. Biotechnol.* **18**, 954–958 (2000). doi: 10.1038/79439; pmid: 10973215
111. M. Nishikawa *et al*., Stepwise renal lineage differentiation of mouse embryonic stem cells tracing in vivo development. *Biochem. Biophys. Res. Commun.* **417**, 897–902 (2012). doi: 10.1016/j.bbrc.2011.12.071; pmid: 22209845
112. S. Mae *et al*., Monitoring and robust induction of nephrogenic intermediate mesoderm from human pluripotent stem cells. *Nat. Commun.* **4**, 1367 (2013). doi: 10.1038/ncomms2378; pmid: 23340407
113. H. Ikeda *et al*., Generation of Rx+/Pax6+ neural retinal precursors from embryonic stem cells. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 11331–11336 (2005). doi: 10.1073/pnas.0500010102; pmid: 16076961
114. D. A. Lamba, M. O. Karl, C. B. Ware, T. A. Reh, Efficient generation of retinal progenitor cells from human embryonic stem cells. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 12769–12774 (2006). doi: 10.1073/pnas.0601990103; pmid: 16908856
115. H. Sobel, The behaviour in vitro of dissociated embryonic pituitary tissue. *J. Embryol. Exp. Morphol.* **6**, 518–526 (1958). pmid: 13575665
116. G. Y. Lee, P. A. Kenny, E. H. Lee, M. J. Bissell, Three-dimensional culture models of normal and malignant breast epithelial cells. *Nat. Methods* **4**, 359–365 (2007). doi: 10.1038/nmeth1015; pmid: 17396127
117. M. F. Orr, Histogenesis of sensory epithelium in reaggregates of dissociated embryonic chick otocysts. *Dev. Biol.* **17**, 39–54 (1968). doi: 10.1016/0012-1606(68)90088-2; pmid: 5642021
118. H. Li, G. Roblin, H. Liu, S. Heller, Generation of hair cells by stepwise differentiation of embryonic stem cells. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 13495–13500 (2003). doi: 10.1073/pnas.2334503100; pmid: 14593207
119. C. Coraux *et al*., Reconstituted skin from murine embryonic stem cells. *Curr. Biol.* **13**, 849–853 (2003). doi: 10.1016/S0960-9822(03)00296-3; pmid: 12747834
120. H. V. Wilson, A new method by which sponges may be artificially reared. *Science* **25**, 912–915 (1907). doi: 10.1126/science.25.649.912; pmid: 17842577
121. J. Holtfreter, Experimental studies on the development of the pronephros. *Rev. Can. Biol.* **3**, 220–250 (1944).
122. G. B. Pierce Jr., E. L. Verney, An in vitro and in vivo study of differentiation in teratocarcinomas. *Cancer* **14**, 1017–1029 (1961). doi: 10.1002/1097-0142(196109/10)14:5<1017::AID-CNCR2820140516>3.0.CO;2-P; pmid: 13735537
123. M. Evans, Origin of mouse embryonal carcinoma cells and the possibility of their direct isolation into tissue culture. *J. Reprod. Fertil.* **62**, 625–631 (1981). doi: 10.1530/jrf.0.0620625; pmid: 7019433
124. G. R. Martin, Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc. Natl. Acad. Sci. U.S.A.* **78**, 7634–7638 (1981). doi: 10.1073/pnas.78.12.7634; pmid: 6950406
125. J. M. Shannon, R. J. Mason, S. D. Jennings, Functional differentiation of alveolar type II epithelial cells in vitro: Effects of cell shape, cell-matrix interactions and cell-cell interactions. *Biochim. Biophys. Acta* **931**, 143–156 (1987). doi: 10.1016/0167-4889(87)90200-X; pmid: 3663713
126. J. A. Thomson *et al*., Embryonic stem cell lines derived from human blastocysts. *Science* **282**, 1145–1147 (1998). doi: 10.1126/science.282.5391.1145; pmid: 9804556

ACKNOWLEDGMENTS

We are grateful to N. Corsini and M. Renner for critical feedback as well as other members of the Knoblich lab for discussions. M.A.L. received funding from the European Molecular Biology Organization (EMBO), the Helen Hay Whitney Foundation, and a Marie Curie post-doctoral fellowship. Work in J.A.K.'s laboratory is supported by the Austrian Academy of Sciences, the Austrian Science Fund (FWF) (projects Z153-B09 and I552-B19), and an advanced grant from the European Research Council.

10.1126/science.1247125

SLICING THE wheat genome

By **Kellye Eversole,[1] Catherine Feuillet,[1,2] Klaus F. X. Mayer,[3] Jane Rogers[1]**

This year celebrates the 100th anniversary of the birth of Norman Borlaug, the Nobel Prize–winning plant geneticist who, through his contribution to the "green revolution," reminds us of the importance of applying scientific knowledge to develop crop varieties. This is even more important today as we face a rapidly expanding global population, climate change, and the need to keep agricultural efforts sustainable while minimizing environmental impacts. Accessing the fundamental information of crop genomes aids in accelerating breeding pipelines and improves our understanding of the molecular basis of agronomically important traits, such as yield and tolerance to abiotic and biotic stresses.

Obtaining a reference sequence of the genome of bread wheat (*Triticum aestivum*), the staple food for 30% of the world's population, is a scientific challenge. Wheat's hexaploid genome was formed from multiple hybridization events between three different progenitor species (comprising three individual subgenomes: A, B, and D). This resulted in a large—five times that of humans—and highly redundant genome with more than 80% of the genome consisting of repeated sequences. For these reasons, a reference sequence—a contiguous sequence ordered along the chromosomes—cannot be generated by using whole-genome shotgun sequencing approaches with current high-throughput short read technologies. To overcome this complexity, the International Wheat Genome Sequencing Consortium (IWGSC) developed a strategy of physical mapping and sequencing the individual chromosomes and chromosome arms of the bread wheat genome. In this special issue of *Science*, four Research Articles are presented in full online (www.sciencemag.org/extra/wheatgenome), with abstracts in print on p. 286 and a News story on p. 251. These papers present major advances toward obtaining a reference sequence and enhancing our understanding of the bread wheat genome.

The IWGSC produced a survey of the gene content and composition of all 21 chromosomes and identified 124,201 gene loci, with more than 75,000 positioned along the chromosomes. Comparing the bread wheat gene sequences with gene repertoires from its closest extant relatives (representing the species that donated the A, B, and D progenitor genomes) showed limited gene loss during the evolution of the hexaploid wheat genome but frequent gene duplications after these genomes came together. Gene expression patterns revealed that none of the subgenomes dominated gene expression.

Triticum aestivum L.

PHOTO: SUSANNE STAMP, ERNST MERZ/ETH ZURICH

Choulet *et al.* describe the sequencing, assembly, annotation, and analysis of the reference sequence of the largest wheat chromosome, 3B, which at nearly 1 gigabase is more than seven times larger than the entire sequence of the model plant *Arabidopsis thaliana*. Relying on a physical map derived from the chromosome 3B–specific bacterial artificial chromosome (BAC) library (*1*), more than 8000 BAC clones were sequenced and assembled into a pseudomolecule—a nearly complete representation of the entire chromosome. This high-quality, ordered sequence revealed a partitioning into distinct regions along the chromosome, including distal segments that are preferential targets for recombination, adaptation, and genomic plasticity. Many inter- and intrachromosomal duplications were also observed, illuminating the structural and functional redundancy of the wheat genome. This sequence, which can be anchored to the genetic and phenotypic maps, will aid breeders by increasing the pace and simplifying the process of identifying and cloning genes underlying agronomic traits.

Marcussen *et al.* used the IWGSC chromosome survey sequences to analyze the timing and phylogenetic origin of the diploid genomes that have come together to form the A, B, and D subgenomes of bread wheat. They unravel ancient hybridization events in the wheat lineage and reveal that the ancestral A and B genomes diverged from a common ancestor ~7 million years ago. They also show that the D genome was formed through homoploid hybrid speciation—hybridization that does not result in a genome duplication event—between relatives of the A and B genomes 1 million to 2 million years later.

Pfeifer *et al.* address inter- and intragenomic gene expression regulation within a polyploid genome by providing an in-depth analysis of the transcriptional landscape of the developing wheat grain. They show that the transcriptional network delineates a complex and highly orchestrated interplay of the individual wheat subgenomes and identify transcriptional active or inactive domains along the chromosomes that might indicate epigenetic control of grain development.

Together, these Research Articles explore multiple dimensions of the 17-gigabase wheat genome and pave the way toward achieving a full reference sequence to underpin wheat research and breeding.

REFERENCE

1. E. Paux *et al.*, *Science* **322**, 101–104 (2008).

10.1126/science.1257983

[1]*International Wheat Genome Sequencing Consortium, 5207 Wyoming Road, Bethesda, MD 20816, USA.* [2]*Bayer CropScience, Innovation Center, 3500 Paramount Parkway, Morrisville, NC 27560, USA.* [3] *Plant Genome and Systems Biology, Helmholtz Zentrum Munich, Ingolstädter Landstrasse 1, 85764 Neuherberg, Germany.*

Published by AAAS

Downloaded from www.sciencemag.org on August 4, 2014

A chromosome-based draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome

The International Wheat Genome Sequencing Consortium (IWGSC)

An ordered draft sequence of the 17-gigabase hexaploid bread wheat (*Triticum aestivum*) genome has been produced by sequencing isolated chromosome arms. We have annotated 124,201 gene loci distributed nearly evenly across the homeologous chromosomes and subgenomes. Comparative gene analysis of wheat subgenomes and extant diploid and tetraploid wheat relatives showed that high sequence similarity and structural conservation are retained, with limited gene loss, after polyploidization. However, across the genomes there was evidence of dynamic gene gain, loss, and duplication since the divergence of the wheat lineages. A high degree of transcriptional autonomy and no global dominance was found for the subgenomes. These insights into the genome biology of a polyploid crop provide a springboard for faster gene isolation, rapid genetic marker development, and precise breeding to meet the needs of increasing food demand worldwide.

Lists of authors and affiliations are available in the full article online.
Corresponding author: K. X. Mayer, e-mail: k.mayer@helmholtz-muenchen.de
Read the full article at http://dx.doi.org/10.1126/science.1251788

Ancient hybridizations among the ancestral genomes of bread wheat

Thomas Marcussen, Simen R. Sandve,* Lise Heier, Manuel Spannagl, Matthias Pfeifer, The International Wheat Genome Sequencing Consortium,† Kjetill S. Jakobsen, Brande B. H Wulff, Burkhard Steuernagel, Klaus F. X. Mayer, Odd-Arne Olsen

The allohexaploid bread wheat genome consists of three closely related subgenomes (A, B, and D), but a clear understanding of their phylogenetic history has been lacking. We used genome assemblies of bread wheat and five diploid relatives to analyze genome-wide samples of gene trees, as well as to estimate evolutionary relatedness and divergence times. We show that the A and B genomes diverged from a common ancestor ~7 million years ago and that these genomes gave rise to the D genome through homoploid hybrid speciation 1 to 2 million years later. Our findings imply that the present-day bread wheat genome is a product of multiple rounds of hybrid speciation (homoploid and polyploid) and lay the foundation for a new framework for understanding the wheat genome as a multilevel phylogenetic mosaic.

The list of author affiliations is available in the full article online.***Corresponding author.**
E-mail: simen.sandve@nmbu.no †The International Wheat Genome Sequencing Consortium (IWGSC) authors and affiliations are listed in the supplementary materials.
Read the full article at http://dx.doi.org/10.1126/science.1250092

Triticum monococcum

Triticum carthlicum

Ancestral wheat

Wheat varieties and species (shown) believed to be the closest living relatives of modern bread wheat (*T. aestivum*). Multiple ancestral hybridizations occurred among most of these species, many of which are cultivated, and along with *T. aestivum* represent a dominant source of global nutrition.

Triticum boeticum

Triticum polonicum L.

Triticum macha

Triticum dicoccoides var. *araraticum*

Published by AAAS

PHOTOS: SUSANNE STAMP, ERNST MERZ/ETH ZURICH

Genome interplay in the grain transcriptome of hexaploid bread wheat

Matthias Pfeifer, Karl G. Kugler, Simen R. Sandve, Bujie Zhan, Heidi Rudi, Torgeir R. Hvidsten, International Wheat Genome Sequencing Consortium,* Klaus F. X. Mayer, Odd-Arne Olsen†

Allohexaploid bread wheat (*Triticum aestivum* L.) provides approximately 20% of calories consumed by humans. Lack of genome sequence for the three homeologous and highly similar bread wheat genomes (A, B, and D) has impeded expression analysis of the grain transcriptome. We used previously unknown genome information to analyze the cell type–specific expression of homeologous genes in the developing wheat grain and identified distinct co-expression clusters reflecting the spatiotemporal progression during endosperm development. We observed no global but cell type– and stage-dependent genome dominance, organization of the wheat genome into transcriptionally active chromosomal regions, and asymmetric expression in gene families related to baking quality. Our findings give insight into the transcriptional dynamics and genome interplay among individual grain cell types in a polyploid cereal genome.

The list of author affiliations is available in the full article online. *The International Wheat Genome Sequencing Consortium (IWGSC) authors and affiliations are listed in the supplementary materials. †**Corresponding author. E-mail: odd-arne.olsen@nmbu.no**

Read the full article at http://dx.doi.org/10.1126/science.1250091

Structural and functional partitioning of bread wheat chromosome 3B

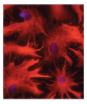
Frédéric Choulet,* Adriana Alberti, Sébastien Theil, Natasha Glover, Valérie Barbe, Josquin Daron, Lise Pingault, Pierre Sourdille, Arnaud Couloux, Etienne Paux, Philippe Leroy, Sophie Mangenot, Nicolas Guilhot, Jacques Le Gouis, Francois Balfourier, Michael Alaux, Véronique Jamilloux, Julie Poulain, Céline Durand, Arnaud Bellec, Christine Gaspin, Jan Safar, Jaroslav Dolezel, Jane Rogers, Klaas Vandepoele, Jean-Marc Aury, Klaus Mayer, Hélène Berges, Hadi Quesneville, Patrick Wincker, Catherine Feuillet

We produced a reference sequence of the 1-gigabase chromosome 3B of hexaploid bread wheat. By sequencing 8452 bacterial artificial chromosomes in pools, we assembled a sequence of 774 megabases carrying 5326 protein-coding genes, 1938 pseudogenes, and 85% of transposable elements. The distribution of structural and functional features along the chromosome revealed partitioning correlated with meiotic recombination. Comparative analyses indicated high wheat-specific inter- and intrachromosomal gene duplication activities that are potential sources of variability for adaption. In addition to providing a better understanding of the organization, function, and evolution of a large and polyploid genome, the availability of a high-quality sequence anchored to genetic maps will accelerate the identification of genes underlying important agronomic traits.

The list of author affiliations is available in the full article online.
***Corresponding author. E-mail: frederic.choulet@clermont.inra.fr**

Read the full article at http://dx.doi.org/10.1126/science.1249721

Published by AAAS

EXTENDED PDF FORMAT
SPONSORED BY

Travel Grants Available
Neuroscience 2014
Application Deadline: July 31st
>Apply Now
R&D systems
a bio-techne brand
www.rndsystems.com

Structural and functional partitioning of bread wheat chromosome 3B

Frédéric Choulet *et al.*

Science **345**, (2014);
DOI: 10.1126/science.1249721

This copy is for your personal, non-commercial use only.

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by clicking here.

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines here.

The following resources related to this article are available online at www.sciencemag.org (this information is current as of August 4, 2014):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:
http://www.sciencemag.org/content/345/6194/1249721.full.html

Supporting Online Material can be found at:
http://www.sciencemag.org/content/suppl/2014/07/16/345.6194.1249721.DC1.html

A list of selected additional articles on the Science Web sites **related to this article** can be found at:
http://www.sciencemag.org/content/345/6194/1249721.full.html#related

This article **cites 121 articles**, 47 of which can be accessed free:
http://www.sciencemag.org/content/345/6194/1249721.full.html#ref-list-1

This article has been **cited by** 1 articles hosted by HighWire Press; see:
http://www.sciencemag.org/content/345/6194/1249721.full.html#related-urls

This article appears in the following **subject collections:**
Botany
http://www.sciencemag.org/cgi/collection/botany
Genetics
http://www.sciencemag.org/cgi/collection/genetics

Downloaded from www.sciencemag.org on August 4, 2014

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published weekly, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. Copyright 2014 by the American Association for the Advancement of Science; all rights reserved. The title *Science* is a registered trademark of AAAS.

A chromosome-based draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome

The International Wheat Genome Sequencing Consortium (IWGSC)

An ordered draft sequence of the 17-gigabase hexaploid bread wheat (*Triticum aestivum*) genome has been produced by sequencing isolated chromosome arms. We have annotated 124,201 gene loci distributed nearly evenly across the homeologous chromosomes and subgenomes. Comparative gene analysis of wheat subgenomes and extant diploid and tetraploid wheat relatives showed that high sequence similarity and structural conservation are retained, with limited gene loss, after polyploidization. However, across the genomes there was evidence of dynamic gene gain, loss, and duplication since the divergence of the wheat lineages. A high degree of transcriptional autonomy and no global dominance was found for the subgenomes. These insights into the genome biology of a polyploid crop provide a springboard for faster gene isolation, rapid genetic marker development, and precise breeding to meet the needs of increasing food demand worldwide.

Lists of authors and affiliations are available in the full article online.
Corresponding author: K. X. Mayer, e-mail: k.mayer@helmholtz-muenchen.de
Read the full article at http://dx.doi.org/10.1126/science.1251788

Ancient hybridizations among the ancestral genomes of bread wheat

Thomas Marcussen, Simen R. Sandve,* Lise Heier, Manuel Spannagl, Matthias Pfeifer, The International Wheat Genome Sequencing Consortium,† Kjetill S. Jakobsen, Brande B. H Wulff, Burkhard Steuernagel, Klaus F. X. Mayer, Odd-Arne Olsen

The allohexaploid bread wheat genome consists of three closely related subgenomes (A, B, and D), but a clear understanding of their phylogenetic history has been lacking. We used genome assemblies of bread wheat and five diploid relatives to analyze genome-wide samples of gene trees, as well as to estimate evolutionary relatedness and divergence times. We show that the A and B genomes diverged from a common ancestor ~7 million years ago and that these genomes gave rise to the D genome through homoploid hybrid speciation 1 to 2 million years later. Our findings imply that the present-day bread wheat genome is a product of multiple rounds of hybrid speciation (homoploid and polyploid) and lay the foundation for a new framework for understanding the wheat genome as a multilevel phylogenetic mosaic.

The list of author affiliations is available in the full article online.***Corresponding author.**
E-mail: simen.sandve@nmbu.no †The International Wheat Genome Sequencing Consortium (IWGSC) authors and affiliations are listed in the supplementary materials.
Read the full article at http://dx.doi.org/10.1126/science.1250092

Triticum monococcum

Triticum carthlicum

Ancestral wheat

Wheat varieties and species (shown) believed to be the closest living relatives of modern bread wheat (*T. aestivum*). Multiple ancestral hybridizations occurred among most of these species, many of which are cultivated, and along with *T. aestivum* represent a dominant source of global nutrition.

Triticum boeticum

Triticum polonicum L.

Triticum macha

Triticum dicoccoides var. *araraticum*

Published by AAAS

PHOTOS: SUSANNE STAMP, ERNST MERZ/ETH ZURICH

Genome interplay in the grain transcriptome of hexaploid bread wheat

Matthias Pfeifer, Karl G. Kugler, Simen R. Sandve, Bujie Zhan, Heidi Rudi, Torgeir R. Hvidsten, International Wheat Genome Sequencing Consortium,* Klaus F. X. Mayer, Odd-Arne Olsen†

Allohexaploid bread wheat (*Triticum aestivum* L.) provides approximately 20% of calories consumed by humans. Lack of genome sequence for the three homeologous and highly similar bread wheat genomes (A, B, and D) has impeded expression analysis of the grain transcriptome. We used previously unknown genome information to analyze the cell type–specific expression of homeologous genes in the developing wheat grain and identified distinct co-expression clusters reflecting the spatiotemporal progression during endosperm development. We observed no global but cell type– and stage-dependent genome dominance, organization of the wheat genome into transcriptionally active chromosomal regions, and asymmetric expression in gene families related to baking quality. Our findings give insight into the transcriptional dynamics and genome interplay among individual grain cell types in a polyploid cereal genome.

The list of author affiliations is available in the full article online. *The International Wheat Genome Sequencing Consortium (IWGSC) authors and affiliations are listed in the supplementary materials. †**Corresponding author. E-mail: odd-arne.olsen@nmbu.no**

Read the full article at http://dx.doi.org/10.1126/science.1250091

Structural and functional partitioning of bread wheat chromosome 3B

Frédéric Choulet,* Adriana Alberti, Sébastien Theil, Natasha Glover, Valérie Barbe, Josquin Daron, Lise Pingault, Pierre Sourdille, Arnaud Couloux, Etienne Paux, Philippe Leroy, Sophie Mangenot, Nicolas Guilhot, Jacques Le Gouis, Francois Balfourier, Michael Alaux, Véronique Jamilloux, Julie Poulain, Céline Durand, Arnaud Bellec, Christine Gaspin, Jan Safar, Jaroslav Dolezel, Jane Rogers, Klaas Vandepoele, Jean-Marc Aury, Klaus Mayer, Hélène Berges, Hadi Quesneville, Patrick Wincker, Catherine Feuillet

We produced a reference sequence of the 1-gigabase chromosome 3B of hexaploid bread wheat. By sequencing 8452 bacterial artificial chromosomes in pools, we assembled a sequence of 774 megabases carrying 5326 protein-coding genes, 1938 pseudogenes, and 85% of transposable elements. The distribution of structural and functional features along the chromosome revealed partitioning correlated with meiotic recombination. Comparative analyses indicated high wheat-specific inter- and intrachromosomal gene duplication activities that are potential sources of variability for adaption. In addition to providing a better understanding of the organization, function, and evolution of a large and polyploid genome, the availability of a high-quality sequence anchored to genetic maps will accelerate the identification of genes underlying important agronomic traits.

The list of author affiliations is available in the full article online.
***Corresponding author. E-mail: frederic.choulet@clermont.inra.fr**

Read the full article at http://dx.doi.org/10.1126/science.1249721

Published by AAAS

WHEAT GENOME

Structural and functional partitioning of bread wheat chromosome 3B

Frédéric Choulet,[1,2]* Adriana Alberti,[3] Sébastien Theil,[1,2] Natasha Glover,[1,2] Valérie Barbe,[3] Josquin Daron,[1,2] Lise Pingault,[1,2] Pierre Sourdille,[1,2] Arnaud Couloux,[3] Etienne Paux,[1,2] Philippe Leroy,[1,2] Sophie Mangenot,[3] Nicolas Guilhot,[1,2] Jacques Le Gouis,[1,2] Francois Balfourier,[1,2] Michael Alaux,[4] Véronique Jamilloux,[4] Julie Poulain,[3] Céline Durand,[3] Arnaud Bellec,[5] Christine Gaspin,[6] Jan Safar,[7] Jaroslav Dolezel,[7] Jane Rogers,[8] Klaas Vandepoele,[9] Jean-Marc Aury,[3] Klaus Mayer,[10] Hélène Berges,[5] Hadi Quesneville,[4] Patrick Wincker,[3,11,12] Catherine Feuillet[1,2]

We produced a reference sequence of the 1-gigabase chromosome 3B of hexaploid bread wheat. By sequencing 8452 bacterial artificial chromosomes in pools, we assembled a sequence of 774 megabases carrying 5326 protein-coding genes, 1938 pseudogenes, and 85% of transposable elements. The distribution of structural and functional features along the chromosome revealed partitioning correlated with meiotic recombination. Comparative analyses indicated high wheat-specific inter- and intrachromosomal gene duplication activities that are potential sources of variability for adaption. In addition to providing a better understanding of the organization, function, and evolution of a large and polyploid genome, the availability of a high-quality sequence anchored to genetic maps will accelerate the identification of genes underlying important agronomic traits.

Bread wheat (*Triticum aestivum* L.) is a staple food for 30% of the world population. It is a hexaploid species (6x = 2*n* = 42, AABBDD) that originates from two interspecific hybridizations estimated to have taken place ~0.5 million and 10,000 years ago (*1*). The predicted closest extant representatives of the ancestral parental diploid species (2*n* = 14) are *Triticum urartu* (A genome), *Aegilops speltoides* (S genome related to the B genome), and *Aegilops tauschii* (D genome). Each of the three ancestral genomes is about 5.5 Gb in size and, therefore, results in a highly redundant 17-Gb hexaploid genome with three homeologous sets of seven chromosomes (1A to 7A, 1B to 7B, and 1D to 7D), each carrying highly similar gene copies. Moreover, most of the genome was shaped by the amplification of transposable elements (TEs) that include highly repeated families and sequences (*2*). This high redundancy has complicated the assembly of a complete and properly ordered reference sequence of the bread wheat genome. A fully sequenced genome enables scientists and breeders to have access to a complete gene set, with the gene order along each chromosome, and to identify candidate genes between markers associated with important traits. It also enables the identification of recent duplicates, which may be involved in species-specific evolution (*3*), and tracing of their evolutionary history. Before obtaining a full genome sequence, the wheat gene space has been investigated through various genome and transcriptome survey sequencing approaches and through microarray hybridizations (*4*–*7*). Recently, whole-genome shotgun sequencing of cultivar Chinese Spring using Roche/454 technology and synteny-driven assembly yielded ~95,000 gene models [N50 = 0.9 kb (*8*)]. Furthermore, the gene space of the diploid wild relatives *Ae. tauschii* (DD) and *T. urartu* (AA) has also been assembled and led to describe 43,150 and 34,879 genes, respectively (*9*, *10*). Although these sequences are useful templates for marker design and comparative analyses, as a result of assembly limitations of short-read-based sequencing (*11*, *12*), they are still very fragmented, and a large fraction of the genes are unanchored to chromosomes. The maize (*Zea mays*) and potato (*Solanum tuberosum*) sequencing projects, both representing species with highly repetitive genomes, were able to avoid overfragmentation by combining multiple sequencing technologies and through the use of DNA libraries with a diversity of insert sizes (*13*, *14*).

The International Wheat Genome Sequencing Consortium (IWGSC) road map focuses on physically mapping and obtaining a high-quality reference sequence of each of the 21 individual wheat chromosomes rather than approaching the hexaploid genome as a whole. This strategy relies on flow-sorting individual chromosomes and/or chromosome arms from ditelosomic lines of the cultivar Chinese Spring to construct bacterial artificial chromosome (BAC) libraries (*15*). The largest chromosome is 3B (~1 Gb). It was the first chromosome for which a BAC library was constructed (*16*) and a physical map achieved (*17*). A pilot sequencing study on 13 contigs (*2*) suggested that genes tend to be mainly clustered into small islands, the presence of a twofold gene density increase from the centromere toward the telomeres, and a high proportion of nonsyntenic genes interspersed within a conserved ancestral grass gene backbone. It provided a proof of principle for this strategy and opened the way for producing a reference sequence of the large and polyploidy wheat genome.

Sequencing and construction of a pseudomolecule

We used a hybrid sequencing and BAC pooling strategy to sequence 8452 BAC clones from the minimal tiling path (MTP) that was established during the construction of the chromosome 3B physical map (*4*, *18*). After the integration of BAC-end sequences, manual curation of the scaffolding, gap filling, and correction of potential sequencing errors (*18*), we obtained a final assembly of 2808 scaffolds representing 833 Mb with a N50 of 892 kb (i.e., half of the chromosome sequence is assembled in scaffolds larger than 892 kb). We estimated that about 6% of the chromosome sequence was not present in the MTP BAC-based assembly through comparison with the 546,922 contigs assembled from whole-chromosome shotgun sequencing of flow-sorted 3B DNA (*19*). This suggests that the size of chromosome 3B is nearly 886 Mb—that is, about 11% smaller than originally predicted (*16*, *20*). We built a pseudomolecule of chromosome 3B by ordering 1358 scaffolds along the chromosome using an ordered set of 2594 anchor single-nucleotide polymorphism (SNP) markers. The pseudomolecule represents 774.4 Mb (93% of the complete sequence), with a scaffold N50 of 949 kb (table S1). The order of markers was determined by linkage analysis of a recombinant inbred line (RIL) population derived from a cross between *T. aestivum* cultivars Chinese Spring (reference sequence) and Renan (a French elite cultivar) and refined by integrating linkage disequilibrium data from two panels and physical BAC contig information (*18*). This sequence corresponds to an annotation-directed improved high-quality draft (*21*) situated between the high-quality finished rice

[1]Institut National de la Recherche Agronomique (INRA) UMR1095, Genetics, Diversity and Ecophysiology of Cereals, 5 Chemin de Beaulieu, 63039 Clermont-Ferrand, France. [2]University Blaise Pascal, UMR1095, Genetics, Diversity and Ecophysiology of Cereals, 5 Chemin de Beaulieu, 63039 Clermont-Ferrand, France. [3]Commissariat à l'Énergie Atomique et aux Énergies Alternatives, Direction des Sciences du Vivant, Institut de Génomique, Genoscope, 2 Rue Gaston Crémieux, 91000 Evry, France. [4]INRA, UR1164 Unité de Recherche Génomique Info Research Unit in Genomics-Info, INRA de Versailles, Route de Saint-Cyr, 78026 Versailles, France. [5]Centre National des Ressources Génomiques Végétales, INRA UPR 1258, 24 Chemin de Borde Rouge, 31326 Castanet-Tolosan, France. [6]Biométrie et Intelligence Artificielle, INRA, Chemin de Borde Rouge, BP 27, 31326 Castanet-Tolosan, France. [7]Centre of the Region Haná for Biotechnological and Agricultural Research, Institute of Experimental Botany, Slechtitelu 31, CZ-78371 Olomouc, Czech Republic. [8]The Genome Analysis Centre, Norwich, Norwich Research Park, Norwich NR4 7UH, UK. [9]Department of Plant Systems Biology (VIB) and Department of Plant Biotechnology and Bioinformatics (Ghent University), Technologiepark 927, 9052 Gent, Belgium. [10]Munich Information Center for Protein Sequences, Institute of Bioinformatics and Systems Biology, Helmholtz Zentrum Muenchen, D-85764 Neuherberg, Germany. [11]CNRS UMR 8030, 2 Rue Gaston Crémieux, 91000 Evry, France. [12]Université d'Evry, CP5706 Evry, France.
*Corresponding author. E-mail: frederic.choulet@clermont.inra.fr

genome sequence and the improved draft maize genome (*13*).

Annotation of genes, transcribed loci, and transposable elements

Gene modeling led to the prediction of 7264 coding loci on the 3B pseudomolecule (Table 1), including 5326 with a functional structure and 1938 (27%) likely corresponding to pseudogenes. An additional 251 gene models and 188 pseudogenes were annotated in unanchored scaffolds. RNA-Seq data revealed that 71.4% of the predicted genes/pseudogenes are transcribed and led to the identification of 3692 unannotated transcribed loci that may encode functional noncoding RNAs or unknown proteins, hereafter referred to as novel transcribed regions (NTRs) (Table 1). In addition, 791 highly conserved noncoding RNA genes involved in RNA maturation and protein synthesis were also predicted (Table 1). Chromosome 3B appears to contain a high number of small nuclear RNA genes (U1 to U6) including nine U1-snRNAs (small nuclear RNAs), seven of which are tandemly duplicated. As a comparison, there are 14 U1-snRNAs in the entire *Arabidopsis thaliana* genome (www.plantgdb.org). The higher number of U1-snRNAs may reflect a higher level of duplication in the wheat genome. We found 53,288 complete and 181,058 truncated copies of TEs, belonging to 485 TE families and representing 85% (640 Mb) of the 3B pseudomolecule, through a similarity-search approach. Further de novo repeat detection (*18*) identified 3.6% putatively new TEs.

We estimated the putative location of the centromeric region by plotting the density of the long terminal repeat retrotransposons (LTR-RTs) CRW (centromeric retrotransposons of wheat) and Quinta along the pseudomolecule. These LTR-RTs are recognized by the centromere-specific histone CenH3 and thus are centromere-functional sequences (*22*). Two major peaks covering a region of 122 Mb (265 to 387 Mb) (fig. S1)—which includes 1 Mb previously shown as interacting with histone CenH3 (*22*) and encompassing the centromere of the orthologous rice chromosome 1 (*23*)—were identified. This region was defined as the centromeric-pericentromeric region of chromosome 3B. A strong correlation has been observed between the size of the centromeres and the chromosomes in grasses (*24*), and it is likely that large chromosomes have centromeres larger than 10 Mb. This may be critical to ensure the structural rigidity of the pericentromeric regions needed for kinetochore co-orientation (*25*). Marker assignation to either the short or the long arm indicated the presence of a break point between 349.4 and 350.0 Mb that might be the position of the core centromere.

Variability in recombination rate and gene density along the chromosome

We found 787 crossover (CO) events on chromosome 3B in the Chinese Spring x Renan population, with on average 2.6 COs per chromosome per individual, which is similar to maize [2.7 to 3.4 (*26*)]. Distribution of meiotic recombination rates revealed extreme variations along the chromosome. Whereas the average recombination rate is 0.16 cM/Mb, actual values range from 0 to 2.30 cM/Mb (per 10-Mb window) (Fig. 1A). Segmentation analysis (*18*) revealed partitioning with the two distal regions of 68 Mb (region R1) and 59 Mb (region R3) on the short and long arms, respectively, showing recombination rates of 0.60 cM/Mb and 0.96 cM/Mb on average, and a large proximal region of 648 Mb (region R2) spanning the centromere with an average recombination rate of 0.05 cM/Mb (Table 2 and Fig. 1A). This provides insight into the actual physical size of the highly recombinogenic regions previously detected at the end of the wheat chromosomes (*27*, *28*). When a narrower window of 1 Mb was used, variations ranged from 0 to 12 cM/Mb (Fig. 1A), a range similar to that observed in maize [0.8 to 11.5 cM/Mb (*26*)] and sorghum [0 to 10 cM/Mb (*29*)]. All crossover events occurred in only 13% of the chromosomes in our population of 305 individuals. The largest region totally deprived of recombination corresponds to 150 Mb and includes the putative 122-Mb centromeric-pericentromeric region. This was

Table 1. General features of the 3B pseudomolecule. LINEs, long interspersed nuclear elements; SINEs, short interspersed nuclear elements; rRNA, ribosomal RNA; snoRNA, small nucleolar RNA; TIRs, terminal inverted repeats.

Pseudomolecule sequence				
	Length (bps)	**774,434,471**		
	G+C content	**46.16%**		
Protein-coding genes		**All**	**Full genes**	**Pseudogenes**
	No. of genes	**7264**	**5326**	**1938**
	Average size (bps) of coding sequences (± standard deviation)	1095 ± 807	1187 ± 821	840 ± 710
	Average number of exons (± standard deviation)	4.2 ± 4.4	4.4 ± 4.6	3.6 ± 3.8
	Gene density (kb^{-1})	107	145	400
	No. of expressed genes	5185	4125	1060
	No. of genes with alternative splicing	3185	2596	589
	% genes with alternative splicing	61	63	56
	Average no. isoforms/expressed gene	5.8	5.8	5.8
	NTRs	3692		
Noncoding RNA genes				
	tRNA	589		
	5S rRNA	85		
	Others (snRNA, snoRNA)	117		
	Total	791		
Transposable elements (TEs)				
Class I	Copia	15.6%		
	Gypsy	46.9%		
	Unclassified LTR-retrotransposons	3.5%		
	LINEs	1.2%		
	SINEs	0.01%		
	Total class I	67.1%		
Class II	CACTA	16.4%		
	Harbinger	0.19%		
	Mariner	0.19%		
	Mutator	0.43%		
	hAT	0.02%		
	Unclassified class II with TIRs	0.22%		
	Unclassified class II	0.10%		
	Helitron	0.01%		
	Total class II	17.6%		
Unclassified repeats		0.81%		
Total TEs		85.5%		

confirmed by the linkage disequilibrium (LD) pattern (fig. S2). Twenty-two regions showed a recombination ratio higher than 1.6 cM/Mb—i.e., >10 times the average for this chromosome—and thus may contain recombination hot spots (Fig. 1A). However, no significant correlation was observed between the recombination rate and gene content, coding DNA, or TE content of these regions.

The 7264 genes are not evenly distributed, and gene density is increasing on both arms along the centromere-telomere axis, correlating with the distance to the centromere (r_s = 0.79, $P < 2.2 \times 10^{-16}$, R^2 = 0.61) (Fig. 1B). Using a 10-Mb window, the average gene density estimate is 9 ± 5 genes/Mb, ranging from 1.3 in the centromeric-pericentromeric region up to 27.9 at the most telomeric end of the short arm, a pattern commonly observed in grass genomes. Variation of the gene density in wheat chromosome 3B is higher than for chromosomes in the more compact rice genome (*30*); lower than in sorghum, where genes are mostly found in the telomeric regions (*31*); and in the same range as in maize, which also contains a high percentage of TEs (*13*). Segmentation analysis revealed five major regions with contrasted gene densities (Fig. 1B) and a fourfold gradient of the gene density—i.e., twice as many as suggested by the pilot study on chromosome 3B (*2*). The distal segments exhibiting the highest gene density (19 genes per Mb) correspond nearly to the highly recombinogenic R1 and R3 regions (Fig. 1A). The R2 region was subdivided into three segments, with the lowest gene density (5 genes per Mb) in a 234-Mb segment encompassing the centromeric-pericentromeric region. As previously suggested (*2*, *4*), there is no large region completely devoid of coding sequence (maximum of 3.7 Mb). We confirmed that the intergenic distances (IGDs) are extremely variable (average 104 ± 190 kb) and that a majority (73%) of the genes are organized in small islands, or insulae (*32*). This suggests that most of the intergenic regions are under selective constraint prevented from TE insertion. Indeed, only 29% of the IGDs are larger than 104 kb, but they account for 81% of the chromosome size, demonstrating that TE-mediated genome expansion likely occurred within a limited number of intergenic regions.

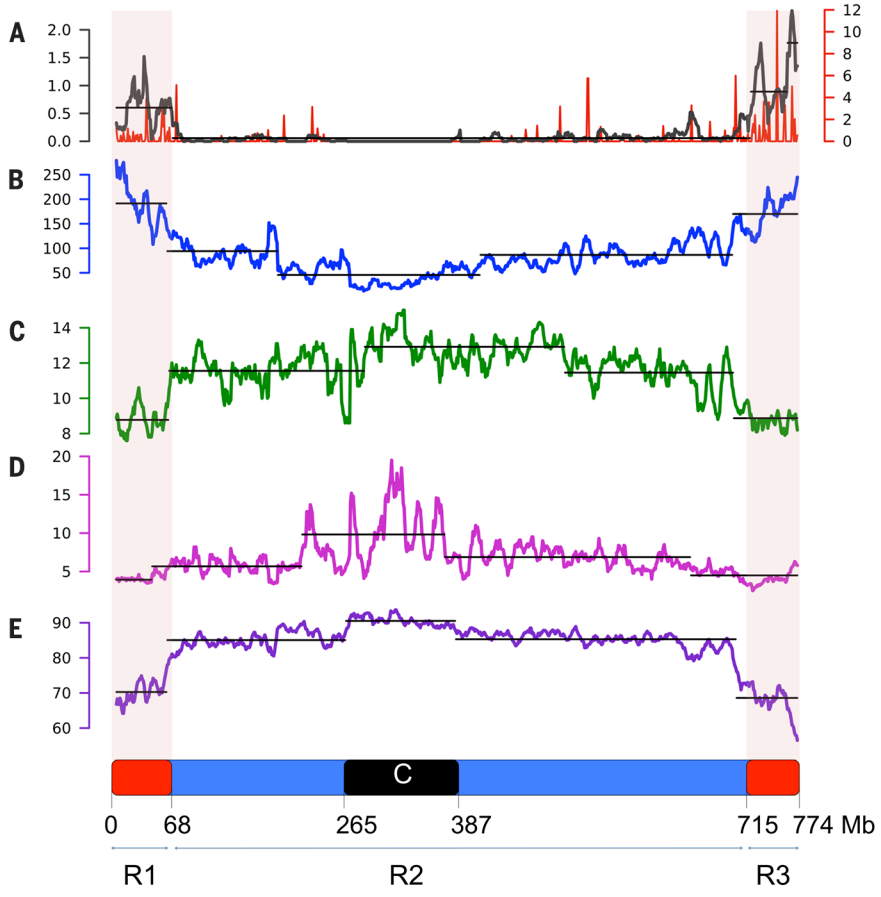


Fig. 1. Structural and functional partitioning of wheat chromosome 3B. Distribution and segmentation analysis of (**A**) meiotic recombination rate (cM/Mb in sliding window of 10 Mb in black and 1 Mb in red); (**B**) gene density (CDS/10 Mb); (**C**) expression breadth; (**D**) average number of alternative spliced transcripts per expressed gene; and (**E**) TE content along the 3B pseudomolecule. Distal regions of the chromosome R1 and R3 are represented in red. In (C), the centromeric/pericentromeric region is in black. The borders of these regions are indicated in Mb. Sliding window size: 10 Mb; step: 1 Mb.

Relationships between gene expression, function, and chromosome location

Of all annotated genes on chromosome 3B, 71.4% are expressed in at least one of the 15 conditions analyzed [five organs at three developmental stages each (table S2)], 33% in all conditions, and 5% in one only (fig. S3). On average, genes are expressed in 10.8 of 15 conditions (considering all predictions), and expressed genes are transcribed into 5.8 alternative transcripts, or isoforms. Both the expression breadth and the average number of isoforms are distributed unevenly along the chromosome, with a clear decrease of the two parameters toward the telomeres (Fig. 1, C and D). Segmentation revealed distal segments with boundaries similar to that of regions R1 and R3 and with genes expressed in fewer conditions than in the proximal region: 8.7 versus 11.7, respectively ($P < 2.2 \times 10^{-16}$, Welch *t* test) (Table 2). Similarly, the average number of alternative transcripts is higher in the proximal (6.5) than in the distal (4.3) regions ($P < 2.2 \times 10^{-16}$, Welch *t* test) (Table 2).

Gene ontology (GO) term enrichment was estimated for the R1, R2, and R3 regions (*18*) (tables S3 to S5). The distal regions are enriched in many GO categories, some being related to adaptation (response to abiotic stimulus or response to stress). Well-known examples of genes related to adaptation are those involved in resistance to pathogens. Chromosome 3B carries 171 genes putatively associated with disease resistance (*18*), and their distribution is highly biased, with 135 (79%) of them located in the distal regions (whereas these regions contain just 33% of the gene set). Such uneven distribution and the correlation with the distribution of crossovers suggest that meiotic recombination acts as a main driver for creating variability in distal regions of chromosome 3B.

To investigate whether such partitioning is a common pattern of large plant genomes, we analyzed the distribution of the gene expression breadth in maize and barley, which both exhibit large genome size (>1 Gb) and increased recombination rates at chromosomal extremities (*33*, *34*). In barley, segmentation analysis of the seven chromosomes based on recombination data identified the same pattern as on chromosome 3B, with two highly recombinogenic distal regions and a large nonrecombinogenic region. Using expression data of eight conditions (*34*), we also observed that the two high-recombination distal regions carry genes expressed in fewer conditions than those carried by the low-recombination proximal regions (5.9 versus 6.7; $P = 2.2 \times 10^{-16}$, Welch *t* test) (fig. S4A). Using GO terms, we found a significant enrichment of these regions in the categories "cell death" and "defense response," which support previous findings that barley disease resistance genes are clustered in the distal regions (*34*). In contrast, in maize, although we

observed partitioning of the recombination rate, no gene-expression partitioning was detected using RNA-Seq data of 18 conditions (*35*). Overall, the expression breadth is 13.2 and 12.7 in high- and low-recombination regions, respectively, with chromosome-specific patterns (fig. S4B). Nevertheless, high-recombination regions are also enriched in GO categories "cell death" and "defense response to fungus and bacteria," suggesting that such genes are consistently found in distal recombinogenic regions in large plant genomes. These results suggest that the partitioning observed on wheat chromosome 3B is conserved in the *Triticeae* and may not reflect a general pattern of large genomes. Alternatively, it is possible that the active rearrangements observed in the maize genome have modified this pattern. Additional evidence will come once other large plant genomes (>1 Gb) are sequenced and analyzed.

Uneven distribution of transposable elements

LTR-RTs represent 66% of the chromosome 3B sequence (gypsy, 47%; copia, 16%; unclassified, 3%) (Table 1), which is slightly lower than the ~75% of LTR-RT identified in the whole maize genome (*36*). Only 4% (3 out of 85) of the LTR-RT families were found in single copies, compared with 41% in maize (*36*) and 48% in the rice genome (*37*). Sixteen percent of the sequence is composed of class II DNA transposons that mostly correspond to CACTA elements (Table 1), compared with 3.2% in the maize genome (*13*). Only six families account for 50% of the wheat chromosome 3B TE fraction, as previously suggested from partial sequence analyses (*2*, *38*) and from observations in other large genomes (*36*). However, in contrast to the maize genome, in which most of the intact elements are found in 1 to 10 copies (*36*), the majority of the TE families annotated on chromosome 3B have a higher number of copies (10 to 1000 copies). Estimated insertion dates for the most abundant LTR-RT families showed a major peak at 1.5 million years (My) but also quite specific patterns of TE activity for each family (fig. S5). Our data support the hypothesis (*2*, *38*) that most of the transposable elements that shaped the B genome were inserted before polyploidization [0.5 million years ago (Ma)] and have been less active since then. Distribution of recently inserted elements revealed that TE insertion occurred at a similar rate in the distal and proximal regions. In contrast, older insertions (>1.5 Ma) were 1.7 times as abundant in the R2 region compared with the R1 and R3 regions, suggesting a higher rate of TE elimination in the distal ends of chromosome 3B.

The TE density distribution was not random (Fig. 1E), with a lower density in the R1 (73%) and R3 (68%) regions compared with the R2 region (88%) (Table 2). The 122-Mb centromeric-pericentromeric region displayed the highest density (93%) of TEs. Beneath the global TE distribution pattern, each superfamily presents its own specificities (fig. S6). For example, CACTA transposons are more abundant in the distal gene-rich regions (Table 2), supporting in situ hybridization findings at the whole-genome level (*39*). In addition, the distribution of TE families varied on the basis of their relative distance to genes (*18*) (fig. S7). DNA transposons Mutator, Harbinger, and MITEs are found close to genes, whereas LTR-RTs and CACTAs tend to be located at much larger distances from the genes. For instance, the 17,479 annotated MITEs were found to be significantly associated with genes ($r = 0.89$; $P < 1 \times 10^{-10}$), as previously observed in plant genomes (*40*).

Synteny between chromosome 3B and related grass genomes

Comparative genomics in grasses has been used to define syntenic relationships between different species (*41*, *42*) and to provide insight into their evolution since the divergence from a common ancestor 50 to 70 Ma (*43*). We compared the wheat chromosome 3B genes (Ta3B) with the closest sequenced relative, *Brachypodium distachyon* [common ancestor, 32 to 39 Ma (*44*)], and with one representative of each of the *Ehrhartoideae* and *Panicoideae grass* subfamilies: *Oryza sativa indica* [rice (*30*)] and *Sorghum bicolor* (*31*), respectively. Wheat chromosomes of group 3 are syntenic with chromosome 1 of rice (Os1), chromosome 3 of sorghum (Sb3), and the distal parts of *B. distachyon* chromosome 2 (Bd2). We first investigated potential gene loss after polyploidization by using the conserved and syntenic genes found on chromosomes Os1, Bd2, and Sb3. These represent the grass core genes that are expected to be present on wheat homeologous group 3, unless they have been lost by fractionation after polyploidization. The finding that 94% of the conserved genes are also present on the 3B sequence (Fig. 2A), which represents 94% of the chromosome (see above), suggests that no major gene loss has occurred in the B subgenome yet. This is confirmed at the whole-genome level by the results of the chromosome survey sequences (*19*). In contrast, 2065 genes on chromosome 3B (34.6%, including pseudogenes) shared similarity with genes on nonorthologous chromosomes in the other grass genomes. This proportion of nonsyntenic genes is much higher than the 5% (between 149 and 207) of nonsyntenic genes found in the other grass species analyzed (Fig. 2A and table S6). It confirms previous results showing substantial modifications and rearrangements of the wheat gene space (*2*). When looking at the conservation of the gene order, collinear genes represent 42 to 68% of the genes present on Os1, Bd2, and Sb3, whereas they represent less than 30% of the Ta3B genes (including pseudogenes) (table S7 and fig. S8). The spatial distribution of syntenic and nonsyntenic genes

Table 2. Distribution of features in the three regions of chromosome 3B as defined from the recombination segmentation along the chromosome.

	R1	R2	R3
Size (Mb)	68	648	59
Recombination rate (cM/Mb)	0.60	0.05	0.96
Genes			
Predicted gene density (Mb^{-1})	19	7	19
Number of predicted genes/pseudogenes	1318	4845	1,101
Full genes (no.)	910 (69%)	3682 (76%)	734 (67%)
Pseudogenes/gene fragments (no.)	408 (31%)	1163 (24%)	367 (33%)
Mean intergenic distance (kb)	49	130	52
Expressed predicted genes (no.)	823 (62%)	3629 (75%)	733 (67%)
Expressed full genes (no.)	621	2963	541
Expressed pseudogenes/fragments (no.)	202	666	192
Average expression breadth (per expressed gene; /15)	8.8	11.7	8.6
Average FPKM (per expressed gene)	141	255	156
Average number of isoforms (per expressed gene)	4.2	6.5	4.4
Proportion of nonsyntenic genes* (%)	44	28	53
Proportion of intrachromosomally duplicated genes* (%)	49	33	42
Proportion of tandemly duplicated genes* (%)	24	14	22
Proportion of dispersed duplicated genes* (%)	26	18	20
Proportion of interchromosomally duplicated genes* (%)	36	33	37
Transposable elements (%)	73.0	88.3	68.4
Copia (%)	14.7	15.8	14.1
Gypsy (%)	31.7	50.3	27.1
CACTA (%)	18.7	15.9	19.5

*Number of duplicated genes (filtered set, including pseudogenes) divided by the total number of genes in each region.

along the 3B pseudomolecule (Fig. 2B) shows an increased proportion of nonsyntenic genes in the R1 (44%) and R3 (53%) regions compared with the R2 region (28%) (Table 2). This supports the hypothesis that accelerated evolution occurred in the wheat lineage compared with other grasses (*2*, *45*, *46*), with insertions of nonsyntenic genes intercalated in the ancestral grass genome backbone through gene duplications or translocations that preferentially occurred in the distal recombining regions.

Origin and evolution of nonsyntenic genes

With such a high proportion of nonsyntenic genes, one key question is whether these genes are under selection pressure or in the process of becoming pseudogenes. On the basis of the coding sequence structure, 32% of the nonsyntenic genes (versus 17% of syntenic genes) were annotated as likely pseudogenes or gene fragments. This ratio is not surprising, given that TE activity can duplicate gene fragments that are dead upon arrival. Expression patterns revealed that a majority of the nonsyntenic genes (69% versus 82% of syntenic genes) are expressed in at least one condition tested (table S8), thereby suggesting that a large fraction of these relocated genes are unlikely to be pseudogenes and may contribute to recent wheat genome evolution and, therefore, to adaptation. Interestingly, a majority (51%) of the genes expressed in a single condition corresponds to nonsyntenic genes, whereas 80% of the genes that are expressed in all 15 conditions are syntenic genes (fig. S9). This suggests that nonsyntenic genes are involved in specific processes that may be related to adaptation, whereas syntenic genes tend to be associated with essential biological processes. This hypothesis is supported by the fact that putative resistance genes identified on chromosome 3B are mainly nonsyntenic genes (*18*). In addition, GO term enrichment of nonsyntenic genes revealed an overrepresentation of genes involved in response to stress (table S9).

The fact that chromosome 3B exhibits a higher number of genes than its orthologs in other grasses and that at least 94% of the ancestral grass gene backbone is conserved indicates that most insertions of nonsyntenic genes result from interchromosomal duplication with retention of the parental copy. To test this hypothesis, we used the sequences of the 18 bread wheat chromosomes nonhomeologous to group 3 chromosomes (*19*) to search for potential parental copies of chromosome 3B genes elsewhere in the genome (*18*) (table S10). A paralog was identified for 87% of the nonsyntenic genes (*18*), with no bias regarding the chromosomal origin of the interchromosomally duplicated genes (fig. S10). Duplications of DNA fragments to different locations in a genome have been shown to result from double-strand break (DSB) repair (in which a copy of the foreign DNA is used as filler to repair the break) or capture by active TEs (*46*, *47*). We analyzed the composition of the regions flanking the syntenic versus nonsyntenic genes (20 kb on each side) and found a high association of nonsyntenic genes with a class II transposon superfamily: 41% more CACTAs were found around nonsyntenic genes than around syntenic genes (fig. S11). CACTA transposons are known to capture genes (*31*, *48*) and may have contributed substantially to interchromosomal gene duplications in wheat.

We also investigated the time since duplication of nonsyntenic genes through the analysis of nucleotide substitution rates (K_s) (*18*). In total, 63% of these duplications were older than 10 My and, thus, are likely shared within other *Triticeae* species, whereas 37% are potentially wheat-specific. Comparison with the barley genome survey sequence data (*34*) showed that at least 29% of the 3B nonsyntenic genes (versus 51% of the syntenic genes) are orthologous with barley chromosome 3H, confirming that part of the nonsyntenic genes were relocated before the divergence of wheat and barley 10 to 14 Ma.

We next asked if the high gene duplication activity is also observed at the intrachromosomal level. We identified 809 gene families with two or more copies comprising 2216 genes on chromosome 3B, which is about three times as much as in rice, *Brachypodium*, and sorghum (table S11). This indicates that, in proportion, more than twice as many genes were duplicated or retained after intrachromosomal duplications in wheat (~37%) compared with the other three grasses (~15 to 18%). About 46% of the duplicated genes of chromosome 3B are found in tandem, whereas 54% are dispersed duplicates (*18*) (table S12). In other grass species, a majority of the duplicated genes are organized in tandem. Given the high interchromosomal duplication activity observed in our analyses (see above), it is possible that some dispersed duplicates on chromosome 3B originated through independent interchromosomal duplications rather than through intrachromosomal duplications, thereby leading to overestimates of the latter. However, even when considering syntenic dispersed duplicates—i.e., those genes that have remained at their ancestral locus and have undergone intrachromosomal duplication—23% of the whole gene set appears to have originated from recent intrachromosomal duplications, which is still higher than in other grass species. Thus, we conclude that both inter- and intrachromosomal

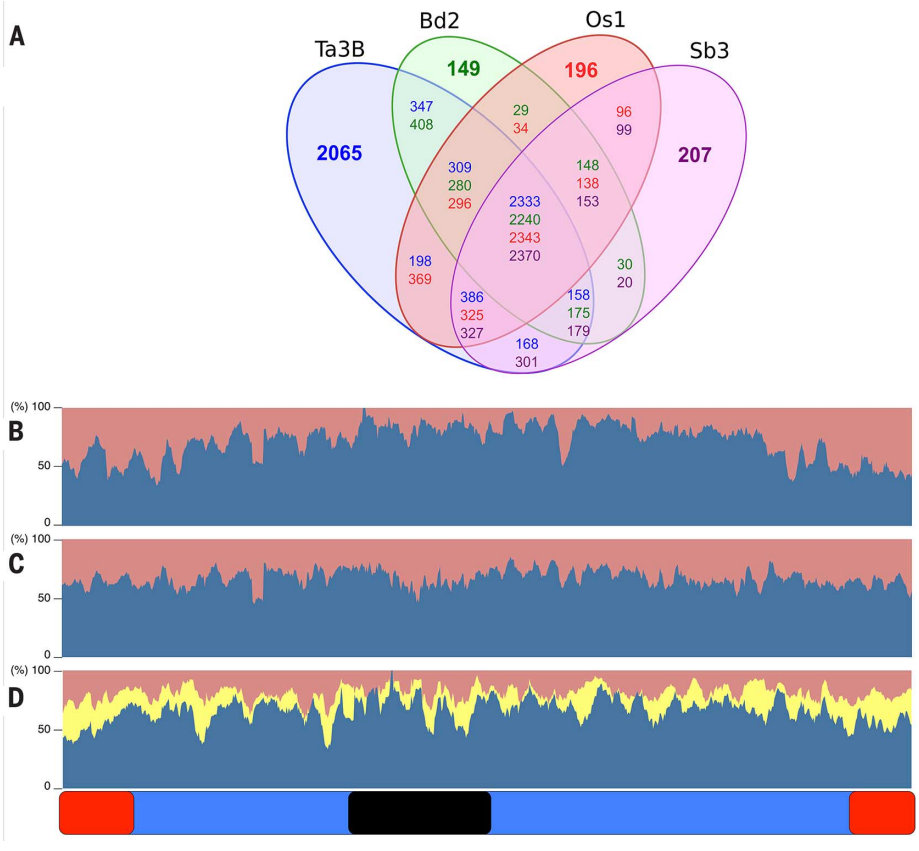


Fig. 2. Inter- and intraspecific comparative analyses of the gene content of wheat chromosome 3B. (**A**) Venn diagram displaying the number of genes conserved between wheat chromosome 3B (Ta3b, blue) and orthologous chromosomes in rice (Os1, red), *Brachypodium* (Bd2, green), and sorghum (Sb3, purple). The number of nonsyntenic genes is indicated in bold for each species. Distribution along the 774 Mb of the chromosome 3B pseudomolecule of the relative proportion of (**B**) syntenic (blue) versus nonsyntenic (red) genes, (**C**) interchromosomal duplications (duplicates in red, group 3–specific genes in blue), and (**D**) tandem (yellow) and dispersed (red) intrachromosomal duplications and singletons (blue). Chromosome 3B is represented at the bottom with distal regions in red and the centromeric/pericentromeric region in black.

rates of duplication are higher in wheat than in the other grass species analyzed so far. Interestingly, interchromosomal duplicates were distributed uniformly along the chromosome, whereas the proportion of tandem duplicates slightly increased in the distal regions (Fig. 2D). This suggests that long-distance and tandem duplications likely arose through different mechanisms. Finally, expression analysis of intrachromosomal duplicated genes indicated that 49% of the families show expression of all copies in at least one condition. Similar to what was observed for the interchromosomal duplicated genes, the intrachromosomal duplicated genes tend to be expressed in fewer conditions as compared with nonduplicated genes (fig. S9 and table S8), suggesting that they may be undergoing subfunctionalization.

QTL mining

As exemplified in rice and other crops, a reference genome sequence provides a resource for gene discovery, marker development, and allele mining in support of crop improvement (*49*). We identified 153,190 insertion site–based polymorphism (ISBP) markers (*50*) and 35,579 microsatellite markers along the 3B chromosome. We also located 121 quantitative trait loci (QTLs) for 50 different traits on chromosome 3B (table S13). Using these data, we conducted a meta-analysis that integrates QTLs defined in independent studies (*51*) and identified 18 metaQTLs with confidence intervals covering between 1.5 and 620 Mb of the chromosome 3B sequence. The largest one encompasses the centromeric region, where recombination is suppressed. Five metaQTLs with small intervals (<10 Mb) that contain between 23 and 266 protein-coding genes and between 511 and 4049 markers are suitable for fine mapping (table S14).

Discussion

We present a reference sequence of chromosome 3B that can be used to precisely delineate structural and functional features along a chromosome and establish correlations between recombination intensity, gene density, gene expression, and evolution rate. Our results indicate that during evolution, regions with distinct features become delineated along chromosome 3B, including relatively small distal regions that are preferential targets for recombination, adaptation, and genomic plasticity. Whether our observations reflect a general pattern for the wheat genome will need to be confirmed by the analysis of other chromosome reference sequences. Already, some of the features—such as the CACTA distribution, the high rate of intrachromosomal duplication, the absence of major gene loss since polyploidization, and the gradient of gene density—have been confirmed at the whole-genome level (*19*, *39*). Moreover, the ordered chromosome 3B sequence allowed us to distinguish duplicated genes and provided evidence for superimposed mechanisms of gene duplications. The high level of gene duplication (allopolyploidy and inter- and intrachromosomal duplications) provides the wheat genome with a vast reservoir of functional genes that likely contribute to wheat adaptation.

On the basis of this work, the IWGSC has already defined an adapted BAC pooling strategy to reach the same sequence quality while reducing sequencing costs for the remaining chromosomes. Although progress in sequencing technologies and cost reduction allows for more cost-efficient sequence production, the challenge of bioinformatics and limitations of current sequencing technologies remain (*12*). Solving these issues and improving methods to efficiently anchor and orient scaffolds within pseudomolecules should make the assembly of high-quality reference sequences of complex genomes routine work in the future. There is no doubt that, as witnessed after the release of the rice genome sequence (*49*), the number of genes cloned from wheat will grow exponentially in the near future, thereby enabling wheat researchers and breeders to cope with the urgent need to improve wheat yield in the face of climate change and food-security challenges (*52*).

Materials and methods

Sequencing, assembly, and scaffolding

A total of 8452 BACs representing the MTP of wheat chromosome 3B were pooled into 922 BAC pools. Each pool was used to create a barcoded Roche/454 8-kb long paired-end library. In total, 150 sequencing runs were performed, leading to an average of 36-fold sequence coverage. After assembly with Newbler (Roche), we integrated 42,551 BAC end sequences to validate and improve scaffolding. Illumina reads generated from sorted DNA of chromosome 3B were used to fill gaps within scaffolds and correct potential sequencing errors remaining in the consensus sequence (*18*).

Anchoring scaffolds

SNP discovery was performed through sequence capture for 52,265 loci flanking TE junctions representing an average density of one locus per 16.2 kb (*18*). Out of 39,077 SNPs distributed along the chromosome, a subset of 3075 evenly distributed (38.2 ± 9.4 SNPs per 10 Mb) SNPs was selected to genotype 1025 lines from recombinant inbred and association panels. An anchor genetic map was built first by linkage analysis and integration of linkage disequilibrium data. A consensus map comprising 5318 markers was also built using 40 different genetic maps to anchor additional scaffolds (*18*). Finally, a position in the pseudomolecule was inferred for scaffolds without marker information but belonging to an anchored physical BAC contig.

Sequence annotation

Gene modeling was performed using an improved version of the TriAnnot pipeline (*53*). Noncoding RNA genes were predicted using three different programs (*18*), and predictions were manually curated. Predictions of TEs and reconstruction of the pattern of nested insertions were performed through the development of a specific program (*18*) that automatically curates similarity-search results obtained with a dedicated databank comprising 4929 known wheat TEs classified into 521 families.

Gene expression analyses

Thirty RNA samples, corresponding to RNAs extracted in duplicates from five organs (root, leaf, stem, spike, and grain) at three developmental stages each from hexaploid wheat cultivar Chinese Spring (*4*), were used for gene expression analyses. RNA-Seq libraries were constructed using the IlluminaTruSeq (Illumina, CA, USA) RNA sample preparation kit and sequenced. An average of 50 ± 11 million paired-end reads per sample were mapped on the chromosome 3B scaffolds and used to reconstruct transcripts and estimate transcript abundance in units of fragments per kb of exon per million mapped reads (FPKM). Regions with FPKM values higher than zero were considered as expressed.

Distribution and segmentation analyses

Distributions of recombination rate, gene and TE densities, and expression breadth were calculated within a sliding window of 10 Mb (and 1 Mb for the recombination rate), with a step of 1 Mb along the chromosome sequence using a homemade Perl script. Segmentation analyses of these distributions were performed using the R package changepoint v1.0.6 (*54*), with Segment Neighborhoods method and Bayesian information criterion penalty on the mean change.

Comparative genomics, gene duplications, and molecular evolution

We performed an all-by-all Basic Local Alignment Search Tool for Proteins (BLASTP) [cutoff expected value (Evalue), 1×10^{-5}] comparison between the amino-acid sequences of predicted genes of wheat chromosome 3B, rice (Michigan State University version 7.0), *Brachypodium* (*Brachypodium* Sequencing Initiative, 2.0), and sorghum (phytozome, version 1.4). We filtered out genes with no homology with at least one other gene in a compared species (cutoff 35% amino acid identity and 35% sequence overlap). Syntenic genes were defined as genes with a best BLAST hit on an orthologous chromosome in at least one other species. Nonsyntenic genes were defined as genes for which the best BLAST hit was on a nonorthologous chromosome in the other species. Clustering of orthologous and paralogous genes was performed using OrthoMCL [Evalue cutoff, 1×10^{-5}; percentage match cutoff, 35% (*55*)]. All 3B genes clustered into the same family were considered intrachromosomal duplicates. 3B genes clustered in a family with wheat gene models annotated on another chromosome (*19*), not including genes from group 3, were considered as interchromosomal duplicates. Tandem duplicates were defined as genes in the same family with five or fewer spacer genes separating them on the pseudomolecule, and dispersed duplicates were defined as having more than five spacer genes. Synonymous (K_s) and nonsynonymous (K_a) substitution rates were calculated based on ClustalW 2.1 (*56*) coding sequence alignments by the Nei and Gojobori method

using codeml [part of the PAML package (*57*)]. Age of gene divergence was estimated by the equation $K_s/2r$, where $r = 6.5 \times 10^{-9}$.

REFERENCES AND NOTES

1. J. Dubcovsky, J. Dvorak, Genome plasticity a key factor in the success of polyploid wheat under domestication. *Science* **316**, 1862–1866 (2007). doi: 10.1126/science.1143986; pmid: 17600208
2. F. Choulet *et al*., Megabase level sequencing reveals contrasted organization and evolution patterns of the wheat gene and transposable element spaces. *Plant Cell* **22**, 1686–1701 (2010). doi: 10.1105/tpc.110.074187; pmid: 20581307
3. J. Zhang, Evolution by gene duplication: An update. *Trends Ecol. Evol.* **18**, 292–298 (2003). doi: 10.1016/S0169-5347(03)00033-8
4. C. Rustenholz *et al*., A 3,000-loci transcription map of chromosome 3B unravels the structural and functional features of gene islands in hexaploid wheat. *Plant Physiol.* **157**, 1596–1608 (2011). doi: 10.1104/pp.111.183921; pmid: 22034626
5. I. D. Wilson *et al*., A transcriptomics resource for wheat functional genomics. *Plant Biotechnol. J.* **2**, 495–506 (2004). doi: 10.1111/j.1467-7652.2004.00096.x; pmid: 17147622
6. P. R. Bhat *et al*., Mapping translocation breakpoints using a wheat microarray. *Nucleic Acids Res.* **35**, 2936–2943 (2007). doi: 10.1093/nar/gkm148; pmid: 17439961
7. L. L. Qi *et al*., A chromosome bin map of 16,000 expressed sequence tag loci and distribution of genes among the three genomes of polyploid wheat. *Genetics* **168**, 701–712 (2004). doi: 10.1534/genetics.104.034868; pmid: 15514046
8. R. Brenchley *et al*., Analysis of the bread wheat genome using whole-genome shotgun sequencing. *Nature* **491**, 705–710 (2012). doi: 10.1038/nature11650; pmid: 23192148
9. J. Jia *et al*., Aegilops tauschii draft genome sequence reveals a gene repertoire for wheat adaptation. *Nature* **496**, 91–95 (2013). doi: 10.1038/nature12028; pmid: 23535592
10. H. Q. Ling *et al*., Draft genome of the wheat A-genome progenitor Triticum urartu. *Nature* **496**, 87–90 (2013). doi: 10.1038/nature11997; pmid: 23535596
11. M. C. Schatz, A. L. Delcher, S. L. Salzberg, Assembly of large genomes using second-generation sequencing. *Genome Res.* **20**, 1165–1173 (2010). doi: 10.1101/gr.101360.109; pmid: 20508146
12. V. Marx, Next-generation sequencing: The genome jigsaw. *Nature* **501**, 263–268 (2013). doi: 10.1038/501261a; pmid: 24025842
13. P. S. Schnable *et al*., The B73 maize genome: Complexity, diversity, and dynamics. *Science* **326**, 1112–1115 (2009). doi: 10.1126/science.1178534; pmid: 19965430
14. X. Xu *et al*., Genome sequence and analysis of the tuber crop potato. *Nature* **475**, 189–195 (2011). doi: 10.1038/nature10158; pmid: 21743474
15. J. Doležel, M. Kubaláková, E. Paux, J. Bartos, C. Feuillet, Chromosome-based genomics in the cereals. *Chromosome Res.* **15**, 51–66 (2007). doi: 10.1007/s10577-006-1106-x; pmid: 17295126
16. J. Safár *et al*., Dissecting large and complex genomes: Flow sorting and BAC cloning of individual chromosomes from bread wheat. *Plant J.* **39**, 960–968 (2004). doi: 10.1111/j.1365-313X.2004.02179.x; pmid: 15341637
17. E. Paux *et al*., A physical map of the 1-gigabase bread wheat chromosome 3B. *Science* **322**, 101–104 (2008). doi: 10.1126/science.1161847; pmid: 18832645
18. Supplementary materials are available on *Science* Online.
19. International Wheat Genome Sequencing Consortium, A chromosome-based draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome. *Science* **345**, 1251788 (2014).
20. B. S. Gill, B. Friebe, T. Endo, Standard karyotype and nomenclature system for description of chromosome bands and structural aberrations in wheat (Triticum aestivum). *Genome* **34**, 830–839 (1991). doi: 10.1139/g91-128
21. P. S. Chain *et al*., Genomics. Genome project standards in a new era of sequencing. *Science* **326**, 236–237 (2009). doi: 10.1126/science.1180614; pmid: 19815760
22. B. Li *et al*., Wheat centromeric retrotransposons: The new ones take a major role in centromeric structure. *Plant J.* **73**, 952–965 (2013). doi: 10.1111/tpj.12086; pmid: 23253213
23. H. Yan *et al*., Intergenic locations of rice centromeric chromatin. *PLOS Biol.* **6**, e286 (2008). doi: 10.1371/journal.pbio.0060286; pmid: 19067486
24. H. Zhang, R. K. Dawe, Total centromere size and genome size are strongly correlated in ten grass species. *Chromosome Res.* **20**, 403–412 (2012). doi: 10.1007/s10577-012-9284-1; pmid: 22552915
25. T. Sakuno, K. Tada, Y. Watanabe, Kinetochore geometry defined by cohesion within the centromere. *Nature* **458**, 852–858 (2009). doi: 10.1038/nature07876; pmid: 19370027
26. Q. Pan, F. Ali, X. Yang, J. Li, J. Yan, Exploring the genetic characteristics of two recombinant inbred line populations via high-density SNP markers in maize. *PLOS ONE* **7**, e52777 (2012). doi: 10.1371/journal.pone.0052777; pmid: 23300772
27. A. J. Lukaszewski, C. A. Curtis, Physical distribution of recombination in B-genome chromosomes of tetraploid wheat. *Theor. Appl. Genet.* **86**, 121–127 (1993). doi: 10.1007/BF00223816; pmid: 24193391
28. C. Saintenac *et al*., Detailed recombination studies along chromosome 3B provide new insights on crossover distribution in wheat (*Triticum aestivum* L.). *Genetics* **181**, 393–403 (2009). doi: 10.1534/genetics.108.097469; pmid: 19064706
29. J. Evans *et al*., Extensive variation in the density and distribution of DNA polymorphism in sorghum genomes. *PLOS ONE* **8**, e79192 (2013). doi: 10.1371/journal.pone.0079192; pmid: 24265758
30. International Rice Genome Sequencing Project, The map-based sequence of the rice genome. *Nature* **436**, 793–800 (2005). doi: 10.1038/nature03895; pmid: 16100779
31. A. H. Paterson *et al*., The Sorghum bicolor genome and the diversification of grasses. *Nature* **457**, 551–556 (2009). doi: 10.1038/nature07723; pmid: 19189423
32. A. Gottlieb *et al*., Insular organization of gene space in grass genomes. *PLOS ONE* **8**, e54101 (2013). doi: 10.1371/journal.pone.0054101; pmid: 23326580
33. M. W. Ganal *et al*., A large maize (Zea mays L.) SNP genotyping array: Development and germplasm genotyping, and genetic mapping to compare with the B73 reference genome. *PLOS ONE* **6**, e28334 (2011). doi: 10.1371/journal.pone.0028334; pmid: 22174790
34. K. F. Mayer *et al*., A physical, genetic and functional sequence assembly of the barley genome. *Nature* **491**, 711–716 (2012). pmid: 23075845
35. R. S. Sekhon *et al*., Maize gene atlas developed by RNA sequencing and comparative evaluation of transcriptomes based on RNA sequencing and microarrays. *PLOS ONE* **8**, e61005 (2013). doi: 10.1371/journal.pone.0061005; pmid: 23637782
36. R. S. Baucom *et al*., Exceptional diversity, non-random distribution, and rapid evolution of retroelements in the B73 maize genome. *PLOS Genet.* **5**, e1000732 (2009). doi: 10.1371/journal.pgen.1000732; pmid: 19936065
37. R. S. Baucom, J. C. Estill, J. Leebens-Mack, J. L. Bennetzen, Natural selection on gene function drives the evolution of LTR retrotransposon families in the rice genome. *Genome Res.* **19**, 243–254 (2009). doi: 10.1101/gr.083360.108; pmid: 19029538
38. M. Charles *et al*., Dynamics and differential proliferation of transposable elements during the evolution of the B and A genomes of wheat. *Genetics* **180**, 1071–1086 (2008). doi: 10.1534/genetics.108.092304; pmid: 18780739
39. E. M. Sergeeva, E. A. Salina, I. G. Adonina, B. Chalhoub, Evolutionary analysis of the CACTA DNA-transposon Caspar across wheat species using sequence comparison and in situ hybridization. *Mol. Genet. Genomics* **284**, 11–23 (2010). doi: 10.1007/s00438-010-0544-5; pmid: 20512353
40. C. Lu *et al*., Miniature inverted-repeat transposable elements (MITEs) have been accumulated through amplification bursts and play important roles in gene expression and species diversity in Oryza sativa. *Mol. Biol. Evol.* **29**, 1005–1017 (2012). doi: 10.1093/molbev/msr282; pmid: 22096216
41. M. D. Gale, K. M. Devos, Comparative genetics in the grasses. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 1971–1974 (1998). doi: 10.1073/pnas.95.5.1971; pmid: 9482816
42. K. M. Devos, M. D. Gale, Genome relationships: The grass model in current research. *Plant Cell* **12**, 637–646 (2000). doi: 10.1105/tpc.12.5.637; pmid: 10810140
43. F. Murat *et al*., Ancestral grass karyotype reconstruction unravels new mechanisms of genome shuffling as a source of plant evolution. *Genome Res.* **20**, 1545–1557 (2010). doi: 10.1101/gr.109744.110; pmid: 20876790
44. International Brachypodium Initiative, Genome sequencing and analysis of the model grass Brachypodium distachyon. *Nature* **463**, 763–768 (2010). doi: 10.1038/nature08747; pmid: 20148030
45. E. D. Akhunov *et al*., Synteny perturbations between wheat homoeologous chromosomes caused by locus duplications and deletions correlate with recombination rates. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 10836–10841 (2003). doi: 10.1073/pnas.1934431100; pmid: 12960374
46. T. Wicker *et al*., Frequent gene movement and pseudogene evolution is common to the large and complex genomes of wheat, barley, and their relatives. *Plant Cell* **23**, 1706–1718 (2011). doi: 10.1105/tpc.111.086629; pmid: 21622801
47. T. Wicker, J. P. Buchmann, B. Keller, Patching gaps in plant genomes results in gene movement and erosion of colinearity. *Genome Res.* **20**, 1229–1237 (2010). doi: 10.1101/gr.107284.110; pmid: 20530251
48. M. Morgante *et al*., Gene duplication and exon shuffling by helitron-like transposons generate intraspecies diversity in maize. *Nat. Genet.* **37**, 997–1002 (2005). doi: 10.1038/ng1615; pmid: 16056225
49. C. Feuillet, J. E. Leach, J. Rogers, P. S. Schnable, K. Eversole, Crop genome sequencing: Lessons and rationales. *Trends Plant Sci.* **16**, 77–88 (2011). doi: 10.1016/j.tplants.2010.10.005; pmid: 21081278
50. E. Paux *et al*., Insertion site-based polymorphism markers open new perspectives for genome saturation and marker-assisted selection in wheat. *Plant Biotechnol. J.* **8**, 196–210 (2010). doi: 10.1111/j.1467-7652.2009.00477.x; pmid: 20078842
51. B. Goffinet, S. Gerber, Quantitative trait loci: A meta-analysis. *Genetics* **155**, 463–473 (2000). pmid: 10790417
52. J. A. Foley *et al*., Solutions for a cultivated planet. *Nature* **478**, 337–342 (2011). doi: 10.1038/nature10452; pmid: 21993620
53. P. Leroy *et al*., TriAnnot: A versatile and high performance pipeline for the automated annotation of plant genomes. *Front. Plant Sci.* **3**, 5 (2012). doi: 10.3389/fpls.2012.00005; pmid: 22645565
54. J. Chen, A. K. Gupta, *Parametric Statistical Change Point Analysis: With Applications to Genetics, Medicine, and Finance* (Birkhäuser, Basel, 2012).
55. L. Li, C. J. Stoeckert Jr., D. S. Roos, OrthoMCL: Identification of ortholog groups for eukaryotic genomes. *Genome Res.* **13**, 2178–2189 (2003). doi: 10.1101/gr.1224503; pmid: 12952885
56. J. D. Thompson, D. G. Higgins, T. J. Gibson, CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673–4680 (1994). doi: 10.1093/nar/22.22.4673; pmid: 7984417
57. Z. Yang, PAML: A program package for phylogenetic analysis by maximum likelihood. *CABIOS* **13**, 555–556 (1997). pmid: 9367129

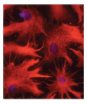
ACKNOWLEDGMENTS

The authors thank the scientific advisory board (P. Schnable, S. Rounsley, D. Ware, J. Rogers, and K. Eversole) of the 3BSEQ project for fruitful discussions; K. Eversole for critical reading and editing of the manuscript; H. Rimbert, N. Cubizolles, and E. Rey for SNP marker discovery and genotyping; M. Kubaláková and J. Vrána for assistance with the preparation of DNA amplified from flow-sorted chromosome 3B; L. Couderc, A. Keliet, and S. Reboux for their support in database and system administration; and C. Poncet and the "Plateforme GENTYANE" for SNP genotyping. This work was supported by a grant from the French National Research Agency (ANR-09-GENM-025 3BSEQ), a grant from France Agrimer, and a grant (project DL-BLE) from the INRA BAP Biologie et Amélioration des Plantes division. N.G. is funded by a grant from the European Commission research training program Marie-Curie Actions (FP7-MC-IIF-NoncollinearGenes). J. Daron is funded by a grant from the French Ministry of Research. L.P. is funded by a grant from the Region Auvergne. K.V. is supported by the Ghent University Multidisciplinary Research Partnership ["Bioinformatics: From nucleotides to networks" (Project 01MR0310W)]. J. Dolezel is supported by the Czech Science Foundation (award no. P501/12/G090). The chromosome 3B BAC library and the pools of the MTP BAC clones are available upon request under a materials transfer agreement at the French Plant Genomic Center, INRA-Centre National de Ressources Génomiques Végétales. Annotation data and browser are available at https://urgi.versailles.inra.fr/gb2/gbrowse/wheat_annot_3B. Sequences and annotations of the reference pseudomolecule and unassigned scaffolds have been deposited in ENA (project PRJEB4376) under accession numbers HG670306 and CBUC010000001 to CBUC010001450, respectively. RNA-Seq data were deposited under accession number ERP004714.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/345/6194/1249721/suppl/DC1
Materials and Methods
Figs. S1 to S11
Tables S1 to S14
References (*58*–*125*)

13 December 2013; accepted 30 May 2014
10.1126/science.1249721

EXTENDED PDF FORMAT
SPONSORED BY

Travel Grants Available
Neuroscience 2014
Application Deadline: July 31st
>Apply Now
R&D systems
a bietechne brand
www.rndsystems.com

Genome interplay in the grain transcriptome of hexaploid bread wheat

Matthias Pfeifer *et al.*

Science **345**, (2014);
DOI: 10.1126/science.1250091

This copy is for your personal, non-commercial use only.

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by clicking here.

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines here.

The following resources related to this article are available online at www.sciencemag.org (this information is current as of August 4, 2014):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:
http://www.sciencemag.org/content/345/6194/1250091.full.html

Supporting Online Material can be found at:
http://www.sciencemag.org/content/suppl/2014/07/16/345.6194.1250091.DC1.html

A list of selected additional articles on the Science Web sites **related to this article** can be found at:
http://www.sciencemag.org/content/345/6194/1250091.full.html#related

This article **cites 67 articles**, 28 of which can be accessed free:
http://www.sciencemag.org/content/345/6194/1250091.full.html#ref-list-1

This article has been **cited by** 2 articles hosted by HighWire Press; see:
http://www.sciencemag.org/content/345/6194/1250091.full.html#related-urls

This article appears in the following **subject collections:**
Botany
http://www.sciencemag.org/cgi/collection/botany
Genetics
http://www.sciencemag.org/cgi/collection/genetics

Downloaded from www.sciencemag.org on August 4, 2014

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published weekly, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. Copyright 2014 by the American Association for the Advancement of Science; all rights reserved. The title *Science* is a registered trademark of AAAS.

A chromosome-based draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome

The International Wheat Genome Sequencing Consortium (IWGSC)

An ordered draft sequence of the 17-gigabase hexaploid bread wheat (*Triticum aestivum*) genome has been produced by sequencing isolated chromosome arms. We have annotated 124,201 gene loci distributed nearly evenly across the homeologous chromosomes and subgenomes. Comparative gene analysis of wheat subgenomes and extant diploid and tetraploid wheat relatives showed that high sequence similarity and structural conservation are retained, with limited gene loss, after polyploidization. However, across the genomes there was evidence of dynamic gene gain, loss, and duplication since the divergence of the wheat lineages. A high degree of transcriptional autonomy and no global dominance was found for the subgenomes. These insights into the genome biology of a polyploid crop provide a springboard for faster gene isolation, rapid genetic marker development, and precise breeding to meet the needs of increasing food demand worldwide.

Lists of authors and affiliations are available in the full article online.
Corresponding author: K. X. Mayer, e-mail: k.mayer@helmholtz-muenchen.de
Read the full article at http://dx.doi.org/10.1126/science.1251788

Ancient hybridizations among the ancestral genomes of bread wheat

Thomas Marcussen, Simen R. Sandve,* Lise Heier, Manuel Spannagl, Matthias Pfeifer, The International Wheat Genome Sequencing Consortium,† Kjetill S. Jakobsen, Brande B. H Wulff, Burkhard Steuernagel, Klaus F. X. Mayer, Odd-Arne Olsen

The allohexaploid bread wheat genome consists of three closely related subgenomes (A, B, and D), but a clear understanding of their phylogenetic history has been lacking. We used genome assemblies of bread wheat and five diploid relatives to analyze genome-wide samples of gene trees, as well as to estimate evolutionary relatedness and divergence times. We show that the A and B genomes diverged from a common ancestor ~7 million years ago and that these genomes gave rise to the D genome through homoploid hybrid speciation 1 to 2 million years later. Our findings imply that the present-day bread wheat genome is a product of multiple rounds of hybrid speciation (homoploid and polyploid) and lay the foundation for a new framework for understanding the wheat genome as a multilevel phylogenetic mosaic.

The list of author affiliations is available in the full article online.***Corresponding author.**
E-mail: simen.sandve@nmbu.no †The International Wheat Genome Sequencing Consortium (IWGSC) authors and affiliations are listed in the supplementary materials.
Read the full article at http://dx.doi.org/10.1126/science.1250092

Triticum monococcum

Triticum carthlicum

Ancestral wheat

Wheat varieties and species (shown) believed to be the closest living relatives of modern bread wheat (*T. aestivum*). Multiple ancestral hybridizations occurred among most of these species, many of which are cultivated, and along with *T. aestivum* represent a dominant source of global nutrition.

Triticum boeticum

Triticum polonicum L.

Triticum macha

Triticum dicoccoides var. *araraticum*

Published by AAAS

PHOTOS: SUSANNE STAMP, ERNST MERZ/ETH ZURICH

Genome interplay in the grain transcriptome of hexaploid bread wheat

Matthias Pfeifer, Karl G. Kugler, Simen R. Sandve, Bujie Zhan, Heidi Rudi, Torgeir R. Hvidsten, International Wheat Genome Sequencing Consortium,* Klaus F. X. Mayer, Odd-Arne Olsen†

Allohexaploid bread wheat (*Triticum aestivum* L.) provides approximately 20% of calories consumed by humans. Lack of genome sequence for the three homeologous and highly similar bread wheat genomes (A, B, and D) has impeded expression analysis of the grain transcriptome. We used previously unknown genome information to analyze the cell type–specific expression of homeologous genes in the developing wheat grain and identified distinct co-expression clusters reflecting the spatiotemporal progression during endosperm development. We observed no global but cell type– and stage-dependent genome dominance, organization of the wheat genome into transcriptionally active chromosomal regions, and asymmetric expression in gene families related to baking quality. Our findings give insight into the transcriptional dynamics and genome interplay among individual grain cell types in a polyploid cereal genome.

The list of author affiliations is available in the full article online. *The International Wheat Genome Sequencing Consortium (IWGSC) authors and affiliations are listed in the supplementary materials. †**Corresponding author. E-mail: odd-arne.olsen@nmbu.no**

Read the full article at http://dx.doi.org/10.1126/science.1250091

Structural and functional partitioning of bread wheat chromosome 3B

Frédéric Choulet,* Adriana Alberti, Sébastien Theil, Natasha Glover, Valérie Barbe, Josquin Daron, Lise Pingault, Pierre Sourdille, Arnaud Couloux, Etienne Paux, Philippe Leroy, Sophie Mangenot, Nicolas Guilhot, Jacques Le Gouis, Francois Balfourier, Michael Alaux, Véronique Jamilloux, Julie Poulain, Céline Durand, Arnaud Bellec, Christine Gaspin, Jan Safar, Jaroslav Dolezel, Jane Rogers, Klaas Vandepoele, Jean-Marc Aury, Klaus Mayer, Hélène Berges, Hadi Quesneville, Patrick Wincker, Catherine Feuillet

We produced a reference sequence of the 1-gigabase chromosome 3B of hexaploid bread wheat. By sequencing 8452 bacterial artificial chromosomes in pools, we assembled a sequence of 774 megabases carrying 5326 protein-coding genes, 1938 pseudogenes, and 85% of transposable elements. The distribution of structural and functional features along the chromosome revealed partitioning correlated with meiotic recombination. Comparative analyses indicated high wheat-specific inter- and intrachromosomal gene duplication activities that are potential sources of variability for adaption. In addition to providing a better understanding of the organization, function, and evolution of a large and polyploid genome, the availability of a high-quality sequence anchored to genetic maps will accelerate the identification of genes underlying important agronomic traits.

The list of author affiliations is available in the full article online.
***Corresponding author. E-mail: frederic.choulet@clermont.inra.fr**

Read the full article at http://dx.doi.org/10.1126/science.1249721

Published by AAAS

WHEAT GENOME

Genome interplay in the grain transcriptome of hexaploid bread wheat

Matthias Pfeifer,[1]* Karl G. Kugler,[1]* Simen R. Sandve,[2] Bujie Zhan,[2] Heidi Rudi,[2] Torgeir R. Hvidsten,[3] International Wheat Genome Sequencing Consortium† Klaus F. X. Mayer,[1] Odd-Arne Olsen[2]‡

Allohexaploid bread wheat (*Triticum aestivum* L.) provides approximately 20% of calories consumed by humans. Lack of genome sequence for the three homeologous and highly similar bread wheat genomes (A, B, and D) has impeded expression analysis of the grain transcriptome. We used previously unknown genome information to analyze the cell type–specific expression of homeologous genes in the developing wheat grain and identified distinct co-expression clusters reflecting the spatiotemporal progression during endosperm development. We observed no global but cell type– and stage-dependent genome dominance, organization of the wheat genome into transcriptionally active chromosomal regions, and asymmetric expression in gene families related to baking quality. Our findings give insight into the transcriptional dynamics and genome interplay among individual grain cell types in a polyploid cereal genome.

The vast majority (90 to 95%) of wheat produced in the world is the allohexaploid bread wheat (*Triticum aestivum* L., $2n = 6x = 42$, AABBDD) accounting for ~20% of calories consumed by humans worldwide. Because of global economic and social trends, wheat-processing industries put great emphasis on enhancing specific grain-quality attributes. However, an allopolyploid genome composition requires regulatory mechanisms within the cell that can orchestrate the complex intergenomic gene expression through genetic or epigenetic modifications. This may result in genomic asymmetry and favored expression of individual genes from specific, single genomes [reviewed in (*1*)]. For allopolyploid wheat, partial or complete genome dominance has been indicated as affecting traits such as grain size and form. However, the extent and characteristics of gene expression divergence between genomes in different tissues at the whole-genome level is largely unknown. The lack of a genome sequence for bread wheat that could enable the measurement of A, B, and D genome-specific transcription has previously hindered examination (*2*) of the genetic control of grain components (*3*, *4*).

The global landscape of endosperm gene expression

We surveyed the genome-wide patterns of cell-type–specific gene expression, providing detailed insights into the contributions of the A, B, and D genomes to the grain transcriptome. The bread wheat endosperm is composed of three main cell types (Fig. 1A) (*5*). The starchy endosperm (SE) accumulates starch and storage proteins and forms the bulk of the grain. The aleurone layer (AL) is a single cell layer surrounding the SE, except on the ventral crease side, where transfer cells (TCs) develop. Aleurone cells accumulate lipids and play an essential role in grain germination, whereas the TCs actively transport sucrose from the photosynthetic tissues to the endosperm and embryo. We applied RNA sequencing (RNA-seq) in order to monitor gene expression in different cell types at three different developmental stages that reflect the progression of starch and storage protein accumulation [10, 20, or 30 days post anthesis (DPA)] (Fig. 1A). The International Wheat Genome Sequencing Consortium (IWGSC) bread wheat genome survey sequence and annotation (*6*) were used as a reference, enabling assignment of the endosperm transcriptome to the three homeologous genomes A, B, and D and allowing the identification of previously unidentified transcripts and isoforms (*7*).

In total, 46,487 out of 85,173 high-confidence (HC) wheat genes [55%; HC1 to HC3 (*6*)] were expressed during endosperm development (Fig. 1B), which is consistent with observations in *Arabidopsis thaliana* and barley (*8*, *9*). The number of genes from any genome preferentially expressed in individual cell types and developmental stages [up-regulated compared with the other samples (*7*)] varied substantially from 136 in TCs 20 DPA to 644 in AL cells 20 DPA (Fig. 1C). Genes that were preferentially expressed in endosperm at 10 DPA were enriched for gene ontology (GO) (*10*) terms related to carbohydrate metabolic processes and glycolysis (*4*, *9*); AL-specific genes were enriched for lipid metabolism, structural development, carbohydrate metabolic processes, and amino acid biosynthesis (*4*); SE-specific genes were enriched for carbohydrate and saccharide metabolism (*11*); and TC-specific genes were enriched for proteolysis and defense-response genes (*12*). Low numbers of preferentially expressed genes in grain development have also been observed in *A. thaliana* (*8*). This suggests that there may be conserved regulatory principles across >100 million years of angiosperm evolution. Moreover, the three wheat genomes contribute about equally to the number of expressed genes in the endosperm as a whole (18 to 19%) (Fig. 1B) in each of the individual cell types and developmental stages (Fig. 1C), as well as to the number of preferentially expressed genes in each cell type and stage (Fig. 1C).

Morphologically, as well as functionally, aleurone and SE cells are widely different in that aleurone cells accumulate lipid bodies and are desiccation-tolerant, whereas SE cells accumulate starch granules and storage proteins and die toward grain maturation (*13*). The difference between aleurone cells and SE cells was also apparent from our whole-genome transcriptome analysis, in which aleurone cells exhibit a separate expression cluster from all other samples (Fig. 1D). In addition, gene expression profiles related to the developmental stages were more similar than that of cell type–expression patterns; samples from the 20-DPA stage cluster together (SE, TC, and W at 20 DPA), whereas the clean SE sample from 30 DPA (SE at 30 DPA) and endosperm from 10 DPA (W at 10 DPA) form separate clusters. Although TCs and SE cells are highly different in that TCs have extended cell walls to facilitate sucrose transport, the TC sample clustered with the SE samples, which most likely is due to tightly attached SE cells. Taken together, these observations fit with characterized endosperm cell function and the progression of endosperm development toward grain maturation (*5*, *13*, *14*).

Spatiotemporal gene expression control

Endosperm development progresses through four phases: the syncytial phase, the cellularization phase, the differentiation phase, and the maturation phase (*13*). We identified one group of genes with spatiotemporal unspecific gene expression (cluster 0) and seven clusters of co-expressed genes (clusters I to VII) (Fig. 2). The latter reflected phases of endosperm differentiation and maturation, in which the industrially important characteristics of the wheat grain are developed. They contained from 2257 to 5369 genes (24,826 in total) without evident genome dominance. Each cluster showed clear and distinct characteristics in terms of gene expression profiles, enrichments of preferentially expressed genes, and gene function enrichment (*7*). For example, cluster I exhibited increased expression in 10 DPA whole endosperm (W) when cell divisions were still occurring in the periphery of the endosperm and the transcription of storage proteins and the accumulation of starch had been

[1]Plant Genome and Systems Biology, Helmholtz Center Munich, 85764 Neuherberg, Germany. [2]Department of Plant Sciences/Centre for Integrative Genetics, The Norwegian University of Life Sciences (NMBU), 1432 Ås, Norway. [3]Department of Chemistry, Biotechnology and Food Science, NMBU, 1432 Ås, Norway.
*These authors contributed equally to this work. †International Wheat Genome Sequencing Consortium (IWGSC) authors and affiliations are listed in the supplementary materials.
‡Corresponding author. E-mail: odd-arne.olsen@nmbu.no

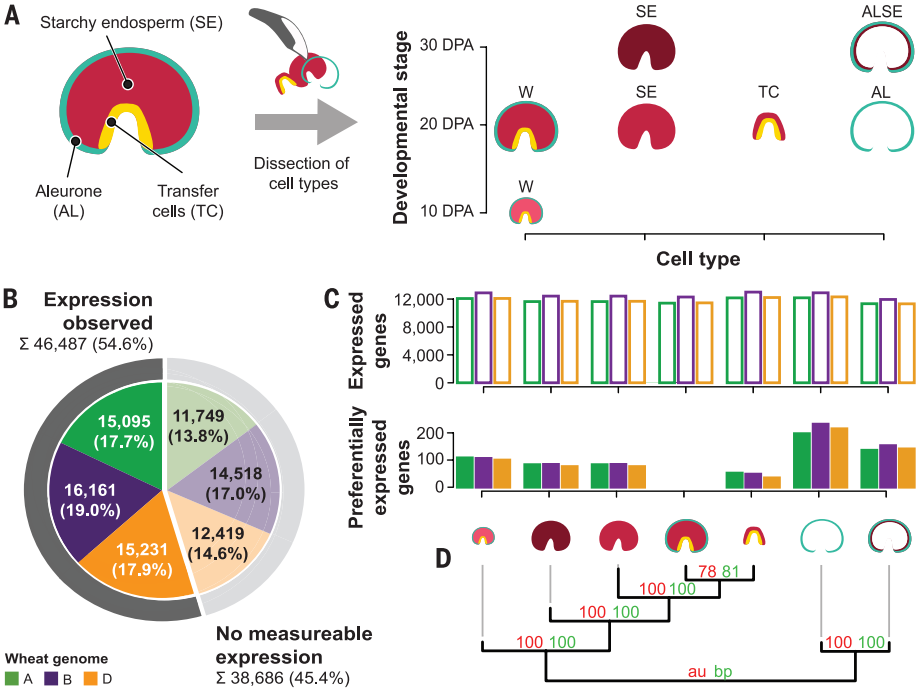


Fig. 1. Overview of the wheat endosperm transcriptome. (**A**) At 10 DPA, we sampled the whole endosperm (W). At 20 DPA, we also separated the SE, the AL, and the TC layer by means of manual dissection. At 30 DPA, the AL adheres tightly to the outermost SE cells, causing aleurone tissue to contain contamination from SE cells (ALSE). (**B**) The number of wheat genes expressed (dark gray outside line) and not expressed (lighter gray outside line) in endosperm (all cell types and developmental stages), distributed across the three genomes (green, A; purple, B; orange, D). (**C**) The number of genes expressed (open bars) and preferentially expressed (solid bars) in each cell type and developmental stage. Genes are considered preferentially expressed in a sample if they are up-regulated in that sample as compared with the other samples. (**D**) Hierarchical clustering of gene expression similarities of the samples across all genes. Bootstrap values indicate the strength of the clustering [red, approximately unbiased (au) *P* value; green, bootstrap probability (bp) *P* value].

initiated. Cluster II was expressed during the early differentiation phase, in which storage protein and starch accumulation reach their maximum. Cluster IV included a significant proportion of genes preferentially expressed in AL cells 20 DPA and was enriched for processes related to catalytic activity, lipid metabolic processes, and carbohydrate metabolism. The remaining clusters were characteristic for specific cell types (IV and VI, AL; III and VII, SE; and V, TCs) in the intermediate (III, IV, and V) and late phase (VI and VII).

We identified 6576 homeologous gene loci that had exactly one representative member from each genome (referred to as homeologous triplets; 6576 × 3 = 19,728 genes). Among these homeologous triplets, 5939 (88%) had at least one homeolog assigned to one of the co-expression clusters (7). We found that only 28% of the homeologous triplets (1663) had all three homeoalleles assigned to the same co-expression cluster. For 41% (2416) of the triplets, two out of three homeoalleles were assigned to the same co-expression cluster, whereas for the remaining 31% (1860), all homeoalleles fell in separate clusters. Co-expressed homeoalleles were uniformly distributed across genome pairs (A and B, 818 genes, 12%; A and D, 794 genes, 12%; B and D, 804 genes, 12%) (7). Therefore, if major expression divergence between A and B homeologs occurred in the ancestral tetraploid, our data suggests that this expression divergence has been reprogrammed in the hexaploid genome. Clusters that are spatiotemporally related (for example, early and intermediate development) often shared significant numbers of homeoalleles from the same triplets, whereas functionally different co-expression clusters only rarely did (one-sided Fisher's exact test with Bonferroni adjusted $P < 0.05$) (Fig. 2). Analyzing homeologous genes with single copies in only two of the genomes gave similar results (7). Therefore, although there is apparent expression divergence among homeologs, most of this divergence is due to nonradical alterations in spatiotemporal expression. Thus subfunctionalization, rather than neo-functionalization, appears to be occurring among the three bread wheat genomes.

Endosperm cell type function and module-associated genome dominance in the grain transcriptional network

We carried out analysis of genome-wide expression patterns of the homeologous triplets using principal component analysis and hierarchical clustering. Samples grouped according to genomes rather than cell types (7)—a finding that is consistent with observations of transcriptomes from different wheat organs (6). This result demonstrated that genome-specific gene expression dominates over tissue-specific gene expression also during endosperm development. To gain deeper insights into the transcriptional dynamics of homeologs, we inferred a co-expression network that was partitioned into 25 modules (Fig. 3A) (7). These transcriptional modules display different spatiotemporal characteristics, as revealed by integrating the information on gene expression at different developmental stages and in different cell types. We found spatiotemporal clustering of the network that separated cell types, but also a clustering of grain developmental stages (7). AL-related modules (Fig. 3A, left, turquoise nodes) formed a large cluster of genes that were expressed at 20 and 30 DPA and were enriched for functions connected to, for example, energy metabolism, vitamin biosynthesis, and hydrolase activity. SE-related modules (Fig. 3A, left, red nodes) were more scattered and could be linked to, for example, polysaccharide catabolism, the glyoxylate cycle, and autophagy. TCs (Fig. 3A, yellow nodes) formed dense, separated clusters enriched for "response to stimulus" functionality. Transcriptional modules enriched for more general functionalities (such as transport or translation) without cell type or developmental phase specificity were also found (Fig. 3A, gray nodes). In total, AL-related modules constituted more than one third of the nodes (2207), whereas the other cell types contributed to a lesser extent (SE, 658 nodes, 11%; TC, 149 nodes, 2%). The remaining nodes grouped either with the early phase of endosperm development or unspecific clusters.

We tested the transcriptional network for genome asymmetry. Different genomes dominated expression for 23 of the modules (92%), and no single genome proved to be overly dominant (Fig. 3A, right) (7). Highly connected genes in the network (hub genes) were significantly more likely to show a genome bias than were non-hub genes (one-sided Fisher's exact test; $P < 0.001$). They were significantly enriched in modules that serve as connecting layers among different regions of the network (one-sided Fisher's exact test; $P < 0.001$) (Fig. 3A, inset, highlighted red). Because hub genes display characteristic expression profiles for network modules (15), our observations suggest that these genes might play an important role in orchestrating genome-specific expression in the grain transcriptome network.

We superimposed the observed genome asymmetry to a semantic aggregation (16) of significantly overrepresented GO categories. Both general cellular functions as well as specific functions during bread wheat grain development were attributable to distinct contributions from individual genomes (Fig. 3B). These findings illustrate at least partial functional compartmentalization of the endosperm transcriptome among the three genomes.

Sequence divergence analysis comparing homeologous gene pairs corroborated the hypothesis

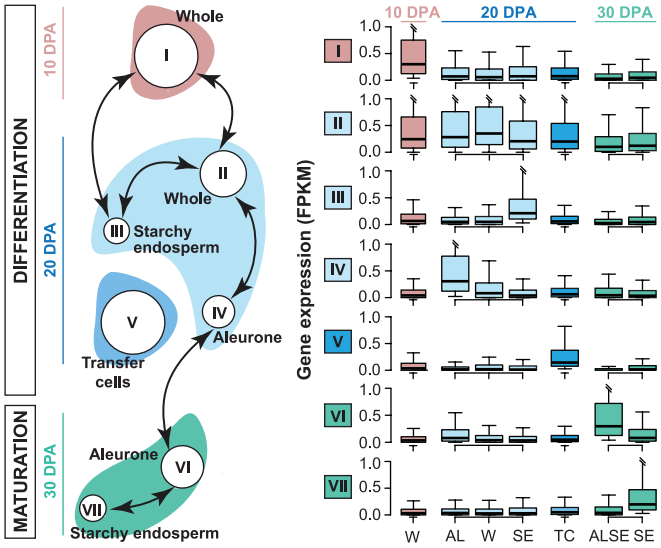


Fig. 2. Spatiotemporal control of endosperm gene expression. Seven co-expression clusters (nodes I to VII) that span the phases of progression of starch and storage protein accumulation were identified with *k*-means clustering and spatially arranged so as to reflect the phases of maturation. For homeologous triplets, spatiotemporal differences of expression patterns between individual (triplet) gene members were analyzed (*7*). Partitioning of expression for individual triplet genes was frequently observed (for example, homeologous A and B genome–encoded genes were located in cluster I, whereas the D genome copy clustered in cluster II). Bidirectional arrows connect pairs of clusters with a significant enrichment for expression partitioning among homeologous triplets. Assignment of different triplet members to different co-expression clusters is biased toward spatiotemporally related clusters (such as clusters IV and VI, late differentiation and maturation phase of aleurone cells). Gene expression profiles across developmental stages and cell types of each cluster are visualized by use of boxplots. Boxes span the data range between the first and the third quartiles, and the median is represented as horizontal lines. Whiskers extend to the most extreme data point, which is no more than 1.5 times the interquartile range away from the first and the third quartile.

of homoploid hybrid speciation of the D genome lineage (*7*, *17*). The numbers of synonymous substitutions per synonymous site between A-D and B-D homeologous genes were similar and significantly smaller than for homeologous gene pairs from the A and B genomes. Expression profiles of homeologs and expression dominance within the co-expression modules were independent of evolutionary history and relatedness (*7*). This suggests other genetic or epigenetic regulatory mechanisms as the cause of genome asymmetry (*18*–*20*).

Regulation of endosperm gene expression linked to chromosomal domains

Although the regulation of gene expression is expected to affect genes at different loci in the genome, epigenetic mechanisms often influence neighboring genes (*21*). To investigate the impact of chromosomal positioning of genes on mRNA abundance, we mapped 57,903 (67%) bread wheat genes to seven Triticeae prototype chromosomes (*7*) that reflect the inferred ancestral linear gene order in the A, B, and D genomes (*22*, *23*). Along all chromosomes, the gene expression oscillated to form chromosomal domains (*24*, *25*) with only minor divergence between the spatiotemporal endosperm conditions (Fig. 4A) (*7*). This spatial distribution of gene expression was to a large extent similar between genomes. However, we also observed chromosomal domains with asynchronous homeologous expression patterns, which indicates local genome asymmetry affecting neighboring genes (Fig. 4B). One of these domains (Fig. 4, blue rectangle) showed elevated gene expression in the D genome. We also found overrepresentation of differentially expressed homeologous triplets that were dominated by the D genome (Fisher's exact test; $P < 0.05$), which excludes gene expression divergence because of variation in local gene content between genomes. Genes located in this chromosomal domain encoded proteins involved in basal cellular processes such as DNA packaging and transport activity. The genome-wide contribution of each genome appears to be balanced, but as well as genetic effects such as variation in gene copy number, there is local genome asymmetry for neighboring genes. This suggests that epigenetic regulatory mechanisms act differently on particular corresponding domains of homeologous chromosomes.

Copy number and homeologous expression in baking-quality genes

Our data set enabled us to discriminate between transcripts from A, B, and D homeologous genes and to identify genes encoding major protein families known to affect dough quality (*7*). Genes identified include the high- and low-molecular-weight glutenin genes (HMW-Glu and LMW-Glu); the α-, γ-, and ω-gliadins (Gli) (*14*); the grain-hardness locus genes encoding puroindoline A (pin-A) and puroindoline B (pin-B) (*26*); and last, the storage protein activators (SPAs) (Fig. 5A) (*27*). For both subunit types of glutenin genes, expression levels varied substantially, reflecting remarkable differences in the contribution of individual genomes to overall gene expression. Whereas the total gene expression of the LMW-Glu was dominated by the B genome (68%), genes of the D genome accounted for two thirds of the total expression of the HMW-Glu subunit. For genes of the A genome, we measured only marginal contribution (2%), which supports the observed inactivation of Glu-A homeoalleles in hexaploid wheat (*28*, *29*). We found dominance of the D genome in the wheat hardness locus, including the pin-A and pin-B genes, which are exclusively encoded on the short arm of chromosome 5D (*26*, *30*) and were expressed at high levels, as well as the three homeologous genes of the pin-B2 locus that were expressed at substantially lower levels (*31*). For the single-copy gene in each genome encoding the SPA (*27*), B genome expression strongly dominated over the expression of the A and the D homeologous gene copies. Although the B and the D genome copies were predominantly expressed at 10 DPA, the A genome–derived SPA allele was almost uniformly expressed at 10 DPA and in SE cells at 20 DPA. For the highly repetitive α-, γ-, and ω-gliadins (*32*–*34*), expression levels varied substantially both in the number of genes and their expression levels between the A, B, and D genomes (Fig. 5B). Whereas gene copies of the B and D genome dominated total gene family expression, copies of the A genome accumulated fewer transcripts. For the α-gliadins encoded on chromosome 6S, we found several gene copies for the A and B genomes, but strikingly, the entire α-Gli locus was absent from the D genome. Most likely, this lack was caused by a previously undescribed deletion of a small segment of chromosome 6DS (comprising ~200 genes) (Fig. 5C). We found the corresponding chromosomal region to be present in *Aegilops tauschii*, which suggests a partial chromosomal deletion that likely occurred after the hybridization event that formed hexaploid bread wheat.

For the inspected gene families, we found large genome-specific variation in the number and abundance of genes, which implies the presence of genomic asymmetry and preferential expression of mostly the B and D genomes. Moreover, we observed discrepancies in genome expression dominance in the absence of homeologous counterparts (α-Gli), potentially indicating trans-regulatory mechanisms and cross-talk between genomes [discussed in (*1*)]. This comprehensive analysis of major grain-quality gene families in Chinese Spring, although not a wheat used for making bread, illustrates the utility of measuring the number and expression levels of homologous genes in order to provide a complete picture of relevant genes for a trait under study. Valuable information of associations

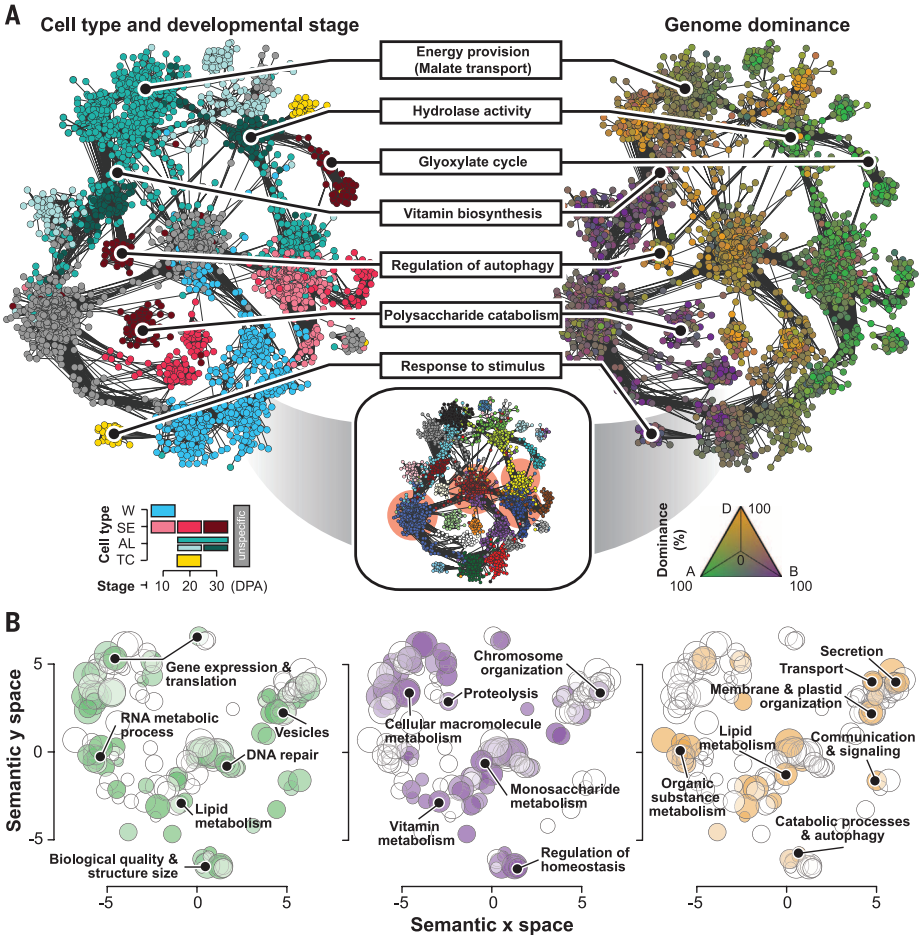


Fig. 3. Analysis of homeologous gene expression. (**A**) The dynamics of homeologous triplet gene expression patterns were captured in a co-expression network. Node coloring in the bottom central inset depicts the segmentation into 25 co-expression network modules. The inset denotes the network, with red circles around those modules that were significantly enriched for hub genes. The transcriptional network is colored for cell type and developmental stage (left) and genome dominance (right) (*7*). Selected GO categories in the center exemplify the relationship between the spatiotemporal layer and genome asymmetry (green, A; purple, B; orange, D). (**B**) Significantly enriched biological process GO categories were projected onto a two-dimensional semantic space (*7*, *16*). Semantic representation of GO categories was colored according to the significant overrepresentation in co-expression modules (green, A; purple, B; orange, D). Color intensity reflects significance of enrichment test, with dark colors corresponding to lower *P* values and white to $P > 0.05$. Circle radiuses depict the size of aggregated GO terms.

between homeologous gene levels and traits of interest for wheat breeding, including baking quality, has to be learned from crosses between appropriate genotypes. This asks for future experimentation, and our work and findings build an important foundation that allows monitoring similarities and differences from genomic, transcriptional, and network-derived angles.

Conclusions

Our findings reveal the complex interplay in gene expression regulation during grain development in the hexaploid bread wheat at several levels. Genome autonomy, asymmetric contribution of genomes to particular functions, as well as indications for regulation of chromosomal domains were found. The study forms an important basis for addressing the inter- and intragenomic regulation within a polyploid genome. It enables studying the functional output in a wide range of wheat cultivars and under different environmental regimes so as to allow the identification of the underlying genetic and epigenetic factors and their interplay. This will affect the improvement of agronomical and industrial traits of one of the world's most important crops and contribute to global food security.

Materials and methods

Plant material and transcriptome sequencing

Grains of bread wheat cultivar Chinese Spring were of the same origin as those used for generating the reference genome sequence (*6*). Plants were grown in two phytotron chambers, and grains were harvested at 10, 20, and 30 DPA and hand-dissected into individual cell types. RNA was isolated and prepared for RNA-seq by using paired-end libraries with 200–base pair inserts for 30 samples with four biological replicates per cell type and time point (two replicates per room) and two additional technical replicates. Libraries were sequenced with Illumina HiSeq. 2000 (Illumina, San Diego, California).

Gene annotation, RNA-seq mapping, and expression profiling

RNA-seq reads were aligned against the repeat-masked chromosome-sorted sequence (CSS) sequence assembly (*6*) with Bowtie2 (*35*) and Tophat2 (*36*). Returned alignments were stringently filtered so as to remove ambiguously mapped reads and read pairs with conflicting alignments. The transcriptome information was used to refine the IWGSC reference gene annotation and to identify endosperm-specific genes and alternative splicing transcripts following the procedure of (*6*) and using a reference-guided gene assembly with Cufflinks (*37*). Gene expression and tests for differentially expressed genes were computed on RNA-seq data by using CuffDiff 2 (*38*), and expression levels were $\log_2$(FPKM+1)–transformed. To exclude small gene fragments and pseudogenes, further statistical analysis was constrained on HC bread wheat genes of the class levels HC1 to HC3 (*6*). Reproducibility of expression measurements were evaluated based on Pearson's correlation of gene expression between biological and technical replicates. The accuracy of computed expression levels was elucidated by means of an in silico simulation of an RNA-seq experiment of comparable size. Hierarchical clustering of gene expression was performed based on Pearson's correlation distance and average linkage method. Preferentially expressed genes were defined by comparing the 95% confidence interval of gene expression and significance tests of differential gene expression (false discovery rate–adjusted $P < 0.05$) between groups of samples.

Identification of homeologous gene triplets and duplets

Homeologous gene triplets were defined by the protein sequence homology between the predicted protein sequences of genes from the A, B, and D genomes [Basic Local Alignment Search Tool, Protein (BLASTP), *e* value $\leq 10^{-5}$, alignment identity ≥90%], requiring strictly persistent best-bidirectional hits in all pairwise combinations. Homeologous duplets (one gene copy is absent in one genome) were determined from a global OrthoMCL (*39*) gene family clustering of individual wheat genomes and *A. thaliana* (*40*) and *Ae. tauschii* (*41*).

Identification of co-expression clusters and analysis of homeologous genes sub-functionalization

k-means clustering was performed by using Pearson's correlation distance on $\log_2$-transformed gene expression levels to group genes that showed

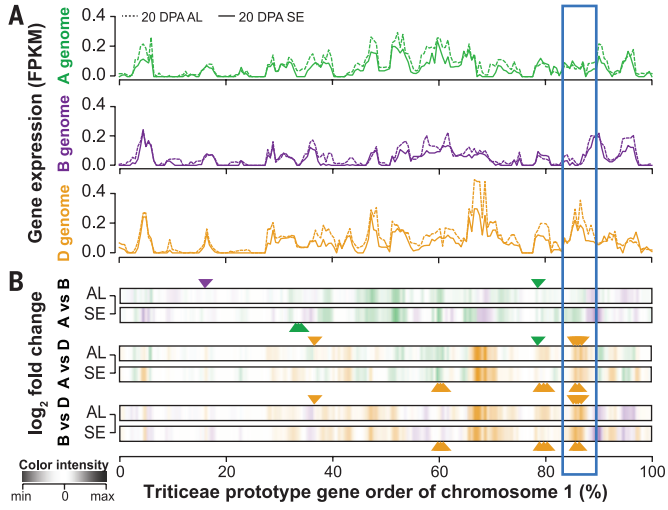


Fig. 4. Local regulatory divergence at chromosomal domains. Distribution of gene expression in 20 DPA AL and 20 DPA SE was analyzed along Triticeae prototype (Tp) chromosome 1 by using a sliding window algorithm (sliding window, size 50 Tp loci and shift 10 Tp loci). (**A**) Line charts show the median gene expression measured in aleurone and SE cells at 20 DPA. We observed a strong correlation between the expression within chromosomal domains among the three genomes. However, in several cases a single genome dominates the expression at specific domains, such as in the region indicated by the rectangle. (**B**) Heat maps display the pairwise $\log_2$-fold changes in gene expression for each window between genomes. Triangles indicate chromosomal regions that are significantly enriched for homeologous triplets up-regulated in a single genome (Fisher's exact test; $P < 0.05$). In the marked area, the D genome expression strongly dominates, indicating at a chromosomal-specific regulation of expression.

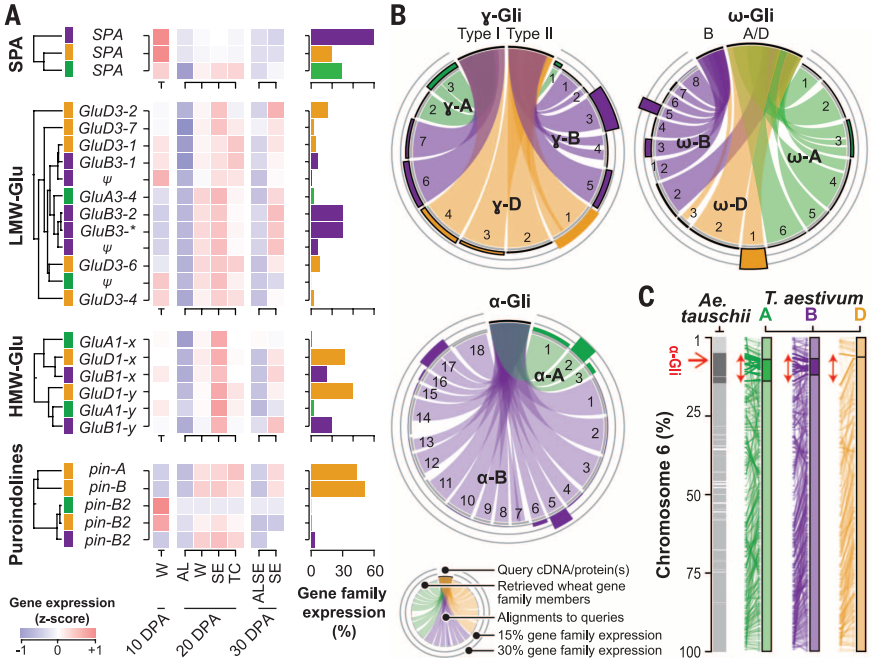


Fig. 5. Genome asymmetry in gene families affecting baking quality. Gene family composition and gene expression was analyzed for major seed storage proteins affecting baking quality of bread wheat. (**A**) Dendrograms depict phylogenetic trees for the SPA, low-molecular-weight glutenin subunit (LMW-Glu), high-molecular-weight glutenin subunit (HMW-Glu), and puroindoline (pin) A, B, and B-2 gene families. ψ indicates genes for which the predicted protein sequence is interrupted by premature stop codons, frame shifts, or repetitive elements (putative pseudogenes). Heat maps depict relative spatiotemporal gene expression, and bar charts visualize the contribution of individual genes and genomes to the total expression measured for each gene family. (**B**) The bread wheat genome and gene annotation were surveyed for α-, γ-, and ω-gliadin genes (*7*). Coverage of query gliadin gene sequences is indicated by the connectors. Bar height in the outer circle indicates the relative contribution of all samples to the overall gene expression of the respective family. (**C**) Comparative analysis between *Ae. tauschii* and bread wheat revealed a previously undescribed deletion of a small segment on chromosome 6DS in bread wheat "Chinese Spring," including the locus encoding the α-gliadins. Lines connect putative orthologous gene pairs of *Ae. tauschii* (*41*) and the three bread wheat genomes as ordered by IWGSC (*6*).

similar spatiotemporal expression profiles. Different values for k were validated, and $k = 10$ was selected as the most appropriate cluster number according to highest overall Silhouette coefficient. Thereby, all co-expression clusters with instable gene-to-cluster assignments and negative Silhouette coefficients were combined to a "zero" cluster. The distribution of homeologous genes among the co-expression clusters was evaluated and tested for a significant number of transitions of homeologs between co-expression clusters by using a one-sided Fisher's exact test (Bonferroni corrected $P < 0.05$).

Analysis of single-copy homeologous genes

For identifying homeologous genes that were significantly differentially expressed between the genomes, we performed a permutation test (1000 iterations) for each condition and genome pairing. To analyze the expression of strict single-copy homeologous genes, expression values from the A, B, and D genomes were concatenated into a triplet expression matrix. A co-expression network was then inferred by using the Weighted Correlation Network Analysis (WGCNA) approach (*42*), and network modules were defined accordingly (*43*). Module-related genome dominance and cell type– and stage-specific expression were assessed by use of the presence of differentially expressed homeologs, module-wise gene expression profiles, and correlation of module eigengenes to predefined spatiotemporal expression patterns. We performed functional analysis for individual modules using the GO terms inferred from the bread wheat gene annotation (*6*) and hypergeometric tests as implemented in the GOstats package ($P < 0.05$) (*44*). These enrichments were summarized and visualized with the REVIGO Web server, which arranges semantically similar GO terms in a two-dimensional space (*16*). For each genome, we colored those terms, which were significantly enriched in the respective genome-dominated co-expression modules, according to the enrichment tests (P values). For sequence-divergence analysis of homeologous genes, we estimated the number of synonymous substitutions per synonymous site (K_S) and the number of nonsynonymous substitutions per nonsynonymous site (K_A) based on the pairwise best-scoring protein alignments (BLASTP) (*45*) using the yn00 module of the PAML 4 suite (*46*).

Gene expression along the Triticeae prototype

The strong conservation of synteny between grass genomes and the availability of high-quality grass genomes allowed for an approximation of linear gene order. To position the A, B, and D genes in a common sequential order, they were mapped onto seven chromosome scaffolds named the Triticeae prototype chromosomes. These artificial scaffolds were generated by integrating syntenic genes from *Brachypodium* (*47*), rice (*48*), and sorghum (*49*) by using gene order information from barley (*50*). Bread wheat genes

were anchored on the scaffolds by using BLASTP and a first best-hit criterion. Gene expression distribution along the chromosomes was analyzed by use of a sliding window approach and median gene expression.

Identification and analysis of wheat grain-quality genes

Grain-quality genes were identified by use of orthologous genes from *Ae. tauschii* (*41*) and publically available wheat sequence information deposited at the National Center for Biotechnology Information. The corresponding genes within the bread wheat genome and gene annotation were defined with BLAST (*45*) and GenomeThreader (*51*) sequence homology searches as well as OrthoMCL gene family clustering (*39*). On the basis of this collective evidence and considering the RNA-seq alignment information, the candidate loci were manually refined. Gene expression for these genes was quantified as reads per kilobase per million (RPKM) (*52*) using HTSeq (www-huber.embl.de/users/anders/HTSeq) and custom python scripts. Phylogenetic trees were inferred by using multiple protein sequence alignments (CLUSTALW algorithm) (*53*), a neighborhood-joining algorithm, and the average percentage identity method. Circoletto (*54*) was used for comparative analysis of bread wheat genome sequences with known gliadin query proteins.

REFERENCES AND NOTES

1. M. Feldman, A. A. Levy, T. Fahima, A. Korol, Genomic asymmetry in allopolyploid plants: Wheat as a model. *J. Exp. Bot.* **63**, 5045–5059 (2012). doi: 10.1093/jxb/ers192; pmid: 22859676
2. A. R. Akhunova, R. T. Matniyazov, H. Liang, E. D. Akhunov, Homoeolog-specific transcriptional bias in allopolyploid wheat. *BMC Genomics* **11**, 505 (2010). doi: 10.1186/1471-2164-11-505; pmid: 20849627
3. T. K. Pellny *et al*., Cell walls of developing wheat starchy endosperm: Comparison of composition and RNA-Seq transcriptome. *Plant Physiol.* **158**, 612–627 (2012). doi: 10.1104/pp.111.189191; pmid: 22123899
4. S. A. Gillies, A. Futardo, R. J. Henry, Gene expression in the developing aleurone and starchy endosperm of wheat. *Plant Biotechnol. J.* **10**, 668–679 (2012). doi: 10.1111/j.1467-7652.2012.00705.x; pmid: 22672716
5. S. Drea *et al*., Systematic spatial analysis of gene expression during wheat caryopsis development. *Plant Cell* **17**, 2172–2185 (2005). doi: 10.1105/tpc.105.034058; pmid: 16006577
6. International Wheat Genome Sequencing Consortium, A chromosome-based draft sequence of the hexaploid bread wheat genome. *Science* **345**, 1251788 (2014).
7. Materials and methods are available as supplementary materials on *Science* Online.
8. M. F. Belmonte *et al*., Comprehensive developmental profiles of gene activity in regions and subregions of the *Arabidopsis* seed. *Proc. Natl. Acad. Sci. U.S.A.* **110**, E435–E444 (2013). doi: 10.1073/pnas.1222061110; pmid: 23319655
9. N. Sreenivasulu *et al*., Barley grain maturation and germination: Metabolic pathway and regulatory network commonalities and differences highlighted by new MapMan/PageMan profiling tools. *Plant Physiol.* **146**, 1738–1758 (2008). doi: 10.1104/pp.107.111781; pmid: 18281415
10. M. Ashburner *et al*., Gene ontology: Tool for the unification of biology. *Nat. Genet.* **25**, 25–29 (2000). doi: 10.1038/75556; pmid: 10802651
11. W. H. Vensel *et al*., Developmental changes in the metabolic protein profiles of wheat endosperm. *Proteomics* **5**, 1594–1611 (2005). doi: 10.1002/pmic.200401034; pmid: 15800972
12. A. Serna *et al*., Maize endosperm secretes a novel antifungal protein into adjacent maternal tissue. *Plant J.* **25**, 687–698 (2001). doi: 10.1046/j.1365-313x.2001.01004.x; pmid: 11319035
13. O. A. Olsen, Nuclear endosperm development in cereals and *Arabidopsis thaliana*. *Plant Cell* **16**, S214–S227 (2004). doi: 10.1105/tpc.017111; pmid: 15010513
14. P. R. Shewry, N. G. Halford, Cereal seed storage proteins: Structures, properties and role in grain utilization. *J. Exp. Bot.* **53**, 947–958 (2002). doi: 10.1093/jexbot/53.370.947; pmid: 11912237
15. S. Horvath, J. Dong, Geometric interpretation of gene coexpression network analysis. *PLOS Comput. Biol.* **4**, e1000117 (2008). doi: 10.1371/journal.pcbi.1000117; pmid: 18704157
16. F. Supek, M. Bošnjak, N. Škunca, T. Šmuc, REVIGO summarizes and visualizes long lists of gene ontology terms. *PLOS ONE* **6**, e21800 (2011). doi: 10.1371/journal.pone.0021800; pmid: 21789182
17. T. Marcussen *et al*., Ancient hybridizations among the ancestral genomes of bread wheat. *Science* **245**, 1250092 (2014).
18. L. Comai, The advantages and disadvantages of being polyploid. *Nat. Rev. Genet.* **6**, 836–846 (2005). doi: 10.1038/nrg1711; pmid: 16304599
19. N. Shitsukawa *et al*., Genetic and epigenetic alteration among three homoeologous genes of a class E MADS box gene in hexaploid wheat. *Plant Cell* **19**, 1723–1737 (2007). doi: 10.1105/tpc.107.051813; pmid: 17586655
20. Z. Hu *et al*., Epigenetic modification contributes to the expression divergence of three TaEXPA1 homoeologs in hexaploid wheat (*Triticum aestivum*). *New Phytol.* **197**, 1344–1352 (2013). doi: 10.1111/nph.12131; pmid: 23360546
21. E. J. Finnegan, C. C. Sheldon, F. Jardinaud, W. J. Peacock, E. S. Dennis, A cluster of *Arabidopsis* genes with a coordinate response to an environmental stimulus. *Curr. Biol.* **14**, 911–916 (2004). doi: 10.1016/j.cub.2004.04.045; pmid: 15186749
22. G. Moore *et al*., Was there a single ancestral cereal chromosome? *Trends Genet.* **11**, 81–82 (1995). doi: 10.1016/S0168-9525(00)89005-8; pmid: 7732576
23. S. Bolot *et al*., The 'inner circle' of the cereal genomes. *Curr. Opin. Plant Biol.* **12**, 119–125 (2009). doi: 10.1016/j.pbi.2008.10.011; pmid: 19095493
24. K. Nakabayashi, M. Okamoto, T. Koshiba, Y. Kamiya, E. Nambara, Genome-wide profiling of stored mRNA in *Arabidopsis thaliana* seed germination: Epigenetic and genetic regulation of transcription in seed. *Plant J.* **41**, 697–709 (2005). doi: 10.1111/j.1365-313X.2005.02337.x; pmid: 15703057
25. C. Rustenholz *et al*., A 3,000-loci transcription map of chromosome 3B unravels the structural and functional features of gene islands in hexaploid wheat. *Plant Physiol.* **157**, 1596–1608 (2011). doi: 10.1104/pp.111.183921; pmid: 22034626
26. C. F. Morris, Puroindolines: The molecular genetic basis of wheat grain hardness. *Plant Mol. Biol.* **48**, 633–647 (2002). doi: 10.1023/A:1014837431178; pmid: 11999840
27. C. Ravel *et al*., Nucleotide polymorphism in the wheat transcriptional activator Spa influences its pattern of expression and has pleiotropic effects on grain protein composition, dough viscoelasticity, and grain hardness. *Plant Physiol.* **151**, 2133–2144 (2009). doi: 10.1104/pp.109.146076; pmid: 19828671
28. P. Payne, G. Lawrence, Catalogue of alleles for the complex gene loci, Glu-A1, Glu-B1, and Glu-D1 which code for high-molecular-weight subunits of glutenin in hexaploid wheat. *Cereal Res. Commun.* **11**, 29–35 (1983).
29. R. D. Thompson, D. Bartels, N. P. Harberd, R. B. Flavell, Characterization of the multigene family coding for HMW glutenin subunits in wheat using cDNA clones. *Theor. Appl. Genet.* **67**, 87–96 (1983). doi: 10.1007/BF00303930; pmid: 24258485
30. N. Chantret *et al*., Molecular basis of evolutionary events that shaped the hardness locus in diploid and polyploid wheat species (*Triticum* and *Aegilops*). *Plant Cell* **17**, 1033–1045 (2005). doi: 10.1105/tpc.104.029181; pmid: 15749759
31. M. Wilkinson *et al*., Identification and genetic mapping of variant forms of puroindoline b expressed in developing wheat grain. *J. Cereal Sci.* **48**, 722–728 (2008). doi: 10.1016/j.jcs.2008.03.007
32. F. M. Dupont, W. H. Vensel, C. K. Tanaka, W. J. Hurkman, S. B. Altenbach, Deciphering the complexities of the wheat flour proteome using quantitative two-dimensional electrophoresis, three proteases and tandem mass spectrometry. *Proteome Sci.* **9**, 10 (2011). doi: 10.1186/1477-5956-9-10; pmid: 21314956
33. O. D. Anderson, N. Huo, Y. Q. Gu, The gene space in wheat: The complete γ-gliadin gene family from the wheat cultivar Chinese Spring. *Funct. Integr. Genomics* **13**, 261–273 (2013). doi: 10.1007/s10142-013-0321-8; pmid: 23564033
34. O. D. Anderson, Y. Q. Gu, X. Kong, G. R. Lazo, J. Wu, The wheat omega-gliadin genes: Structure and EST analysis. *Funct. Integr. Genomics* **9**, 397–410 (2009). doi: 10.1007/s10142-009-0122-2; pmid: 19367421
35. B. Langmead, C. Trapnell, M. Pop, S. L. Salzberg, Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009). doi: 10.1186/gb-2009-10-3-r25; pmid: 19261174
36. C. Trapnell, L. Pachter, S. L. Salzberg, TopHat: Discovering splice junctions with RNA-Seq. *Bioinformatics* **25**, 1105–1111 (2009). doi: 10.1093/bioinformatics/btp120; pmid: 19289445
37. C. Trapnell *et al*., Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* **28**, 511–515 (2010). doi: 10.1038/nbt.1621; pmid: 20436464
38. C. Trapnell *et al*., Differential analysis of gene regulation at transcript resolution with RNA-seq. *Nat. Biotechnol.* **31**, 46–53 (2013). doi: 10.1038/nbt.2450; pmid: 23222703
39. L. Li, C. J. Stoeckert Jr., D. S. Roos, OrthoMCL: Identification of ortholog groups for eukaryotic genomes. *Genome Res.* **13**, 2178–2189 (2003). doi: 10.1101/gr.1224503; pmid: 12952885
40. The Arabidopsis Genome Initiative, Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* **408**, 796–815 (2000). doi: 10.1038/35048692; pmid: 11130711
41. J. Jia *et al*., *Aegilops tauschii* draft genome sequence reveals a gene repertoire for wheat adaptation. *Nature* **496**, 91–95 (2013). doi: 10.1038/nature12028; pmid: 23535592
42. P. Langfelder, S. Horvath, WGCNA: An R package for weighted correlation network analysis. *BMC Bioinformatics* **9**, 559 (2008). doi: 10.1186/1471-2105-9-559; pmid: 19114008
43. P. Langfelder, B. Zhang, S. Horvath, Defining clusters from a hierarchical cluster tree: The Dynamic Tree Cut package for R. *Bioinformatics* **24**, 719–720 (2008). doi: 10.1093/bioinformatics/btm563; pmid: 18024473
44. S. Falcon, R. Gentleman, Using GOstats to test gene lists for GO term association. *Bioinformatics* **23**, 257–258 (2007). doi: 10.1093/bioinformatics/btl567; pmid: 17098774
45. S. F. Altschul, W. Gish, W. Miller, E. W. Myers, D. J. Lipman, Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990). doi: 10.1016/S0022-2836(05)80360-2; pmid: 2231712
46. Z. Yang, PAML 4: Phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **24**, 1586–1591 (2007). doi: 10.1093/molbev/msm088; pmid: 17483113
47. International Brachypodium Initiative, Genome sequencing and analysis of the model grass Brachypodium distachyon. *Nature* **463**, 763–768 (2010). doi: 10.1038/nature08747; pmid: 20148030
48. International Rice Genome Sequencing Project, The map-based sequence of the rice genome. *Nature* **436**, 793–800 (2005). doi: 10.1038/nature03895; pmid: 16100779
49. A. H. Paterson *et al*., The Sorghum bicolor genome and the diversification of grasses. *Nature* **457**, 551–556 (2009). doi: 10.1038/nature07723; pmid: 19189423
50. K. F. Mayer *et al*., Unlocking the barley genome by chromosomal and comparative genomics. *Plant Cell* **23**, 1249–1263 (2011). doi: 10.1105/tpc.110.082537; pmid: 21467582
51. G. Gremme, V. Brendel, M. E. Sparks, S. Kurtz, Engineering a software tool for gene structure prediction in higher organisms. *Inf. Softw. Technol.* **47**, 965–978 (2005). doi: 10.1016/j.infsof.2005.09.005
52. A. Mortazavi, B. A. Williams, K. McCue, L. Schaeffer, B. Wold, Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. Methods* **5**, 621–628 (2008). doi: 10.1038/nmeth.1226; pmid: 18516045
53. M. A. Larkin *et al*., Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947–2948 (2007). doi: 10.1093/bioinformatics/btm404; pmid: 17846036
54. N. Darzentas, Circoletto: Visualizing sequence similarity with Circos. *Bioinformatics* **26**, 2620–2621 (2010). doi: 10.1093/bioinformatics/btq484; pmid: 20736339

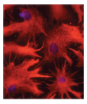
ACKNOWLEDGMENTS

This work was supported by research grant 199387 from the Norwegian Research Council to the Norwegian University of Life Sciences and Graminor AS to O.A.O. K.F.X.M. acknowledges generous funding by the Deutsche Forschungsgemeinschaft (DFG) to project SFB924 "Molecular mechanisms regulating yield and yield stability in plants," the German Federal Ministry of Education and Research (BMBF) in frame of the Synbreed project, and the European Commission's 7th Framework Program in frame of the TransPLANT project. The data reported in this paper are tabulated in the supplementary materials and archived at the ArrayExpress database www.ebi.ac.uk/arrayexpress, accession no. E-MTAB-2137. No conflict of interest is known for the authors of this manuscript regarding data reported herein. S. I. Chi, E. I. Ako, A. Maharajan, and Ø. Jørgensen are thankfully acknowledged for grain dissection and plant cultivation. We thank P. Langridge, P. Shewry, B. Keller, and two anonymous reviewers for helpful suggestions on the manuscript. The RNA-seq was carried out at the Norwegian Sequencing Centre of University of Oslo. Graminor AS is acknowledged for generous support throughout this work.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/345/6194/1250091/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S25
Tables S1 to S31
References (*55*–*69*)
IWGSC Author List

23 December 2013; accepted 21 May 2014
10.1126/science.1250091

EXTENDED PDF FORMAT
SPONSORED BY

Travel Grants Available
Neuroscience 2014
Application Deadline: July 31st
>Apply Now
R&D systems
a bio-techne brand
www.rndsystems.com

Ancient hybridizations among the ancestral genomes of bread wheat

Thomas Marcussen *et al.*

Science **345**, (2014);
DOI: 10.1126/science.1250092

This copy is for your personal, non-commercial use only.

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by clicking here.

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines here.

The following resources related to this article are available online at www.sciencemag.org (this information is current as of August 4, 2014):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:
http://www.sciencemag.org/content/345/6194/1250092.full.html

Supporting Online Material can be found at:
http://www.sciencemag.org/content/suppl/2014/07/16/345.6194.1250092.DC1.html

A list of selected additional articles on the Science Web sites **related to this article** can be found at:
http://www.sciencemag.org/content/345/6194/1250092.full.html#related

This article **cites 43 articles**, 22 of which can be accessed free:
http://www.sciencemag.org/content/345/6194/1250092.full.html#ref-list-1

This article has been **cited by** 1 articles hosted by HighWire Press; see:
http://www.sciencemag.org/content/345/6194/1250092.full.html#related-urls

This article appears in the following **subject collections:**
Evolution
http://www.sciencemag.org/cgi/collection/evolution
Genetics
http://www.sciencemag.org/cgi/collection/genetics

Downloaded from www.sciencemag.org on August 4, 2014

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published weekly, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. Copyright 2014 by the American Association for the Advancement of Science; all rights reserved. The title *Science* is a registered trademark of AAAS.

A chromosome-based draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome

The International Wheat Genome Sequencing Consortium (IWGSC)

An ordered draft sequence of the 17-gigabase hexaploid bread wheat (*Triticum aestivum*) genome has been produced by sequencing isolated chromosome arms. We have annotated 124,201 gene loci distributed nearly evenly across the homeologous chromosomes and subgenomes. Comparative gene analysis of wheat subgenomes and extant diploid and tetraploid wheat relatives showed that high sequence similarity and structural conservation are retained, with limited gene loss, after polyploidization. However, across the genomes there was evidence of dynamic gene gain, loss, and duplication since the divergence of the wheat lineages. A high degree of transcriptional autonomy and no global dominance was found for the subgenomes. These insights into the genome biology of a polyploid crop provide a springboard for faster gene isolation, rapid genetic marker development, and precise breeding to meet the needs of increasing food demand worldwide.

Lists of authors and affiliations are available in the full article online.
Corresponding author: K. X. Mayer, e-mail: k.mayer@helmholtz-muenchen.de
Read the full article at http://dx.doi.org/10.1126/science.1251788

Ancient hybridizations among the ancestral genomes of bread wheat

Thomas Marcussen, Simen R. Sandve,* Lise Heier, Manuel Spannagl, Matthias Pfeifer, The International Wheat Genome Sequencing Consortium,† Kjetill S. Jakobsen, Brande B. H Wulff, Burkhard Steuernagel, Klaus F. X. Mayer, Odd-Arne Olsen

The allohexaploid bread wheat genome consists of three closely related subgenomes (A, B, and D), but a clear understanding of their phylogenetic history has been lacking. We used genome assemblies of bread wheat and five diploid relatives to analyze genome-wide samples of gene trees, as well as to estimate evolutionary relatedness and divergence times. We show that the A and B genomes diverged from a common ancestor ~7 million years ago and that these genomes gave rise to the D genome through homoploid hybrid speciation 1 to 2 million years later. Our findings imply that the present-day bread wheat genome is a product of multiple rounds of hybrid speciation (homoploid and polyploid) and lay the foundation for a new framework for understanding the wheat genome as a multilevel phylogenetic mosaic.

The list of author affiliations is available in the full article online.***Corresponding author.**
E-mail: simen.sandve@nmbu.no †The International Wheat Genome Sequencing Consortium (IWGSC) authors and affiliations are listed in the supplementary materials.
Read the full article at http://dx.doi.org/10.1126/science.1250092

Triticum monococcum

Triticum carthlicum

Ancestral wheat

Wheat varieties and species (shown) believed to be the closest living relatives of modern bread wheat (*T. aestivum*). Multiple ancestral hybridizations occurred among most of these species, many of which are cultivated, and along with *T. aestivum* represent a dominant source of global nutrition.

Triticum boeticum

Triticum polonicum L.

Triticum macha

Triticum dicoccoides var. *araraticum*

Published by AAAS

PHOTOS: SUSANNE STAMP, ERNST MERZ/ETH ZURICH

Genome interplay in the grain transcriptome of hexaploid bread wheat

Matthias Pfeifer, Karl G. Kugler, Simen R. Sandve, Bujie Zhan, Heidi Rudi, Torgeir R. Hvidsten, International Wheat Genome Sequencing Consortium,* Klaus F. X. Mayer, Odd-Arne Olsen†

Allohexaploid bread wheat (*Triticum aestivum* L.) provides approximately 20% of calories consumed by humans. Lack of genome sequence for the three homeologous and highly similar bread wheat genomes (A, B, and D) has impeded expression analysis of the grain transcriptome. We used previously unknown genome information to analyze the cell type–specific expression of homeologous genes in the developing wheat grain and identified distinct co-expression clusters reflecting the spatiotemporal progression during endosperm development. We observed no global but cell type– and stage-dependent genome dominance, organization of the wheat genome into transcriptionally active chromosomal regions, and asymmetric expression in gene families related to baking quality. Our findings give insight into the transcriptional dynamics and genome interplay among individual grain cell types in a polyploid cereal genome.

The list of author affiliations is available in the full article online. *The International Wheat Genome Sequencing Consortium (IWGSC) authors and affiliations are listed in the supplementary materials. †**Corresponding author. E-mail: odd-arne.olsen@nmbu.no**

Read the full article at http://dx.doi.org/10.1126/science.1250091

Structural and functional partitioning of bread wheat chromosome 3B

Frédéric Choulet,* Adriana Alberti, Sébastien Theil, Natasha Glover, Valérie Barbe, Josquin Daron, Lise Pingault, Pierre Sourdille, Arnaud Couloux, Etienne Paux, Philippe Leroy, Sophie Mangenot, Nicolas Guilhot, Jacques Le Gouis, Francois Balfourier, Michael Alaux, Véronique Jamilloux, Julie Poulain, Céline Durand, Arnaud Bellec, Christine Gaspin, Jan Safar, Jaroslav Dolezel, Jane Rogers, Klaas Vandepoele, Jean-Marc Aury, Klaus Mayer, Hélène Berges, Hadi Quesneville, Patrick Wincker, Catherine Feuillet

We produced a reference sequence of the 1-gigabase chromosome 3B of hexaploid bread wheat. By sequencing 8452 bacterial artificial chromosomes in pools, we assembled a sequence of 774 megabases carrying 5326 protein-coding genes, 1938 pseudogenes, and 85% of transposable elements. The distribution of structural and functional features along the chromosome revealed partitioning correlated with meiotic recombination. Comparative analyses indicated high wheat-specific inter- and intrachromosomal gene duplication activities that are potential sources of variability for adaption. In addition to providing a better understanding of the organization, function, and evolution of a large and polyploid genome, the availability of a high-quality sequence anchored to genetic maps will accelerate the identification of genes underlying important agronomic traits.

The list of author affiliations is available in the full article online.
***Corresponding author. E-mail: frederic.choulet@clermont.inra.fr**

Read the full article at http://dx.doi.org/10.1126/science.1249721

Published by AAAS

WHEAT GENOME

Ancient hybridizations among the ancestral genomes of bread wheat

Thomas Marcussen,[1]* Simen R. Sandve,[1]*† Lise Heier,[2] Manuel Spannagl,[3] Matthias Pfeifer,[3] The International Wheat Genome Sequencing Consortium,‡ Kjetill S. Jakobsen,[4] Brande B. H. Wulff,[5] Burkhard Steuernagel,[5] Klaus F. X. Mayer,[3] Odd-Arne Olsen[1]

The allohexaploid bread wheat genome consists of three closely related subgenomes (A, B, and D), but a clear understanding of their phylogenetic history has been lacking. We used genome assemblies of bread wheat and five diploid relatives to analyze genome-wide samples of gene trees, as well as to estimate evolutionary relatedness and divergence times. We show that the A and B genomes diverged from a common ancestor ~7 million years ago and that these genomes gave rise to the D genome through homoploid hybrid speciation 1 to 2 million years later. Our findings imply that the present-day bread wheat genome is a product of multiple rounds of hybrid speciation (homoploid and polyploid) and lay the foundation for a new framework for understanding the wheat genome as a multilevel phylogenetic mosaic.

The rise of modern agriculture and wheat domestication in the Fertile Crescent ~10,000 years ago (*1*–*4*) was pivotal in shaping modern human history. Early farming practices made use of wild diploid wheat species (i.e., *Aegilops* and *Triticum* species), but as agriculture evolved, wild crops were gradually substituted with domesticated diploid and polyploid wheat varieties (*3*, *4*). Presently, the allohexaploid bread wheat (*Triticum aestivum*, $2n = 6x = 42$ chromosomes; genomic code AABBDD) dominates global wheat production. Because of its economic value and the desire for its genetic improvement, questions concerning the evolution and domestication of wheat have been under intense scientific scrutiny (*5*, *6*).

The bread wheat subgenomes A, B, and D were originally derived from three diploid ($2x$; $2n = 14$) species within tribe Triticeae [see figure 1 in (*7*)]: *Triticum urartu* (AA), an unknown close relative of *Aegilops speltoides* (BB), and *Ae. tauschii* (DD) (*4*, *8*). The initial allopolyploidization event is hypothesized to have involved the A and B genome donors, resulting in the extant tetraploid emmer wheat (*T. turgidum*; AABB). This species subsequently hybridized with the D genome donor to form modern hexaploid bread wheat (AABBDD) (*4*, *8*).

Tetraploid emmer wheat is believed to have originated within the past few hundred thousand years (*9*), whereas hexaploid bread wheat is thought to have originated with modern agriculture ~10,000 years ago (*4*). The time of origin for hexaploid bread wheat is currently supported solely by archeological evidence (*2*, *3*) and the apparent absence of hexaploid wheats in wild populations (*4*). Although the relatedness between the bread wheat subgenomes and diploid wheat species has been well documented (*8*, *10*), a clear understanding of the phylogenetic history and divergence times among the three A, B, and D genome lineages is still lacking (*9*, *11*–*13*). This knowledge gap is mainly a consequence of the paucity of Triticeae fossils (*14*), which has prevented investigations of diversification through time; extensive topological discordance between wheat gene trees (*15*); and, most importantly, the lack of genome sequences of the hexaploid bread wheat and its close diploid relatives. Improved understanding of the phylogenetic relationships among the diploid species of wheat and the bread wheat subgenomes is important for understanding genome function and for future agricultural crop improvement in light of a changing global climate (*16*).

Gene tree topology analyses

We used the genome sequences of hexaploid bread wheat subgenomes (denoted TaA, TaB, and TaD) and five diploid relatives (*T. monococcum*, *T. urartu*, *Ae. sharonensis*, *Ae. speltoides*, and *Ae. tauschii*) (*7*, *17*, *18*) to generate a genome-wide sample of 275 gene trees and to estimate the phylogenetic history of the A, B, and D genome lineages. Barley (*Hordeum vulgare*), *Brachypodium distachyon*, and rice (*Oryza sativa*) were used as outgroup species. To generate multiple alignments of ortholog genes, we employed a phylogeny-aware strategy (*19*), which simultaneously filters alignments for unreliably aligned codon sites and putative erroneously predicted ortholog sequences (fig. S1 and supplementary materials and methods). Finally, we used BEAST (*20*) to calculate gene trees topologies.

We found that the basal relatedness among the three lineages A, B, and D varied substantially among the 275 gene trees, with the lineage topologies A(B,D) and B(A,D) each being about twice as common as D(A,B) (Fig. 1A and Table 1). Stochastic population genetic processes typically cause incomplete lineage sorting (ILS), which results in topological discordance (i.e., variation in topology) among individual gene trees. For three taxa under ILS alone, the gene tree topology that equals the species tree topology is expected to be more common than the other

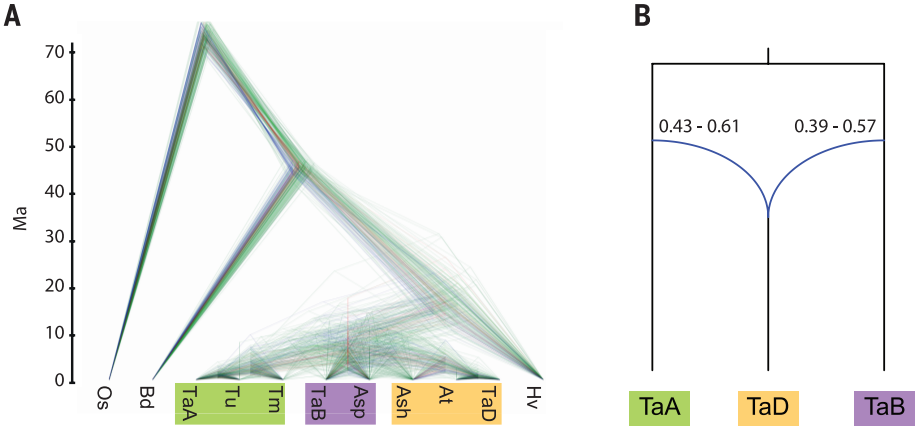


Fig. 1. Analyses of gene tree topologies. (**A**) Superimposed ultrametric gene trees in a consensus DensiTree plot. The branch color changes for every 100 trees plotted. (**B**) Topology-based species phylogeny, assuming incomplete lineage sorting using a data set of 2269 gene trees inferred by PhyloNet. The results presented represent analyses of all gene trees (2269). The numbers on the branches represent estimates of parental contributions to the hybrid. Range estimates of parental contribution are extrapolated from results reported in Table 1. Species names are abbreviated as follows: Ash, *Aegilops sharonensis*; Asp, *Ae. speltoides*; At, *Ae. tauschii*; Tm, *Triticum monococcum*; Tu, *T. urartu*; Hv, *Hordeum vulgare*; TaA, *T. aestivum* A subgenome; TaB, *T. aestivum* B subgenome; TaD, *T. aestivum* D subgenome; Bd, *Brachypodium distachyon*; Os, *Oryza sativa*.

[1]Department of Plant Sciences, Norwegian University of Life Sciences, 1432 Ås, Norway. [2]Strømsveien 78 B, 0663 Oslo, Norway. [3]Plant Genome and Systems Biology, Helmholtz Center Munich, Ingolstädter Landstrasse 1, 85764 Neuherberg, Germany. [4]Centre for Ecological and Evolutionary Synthesis, Department of Biosciences, University of Oslo, 0316 Oslo, Norway. [5]The Sainsbury Laboratory, Norwich Research Park, Norwich NR4 7UH, UK.
*These authors contributed equally to this work. **†Corresponding author. E-mail: simen.sandve@nmbu.no** ‡The International Wheat Genome Sequencing Consortium (IWGSC) authors and affiliations are listed in the supplementary materials.

Table 1. Distribution of ABD lineage topologies in gene trees. Analyses were made on different data subsets: diploid genomes only (2*x*), hexaploid genomes only (6*x*), and either whole genomes or gene trees from individual chromosomes (Chr.). Interrelationships within the A, D, and B clades are not considered in the data set that includes diploids. Topologies including diploids were estimated with Bayesian MCMC sampling using the HKY+G nucleotide substitution model, whereas topologies excluding diploids were taken from IWGSC (*7*) and represent maximum likelihood topologies under the GTR+I+G model. Bold numbers represent the largest topology group. The likelihood ratio test was used to test the probability (*P*) of observing the data under the model of multispecies coalescent and the (conservative) assumption that the most common observed tree topology equaled the species tree topology.

Genomes	Genes	Sample	Observed topologies			Proportion of genes with deep coalescence	Parental contributions in D		Likelihood ratio test *P*
			A,(B,D)	B,(A,D)	D,(A,B)		A	B	
2*x*	275	Whole genome*	**112**	100	63	0.69	0.43	0.57	0.0036‡
	275	Whole genome†	**109**	101	65	0.71	0.45	0.55	0.0050‡
6*x*	275	Whole genome	**107**	105	63	0.69	0.49	0.51	0.0058‡
	2269	Whole genome	786	**909**	574	0.76	0.61	0.39	8.4×10^{-9}‡
	324	Chr. 1	109	**137**	78	0.72	0.66	0.34	0.023§
	428	Chr. 2	131	**191**	106	0.74	0.77	0.23	0.10
	321	Chr. 3	111	**141**	69	0.65	0.63	0.37	0.0016‡
	305	Chr. 4	**121**	102	82	0.81	0.34	0.67	0.14
	347	Chr. 5	127	**129**	91	0.79	0.52	0.49	0.015§
	290	Chr. 6	99	**117**	74	0.77	0.63	0.37	0.057
	254	Chr. 7	88	**92**	74	0.87	0.56	0.44	0.27

*A lineage represented by *T. monococcum*, excluding *T. urartu*. †A lineage represented by *T. urartu*, excluding *T. monococcum*. ‡Significant at $P < 0.01$. §Significant at $P < 0.05$.

two. However, for our data the observed lineage topology ratios differed significantly [$P < 0.01$; likelihood ratio test; df = 1 (Table 1)] from this expectation. This suggests the presence of phylogenetic signals additional to ILS in the data. Except in rare instances of deep coalescence, bread wheat homeologs consistently formed monophyletic clades with orthologs of their close diploid relatives (Fig. 1A), and never with each other. This rules out nonhomologous gene conversion and nonhomologous recombination as explanations for the observed topological discordance. To enable analyses of individual chromosomes, we also made use of a considerably larger data set of 2269 maximum likelihood gene tree topologies taken from (*7*), which did not include the diploid *Triticum* and *Aegilops* species. These gene trees support genome-wide signals of significantly skewed topology frequencies, with B(A,D) and A(D,B) topologies being most common (Table 1). Taken together, these results suggest that lineages A and B are more closely related to D individually than to each other, which agrees with a model of hybrid origin of the D lineage (Fig. 1B).

Under the assumption of ILS and a hybrid origin of the D lineage from the A and B lineages, analysis of tree topology frequencies revealed roughly equal contributions of each parental lineage to the D lineage (Table 1). There was considerable stochasticity among the different data subsets, and between 65 and 87% of the genes displayed deep coalescence for the target nodes. Similar results were obtained with a topology-based parsimony approach to estimate phylogenetic genome networks under the assumption of a single homoploid hybridization at the genome-wide level (Fig. 1B) and for the majority of the chromosome-specific analyses (table S4). Finally, ~80% of the genes could be anchored to a chromosome position in the hexaploid genome using the in silico gene order predictions from the bread wheat genome sequence (*7*). Such positional information can be used to investigate whether different regions of the genome have distinct phylogenetic signals—that is, conserved chromosome blocks from the parental genomes. Homeologs within gene trees showed highly conserved syntenic relationships (fig. S3); however, anchoring of gene tree topologies to chromosome positions in bread wheat did not support the presence of larger chromosome blocks with a single parental origin (*7*), indicating a relatively homogeneous hybrid signal throughout the D subgenome.

Genome divergence times

Because topology-based phylogenetic analyses do not consider the temporal scale, we used a Bayesian hierarchical model to further estimate the genome divergence times under the multispecies coalescent (*21*) from pairwise ortholog coalesent distributions. This approach does not assume a treelike species phylogeny and can handle very large data sets.

Concurrent with the topology-based analyses (Fig. 2, A and B, and fig. S4), genome divergence estimates showed no signs of being affected by nonhomologous gene conversion or recombination within the hexaploid genome, which would have resulted in shallow coalescence of bread wheat subgenomes (figs. S4 and S5). The basal divergence in *Aegilops*/*Triticum* was estimated to have occurred between the A and the B genome lineages ~7 million years ago (Ma) (Table 2). Both A-D and B-D divergence times overlap and are estimated to be 1 to 2 million years younger than the A-B divergence (Fig. 2, A and B). This contradicts a treelike phylogeny for the three subgenome lineages (Table 2 and Fig. 2A) and favors a model of hybridization between the A and B genomes, giving rise to the D genome. Genome divergence times did not support the more complex models of hybridization patterns, as suggested by the topology analyses assuming two hybridization events (table S4). Furthermore, the majority of the analyses produced slightly younger divergence of A and D lineages compared with B and D lineages (Fig. 2A and Table 2), indicating that gene flow from A to D may have persisted after gene flow from B to D had ceased.

The identification of hybridization events in phylogenies strongly depends on taxon sampling. Nevertheless, given that the hybridization event happened basally in the Triticum/Aegilops clade and that the 15 extant diploid species all seem to fall within one of the three lineages A, B, and D (fig. S6), the hybridization pattern is likely to remain unaltered, even with a denser sampling of wheat species. Beyond the phylogenetic evidence presented here, support for a homoploid hybrid origin of the D genome is found in independent analyses using the genome sequence of bread wheat. Both at the base-pair level (*7*, *22*) as well as in gene content (*7*), the A and B lineages are more similar to the D genome lineage than they are to each other.

Although the existence of homoploid hybrid speciation has been acknowledged for well more than a century (*23*), it has only recently been recognized as a relatively common phenomenon (*24*–*26*). Our data support a homoploid hybrid origin of the bread wheat D genome lineage ancestor more than 5 Ma, which is among the oldest cases of homoploid hybrid speciation reported to date.

Divergence of polyploid genomes from diploid relatives

Genome divergence estimates showed that *T. monococcum* and *T. urartu* are successive

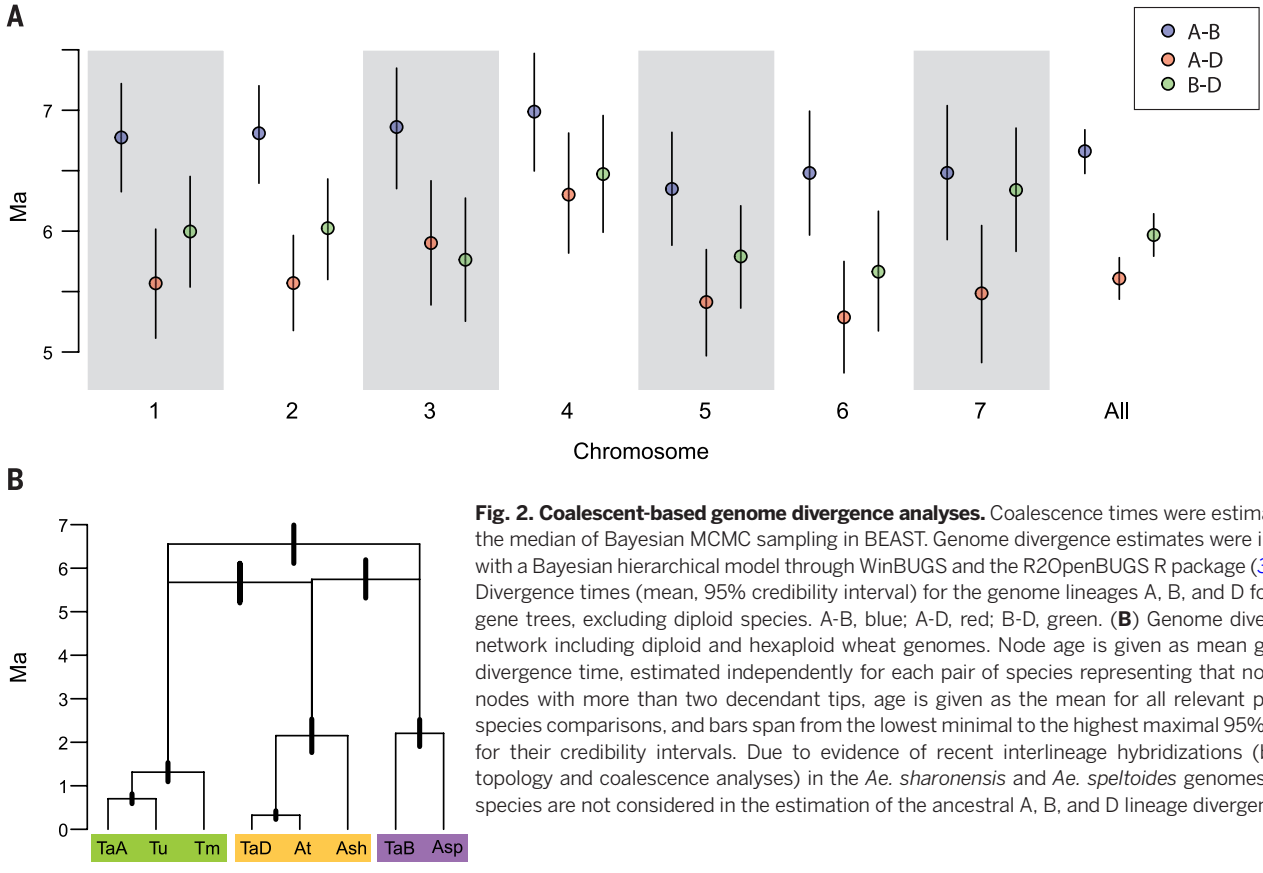


Fig. 2. Coalescent-based genome divergence analyses. Coalescence times were estimated as the median of Bayesian MCMC sampling in BEAST. Genome divergence estimates were inferred with a Bayesian hierarchical model through WinBUGS and the R2OpenBUGS R package (*35*). (**A**) Divergence times (mean, 95% credibility interval) for the genome lineages A, B, and D for 2269 gene trees, excluding diploid species. A-B, blue; A-D, red; B-D, green. (**B**) Genome divergence network including diploid and hexaploid wheat genomes. Node age is given as mean genome divergence time, estimated independently for each pair of species representing that node. For nodes with more than two decendant tips, age is given as the mean for all relevant pairwise species comparisons, and bars span from the lowest minimal to the highest maximal 95% bound for their credibility intervals. Due to evidence of recent interlineage hybridizations (both in topology and coalescence analyses) in the *Ae. sharonensis* and *Ae. speltoides* genomes, these species are not considered in the estimation of the ancestral A, B, and D lineage divergence.

Table 2. Estimated genome divergence times. All age estimates are given in units of million years ago as 95% credibility intervals (CIs). The CI of the Tm-TaA divergence and the CI of the At-TaD divergence represent the summarized CI ranges of two hierarchical Bayesian models using median plus median. The Tm-TaA divergence and the At-TaD divergence are expected to be overestimates of the actual polyploidization times due to the fact that the true ancestral populations to the A and D subgenomes in bread wheat were not sampled. Species names are abbreviated as follows: At, *Aegilops tauschii*; Tm, *Triticum monococcum*; TaA, *T. aestivum* A subgenome; TaD, *T. aestivum* D subgenome. Dashes indicate no data.

Data set	TMRCA				
	A-B	A-D	B-D	Tm-TaA	At-TaD
275 gene trees	6.11–6.99	5.20–6.11	5.31–6.19	0.59–0.82	0.23–0.43
2269 gene trees	6.47–6.83	5.43–5.78	5.79–6.14	–	–

sisters to TaA, *Ae. speltoides* is sister to TaB, and *Ae. sharonensis* and *Ae. tauschii* are successive sisters to TaD (Fig. 2B and table S5). To elucidate the timing of the polyploid speciations giving rise to emmer and bread wheat, we analyzed the pairwise coalescence distributions of the TaA and TaD genomes and the closest related diploid species, *T. urartu* and *Ae. tauschii*, respectively (Fig. 2B). We estimated the *T. urartu*-TaA divergence to 0.58 to 0.82 Ma and the *Ae. tauschii*–TaD divergence to 0.23 to 0.43 Ma (Fig. 2B and Table 1), suggesting that the age of these two polyploidization events could be older than previously suggested (*4*, *9*). It is however important to note that this study includes only single-genome samples from *T. urartu* and *Ae. tauschii* and that these samples are not likely to represent plants from the actual ancestor populations that gave rise to the A and D genome constituents. Hence, our genome divergence estimates between diploid and hexaploid genomes are likely to reflect population divergence events before the actual polyploidization times, and additional analyses of broad samples of both *T. urartu* and *Ae. tauschii* populations are needed to further improve the accuracy of the polyploidization time estimates (*27*).

Conclusions

Our study exemplifies how analyses of whole-genome data sets can aid in resolving convoluted patterns of genome evolution caused by ancient hybridization events. We elucidate genome-wide signatures of hybrid ancestry of the wheat D lineage from A and B lineage ancestors (Fig. 3). Not only is bread wheat a product of hybridization and allopolyploidization involving the A, B, and D genomes, but also the ancestral lineages of these three genomes are the result of ancestral hybridization events among themselves. Our findings could have broad implications for understanding genome function and, thus, cultivar improvement in bread wheat.

Methods

Sequence data for *T. urartu* and *Ae. tauschii* were downloaded from GenBank (accession numbers AOTI00000000 and AOCO00000000.1, respectively). All other sequences were downloaded from the International Wheat Genome Sequencing Consortium (IWGSC) sequence repository at Institut National de la Recherche Agronomique (http://wheat-urgi.versailles.inra.fr/Seq-Repository/Genes-annotations). Preliminary ortholog predictions were carried out with OrthoMCL (*28*). Only orthologous gene sets containing single-gene copies from all species were used. A phylogeny-aware sequence alignment pipeline modified from (*19*) was implemented in R language using MAFFT (*29*) to construct multiple sequence alignments, GUIDANCE (*30*) to assess sequence alignment quality, and phangorn (*31*) to construct maximum likelihood topologies, which were used to evaluate ortholog relationships within gene trees. Ultrametric gene trees were estimated in BEAST (*20*), based on a secondary calibration on the

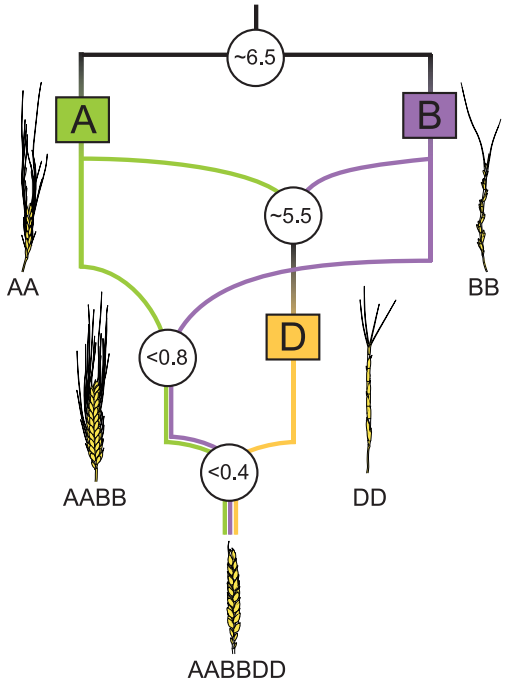


Fig. 3. Model of the phylogenetic history of bread wheat (*Triticum aestivum*; AABBDD). Approximate dates for divergence and the three hybridization events are given in white circles in units of million years ago. Differentiation of the wheat lineage (*Triticum* and *Aegilops*) from a common ancestor into the A and B genome lineages began ~6.5 Ma. The first hybridization occurred ~5.5 Ma between the A and B genome lineages and led to the origin of the D genome lineage by homoploid hybrid speciation. The second hybridization, between a close relative (BB) of *Ae. speltoides* and *T. urartu* (AA), gave rise to the allotetraploid emmer wheat (*T. turgidum*; AABB) by polyploidization. Bread wheat originated by allopolyploidization from a third hybridization, between emmer wheat and *Ae. tauschii* (DD). The three diploid lineages are indicated with color and labels. Inflorescences (spikes) illustrate extant species closely related to those involved in the polyploidizations.

Brachypodium crown node (fig. S2). All analyses of topologies were carried out with the APE R package (*32*). Estimations of parental genome contributions and species networks from gene tree topologies were calculated using parsimonious inference of hybridization in the presence of ILS (*33*) implemented in PhyloNet (*34*) and with a coalescent-based method (described in the supplementary materials). Topology-independent genome divergence times were estimated under the multispecies coalescent model with hierarchical modeling using BUGS (Bayesian inference Using Gibbs Sampling) (www.openbugs.net/w/FrontPage) through the R2OpenBUGS R package (*35*).

REFERENCES AND NOTES

1. M. Heun *et al*., Site of einkorn wheat domestication identified by DNA fingerprinting. *Science* **278**, 1312–1314 (1997). doi: 10.1126/science.278.5341.1312
2. S. Lev-Yadun, A. Gopher, S. Abbo, The cradle of agriculture. *Science* **288**, 1602–1603 (2000). doi: 10.1126/science.288.5471.1602; pmid: 10858140
3. S. Riehl, M. Zeidi, N. J. Conard, Emergence of agriculture in the foothills of the Zagros Mountains of Iran. *Science* **341**, 65–67 (2013). doi: 10.1126/science.1236743; pmid: 23828939
4. F. Salamini, H. Özkan, A. Brandolini, R. Schäfer-Pregl, W. Martin, Genetics and geography of wild cereal domestication in the near east. *Nat. Rev. Genet.* **3**, 429–441 (2002). pmid: 12042770
5. E. S. McFadden, E. R. Sears, The origin of *Triticum spelta* and its free-threshing hexaploid relatives. *J. Hered.* **37**, 81–107, 107 (1946). pmid: 20985728
6. J. Dubcovsky, J. Dvorak, Genome plasticity a key factor in the success of polyploid wheat under domestication. *Science* **316**, 1862–1866 (2007). doi: 10.1126/science.1143986; pmid: 17600208
7. International Wheat Genome Sequencing Consortium, A chromosome-based draft sequence of the hexaploid bread wheat genome. *Science* **345**, 1251788 (2014).
8. G. Petersen, O. Seberg, M. Yde, K. Berthelsen, Phylogenetic relationships of *Triticum* and *Aegilops* and evidence for the origin of the A, B, and D genomes of common wheat (*Triticum aestivum*). *Mol. Phylogenet. Evol.* **39**, 70–82 (2006) and references therein. doi: 10.1016/j.ympev.2006.01.023; pmid: 16504543
9. S. Huang *et al*., Genes encoding plastid acetyl-CoA carboxylase and 3-phosphoglycerate kinase of the *Triticum/Aegilops* complex and the evolutionary history of polyploid wheat. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 8133–8138 (2002). doi: 10.1073/pnas.072223799; pmid: 12060759
10. E. D. Akhunov, A. R. Akhunova, J. Dvorák, BAC libraries of *Triticum urartu*, *Aegilops speltoides* and *Ae. tauschii*, the diploid ancestors of polyploid wheat. *Theor. Appl. Genet.* **111**, 1617–1622 (2005) and references therein. doi: 10.1007/s00122-005-0093-1; pmid: 16177898
11. D. Chalupska *et al*., *Acc* homoeoloci and the evolution of wheat genomes. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 9691–9696 (2008). doi: 10.1073/pnas.0803981105; pmid: 18599450
12. J. Dvorak, E. D. Akhunov, Tempos of gene locus deletions and duplications and their relationship to recombination rate during diploid and polyploid evolution in the Aegilops-Triticum alliance. *Genetics* **171**, 323–332 (2005). doi: 10.1534/genetics.105.041632; pmid: 15996988
13. X. Fan *et al*., Phylogenetic reconstruction and diversification of the Triticeae (Poaceae) based on single-copy nuclear *Acc1* and *Pgk1* gene data. *Biochem. Syst. Ecol.* **50**, 346–360 (2013). doi: 10.1016/j.bse.2013.05.010
14. C. A. E. Strömberg, Evolution of grasses and grassland ecosystems. *Annu. Rev. Earth Planet. Sci.* **39**, 517–544 (2011). doi: 10.1146/annurev-earth-040809-152402
15. J. S. Escobar *et al*., Multigenic phylogeny and analysis of tree incongruences in Triticeae (Poaceae). *BMC Evol. Biol.* **11**, 181–198 (2011). doi: 10.1186/1471-2148-11-181; pmid: 21702931
16. D. B. Lobell, W. Schlenker, J. Costa-Roberts, Climate trends and global crop production since 1980. *Science* **333**, 616–620 (2011). doi: 10.1126/science.1204531; pmid: 21551030
17. H. Q. Ling *et al*., Draft genome of the wheat A-genome progenitor *Triticum urartu*. *Nature* **496**, 87–90 (2013). doi: 10.1038/nature11997; pmid: 23535596
18. J. Jia *et al*., *Aegilops tauschii* draft genome sequence reveals a gene repertoire for wheat adaptation. *Nature* **496**, 91–95 (2013). doi: 10.1038/nature12028; pmid: 23535592
19. M. D. Vigeland *et al*., Evidence for adaptive evolution of low-temperature stress response genes in a Pooideae grass ancestor. *New Phytol.* **199**, 1060–1068 (2013). doi: 10.1111/nph.12337; pmid: 23701123
20. A. J. Drummond, M. A. Suchard, D. Xie, A. Rambaut, Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol. Biol. Evol.* **29**, 1969–1973 (2012). doi: 10.1093/molbev/mss075; pmid: 22367748
21. J. H. Degnan, N. A. Rosenberg, Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends Ecol. Evol.* **24**, 332–340 (2009). doi: 10.1016/j.tree.2009.01.009; pmid: 19307040
22. M. Pfeifer *et al*., Genome interplay in the grain transcriptome of hexploid bread wheat. *Science* **345**, 1250091 (2014).
23. J. P. Lotsy, *Evolution by Means of Hybridization* (Martinus Nijhoff, The Hague, Netherlands, 1916).
24. R. Abbott *et al*., Hybridization and speciation. *J. Evol. Biol.* **26**, 229–246 (2013). doi: 10.1111/j.1420-9101.2012.02599.x; pmid: 23323997
25. F. Eroukhmanoff, R. I. Bailey, G.-P. Sætre, Hybridization and genome evolution I: The role of contingency during hybrid speciation. *Curr. Zool.* **59**, 667–674 (2013).
26. J. Mallet, Hybrid speciation. *Nature* **446**, 279–283 (2007). doi: 10.1038/nature05706; pmid: 17361174
27. J. J. Doyle, A. N. Egan, Dating the origins of polyploidy events. *New Phytol.* **186**, 73–85 (2010). doi: 10.1111/j.1469-8137.2009.03118.x; pmid: 20028472
28. L. Li, C. J. Stoeckert Jr., D. S. Roos, OrthoMCL: Identification of ortholog groups for eukaryotic genomes. *Genome Res.* **13**, 2178–2189 (2003). doi: 10.1101/gr.1224503; pmid: 12952885
29. K. Katoh, G. Asimenos, H. Toh, Multiple alignment of DNA sequences with MAFFT. *Methods Mol. Biol.* **537**, 39–64 (2009). doi: 10.1007/978-1-59745-251-9_3; pmid: 19378139
30. O. Penn, E. Privman, G. Landan, D. Graur, T. Pupko, An alignment confidence score capturing robustness to guide tree uncertainty. *Mol. Biol. Evol.* **27**, 1759–1767 (2010). doi: 10.1093/molbev/msq066; pmid: 20207713
31. K. P. Schliep, phangorn: Phylogenetic analysis in R. *Bioinformatics* **27**, 592–593 (2011). doi: 10.1093/bioinformatics/btq706; pmid: 21169378
32. E. Paradis, J. Claude, K. Strimmer, APE: Analyses of Phylogenetics and Evolution in R language. *Bioinformatics* **20**, 289–290 (2004). doi: 10.1093/bioinformatics/btg412; pmid: 14734327
33. Y. Yu, R. M. Barnett, L. Nakhleh, Parsimonious inference of hybridization in the presence of incomplete lineage sorting. *Syst. Biol.* **62**, 738–751 (2013). doi: 10.1093/sysbio/syt037; pmid: 23736104
34. C. Than, D. Ruths, L. Nakhleh, PhyloNet: A software package for analyzing and reconstructing reticulate evolutionary relationships. *BMC Bioinformatics* **9**, 322 (2008). doi: 10.1186/1471-2105-9-322; pmid: 18662388
35. S. Sturtz, U. Ligges, A. Gelman, R2WinBUGS: A package for running WinBUGS from R. *J. Stat. Softw.* **12**, 1–16 (2005).

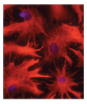
ACKNOWLEDGMENTS

This work was financed by Norwegian Research Council grant 199387 to Norwegian University of Life Sciences and Graminor AS to O.-A.O. The phylogenetic analyses were run, in part, on the Bioportal supercomputing facility at the University of Oslo. We thank C.-P. Antoine and two anonymous reviewers for valuable comments on the manuscript and S. Sæbø for help on implementation of the OpenBUGS analyses in R. Ortholog alignments, gene trees, and the OpenBUGS model used for multispecies coalescent modeling can be downloaded from Dryad (data available from the Dryad Digital Repository: http://doi.org/10.5061/dryad.f6c34).

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/345/6194/1250092/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S6
Tables S1 to S5
References (*36*–*45*)
IWGSC Author List

23 December 2013; accepted 20 May 2014
10.1126/science.1250092

EXTENDED PDF FORMAT
SPONSORED BY

Travel Grants Available
Neuroscience 2014
Application Deadline: July 31st
>Apply Now
R&D systems
a bio-techne brand
www.rndsystems.com

A chromosome-based draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome

The International Wheat Genome Sequencing Consortium (IWGSC)

Science **345**, (2014);
DOI: 10.1126/science.1251788

This copy is for your personal, non-commercial use only.

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by clicking here.

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines here.

The following resources related to this article are available online at www.sciencemag.org (this information is current as of August 4, 2014):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:
http://www.sciencemag.org/content/345/6194/1251788.full.html

Supporting Online Material can be found at:
http://www.sciencemag.org/content/suppl/2014/07/16/345.6194.1251788.DC1.html

A list of selected additional articles on the Science Web sites **related to this article** can be found at:
http://www.sciencemag.org/content/345/6194/1251788.full.html#related

This article **cites 155 articles**, 62 of which can be accessed free:
http://www.sciencemag.org/content/345/6194/1251788.full.html#ref-list-1

This article has been **cited by** 3 articles hosted by HighWire Press; see:
http://www.sciencemag.org/content/345/6194/1251788.full.html#related-urls

This article appears in the following **subject collections:**
Botany
http://www.sciencemag.org/cgi/collection/botany
Genetics
http://www.sciencemag.org/cgi/collection/genetics

Downloaded from www.sciencemag.org on August 4, 2014

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published weekly, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. Copyright 2014 by the American Association for the Advancement of Science; all rights reserved. The title *Science* is a registered trademark of AAAS.

A chromosome-based draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome

The International Wheat Genome Sequencing Consortium (IWGSC)

An ordered draft sequence of the 17-gigabase hexaploid bread wheat (*Triticum aestivum*) genome has been produced by sequencing isolated chromosome arms. We have annotated 124,201 gene loci distributed nearly evenly across the homeologous chromosomes and subgenomes. Comparative gene analysis of wheat subgenomes and extant diploid and tetraploid wheat relatives showed that high sequence similarity and structural conservation are retained, with limited gene loss, after polyploidization. However, across the genomes there was evidence of dynamic gene gain, loss, and duplication since the divergence of the wheat lineages. A high degree of transcriptional autonomy and no global dominance was found for the subgenomes. These insights into the genome biology of a polyploid crop provide a springboard for faster gene isolation, rapid genetic marker development, and precise breeding to meet the needs of increasing food demand worldwide.

Lists of authors and affiliations are available in the full article online.
Corresponding author: K. X. Mayer, e-mail: k.mayer@helmholtz-muenchen.de
Read the full article at http://dx.doi.org/10.1126/science.1251788

Ancient hybridizations among the ancestral genomes of bread wheat

Thomas Marcussen, Simen R. Sandve,* Lise Heier, Manuel Spannagl, Matthias Pfeifer, The International Wheat Genome Sequencing Consortium,† Kjetill S. Jakobsen, Brande B. H Wulff, Burkhard Steuernagel, Klaus F. X. Mayer, Odd-Arne Olsen

The allohexaploid bread wheat genome consists of three closely related subgenomes (A, B, and D), but a clear understanding of their phylogenetic history has been lacking. We used genome assemblies of bread wheat and five diploid relatives to analyze genome-wide samples of gene trees, as well as to estimate evolutionary relatedness and divergence times. We show that the A and B genomes diverged from a common ancestor ~7 million years ago and that these genomes gave rise to the D genome through homoploid hybrid speciation 1 to 2 million years later. Our findings imply that the present-day bread wheat genome is a product of multiple rounds of hybrid speciation (homoploid and polyploid) and lay the foundation for a new framework for understanding the wheat genome as a multilevel phylogenetic mosaic.

The list of author affiliations is available in the full article online.***Corresponding author.**
E-mail: simen.sandve@nmbu.no †The International Wheat Genome Sequencing Consortium (IWGSC) authors and affiliations are listed in the supplementary materials.
Read the full article at http://dx.doi.org/10.1126/science.1250092

Triticum monococcum

Triticum carthlicum

Ancestral wheat

Wheat varieties and species (shown) believed to be the closest living relatives of modern bread wheat (*T. aestivum*). Multiple ancestral hybridizations occurred among most of these species, many of which are cultivated, and along with *T. aestivum* represent a dominant source of global nutrition.

Triticum boeticum

Triticum polonicum L.

Triticum macha

Triticum dicoccoides var. *araraticum*

Published by AAAS

PHOTOS: SUSANNE STAMP, ERNST MERZ/ETH ZURICH

Genome interplay in the grain transcriptome of hexaploid bread wheat

Matthias Pfeifer, Karl G. Kugler, Simen R. Sandve, Bujie Zhan, Heidi Rudi, Torgeir R. Hvidsten, International Wheat Genome Sequencing Consortium,* Klaus F. X. Mayer, Odd-Arne Olsen†

Allohexaploid bread wheat (*Triticum aestivum* L.) provides approximately 20% of calories consumed by humans. Lack of genome sequence for the three homeologous and highly similar bread wheat genomes (A, B, and D) has impeded expression analysis of the grain transcriptome. We used previously unknown genome information to analyze the cell type–specific expression of homeologous genes in the developing wheat grain and identified distinct co-expression clusters reflecting the spatiotemporal progression during endosperm development. We observed no global but cell type– and stage-dependent genome dominance, organization of the wheat genome into transcriptionally active chromosomal regions, and asymmetric expression in gene families related to baking quality. Our findings give insight into the transcriptional dynamics and genome interplay among individual grain cell types in a polyploid cereal genome.

The list of author affiliations is available in the full article online. *The International Wheat Genome Sequencing Consortium (IWGSC) authors and affiliations are listed in the supplementary materials. †**Corresponding author. E-mail: odd-arne.olsen@nmbu.no**

Read the full article at http://dx.doi.org/10.1126/science.1250091

Structural and functional partitioning of bread wheat chromosome 3B

Frédéric Choulet,* Adriana Alberti, Sébastien Theil, Natasha Glover, Valérie Barbe, Josquin Daron, Lise Pingault, Pierre Sourdille, Arnaud Couloux, Etienne Paux, Philippe Leroy, Sophie Mangenot, Nicolas Guilhot, Jacques Le Gouis, Francois Balfourier, Michael Alaux, Véronique Jamilloux, Julie Poulain, Céline Durand, Arnaud Bellec, Christine Gaspin, Jan Safar, Jaroslav Dolezel, Jane Rogers, Klaas Vandepoele, Jean-Marc Aury, Klaus Mayer, Hélène Berges, Hadi Quesneville, Patrick Wincker, Catherine Feuillet

We produced a reference sequence of the 1-gigabase chromosome 3B of hexaploid bread wheat. By sequencing 8452 bacterial artificial chromosomes in pools, we assembled a sequence of 774 megabases carrying 5326 protein-coding genes, 1938 pseudogenes, and 85% of transposable elements. The distribution of structural and functional features along the chromosome revealed partitioning correlated with meiotic recombination. Comparative analyses indicated high wheat-specific inter- and intrachromosomal gene duplication activities that are potential sources of variability for adaption. In addition to providing a better understanding of the organization, function, and evolution of a large and polyploid genome, the availability of a high-quality sequence anchored to genetic maps will accelerate the identification of genes underlying important agronomic traits.

The list of author affiliations is available in the full article online.
***Corresponding author. E-mail: frederic.choulet@clermont.inra.fr**

Read the full article at http://dx.doi.org/10.1126/science.1249721

Published by AAAS

WHEAT GENOME

A chromosome-based draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome

The International Wheat Genome Sequencing Consortium (IWGSC)*†

An ordered draft sequence of the 17-gigabase hexaploid bread wheat (*Triticum aestivum*) genome has been produced by sequencing isolated chromosome arms. We have annotated 124,201 gene loci distributed nearly evenly across the homeologous chromosomes and subgenomes. Comparative gene analysis of wheat subgenomes and extant diploid and tetraploid wheat relatives showed that high sequence similarity and structural conservation are retained, with limited gene loss, after polyploidization. However, across the genomes there was evidence of dynamic gene gain, loss, and duplication since the divergence of the wheat lineages. A high degree of transcriptional autonomy and no global dominance was found for the subgenomes. These insights into the genome biology of a polyploid crop provide a springboard for faster gene isolation, rapid genetic marker development, and precise breeding to meet the needs of increasing food demand worldwide.

Rich in protein, carbohydrates, and minerals, bread wheat (*Triticum aestivum* L.) is one of the world's most important cereal grain crops, serving as the staple food source for 30% of the human population. Between 2000 and 2008, wheat production fell by 5.5% primarily because of climatic trends (*1*), and, in 5 of the past 10 years, worldwide wheat production was not sufficient to meet demand (*2*). With the global population projected to exceed 9 billion by 2050, researchers, breeders and growers are facing the challenge of increasing wheat production by about 70% to meet future demands (*3*, *4*). Concurrently, growers are facing rising fertilizer and other input costs, weather extremes resulting from climate change, increasing competition between food and nonfood uses, and declining annual yield growth (*5*). A rapid paradigm shift in science-based advances in wheat genetics and breeding, comparable to the first green revolution of the 1960s, will be essential to meet these challenges. As for other major cereal crops (rice, maize, and sorghum), new knowledge and molecular tools using a reference genome sequence of wheat are needed to underpin breeding to accelerate the development of new wheat varieties.

One key factor in the success of wheat as a global food crop is its adaptability to a wide range of climatic conditions. This is attributable, in part, to its allohexaploid genome structure, which arose as a result of two polyploidization events (Fig. 1). The first of these is estimated to have occurred several hundred thousand years ago and brought together the genomes of two diploid species related to the wild species *Triticum urartu* ($2n$ = $2x$ = 14; AA; $2n$ is the number of chromosomes in each somatic cell and $2x$ is the basic chromosome number) and a species from the Sitopsis section of *Triticum* that is believed to be related to *Aegilops speltoides* ($2n$ = 14; SS) (*6*). This hybridization formed the allotetraploid *Triticum turgidum* ($2n$ = $4x$ = 28; AABB), an ancestor of wild emmer wheat cultivated in the Middle East and *T. turgidum* sp. *durum* grown for pasta today. A second hybridization event between *T. turgidum* and a diploid grass species, *Aegilops tauschii* (DD), produced the ancestral allohexaploid *T. aestivum* ($2n$ = $6x$ = 42, AABBDD) (*6*, *7*), which has since been cultivated as bread wheat and accounts for over 95% of the wheat grown worldwide.

With 21 pairs of chromosomes, bread wheat is structurally an allopolyploid with three homeologous sets of seven chromosomes in each of the A, B, and D subgenomes. Genetically, however, it behaves as a diploid because homeologous pairing is prevented through the action of *Ph* genes (*8*). Each of the subgenomes is large, about 5.5 Gb in size and carries, in addition to related sets of genes, a high proportion (>80%) of highly repetitive transposable elements (TEs) (*9*, *10*).

The large and repetitive nature of the genome has hindered the generation of a reference genome sequence for bread wheat. Early work focused primarily on coding sequences that represent less than 2% of the genome. Coordinated efforts generated over 1 million expressed sequence tags (ESTs), 40,000 unigenes (www.ncbi.nlm.nih.gov/dbEST/dbEST_summary.html), and 17,000 full-length complementary DNA (cDNA) sequences (*11*). These resources have enabled studies of individual genes and facilitated the development of microarrays and marker sets for targeted gene association and expression studies (*12*–*14*). At least 7000 ESTs have been assigned to chromosome-specific bins (*15*), providing an initial view of subgenome localization and chromosomal organization and facilitating low-resolution mapping of traits. More recently, high-throughput low-cost sequencing technologies have been applied to assemble the gene space of *T. urartu* (*16*) and *Ae. tauschii* (*17*), two diploid species related to bread wheat (Fig. 1). About 60,000 genic sequences were also putatively assigned to the bread wheat A, B, or D subgenomes by using assembled Illumina (Illumina, Incorporated, San Diego, CA) sequence data for *Triticum monococcum* and *Ae. tauschii* and cDNAs from *Ae. speltoides* to guide gene assemblies of five-fold whole-genome sequence reads from *T. aestivum* 'Chinese Spring' (*18*). These resources have contributed information about the genes of hexaploid wheat and its wild diploid relatives and have underpinned the

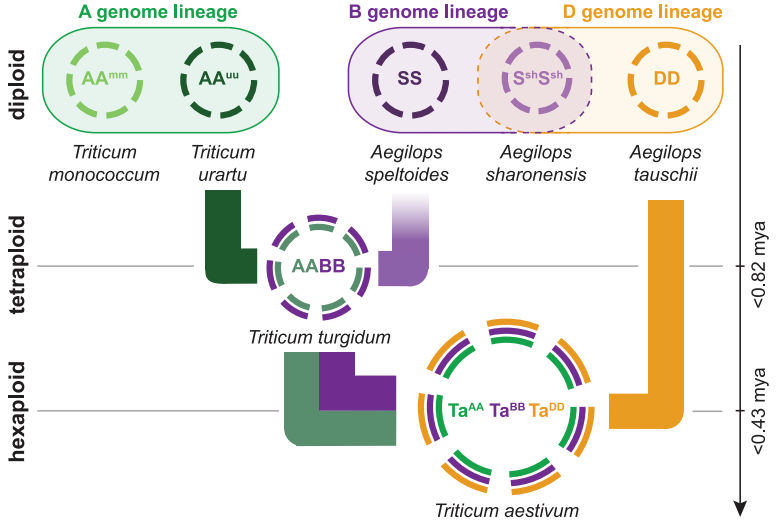


Fig. 1. Schematic diagram of the relationships between wheat genomes with polyploidization history and genealogy. Names and nomenclature for the genomes are indicated within circles that provide a schematic representation of the chromosomal complement for each species. Time estimates are from Marcussen *et al.* (*45*). mya, million years ago.

*All authors with their affiliations appear at the end of this paper.
†Corresponding author: K. X. Mayer (k.mayer@helmholtz-muenchen.de)

development of large sets of single-nucleotide polymorphism (SNP) markers (*19*–*21*). To date, however, relatively little is known about the position and distribution of genes on each of the bread wheat chromosomes and their evolution during the polyploidization events that resulted in the emergence of the hexaploid species.

Survey sequencing the bread wheat genome

We used aneuploid bread wheat lines derived from double ditelosomic stocks of the hexaploid wheat cultivar Chinese Spring (*22*) to isolate, sequence, and assemble de novo each individual chromosome arm [except for 3B, which was isolated and sequenced as a complete chromosome (*23*)]. This approach reduced the complexity of assembling a highly redundant genome and enabled the differentiation of genes present in multiple copies and highly conserved homologs. Each chromosome arm, representing between 1.3 and 3.3% of the genome (*24*), was purified by flow-cytometric sorting and sequenced to a depth of between 30× and 241× with Illumina technology platforms (*25*). The paired end sequence reads were assembled with the short-read de novo assembly tool ABySS (*25*, *26*). A high proportion of reads assembled into contigs of repetitive sequence less than 200 base pairs (bp) and were excluded from the final assembly of 10.2 Gb. The quality of the assemblies and purity of chromosome arm preparations were assessed by using alignment to bin-mapped ESTs (*15*) and to the virtual barley genome (*27*). Summary statistics for the chromosome arm assemblies are shown in Tables 1 to 3. Compared with cytogenetically estimated chromosome sizes (*24*), the sequence assemblies represent 61% of the genome sequence, with the L50 of repeat-masked assemblies ranging from 1.7 to 8.9 kb.

Repetitive DNA

We assessed the TE and sequence repeat space across the whole wheat genome and compared the repeat content of the A, B, and D subgenomes (*25*). From the frequency of mathematically defined repeats (MDRs; 20mers) (*28*), we estimated that 24 to 26% of the sequence reads contain high copy number repeats, represented by 20mers with more than 1000 copies. In total, 81% of raw reads and 76.6% of assembled sequences contained repeats, the latter showing reduced representation of Gypsy long terminal repeat (LTR) retrotransposons, as well as Mutator and Mariner-type DNA transposons.

Analysis of the distribution of transposons across the three subgenomes revealed that class I elements (retroelements) were more abundant in the A genome chromosomes relative to B or D (A > B > D), whereas class II elements (DNA transposons) showed the reverse (D > B > A). The most pronounced differences were observed between deteriorated and thus unclassifiable LTR retrotransposons, which showed a gradient of abundance across the subgenomes (A > D > B) distinct from other class I or class II elements. We hypothesize that unclassifiable LTR retrotransposons represent older (and thus more deteriorated) elements that were modified through polyploidization and ongoing TE amplification or degeneration. Assuming the amplification/degeneration dynamics are similar within each genome, the distribution of LTR retrotransposons across the three subgenomes suggest that the B genome progenitor contained a lower number of LTR retroelements and that transposon activity post-polyploidization has introduced a higher proportion of more recent amplifications into the B genome.

We observed a substantial reduction (down to 19.6%) in the TE content associated with the 0.8% (615 Mb) of the chromosomal survey sequences (CSSs) representing contigs containing high-confidence genes (for definition see below) (*25*). The analysis revealed a marked depletion of all class I elements in the neighborhood of genes, with the exception of non-LTR retrotransposons, which were enriched twofold. CACTA transposons accounted for the greatest proportion of the observed 67% reduction in class II elements, whereas minor components, especially Harbinger and miniature inverted-repeat TEs, were enriched. Selective exclusion of high-copy transposons that undergo epigenetic silencing and reduce expression by heterochromatin spreading (*29*) may result in depletion of repeat element types in the vicinity of genes.

miRNAs

A total of 270 different putative microRNA molecules (miRNAs) (49 not previously reported)

Table 1. Sequencing, assembly, and GenomeZipper statistics for wheat A genome chromosome arms. Sequence indicates the total assembled sequence (>200 bp); number of contigs is after filtering of highly repetitive sequence assemblies; syntenic loci is the number of gene loci anchored to reference gene; and the last row is the number of total anchored gene loci. Blank entries in all tables indicate data not applicable; fl-cDNA, full-length cDNA; nonred., nonredundant.

	1AS	1AL	2AS	2AL	3AS	3AL	4AS	4AL	5AS	5AL	6AS	6AL	7AS	7AL	Σ
Assembly															
Chromosome size (Mbp)	275	523	391	508	360	468	317	539	295	532	336	369	407	407	5,727
Sequence (Mbp)	178.1	250	255.2	328.2	201.8	247.2	282.3	362	198.8	318.1	219.2	214.4	198	252.4	3,505.7
Coverage (*x*-fold)	0.65	0.48	0.65	0.65	0.56	0.53	0.89	0.67	0.67	0.60	0.65	0.58	0.49	0.62	0.62
L50 (bp)	2,242	2,639	2,398	2,688	1,404	1,346	2,782	3,053	3,509	2,078	2,669	2,154	1,470	2,271	
Repeat															
No. of contigs	34,793	26,746	34,722	45,893	33,943	43,823	32,079	64,364	19,719	47,572	28,041	34,030	44,175	35,586	542,486
L50	4,769	6,369	6,678	6,677	3,846	3,789	7,499	6,601	8,713	5,355	7,091	6,589	4,397	5,849	
GenomeZipper															
No. of markers	147	380	139	278	106	332	167	200	150	309	174	286	169	278	3,115
No. of wheat fl-cDNAs	95	241	162	258	134	240	153	189	54	231	94	181	178	155	2,365
No. of nonred. contigs	937	1,750	1,673	2,499	1,323	2,300	848	2,613	574	2,495	811	1,422	2,100	1,600	22,945
No. of syntenic gene loci	544	1,515	1,155	1,816	850	1,628	842	1,642	405	1,821	647	1,073	1,228	1,049	16,215
No. of anchored gene loci	649	1,811	1,262	2,032	929	1,864	948	1,777	522	2,050	794	1,279	1,349	1,269	18,535
POP-Seq Positioning															
No. of contigs	38,940	45,649	34,853	32,941	31,094	49,586	25,068	27,248	5,578	35,333	28,234	30,828	31,628	32,435	449,415
No. of anchored gene loci	972	1,720	1,452	1,913	788	1,302	883	1,702	137	1,579	1,145	1,305	1,305	1,094	17,297
No. of anchored gene loci	618	1,257	1,408	1,903	769	1,469	778	1,116	678	2,432	995	1,458	1,405	1,711	17,997

were identified corresponding to 98,068 predicted miRNA-coding loci (*25*). Only 1668 loci (1.7%) evidenced expression on the basis of publicly available ESTs and of RNA sequencing (RNA-seq) data reported in this work, consistent with previous analyses in wheat (*30*, *31*). Similarly, we observed that class II DNA transposons, specifically TcMar transposons, were predominantly found in miRNAs. For 87% of the putative miRNA-coding loci, at least one putative target gene was identified in the wheat CSS. A total of 6615 predicted miRNA-coding sequences (44 with evidence of expression) were characterized by at least one mature sequence and one target site covered by the same repeat element. This suggests that an active miRNA could arise when an advantageous regulatory niche evolves from a series of random

Table 2. Sequencing, assembly, and GenomeZipper statistics for wheat B genome chromosome arms. Sequence indicates the total assembled sequence (>200 bp); number of contigs is after filtering of highly repetitive sequence assemblies; syntenic loci is the number of gene loci anchored to reference gene; and the last row is the number of total anchored gene loci.

	1BS	1BL	2BS	2BL	3B	4BS	4BL	5BS	5BL	6BS	6BL	7BS	7BL	Σ
							Assembly							
Chromosome size (Mbp)	314	535	422	506	993	391	430	290	580	415	498	360	540	6,274
Sequence (Mbp)	212.8	299.4	292	404.5	638.6	308.2	248.7	174.5	415.2	210.2	257.4	206.1	259.6	3,927.2
Coverage (*x*-fold)	0.68	0.56	0.69	0.80	0.64	0.79	0.58	0.60	0.72	0.51	0.52	0.57	0.48	0.63
L50 (bp)	3,287	3,120	3,711	2,941	2,655	3,463	1,974	3,315	2,924	2,366	2,031	2,428	1,556	
							Repeat							
No. of contigs	26,050	29,783	35,743	75,879	75,022	38,515	46,576	18,001	75,887	29,566	35,727	24,119	58,554	569,422
L50	7,413	7,151	8,069	6,890	6,855	8,755	5,883	7,365	7,537	4,972	4,824	6,435	4,144	
							GenomeZipper							
No. of markers	78	348	278	428	500	46	145	167	404	217	245	140	198	3,194
No. of wheat fl-cDNAs	78	219	155	268	479	97	170	66	360	88	147	109	137	2,373
No. of nonred. contigs	776	1,927	1,859	3,079	5,011	893	1,634	576	3,296	915	1,525	1,172	1,890	24,553
No. of syntenic gene loci	485	1,485	1,181	1,973	3,123	788	1,155	426	2,315	565	1,003	733	1,050	16,282
No. of anchored gene loci	546	1,745	1,388	2,265	3,490	819	1,243	565	2,600	728	1,177	838	1,203	18,607
							POP-Seq Anchoring							
No. of contigs	31,038	50,219	33,603	54,522	99,341	50,927	41,135	19,794	49,140	30,962	38,064	48,514	50,397	597,656
No. of anchored gene loci	956	1,881	1,588	2,389	3,772	1,365	1,433	727	2,857	831	996	1,055	1251	21,101

Table 3. Sequencing, assembly, and GenomeZipper statistics for wheat D genome chromosome arms. Sequence indicates the total assembled sequence (>200 bp); number of contigs is after filtering of highly repetitive sequence assemblies; syntenic loci is the number of gene loci anchored to reference gene; and the last row is the number of total anchored gene loci.

	1DS	1DL	2DS	2DL	3DS	3DL	4DS	4DL	5DS	5DL	6DS	6DL	7DS	7DL	Σ
								Assembly							
Chromosome size (Mbp)	224	381	317	412	321	450	232	417	259	491	324	389	381	347	4,937
Sequence (Mbp)	128.2	254.4	166	261.6	145.4	186.5	142.1	347.6	148	236.8	156.6	199.8	209.1	222.9	2,805
Coverage (*x*-fold)	0.57	0.67	0.52	0.63	0.45	0.41	0.61	0.83	0.57	0.48	0.48	0.51	0.55	0.64	0.57
L50 (bp)	2,850	2561	1241	701	515	967	3278	1013	2,353	2,647	4,297	2,077	1,967	3,638	
								Repeat							
No. of contigs	17,725	35,770	43,044	110,446	46,795	69,259	18,245	197,398	22,449	34,622	16,077	26,236	36,701	26,737	701,504
L50	6,622	6,297	4,635	3,247	1,697	2941	7428	1855	5945	7049	8,904	6,821	5,031	7,399	
								GenomeZipper							
No. of markers	258	653	457	739	379	633	269	498	225	744	297	411	579	515	6,657
No. of wheat fl-cDNAs	89	251	177	323	128	244	130	255	99	375	103	208	200	212	2,794
No. of nonred. contigs	968	2,797	3,023	5,804	2,933	3,712	1,231	3,174	890	3,436	973	1,923	3,006	2,083	35,953
No. of syntenic gene loci	474	1,483	1,197	2,141	799	1,575	779	1,277	454	2,073	538	1,117	1,222	1,099	16,228
No. of anchored gene loci	642	1,882	1,475	2,542	1,051	1,923	912	1,582	598	2,482	758	1,347	1,592	1,423	20,209
								POP-Seq Anchoring							
No. of contigs	7,686	24,149	24,652	31,359	26,447	37,874	14,198	23,842	14,458	29,604	18,701	23,763	41,796	31,832	350,361

TE insertions and may represent a means by which a network of putative miRNAs and target genes may develop, even before miRNA activation (*32*).

Protein-coding genes

Annotation of protein-coding gene sequences in the CSS assemblies had its basis in comparisons to annotated genes in related grasses [*Brachypodium distachyon* (*33*), *Oryza sativa* (*34*), *Sorghum bicolor* (*35*), and *Hordeum vulgare* (*27*)], as well as publically available wheat full-length cDNAs (fl-cDNAs) (*11*) and RNA-seq data generated from five tissues of a Chinese Spring cultivar at three different developmental stages. Briefly, the reference grass coding sequences and wheat transcript resources were mapped separately to assembled CSS contigs, and the alignments were merged to define the exact coordinates of gene loci, alternative splicing forms, and transcripts with no similarity to related grass genes (*25*).

This analysis identified 976,962 loci with 1,265,548 distinct splicing variants. A total of 133,090 loci showing homology to related grass genes were classified as high confidence (HC) gene calls. These were further subdivided into four groups (HC1 to HC4) on the basis of the proportion of the length of the reference gene covered by a predicted locus. Of these, 124,201 (93.3%) genes were annotated on individual chromosome arm sequences, and the remaining 6.7% corresponded to wheat transcripts, which were not detected in the CSS assemblies (Fig. 2A). In total, 55,249 (44%) of the loci assigned to chromosomes were classified as HC1, that is, representing functional genes spanning at least 70% of the length of the supporting evidence (Table 4). The remaining 56% of HC genes comprised genes that were fragmented in the assembly and thus could only be partially structurally defined or were classified as gene fragments and pseudogenes. We expect that many of these will be merged as further sequencing improves the coverage and quality of genic sequences. On the basis of the level of completion of the assembly and the detection rate of HC1 genes (*25*), we estimated that the wheat genome contains 106,000 functional protein-coding genes. This supports gene number estimates ranging between 32,000 and 38,000 for each diploid subgenome in hexaploid wheat and is consistent with findings in related diploid species (*16*–*18*, *20*, *36*).

Consistent with observations of high levels of non–protein-coding loci in both plants (*27*, *37*) and animals (*38*), 890,576 loci did not share any, or only low, similarity with related grass genes. Loci with low sequence similarity (88,998) were defined as low-confidence (LC) genes, and the remainder were classified as repeat-associated, noncoding, or non–homology-supported loci (*25*). More than 96% of public wheat ESTs (HarvEST) mapped to the CSS gene sets (BLASTN; *E* value $<10^{-10}$), including 89% that correspond to HC gene-coding loci, demonstrating that the CSS assemblies contain a high representation of the current gene inventory of the bread wheat genome.

Our analysis revealed that 49% of the HC genes exhibit alternative splicing (AS) with an average of 2.6 transcripts per locus. This may be an underestimation, because 69% of the most complete gene loci (HC1) were alternatively spliced with an average of 3.5 transcripts per locus. Evidence that additional AS variants will be identified has already emerged from a preliminary assessment of gene structure prediction using proteomics analyses. In a study of 63 genes, 50 (81%) structures were confirmed, 8 (13%) provided evidence for alternative gene structures, whereas 5 were absent in the structural gene calls. Extrapolating these data to the whole genome, we estimate that hexaploid bread wheat encodes more than 300,000 distinctive protein-coding transcripts. The proportion of genes exhibiting AS appeared to be similar in all three subgenomes and is consistent with the transcriptional complexity reported for plant species such as *Arabidopsis thaliana* (*39*) and *H. vulgare* (*27*).

Gene distribution and order

Analysis of the gene distribution across the three subgenomes revealed a higher number of gene loci on the B subgenome (44,523; 35%) compared with the A and D subgenomes, which contained 40,253 (33%) and 39,425 (32%), respectively (Fig. 2A). This distribution was not consistent at the chromosomal level. For example, the gene distribution across homeologous group 3 chromosomes is 30% 3A, 42% 3B, and 28% 3D, whereas in homeologous group 7 the D genome contains the highest proportion of genes. These observations may reflect preexisting differences in the subgenomes before polyploidization or indicate that drivers determining the composition of the genome do not act at the subgenome level but regionally.

Up to 2.4-fold variation in gene density was observed on different chromosome arms, ranging from 4.4 loci per Mb (5AS) up to 10.4 loci per Mb (2DL) (Fig. 2B). Consistent with observations in rye (*40*) and the complete sequence of wheat chromosome 3B (*23*), on average 53.2% of the

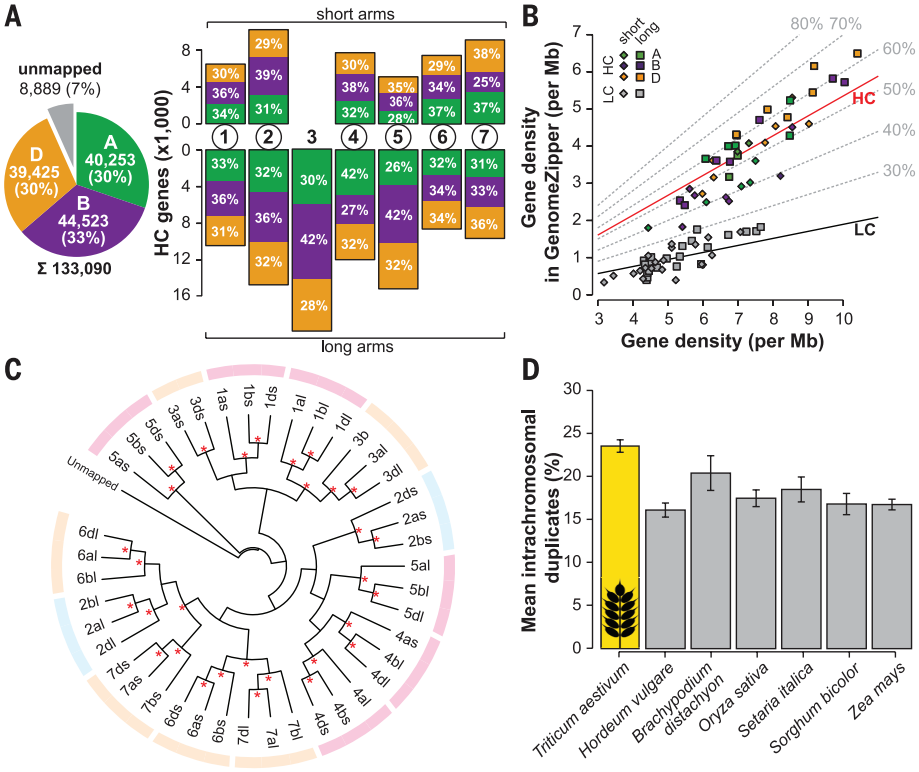


Fig. 2. Gene content, density, synteny, structural conservation, and tandemly duplicated genes. (**A**) Total number of HC bread wheat genes identified on the A (green), B (purple), and D (orange) subgenomes (left) and their distribution on individual chromosome arms or chromosomes (in the case of group 3) (right). (**B**) Syntenic conservation of HC and LC genes for each chromosome arm defined by the ratio of the number of genes anchored in the GenomeZipper and the number of annotated genes normalized per Mb of physical chromosome(-arm) size. Solid lines visualize average syntenic conservation for LC (black) and HC (red) genes, and dashed lines give isochores for different percentages of synteny. (**C**) Conservation of gene family composition between single chromosome arms. Color-coding in the outer ring indicates relatedness of the respective branches (A/D > B, light orange; A/B > D, light blue; B/D > A, light red). Red asterisks mark edges with boot-strapping values > 0.95. (**D**) Proportion of lineage-specific, intrachromosomally duplicated genes in the wheat genome compared with other grass genomes. Error bars indicate deviations among individual chromosomes.

Table 4. Characteristics of HC bread wheat genes. Distinct exons means that exons of two or more transcripts were counted once if they had identical start and stop positions; mean transcripts and mean exons are transcripts per locus and exons per locus, respectively; the second mean exons row shows exons per transcript.

	HC1	HC2	HC3	HC4	Σ
Gene loci	55,249	14,367	15,475	39,110	124,201
Single exon	9,181 (17%)	3,230 (22%)	4,906 (32%)	20,375 (52%)	37,692 (30%)
Multiple exon	46,068 (83%)	11,137 (78%)	10,569 (68%)	18,735 (48%)	86,509 (70%)
Alternatively spliced	38,059 (69%)	7,916 (55%)	6,465 (42%)	8,728 (22%)	61,168 (49%)
Mean size (bp)	3,319	2,204	1,608	901	2,216
Transcripts	194,624	37,116	31,957	61,450	325,147
Mean transcripts	3.52	2.58	2.07	1.57	2.62
Distinct exons	538,250	94,864	74,630	117,530	825,274
Mean exons	9.74	6.60	4.82	3.01	6.64
Mean exons^3	6.29	4.45	3.52	2.56	5.1
Mean size (bp)	321	315	314	281	314

HC genes were located on syntenic chromosomes compared to *B. distachyon* (Bd), *O. sativa* (Os), and *S. bicolor* (Sb). The average level of synteny for genes located on the D genome chromosomes (58%) was higher than the average for those on the A (51%) and the B (50%) chromosomes. Sequence conservation in LC genes is low, and, in comparison to HC genes, reduced syntenic conservation is observed. Thus, although the majority of LC genes are likely to result from the frequent generation of gene fragments by double-strand repair mechanisms or are deteriorated (pseudo)genes that were fragmented after the divergence from the other sequenced grass genomes (*10*), the retained synteny to other grass genomes suggests that some LC genes may be functional.

To determine the extent of gene conservation across homeologous chromosomes, we clustered the HC genes into protein families by sequence similarity (Fig. 2C) (*25*). With the exception of chromosome 4AL, the genes on all chromosome arms clustered with their corresponding homologs. The pattern of clustering observed for 4A is consistent with a known pericentromeric inversion and two translocations of segments from chromosome arms 5AL and 7BS (*41*, *42*). All possible cluster topologies were found between genes on the A, B, and D genomes. Overall, the patterns of conservation suggest that the gene content of the A and B homeologous chromosomes is more similar to the D genome chromosomes than to each other. This observation contradicts a model of bifurcating evolutionary relationships between the A, B, and D genomes but is consistent with models of interlineage hybridization (i.e., reticulate evolution) in the Triticeae (*43*, *44*) and corroborate phylogenomic analyses that suggest that the D genome is a product of homoploid hybrid speciation between A and B genome ancestors >5 million years ago (*45*). Although the potential for preexisting differences needs to be considered, the preservation of gene copies in each of the A, B, and D genomes provides evidence for their structural autonomy, a likely consequence of independent pairing during meiosis (*46*). A high degree of subgenome autonomy was also reflected in the observed patterns of gene expression (see below).

We used two independent but complementary approaches to generate an order for the many small contigs that comprise the chromosome arm assemblies (*25*). The GenomeZipper approach (*47*) combines the syntenic conservation of gene order in grasses (*48*) and the known gene orders of fully sequenced grass genomes (*33*–*35*) with high-density SNP-based genetic maps (*21*, *49*) to create a virtual gene order in wheat. The number of genes anchored per chromosome (chr.) ranged from 2125 (chr. 6B) to 4404 (chr. 2D) (Table 1). Overall, the GenomeZipper inferred positions of 21,221, 22,051, and 22,813 genes, respectively, in the A, B, and D genomes. To complement this, the POPSEQ approach (*50*) was used to build an ultradense genetic map comprising 13.3 million SNPs identified after shallow-coverage whole-genome sequencing of 90 doubled haploid individuals of the synthetic W7984 × Opata M85 population (*51*). This map assigned a partially overlapping set of 17,297, 21,101, and 17,997 HC genes, respectively, to the individual chromosomes of the A, B, and D genomes. The POPSEQ genetic map showed concordance with the gene assignments to flow-sorted chromosomes (99.4%) and the GenomeZipper (99.8%). The two inferred gene orders along chromosomes were also largely collinear (Spearman's correlation coefficient = 0.85). From both anchored data sets, we were able to position a nonredundant set of 75,183 HC genes on the 21 chromosomes of bread wheat by genetic mapping and/or syntenic conservation.

Gene duplication is frequently observed in plant genomes, arising from polyploidization or through tandem or segmental duplication associated with replication (*52*). For each wheat chromosome, the percentage of genes that have undergone lineage-specific intrachromosomal duplication was determined with OrthoMCL (*53*). By using the HC1 genes, we estimated that between 19.1% (chr. 7B) and 29.7% (chr. 2B) (23.6% average for all chromosomes) of the genes are duplicated on each chromosome (*25*). Comparison of the number of duplicated genes identified by this analysis for chr. 3B (25.3% of HC1 genes) with the 3B reference pseudomolecule (37% duplicated genes) (*23*) indicated that we are likely underestimating the number of duplicated genes. This is due to the fragmented nature of the assemblies obtained from whole-genome or chromosome-shotgun sequences that collapse highly conserved duplicates. No significant differences in the proportion of duplicates were observed between the three subgenomes (χ^2 test, $\chi^2 = 3.8$, $P = 0.15$).

For each chromosome, an average of 73% of the duplicates are located on one of the chromosome arms, suggesting that they may be tandem duplicates that arise through unequal crossing-over and replication-dependent chromosome breakage (*54*) or through the activity of transposable elements. When compared with the percentage of intrachromosomal duplicates found in rice, sorghum, barley, maize, and foxtail millet (17 to 20%) (*27*, *33*–*35*, *55*, *56*), the proportion of gene duplications in wheat was significantly higher (Fig. 2D; Tukey's honest significant difference, pairwise $P < 0.007$).

Comparisons with related species

We assembled sequence data from seven species related to progenitors of the bread wheat A, B, and D subgenomes (*25*). Illumina whole-genome sequence data and assemblies were generated from two tetraploid wheat cultivars (AABB) *T. turgidum* 'Cappelli' (originating from Italy) and *T. turgidum* 'Strongfield' (originating from Canada) as well as from the diploid genome of *Ae. speltoides* (SS). These data were combined with whole-genome sequence data from *T. urartu* (AA^{uu}) (*16*), *T. monococcum* (AA^{mm}), *Ae. tauschii* (DD) (*17*), and *Aegilops sharonensis* ($S^{sh}S^{sh}$). For the unannotated genomes of *T. turgidum*, *T. monococcum*, *Ae. speltoides*, and *Ae. sharonensis*, proteins of annotated grass genomes (*27*, *33*, *35*, *57*) and *T. aestivum* gene models were projected on the sequence assemblies.

Genes and gene families in the hexaploid, tetraploid, and diploid genomes were then compared to assess the dynamics of gene retention or loss after polyploidization and to define the core wheat genes. When comparing the sizes of gene families in *Ae. tauschii* (*17*) and *T. urartu* (*16*) diploid genomes with the individual subgenomes of hexaploid wheat (Fig. 3, A and B), we found that gene loss mainly affected genes belonging to expanded families, consistent with previous observations (*18*). In contrast, singletons (i.e., genes without paralogous copies within the same genome) were not usually subject to gene loss after polyploidization. Pronounced variations of gene copy retention or loss patterns were observed depending on the gene family considered.

Highly similar gene retention rates were found for all bread wheat subgenomes in comparison to *Ae. tauschii* and *T. urartu* [0.91 (A), 0.94 (B), and 0.89 (D) versus *Ae. tauschii* and 0.91 (A), 0.96 (B), and 0.91 (D) versus *T. urartu*] (Fig. 3, A and B). The extent of gene loss in the D subgenome, the most recent addition to the hexaploid genome, appeared slightly lower than the more ancient A and B subgenomes. Thus, as observed for

the gene content and structural similarities between individual chromosome arms, we found no evidence for a gradual gene loss induced by polyploidization. This may indicate that gene loss occurred rapidly after polyploid formation, followed by stabilization of gene content consistent with observations in newly created polyploids (*58*, *59*) and gene retention in cotton (*60*).

We conducted a clustering analysis of gene families and determined the number of genes in the bread wheat subgenomes that have an ortholog in the genomes from the A genome lineage (*T. urartu* and *T. monococcum*), the closest known relatives for the B lineage (*Ae. sharonensis* and *Ae. speltoides*), the D lineage (*Ae. tauschii*), as well as in the tetraploid *T. turgidum* genome (Fig. 3C). We found that the A, B, and D subgenomes contain very similar proportions of genes (60.1 to 61.3%) with orthologs in all the related diploid genomes. We also estimated the contribution of unique genes of the three subgenomes to the bread wheat genome. Because the absence of a particular gene in a single species could be due to incomplete sequence coverage or assembly errors, only lineage-specific gene family absence was considered in the analysis. Only a small fraction of the genes (1.3 to 1.7%) were specific to the A, B, or D lineages, demarcating the likely upper estimate of unique genes or gene families added to the bread wheat gene complement by the individual subgenomes.

High sequence similarity between genes in the bread wheat subgenomes impedes efficient marker development and the identification of nonsynonymous sequence variations that can potentially affect gene or protein functionality. We delineated single-nucleotide variations (SNVs) between the bread wheat genes and the diploid and tetraploid related genomes and reconstructed phylogenetic relationships by using unrooted parsimony (Fig. 4A) (*25*). In total, 11,435 SNVs within 6498 genes were specific to bread wheat and thus have likely been introduced after the second polyploidization event. Although most relationships support the known phylogeny of wheat, *Ae. sharonensis* was placed closer to the bread wheat D subgenome and *Ae. tauschii* than to *Ae. speltoides* and the B genome branch. This suggests that the Sitopsis group, which includes *Ae. sharonensis* and *Ae. speltoides*, is deeply furcated and related to both D and B genome branches.

The potential impact of all SNVs detected on proteins was measured by using Grantham amino acid substitution matrix scores (*25*, *61*). Most of the substitutions (80.8%) in gene sequences were conservative or moderately conservative and were randomly distributed across all chromosomes. However, bread wheat genes contained a higher proportion of substitutions with a predicted large impact on the protein functionality (i.e., moderately radical and radical changes) compared with their closest diploid or tetraploid relatives. This points to gene redundancy in hexaploid bread wheat enabling accelerated sequence evolution and potentially the evolution of novel protein functions.

We used the bread wheat gene annotation to analyze the introduction of likely premature stop codons in diploid and tetraploid related genomes as a measure for the rate and degree of pseudogenization (Fig. 4B). Using only the highest confidence genes (HC1), 290 (1.6%; *T. turgidum* A genome versus *T. aestivum* A genome) to 636 (3.6%; *Ae. sharonensis* versus *T. aestivum* D genome) gene loci had characteristics of pseudogenization in the respective related diploid genomes compared with the respective bread wheat A, B, and D subgenomes. Most of these likely pseudogenized loci were specific to the respective genomes, although overlapping candidate pseudogenized loci were also observed. However, the numbers of genes in these categories were small, ranging from 0.1 to 0.7%. Similar inferred pseudogenization rates were found in the A and B subgenomes of *T. turgidum* [290 (1.6%) in the A genome and 395 (2.0%) in the B genome, respectively], indicating no preferential pseudogenization or gene loss in any of the subgenomes. The number of pseudogenes observed in the D genome was similar to that of the A and B subgenomes and their diploid relatives, suggesting a rapid elimination process for pseudogenes. These findings are consistent with those from other plants, notably among *Arabidopsis* ecotypes (*62*), and smaller-scale analysis of pseudogenization dynamics within the bread wheat genome (*63*).

Earlier studies showed a high degree of gene sequence similarity between A, B, and D bread wheat subgenomes and their related diploid species (*6*). We analyzed the sequence conservation in bread wheat chromosomes compared to their diploid and tetraploid relatives to test for intergenomic translocations or introgressions (Fig. 4C). The sequences of genes were highly conserved, exceeding 99% identity, between the hexaploid subgenomes and their respective diploid relatives. High levels of conservation, averaging 97%, were also found between the A, B, and D lineages.

No gradients in sequence conservation were apparent along the chromosomes for the most closely related genomes. However, when comparing more distant genomes (e.g., *T. aestivum* D genome versus *T. urartu*), higher levels of sequence conservation were observed in genes located in proximal, pericentromeric, and centromeric regions. These results are consistent with findings for the 3B pseudomolecule analysis that demonstrated a partitioning of the chromosome with variable telomeric regions and a more conserved central chromosomal region (*23*). The most pronounced deviation in gene sequence similarity from the overall distribution is found for chr. 4A, which has undergone a recent inversion and translocations from chrs. 5A and 7B (*41*, *42*) (Fig. 4C). Other, smaller regions showing altered similarity profiles were also observed on other chromosomes (e.g., chrs. 2A and 7B) (*25*) suggesting the presence of further small translocations or introgressions that may have occurred after hybridization.

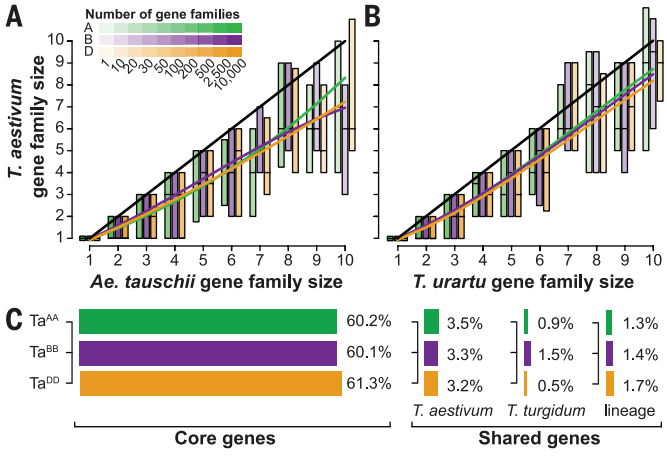


Fig. 3. Gene conservation and the wheat pan- and core genes. (**A** and **B**) Relationship between gene family sizes in diploid *Ae. tauschii* (A) and *T. urartu* (B) and each subgenome of hexaploid bread wheat (colors as in Fig. 2A). Boxes visualize the lower and upper quartiles of gene family sizes. Color intensity indicates the number of gene families in the respective bin. The black line shows a 1:1 gene copy number relationship for bread wheat, *Ae. tauschii*, and *T. urartu*, and colored lines show the regression fit for observed gene family size in the wheat subgenomes. (**C**) Percentages of genes of the bread wheat subgenomes that show significant sequence similarity to other genomes: Core genes correspond to genes with hits to all subgenomes as well as to *T. turgidum* and all diploid related progenitor genomes; shared genes–*T. aestivum* are genes with hits to any other *T. aestivum* subgenome but not to *T. turgidum* or any of the closest diploid relatives; shared genes–*T. turgidum* correspond to genes with hits to *T. turgidum* but not to any of the closest diploid relatives; shared genes–lineage, with hits to the subgenome's closest relative genome but not to *T. turgidum* or any of the other closest related genomes.

Hexaploid genome phylogeny

To further test the relatedness of the A, B, and D subgenomes across the entire wheat genome, we used syntenic gene alignments to estimate maximum likelihood phylogenetic trees. We obtained 2269 trees and analyzed them for topological variation. Across all chromosome groups, 40, 35, and 25% of the gene phylogenies supported AD,

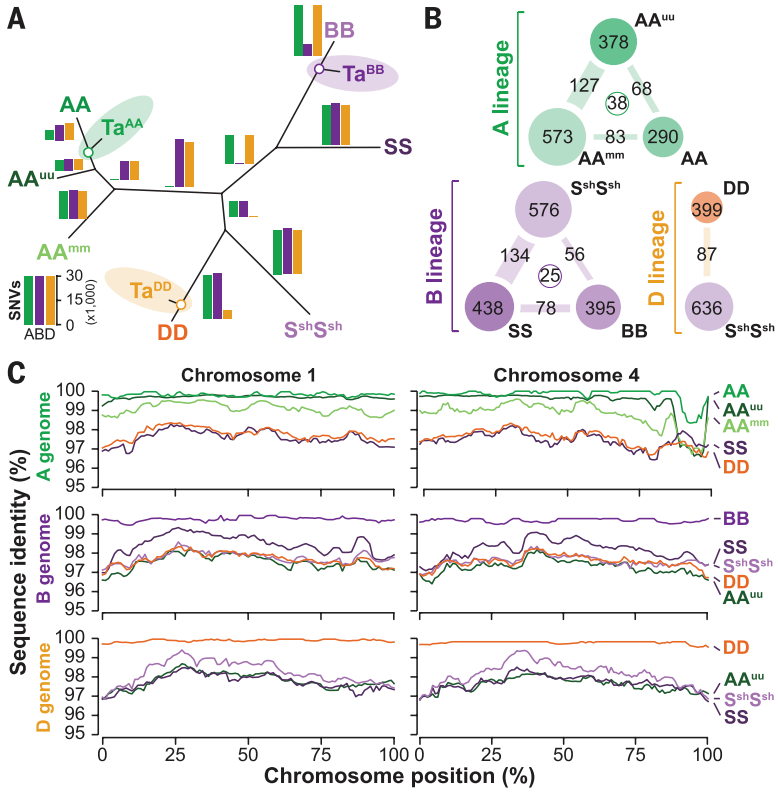


Fig. 4. Molecular evolution of the wheat lineage. SNVs were identified for coding sequences of bread wheat genes (TaAA, TaBB, and TaDD) against diploid *T. monococcum* (AA^{mm}), *T. urartu* (AA^{uu}), *Ae. speltoides* (SS), *Ae. sharonensis* (SshSsh), *Ae. tauschii* (DD), and tetraploid *T. turgidum* (AABB). (**A**) Unrooted phylogeny constructed on the basis of SNVs between bread wheat and its diploid or tetraploid relatives. The respective number of SNVs in each phylogenetic internodes is indicated with bar charts (scale at bottom left corner); colors indicate the respective bread wheat subgenome as in Fig. 2A. (**B**) Genes with stop codons in the respective related diploid genomes in comparison to the bread wheat A, B, and D subgenomes. Numbers in node connectors or in the center correspond to the number of introduced stop codons found in two (node connectors) or all (center) related genomes. (**C**) Chromosomal distribution of sequence identity between bread wheat genes and the diploid and tetraploid relatives for homeologous chromosomes.

BD, and AB as the closest pairs, respectively. This genome-wide observation supports previous findings of discordant phylogenetic signals within *Aegilops* and *Triticum* genera (*6*, *43*, *45*). Some variation in genome relationships was found among chromosomes: On group 4 chromosomes, most gene trees supported BD as closest pairs, whereas group 5 chromosomes had similar numbers of AD and BD topologies (AD = BD > AB). Distribution of variation in phylogenetic signals across homeologous chromosomes can help to better understand the nature of the evolutionary processes underlying such phylogenetic incongruence. Under incomplete lineage sorting and stochastic coalescence, levels of phylogenetic incongruence will be correlated with recombination rates, whereas single introgression events and limited recombination are expected to generate local chromosome blocks of homogenous phylogenetic signals. We used the inferred gene orders from the GenomeZipper to test for nonrandom distribution of phylogenetic signals along chromosomes. We were unable to consistently identify block structures larger than would be expected by chance. However, it is possible that the limitations of the inferred gene order hamper the ability to detect such patterns.

Gene expression

Our study did not reveal any pronounced bias in gene content, structure, or composition between the different wheat subgenomes. In paleopolyploid maize and soybean, transcriptional dominance of genes derived from one progenitor genome has been described (*64*–*66*). Previous analyses have shown that rapid initiation of differential expression of homeologous wheat genes occurs upon polyploidization with a predominantly additive mode (*13*, *67*). Sets of homeologous wheat genes with only one copy present in each of the subgenomes (triads) were used to test for differential expression at a genome-wide scale. Expression correlations were calculated for 6219 triads (18,657 genes) by using RNA-seq data from five organs (leaf, root, grain, spike, and stem) (Fig. 5A) (*25*). Whereas root-derived expression clustered separately, genes expressed in stem, leaves, grain, and spike clustered in a subgenome-specific manner. This indicates that the individual subgenomes exhibit a high degree of regulatory and transcriptional autonomy, with limited trans (inter-subgenome) regulation (*68*). At a global level, the overall pairwise expression correlation between subgenomes was very similar (Fig. 5B), and no evidence for genome-wide transcriptional dominance of an individual subgenome was observed.

By using hierarchical cluster analysis, we aggregated expressed genes into 13 distinct groups. These groups show predominant expression in particular organs (e.g., groups III and XIII in Fig. 5A) or in one of the subgenomes (e.g., groups II, IX, and X in Fig. 5A). Pairwise comparisons of individual expressed homeologous genes in the groups revealed abundant transcriptional dominance from specific subgenomes (Fig. 5B). Overall, 1333 (21%) of the homeologous gene triads showed an expression bias in one of the pairwise comparisons, and we detected a similar number of preferentially transcribed genes (378 to 393) in each subgenome (permutation test; $P < 0.05$). For the individual transcriptional groups, however, between 2% (groups I, IV, and V) and 20% (groups II and VI) of the genes were found to be transcriptionally dominant.

These patterns of gene expression across the three genomes contrast with patterns of gene expression reported in allopolyploid cotton (*69*, *70*); mesopolyploid *Brassica rapa* (*71*); synthetic allotetraploid *Arabidopsis* (*72*); and the paleopolyploid maize genome (*64*), where one of the genomes is more transcriptionally active than others. The apparent autonomy of the three wheat subgenomes may be explained by the relatively recent polyploidization. It may also be related to regulatory mechanisms that control the transcriptional interplay of homeologous genomes to balance expression of individual and groups of genes. While maintaining subgenome-specific expression profiles, a high degree of orchestration and functional partitioning between homeologous genes was also reported in grain development of bread wheat (*68*) and has been attributed to the rapid evolution of cis elements coupled to epigenetic mechanisms controlling gene expression (*68*, *73*, *74*).

Gene family size variation

The relationship between genes important to wheat adaptation, disease resistance, and end-use functionality in hexaploid wheat and its diploid relatives was examined for signs of adaptive evolution. These analyses identified three distinct patterns: gene expansion, gene loss, or independent gene evolution that may or may not include expansion or loss. In some cases, such as the genes containing a NB-ARC domain characteristic of many plant disease-resistance genes (*75*), we observed an expansion within a single subgenome (Fig. 6A). Indeed, a substantial expansion in *Ae. tauschii*, compared with the other diploid species and the D genome of hexaploid wheat, is consistent with the rich reservoir of disease-resistance genes known in this species

(*17*). In genes coding for the cysteine-rich gliadin domain, a functional domain characteristic of storage proteins, we observed a similar number of genes in all diploid genomes (except *T. monococcum*) that is higher than the number of genes found in each of the three hexaploid wheat subgenomes (Fig. 6B). This may indicate that gene loss occurred in hexaploid wheat and that there is a trend for the gliadin gene family to maintain some homeostasis with a similar global number of genes in polyploid and diploid wheat. In other cases, the patterns observed suggested independent evolution of gene families within the different genomes and subgenomes of wheat. This was seen for genes associated with abiotic stress tolerance. For example, for genes encoding the Apetala2 (AP2) DNA binding domain, associated with drought, heat, salinity, and cold stress–tolerance responses, we observed fewer AP2 genes in the A and D genomes of Chinese Spring compared with the diploid relatives or the B subgenome (Fig. 6C). Likewise, genes coding for MYB transcription factors, which have also been involved in abiotic stress response in plants (*76*), were underrepresented in the A subgenome of hexaploid wheat and *T. monococcum*, whereas a higher frequency was observed in *Ae. tauschii* (*17*) and *T. urartu* (*16*) (Fig. 6D).

In contrast, there was no evidence of expansion or loss of genes underlying phenology, such as the vernalization (*Vrn1*) and photoperiod response regulator (*Ppd1*) genes that differentiate spring and winter growth habits and sensitivity to day length, respectively. Similar numbers of genes were found in the diploids and hexaploid subgenomes coding for the two functional domains of *Vrn1*, a MADS-box and K-box domain (*77*) (Fig. 6E), and for genes containing the response regulator domain and CCT motif typical of cereal *Ppd* genes (*78*) (Fig. 6F). We identified an additional copy of a *Vrn1*-like gene in the hexaploid Chinese Spring A and D genomes and *T. urartu* (*16*) when compared with the remaining diploid species. An additional copy of a *Ppd1*-like gene was also identified in the Chinese Spring B genome relative to *Ae. sharonesis* and *Ae. speltoides* (Fig. 6F). Although only small differences were observed, small increases in copy number variation of *Vrn-A1* (A genome) and *Ppd-B1* (B genome) have been associated with longer periods of vernalization to potentiate flowering and an early flowering day neutral phenotype, respectively (*79*). Thus, the relative distribution of such patterns in ontology of these two genes is likely to reflect important factors that have allowed wheat to adjust its flowering time to adapt to a range of environmental conditions.

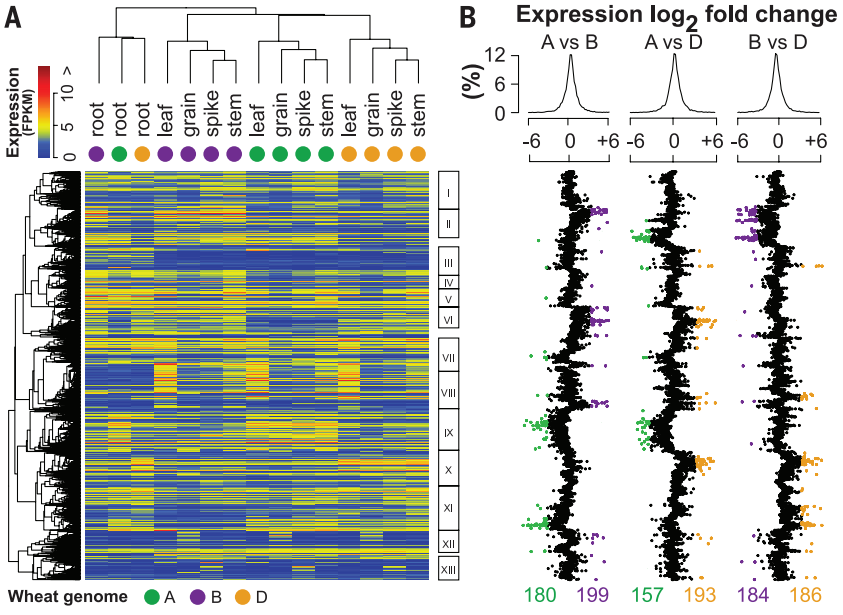


Fig. 5. Subgenome transcriptional profiling for individual wheat tissues. (**A**) Two-dimensional hierarchical cluster analysis of single-copy wheat homeologous gene expression (colors as in Fig. 2A) compared with organ-specific gene expression. (**B**) Analysis of $\log_2$-fold changes in pairwise gene expression between homeologous genes (averaged across organs). Top graphs depict the distributions of $\log_2$ fold changes. Dot plots show the fold changes for each triplet ordered as shown in the *y* axis in (A). Colored dots highlight homologs that show significant differential expression ($P < 0.05$). The numbers of differentially expressed triplets across all organs are shown at the bottom of the figure.

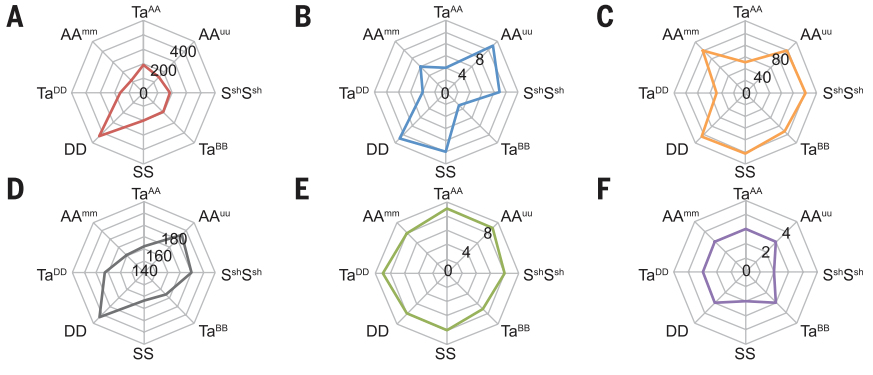


Fig. 6. Sizes of selected gene families and protein domains among hexaploid wheat and diploid relatives. (**A**) NB-ARC domain, (**B**) cysteine-rich gliadin domain, (**C**) AP2 domain, (**D**) MYB domain, (**E**) *Vrn1* (MADS-box/K-box domain), and (**F**) *Ppd* (photoperiod response regulator/CCT domain).

Molecular markers

Wheat improvement relies in part on the use of molecular markers to improve selection efficiencies and to allow the precise transfer of genes and QTL between different genetic backgrounds. To enhance the CSS as a genomic resource for the wheat genetics and breeding community, we anchored all publicly available DNA markers that are routinely used for genetic mapping and marker-assisted breeding in wheat. Because the majority of these markers are anchored to phenotypic maps, anchoring them to the CSS allows immediate association of CSS to traits targeted by breeders. In addition, insertion site–based polymorphism (ISBP) and SNP markers identified from recent whole-genome shotgun and transcriptome sequencing (*19*) and genotyping by sequencing (GBS) tags identified by using DArTSeq (Diversity Arrays Technology, Bruce, Australia) technology were also anchored. In total, over 3.6 million marker loci were anchored to the CSS, including 1,347,669 marker loci and 2,310,988 SNPs (Table 5).

Most marker types showed a distribution gradient across subgenomes, with the highest number associated with the B genome chromosomes and the lowest with the D genome, reflecting the differences in the level of polymorphism in these subgenomes. The proportions of ISBPs, SNPs detected from cultivar sequencing and GBS tags localized to the D genome ranged between 9.3 and 12%, with the lowest numbers mapping to the group 4 chromosomes (Table 5). Two hundred and ninety-two of 1867 simple sequence repeat (SSR) loci were successfully anchored to the CSS survey sequence. This low number is not surprising, given that these loci derive from repetitive AT- and GC-rich sequences that may be collapsed or

Table 5. Number and type of molecular markers mapped on individual chromosomes of the bread wheat genome.

	Bin mapped ESTs	EST-SSRs	Genomic SSRs	DArT Probes	Cereals DB	90K iSelect SNPs (87)	DArT Seq	ISBPs	Genic SNPs	Intergenic SNPs	Σ
Queries	18,771	2,926	1,867	7,552	7,228	81,987	29,375	Derived from cultivar sequencing			-
Mapped queries	16,876	2,435	282	5,228	5,136	80,820	18,515				
1A	1,325	156	8	414	479	13,093	1,371	68,074	13,980	127,663	226,563
2A	1,614	257	28	356	544	17,502	1,378	84,440	18,349	148,204	272,672
3A	1,136	75	14	252	302	12,172	1,008	44,740	10,770	94,975	165,444
4A	1,766	266	27	331	357	14,043	1,530	39,483	10,367	86,543	154,713
5A	1,189	155	46	256	343	13,099	893	62,193	12,624	115,085	205,883
6A	1,150	132	63	418	421	12,072	1,127	60,169	15,884	110,850	202,286
7A	1,240	146	120	321	326	13,168	1,474	71,597	15,516	154,748	258,656
Σ A genome	9,420	1,187	306	2,348	2,772	95,149	8,781	430,696	97,490	838,068	1,486,217
1B	1,379	226	15	378	618	13,776	1,846	66,994	14,447	131,682	231,361
2B	1,810	367	39	466	606	18,352	2,557	90,852	23,958	162,335	301,342
3B	1,845	188	29	406	444	14,471	2,294	108,810	22,032	208,306	358,825
4B	1,401	188	42	278	294	11,019	856	36,937	7,506	59,175	117,696
5B	1,911	343	86	399	527	17,087	2,112	84,179	21,389	159,359	287,392
6B	978	43	139	320	313	12,448	1,171	65,982	11,974	130,463	223,831
7B	999	107	151	270	205	11,635	1,123	72,307	10,997	136,932	234,726
Σ B genome	10,323	1,462	501	2,517	3,007	98,788	11,959	526,061	112,303	988,252	1,755,173
1D	1,165	149	13	378	380	12,093	660	17,366	5,004	36,457	73,665
2D	1,309	199	22	414	331	16,978	609	19,532	6,745	34,967	81,106
3D	854	104	14	428	151	11,699	420	10,920	1,403	18,078	44,071
4D	1,221	239	27	245	196	10,198	307	10,097	1,108	13,249	36,887
5D	1,584	408	78	400	289	13,308	488	13,629	3,582	22,957	56,723
6D	1,132	91	135	289	240	10,504	417	12,042	3,609	23,341	51,800
7D	1,461	230	139	862	243	12,826	767	18,174	3,969	34,344	73,015
Σ D genome	8,726	1,420	428	3,016	1,830	87,606	3,668	101,760	25,420	183,393	417,267
Σ	28,469	4,069	1,235	7,881	7,609	281,543	24,408	1,058,517	235,213	2,009,713	3,658,657

represented by uneven read coverage in Illumina sequences (*80*).

Well over 70 DNA markers are routinely deployed by breeders for agronomic, pest resistance, and end-use quality, and most are available in the public domain (http://maswheat.ucdavis.edu). Anchoring of these to the CSS would facilitate identification of SNP markers for development of high-density marker maps, as a resource of correlated markers, and to aid map-based cloning of genes underlying important traits. In total, we anchored 68 of these markers to 74 contigs in the CSS. The application of the CSS in marker improvement was demonstrated with the CAPS (cleaved amplified polymorphic sequence) marker *Usw47*, which is linked to *Cdu-B1*, a gene responsible for reduced grain cadmium content in tetraploid wheat (*81*, *82*). Although *Usw47* is routinely used in marker-assisted selection, it is not amenable to high-throughput genotyping. Alignment of the *Usw47* sequence against the CSS mapped it to contig 5BL-10759151. This and eight neighboring contigs in the GenomeZipper contained 33 SNP markers, of which 5 were polymorphic in a doubled haploid mapping population used previously to localize *Cdu-B1*. Of the five SNP markers, two co-segregated, and the remainder flanked the gene by a single recombination event. These SNP markers can be readily implemented now in a high-throughput fashion to select for reduced grain cadmium content within breeding programs.

Conclusion

We present the ordered and structured draft sequence of the bread wheat genome as well as a comparison between eight related wheat genomes. We defined a gene catalog for each of the 21 bread wheat chromosomes and positioned more than 75,000 genes along the chromosomes by using a combination of high-density wheat SNP mapping and synteny to sequenced grass genomes. In contrast to other species (*83*), polyploidization events in wheat did not cause a "genome shock" with subsequent rapid genome changes or functional dominance of one subgenome over the others. Intraspecific comparative analyses revealed a dynamic wheat genome with a high level of plasticity and a changing gene repertoire shaped by gene losses and gene-family expansions in all wheat genomes and subgenomes, with only a few species-specific genes. Through interspecific comparisons, we observed a higher abundance of intrachromosomal gene duplications in wheat compared with other grass genomes, which may be a mechanism for functional adaptation and underlie the global success of wheat as a cultivated crop.

The detection, chromosomal assignment, and description of a large proportion of the gene complement of bread wheat and their positional assignment on chromosome arms is a major milestone in facilitating the isolation of genes underlying agronomically important traits, providing a reference for future integration into systems biology, and improving wheat breeding efficiency. Already, the resources developed in this work have been used to support the analysis of selected wheat chromosomes (*20*, *41*, *84*–*86*). Last, as demonstrated by the completion of the reference sequence for chr. 3B (*23*), this draft genome sequence and complementary resources will support the assembly and annotation of the physical map–based reference sequences for the 21 bread wheat chromosomes.

REFERENCES AND NOTES

1. D. B. Lobell, W. Schlenker, J. Costa-Roberts, Climate trends and global crop production since 1980. *Science* **333**, 616–620 (2011). doi: 10.1126/science.1204531; pmid: 21551030
2. Food and Agriculture Organization (FAO) of the United Nations, FAO cereal supply and demand brief (2013); www.fao.org/worldfoodsituation/csdb/en/.
3. D. Tilman, K. G. Cassman, P. A. Matson, R. Naylor, S. Polasky, Agricultural sustainability and intensive production practices. *Nature* **418**, 671–677 (2002). doi: 10.1038/nature01014; pmid: 12167873
4. J. A. Foley *et al.*, Solutions for a cultivated planet. *Nature* **478**, 337–342 (2011). doi: 10.1038/nature10452; pmid: 21993620
5. Organisation for Economic Cooperation and Development (OECD)/FAO, *OECD-FAO Agricultural Outlook 2013* (OECD, Paris, 2013); doi: 10.1787/agr_outlook-2013-en.
6. G. Petersen, O. Seberg, M. Yde, K. Berthelsen, Phylogenetic relationships of *Triticum* and *Aegilops* and evidence for the

origin of the A, B, and D genomes of common wheat (*Triticum aestivum*). *Mol. Phylogenet. Evol.* **39**, 70–82 (2006). doi: 10.1016/j.ympev.2006.01.023; pmid: 16504543
7. M. Nesbitt, D. Samuel, "From staple crop to extinction? The archaeology and history of the hulled wheats," in *Hulled Wheat: Proceedings of the First International Workshop on Hulled Wheats*, S. Padulosi, K. Hammer, J. Heller, Eds. (International Plant Genetic Resources Institute, Rome, 1995), pp. 41–102.
8. E. Martinez-Perez, P. Shaw, G. Moore, The Ph1 locus is needed to ensure specific somatic and meiotic centromere association. *Nature* **411**, 204–207 (2001). doi: 10.1038/35075597; pmid: 11346798
9. T. Eilam *et al*., Genome size and genome evolution in diploid Triticeae species. *Genome* **50**, 1029–1037 (2007). doi: 10.1139/G07-083; pmid: 18059548
10. T. Wicker *et al*., Frequent gene movement and pseudogene evolution is common to the large and complex genomes of wheat, barley, and their relatives. *Plant Cell* **23**, 1706–1718 (2011). doi: 10.1105/tpc.111.086629; pmid: 21622801
11. K. Mochida, T. Yoshida, T. Sakurai, Y. Ogihara, K. Shinozaki, TriFLDB: A database of clustered full-length coding sequences from Triticeae with applications to comparative grass genomics. *Plant Physiol.* **150**, 1135–1146 (2009). doi: 10.1104/pp.109.138214; pmid: 19448038
12. A. N. Bernardo *et al*., Discovery and mapping of single feature polymorphisms in wheat using Affymetrix arrays. *BMC Genomics* **10**, 251 (2009). doi: 10.1186/1471-2164-10-251; pmid: 19480702
13. H. Chelaifa *et al*., Prevalence of gene expression additivity in genetically stable wheat allohexaploids. *New Phytol.* **197**, 730–736 (2013). doi: 10.1111/nph.12108; pmid: 23278496
14. T. E. Coram, M. L. Settles, M. Wang, X. Chen, Surveying expression level polymorphism and single-feature polymorphism in near-isogenic wheat lines differing for the Yr5 stripe rust resistance locus. *Theor. Appl. Genet.* **117**, 401–411 (2008). doi: 10.1007/s00122-008-0784-5; pmid: 18470504
15. L. L. Qi *et al*., A chromosome bin map of 16,000 expressed sequence tag loci and distribution of genes among the three genomes of polyploid wheat. *Genetics* **168**, 701–712 (2004). doi: 10.1534/genetics.104.034868; pmid: 15514046
16. H. Q. Ling *et al*., Draft genome of the wheat A-genome progenitor *Triticum urartu*. *Nature* **496**, 87–90 (2013). doi: 10.1038/nature11997; pmid: 23535596
17. J. Jia *et al*., *Aegilops tauschii* draft genome sequence reveals a gene repertoire for wheat adaptation. *Nature* **496**, 91–95 (2013). doi: 10.1038/nature12028; pmid: 23535592
18. R. Brenchley *et al*., Analysis of the bread wheat genome using whole-genome shotgun sequencing. *Nature* **491**, 705–710 (2012). doi: 10.1038/nature11650; pmid: 23192148
19. A. M. Allen *et al*., Discovery and development of exome-based, co-dominant single nucleotide polymorphism markers in hexaploid wheat (*Triticum aestivum* L.). *Plant Biotechnol. J.* **11**, 279–295 (2013). doi: 10.1111/pbi.12009; pmid: 23279710
20. K. V. Krasileva *et al*., Separating homeologs by phasing in the tetraploid wheat transcriptome. *Genome Biol.* **14**, R66 (2013). doi: 10.1186/gb-2013-14-6-r66; pmid: 23800085
21. C. Saintenac, D. Jiang, S. Wang, E. Akhunov, Sequence-based mapping of the polyploid wheat genome. *G3* **3**, 1105–1114 (2013).
22. E. Sears, L. Sears, "The telocentric chromosomes of common wheat," in *Proceedings 5th International Wheat Genetics Symposium*, S. Ramanujam, Ed. (Indian Agricultural Research Institute, New Delhi, 1978) vol. 1, pp. 389–407.
23. F. Choulet *et al*., A reference sequence of wheat chromosome 3B reveals structural and functional compartmentalization. *Science* **345**, 1249721 (2014).
24. J. Šafář *et al*., Development of chromosome-specific BAC resources for genomics of bread wheat. *Cytogenet. Genome Res.* **129**, 211–223 (2010). doi: 10.1159/000313072; pmid: 20501977
25. Materials and methods are available as supporting materials on *Science* Online.
26. J. T. Simpson *et al*., ABySS: A parallel assembler for short read sequence data. *Genome Res.* **19**, 1117–1123 (2009). doi: 10.1101/gr.089532.108; pmid: 19251739
27. K. F. Mayer *et al*., A physical, genetic, and functional sequence assembly of the barley genome. *Nature* **491**, 711–716 (2012). pmid: 23075845
28. S. Kurtz, A. Narechania, J. C. Stein, D. Ware, A new method to compute K-mer frequencies and its application to annotate large repetitive plant genomes. *BMC Genomics* **9**, 517 (2008). doi: 10.1186/1471-2164-9-517; pmid: 18976482
29. J. D. Hollister, B. S. Gaut, Epigenetic silencing of transposable elements: A trade-off between reduced transposition and deleterious effects on neighboring gene expression. *Genome Res.* **19**, 1419–1428 (2009). doi: 10.1101/gr.091678.109; pmid: 19478138
30. M. Kantar *et al*., Subgenomic analysis of microRNAs in polyploid wheat. *Funct. Integr. Genomics* **12**, 465–479 (2012). doi: 10.1007/s10142-012-0285-0; pmid: 22592659
31. S. J. Lucas, H. Budak, Sorting the wheat from the chaff: Identifying miRNAs in genomic survey sequences of *Triticum aestivum* chromosome 1AL. *PLOS ONE* **7**, e40859 (2012). doi: 10.1371/journal.pone.0040859; pmid: 22815845
32. G. M. Borchert *et al*., Comprehensive analysis of microRNA genomic loci identifies pervasive repetitive-element origins. *Mob. Genet. Elements* **1**, 8–17 (2011). doi: 10.4161/mge.1.1.15766; pmid: 22016841
33. International Brachypodium Initiative, Genome sequencing and analysis of the model grass *Brachypodium distachyon*. *Nature* **463**, 763–768 (2010). doi: 10.1038/nature08747; pmid: 20148030
34. International Rice Genome Sequencing Project, The map-based sequence of the rice genome. *Nature* **436**, 793–800 (2005). doi: 10.1038/nature03895; pmid: 16100779
35. A. H. Paterson *et al*., The *Sorghum bicolor* genome and the diversification of grasses. *Nature* **457**, 551–556 (2009). doi: 10.1038/nature07723; pmid: 19189423
36. F. Choulet *et al*., Megabase level sequencing reveals contrasted organization and evolution patterns of the wheat gene and transposable element spaces. *Plant Cell* **22**, 1686–1701 (2010). doi: 10.1105/tpc.110.074187; pmid: 20581307
37. T. Lu *et al*., Function annotation of the rice transcriptome at single-nucleotide resolution by RNA-seq. *Genome Res.* **20**, 1238–1249 (2010). doi: 10.1101/gr.106120.110; pmid: 20627892
38. Y. Okazaki *et al*., Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs. *Nature* **420**, 563–573 (2002). doi: 10.1038/nature01266; pmid: 12466851
39. Y. Marquez, J. W. Brown, C. Simpson, A. Barta, M. Kalyna, Transcriptome survey reveals increased complexity of the alternative splicing landscape in *Arabidopsis*. *Genome Res.* **22**, 1184–1195 (2012). doi: 10.1101/gr.134106.111; pmid: 22391557
40. M. M. Martis *et al*., Reticulate evolution of the rye genome. *Plant Cell* **25**, 3685–3698 (2013). doi: 10.1105/tpc.113.114553; pmid: 24104565
41. P. Hernandez *et al*., Next-generation sequencing and syntenic integration of flow-sorted arms of wheat chromosome 4A exposes the chromosome structure and gene content. *Plant J.* **69**, 377–386 (2012). doi: 10.1111/j.1365-313X.2011.04808.x; pmid: 21974774
42. J. Ma *et al*., Sequence-based analysis of translocations and inversions in bread wheat (*Triticum aestivum* L.). *PLOS ONE* **8**, e79329 (2013). doi: 10.1371/journal.pone.0079329; pmid: 24260197
43. J. S. Escobar *et al*., Multigenic phylogeny and analysis of tree incongruences in Triticeae (Poaceae). *BMC Evol. Biol.* **11**, 181 (2011). doi: 10.1186/1471-2148-11-181; pmid: 21702931
44. P. Cíváň, Z. Ivaničová, T. A. Brown, Reticulated origin of domesticated emmer wheat supports a dynamic model for the emergence of agriculture in the fertile crescent. *PLOS ONE* **8**, e81955 (2013). doi: 10.1371/journal.pone.0081955; pmid: 24312385
45. T. Marcussen *et al*., Ancient hybridizations among the ancestral genomes of bread wheat. *Science* **345**, 1250092 (2014).
46. S. Griffiths *et al*., Molecular characterization of Ph1 as a major chromosome pairing locus in polyploid wheat. *Nature* **439**, 749–752 (2006). doi: 10.1038/nature04434; pmid: 16467840
47. K. F. X. Mayer *et al*., Gene content and virtual gene order of barley chromosome 1H. *Plant Physiol.* **151**, 496–505 (2009). doi: 10.1104/pp.109.142612; pmid: 19692534
48. G. Moore, K. M. Devos, Z. Wang, M. D. Gale, Cereal genome evolution. Grasses, line up and form a circle. *Curr. Biol.* **5**, 737–739 (1995). doi: 10.1016/S0960-9822(95)00148-5; pmid: 7583118
49. M. C. Luo *et al*., A 4-gigabase physical map unlocks the structure and evolution of the complex genome of *Aegilops tauschii*, the wheat D-genome progenitor. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 7940–7945 (2013). doi: 10.1073/pnas.1219082110; pmid: 23610408
50. M. Mascher *et al*., Anchoring and ordering NGS contig assemblies by population sequencing (POPSEQ). *Plant J.* **76**, 718–727 (2013). doi: 10.1111/tpj.12319; pmid: 23998490
51. M. E. Sorrells *et al*., Reconstruction of the synthetic W7984 x Opata M85 wheat reference population. *Genome* **54**, 875–882 (2011). doi: 10.1139/g11-054; pmid: 21999208
52. J. Zhang, Evolution by gene duplication: An update. *Trends Ecol. Evol.* **18**, 292–298 (2003). doi: 10.1016/S0169-5347(03)00033-8
53. L. Li, C. J. Stoeckert Jr., D. S. Roos, OrthoMCL: Identification of ortholog groups for eukaryotic genomes. *Genome Res.* **13**, 2178–2189 (2003). doi: 10.1101/gr.1224503; pmid: 12952885
54. R. Koszul, S. Caburet, B. Dujon, G. Fischer, Eucaryotic genome evolution through the spontaneous duplication of large chromosomal segments. *EMBO J.* **23**, 234–243 (2004). doi: 10.1038/sj.emboj.7600024; pmid: 14685272
55. J. L. Bennetzen *et al*., Reference genome sequence of the model plant *Setaria*. *Nat. Biotechnol.* **30**, 555–561 (2012). doi: 10.1038/nbt.2196; pmid: 22580951
56. P. S. Schnable *et al*., The B73 maize genome: Complexity, diversity, and dynamics. *Science* **326**, 1112–1115 (2009). doi: 10.1126/science.1178534; pmid: 19965430
57. T. Tanaka *et al*., The Rice Annotation Project Database (RAP-DB): 2008 update. *Nucleic Acids Res.* **36**, D1028–D1033 (2008).pmid: 18089549
58. H. Ozkan, A. A. Levy, M. Feldman, Allopolyploidy-induced rapid genome evolution in the wheat (*Aegilops-Triticum*) group. *Plant Cell* **13**, 1735–1747 (2001). doi: 10.1105/tpc.13.8.1735; pmid: 11487689
59. R. J. Buggs *et al*., Rapid, repeated, and clustered loss of duplicate genes in allopolyploid plant populations of independent origin. *Curr. Biol.* **22**, 248–252 (2012). doi: 10.1016/j.cub.2011.12.027; pmid: 22264605
60. A. H. Paterson *et al*., Repeated polyploidization of *Gossypium* genomes and the evolution of spinnable cotton fibres. *Nature* **492**, 423–427 (2012). doi: 10.1038/nature11798; pmid: 23257886
61. R. Grantham, Amino acid difference formula to help explain protein evolution. *Science* **185**, 862–864 (1974). doi: 10.1126/science.185.4154.862; pmid: 4843792
62. J. Cao *et al*., Whole-genome sequencing of multiple *Arabidopsis thaliana* populations. *Nat. Genet.* **43**, 956–963 (2011). doi: 10.1038/ng.911; pmid: 21874002
63. E. D. Akhunov *et al*., Comparative analysis of syntenic genes in grass genomes reveals accelerated rates of gene structure and coding sequence evolution in polyploid wheat. *Plant Physiol.* **161**, 252–265 (2013). doi: 10.1104/pp.112.205161; pmid: 23124323
64. J. C. Schnable, N. M. Springer, M. Freeling, Differentiation of the maize subgenomes by genome dominance and both ancient and ongoing gene loss. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 4069–4074 (2011). doi: 10.1073/pnas.1101368108; pmid: 21368132
65. R. A. Rapp, J. A. Udall, J. F. Wendel, Genomic expression dominance in allopolyploids. *BMC Biol.* **7**, 18 (2009). doi: 10.1186/1741-7007-7-18; pmid: 19409075
66. B. Chaudhary *et al*., Reciprocal silencing, transcriptional bias and functional divergence of homeologs in polyploid cotton (Gossypium). *Genetics* **182**, 503–517 (2009). doi: 10.1534/genetics.109.102608; pmid: 19363125
67. M. Pumphrey, J. Bai, D. Laudencia-Chingcuanco, O. Anderson, B. S. Gill, Nonadditive expression of homoeologous genes is established upon polyploidization in hexaploid wheat. *Genetics* **181**, 1147–1157 (2009). doi: 10.1534/genetics.108.096941; pmid: 19104075
68. M. Pfeifer *et al*., Genome interplay in the grain transcriptome of hexaploid bread wheat. *Science* **345**, 1250091 (2014).
69. M. J. Yoo, E. Szadkowski, J. F. Wendel, Homoeolog expression bias and expression level dominance in allopolyploid cotton. *Heredity* **110**, 171–180 (2013). doi: 10.1038/hdy.2012.94; pmid: 23169565
70. K. L. Adams, R. Cronn, R. Percifield, J. F. Wendel, Genes duplicated by polyploidy show unequal contributions to the transcriptome and organ-specific reciprocal silencing. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 4649–4654 (2003). doi: 10.1073/pnas.0630618100; pmid: 12665616
71. F. Cheng *et al*., Biased gene fractionation and dominant gene expression among the subgenomes of *Brassica rapa*. *PLOS ONE* **7**, e36442 (2012). doi: 10.1371/journal.pone.0036442; pmid: 22567157
72. J. Wang *et al*., Stochastic and epigenetic changes of gene expression in *Arabidopsis* polyploids. *Genetics* **167**, 1961–1973 (2004). doi: 10.1534/genetics.104.027896; pmid: 15342533
73. Z. J. Chen, Genetic and epigenetic mechanisms for gene expression and phenotypic variation in plant polyploids. *Annu. Rev. Plant Biol.* **58**, 377–406 (2007). doi: 10.1146/annurev.arplant.58.032806.103835; pmid: 17280525
74. K. L. Adams, Evolution of duplicate gene expression in polyploid and hybrid plants. *J. Hered.* **98**, 136–141 (2007). doi: 10.1093/jhered/esl061; pmid: 17208934
75. G. van Ooijen *et al*., Structure-function analysis of the NB-ARC domain of plant disease resistance proteins. *J. Exp. Bot.* **59**, 1383–1397 (2008). doi: 10.1093/jxb/ern045; pmid: 18390848
76. A. Katiyar *et al*., Genome-wide classification and expression analysis of MYB transcription factor families in rice and

Arabidopsis. *BMC Genomics* **13**, 544 (2012). doi: 10.1186/1471-2164-13-544; pmid: 23050870
77. L. Yan *et al*., Positional cloning of the wheat vernalization gene VRN1. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 6263–6268 (2003). doi: 10.1073/pnas.0937399100; pmid: 12730378
78. A. Turner, J. Beales, S. Faure, R. P. Dunford, D. A. Laurie, The pseudo-response regulator Ppd-H1 provides adaptation to photoperiod in barley. *Science* **310**, 1031–1034 (2005). doi: 10.1126/science.1117619; pmid: 16284181
79. A. Díaz, M. Zikhali, A. S. Turner, P. Isaac, D. A. Laurie, Copy number variation affecting the Photoperiod-B1 and Vernalization-A1 genes is associated with altered flowering time in wheat (*Triticum aestivum*). *PLOS ONE* **7**, e33234 (2012). doi: 10.1371/journal.pone.0033234; pmid: 22457747
80. S. O. Oyola *et al*., Optimizing Illumina next-generation sequencing library preparation for extremely AT-biased genomes. *BMC Genomics* **13**, 1 (2012). doi: 10.1186/1471-2164-13-1; pmid: 22214261
81. R. E. Knox *et al*., Chromosomal location of the cadmium uptake gene (*Cdu1*) in durum wheat. *Genome* **52**, 741–747 (2009). doi: 10.1139/G09-042; pmid: 19935921
82. K. Wiebe *et al*., Targeted mapping of Cdu1, a major locus regulating grain cadmium concentration in durum wheat (*Triticum turgidum* L. var durum). *Theor. Appl. Genet.* **121**, 1047–1058 (2010). doi: 10.1007/s00122-010-1370-1; pmid: 20559817
83. L. Comai, The advantages and disadvantages of being polyploid. *Nat. Rev. Genet.* **6**, 836–846 (2005). doi: 10.1038/nrg1711; pmid: 16304599
84. P. J. Berkman *et al*., Sequencing and assembly of low copy and genic regions of isolated *Triticum aestivum* chromosome arm 7DS. *Plant Biotechnol. J.* **9**, 768–775 (2011). doi: 10.1111/j.1467-7652.2010.00587.x; pmid: 21356002
85. P. J. Berkman *et al*., Sequencing wheat chromosome arm 7BS delimits the 7BS/4AL translocation and reveals homoeologous gene conservation. *Theor. Appl. Genet.* **124**, 423–432 (2012). doi: 10.1007/s00122-011-1717-2; pmid: 22001910
86. T. Tanaka *et al*., Next-generation survey sequencing and the molecular organization of wheat chromosome 6B. *DNA Res.* **21**, 103–114 (2013). pmid: 24086083
87. S. Wang *et al*., Characterization of polyploid wheat genomic diversity using a high-density 90, 000 single nucleotide polymorphism array. *Plant Biotechnol. J.* (2014). doi: 10.1111/pbi.12183; pmid: 24646323

ACKNOWLEDGMENTS

The authors would like to thank Graminor AS; Biogemma; Institut National de la Recherche Agronomique (INRA); International Center for Agricultural Research in the Dry Areas; Department of Biotechnology, Ministry of Science and Technology, Government of India (chr. 2A; grant no. BT/IWGSC/03/TF/2008); and the Biotechnology and Biological Sciences Research Council (BBSRC UK) for funding the chromosome sequencing at the Genome Analysis Centre. Chromosome sequencing at other centers was funded by the following: chr. 3A—U.S. Department of Agriculture Agriculture and Food Research Initiative (USDA AFRI) Triticeae-CAP (2011-68002-30029) and the Kansas Wheat Commission; chr. 3B—grants from the French National Research Agency (ANR-09- GENM-025 3BSEQ) and France Agrimer; chr. 6B—grants from the Ministry of Agriculture, Forestry and Fisheries of Japan "Genomics for agricultural innovation KGS-1003,1004", "Genomics based technology for agricultural improvement, NGB- 1003," and Nisshin Flour Milling Incorporated; chr. 6D and *Triticum durum* cv. Strongfield—grants from Genome Canada, Genome Prairie, University of Saskatchewan Ministry of Agriculture, Western Grains Research Foundation; chr. 7B—grant no. 199387 from the Norwegian Research Council and from Graminor AS; chr. 7A and 7D sequence reads were provided by D.E.. Chromosome flow sorting and DNA preparation was supported through grants P501/12/G090 and P501/12/2554 from the Czech Science foundation. Chromosome sequence assembly was supported by the BBSRC (UK). K.F.X.M. acknowledges grants from the German Ministry for Education and Research (BMBF) Plant2030, TRITEX, Deutsche Forschungsgemeinschaft (DFG) SFB 924, and EC Transplant. K.E. and J.R. are supported by sponsors of the IWGSC, which include Arcadia Biosciences, Australian Centre for Plant Functional Genomics, Biogemma, Bayer CropScience, Commonwealth Science and Industrial Research Organisation, Centro Internacional de Mejoramiento de Maíz y Trigo, Céréales Vallée, Dow AgroSciences, Dupont, Evogene, Florimond Desprez, Grains Research and Development Corporation, Graminor, Heartland Plant Innovation, INRA, KWS, Kansas Wheat Commission, Limagrain, Monsanto, RAGT, and Syngenta. N.G. is supported by European Commission Marie Curie Actions (FP7-MC-IIF-Noncollinear Genes). T.W. is supported by the Swiss National Foundation and P.F., M.C., A.M.S., and L.C. are supported by the Italian Ministry of Agriculture special project "MAPPA-5A." H.B. acknowledges funding from Sabanci University and the Scientific and Technological Research Council of Turkey. B.W. and B.S. were funded by the Gatsby Charitable Foundation and the BBSRC (UK) Grant BB/J003166/1. R.W. is a Trustee Director of TGAC, Norwich, UK, and A.K. is a shareholder of Diversity Arrays Technology Pty Ltd. The POPSeq analysis carried out by the U.S. Department of Energy Joint Genome Institute was supported by the Office of Science of the U.S. Department of Energy under contract no. DE-AC02-05CH11231. Additional support for the work was funded from the Triticeae-CAP, USDA AFRI (2011-68002-30029) to G.J.M.; the Scottish Government Rural and Environment Science and Analytical Services Division Research Programme to R.W.; and the German Ministry of Research and Education (BMBF TRITEX 0315954) to N.S. Sequence reads and assembled sequences are available at European Molecular Biology Laboratory/GenBank/DNA Data Bank of Japan short read archives and sequence repositories, respectively (PRJEB3955—whole-genome sequences of *T. aestivum* 'Chinese Spring,' *T. urartu*, *Ae. speltoides*, *Ae. tauschii*, *T. turgidum*; SRP004490.3—whole-genome sequencing of *T. monococcum*; SRP004490—whole-genome sequencing of *Ae. tauschii*; PRJEB4849—whole-genome sequences of *Ae. sharonensis*; PRJEB4750—*T. aestivum* RNA-seq data; SRP037990—*T. aestivum* SynOpDH mapping population; SRP037781—*T. aestivum* synthetic opata M85; SRP037994—*T. aestivum* synthetic W7984). All data can be accessed via the IWGSC repository at Unité de Rercherche Génomique Info: http://wheat-urgi.versailles.inra.fr/Seq-Repository/.

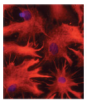
The International Wheat Genome Sequencing Consortium (IWGSC)
Authorship of this paper should be cited as "International Wheat Genome Sequencing Consortium." Participants are arranged by working group. Corresponding authors (*), major contributors (†), and equally contributing authors (‡) are indicated.
Principal Investigators: Klaus F. X. Mayer[1]* (k.mayer@helmholtz-muenchen.de), Jane Rogers[2]* (janerogersh@gmail.com), Jaroslav Doležel[3]* (dolezel@ueb.cas.cz), Curtis Pozniak[4]* (curtis.pozniak@usask.ca), Kellye Eversole[2]* (eversole@eversoleassociates.com), Catherine Feuillet[5]* (catherine.feuillet@bayer.com)
Provision of seed material for ditelosomic wheat lines: Bikram Gill,[6] Bernd Friebe,[6] Adam J. Lukaszewski,[7] Pierre Sourdille,[14] Takashi R Endo[8]
Chromosome sorting and DNA preparation: Jaroslav Doležel,[3]† Marie Kubaláková,[3] Jarmila Číhalíková,[3] Zdeňka Dubská,[3] Jan Vrána,[3] Romana Šperková,[3] Hana Šimková[3]
DNA sequencing: Jane Rogers,[2]† Melanie Febrer,[9] Leah Clissold,[10] Kirsten McLay,[10] Kuldeep Singh,[11] Parveen Chhuneja,[11] Nagendra K. Singh,[12] Jitendra Khurana,[13] Eduard Akhunov,[6] Frédéric Choulet,[14] Pierre Sourdille,[14] Catherine Feuillet,[5] Adriana Alberti,[15] Valérie Barbe,[15] Patrick Wincker,[15] Hiroyuki Kanamori,[16] Fuminori Kobayashi,[16] Takeshi Itoh,[16] Takashi Matsumoto,[16] Hiroaki Sakai,[16] Tsuyoshi Tanaka,[16] Jianzhong Wu,[16] Yasunari Ogihara,[17] Hirokazu Handa,[16] Curtis Pozniak,[4] P. Ron Maclachlan,[4] Andrew Sharpe,[18] Darrin Klassen,[18] David Edwards,[19] Jacqueline Batley,[19] Odd-Arne Olsen,[20,21] Simen Rød Sandve,[20] Sigbjørn Lien,[37] Burkhard Steuernagel,[22] Brande Wulff[22]
DNA sequence assembly: Mario Caccamo,[10]† Sarah Ayling,[10] Ricardo H. Ramirez-Gonzalez,[10] Bernardo J. Clavijo,[10] Burkhard Steuernagel,[22] Jonathan Wright[10]
Gene annotation: Matthias Pfeifer,[1] Manuel Spannagl,[1] Klaus F. X. Mayer[1]†
Genome Zipping: Mihaela M. Martis,[1] Eduard Akhunov,[6] Frédéric Choulet,[14] Klaus F. X. Mayer[1]†
POPSEQ analysis: Martin Mascher,[23] Jarrod Chapman,[24] Jesse A. Poland,[25] Uwe Scholz,[23] Kerrie Barry,[24] Robbie Waugh,[26] Daniel S. Rokhsar,[24] Gary J. Muehlbauer,[27] Nils Stein[28]
Repetitive DNA analysis: Heidrun Gundlach,[1] Matthias Zytnicki,[29] Véronique Jamilloux,[29] Hadi Quesneville,[29] Thomas Wicker,[30] Klaus F. X. Mayer[1]
miRNAs: Primetta Faccioli,[31]‡ Moreno Colaiacovo,[31]‡ Matthias Pfeifer,[1]‡ Antonio Michele Stanca,[31] Hikmet Budak,[32] Luigi Cattivelli[31]†
Genome structure and duplications: Natasha Glover,[14] Mihaela M. Martis,[1] Frédéric Choulet,[14] Catherine Feuillet,[5] Klaus F. X. Mayer[1]
Transcriptome sequencing and expression analysis: Matthias Pfeifer,[1] Lise Pingault,[14] Klaus F. X. Mayer,[1]† Etienne Paux[14]†
Gene family analysis: Manuel Spannagl,[1] Sapna Sharma,[1] Klaus F. X. Mayer,[1]† Curtis Pozniak[4]†
Proteogenomics analysis: Rudi Appels,[33]† Matthew Bellgard,[33] Brett Chapman,[33] Matthias Pfeifer[1]
Comparative analysis of diploid, tetraploid and hexaploid wheat: Matthias Pfeifer,[1] Simen Rød Sandve,[20] Thomas Nussbaumer,[1] Kai Christian Bader,[1] Frédéric Choulet,[14] Catherine Feuillet,[5] Klaus F. X. Mayer[1]†
Development and mapping of marker sets: Eduard Akhunov,[6] Etienne Paux,[14] Hélène Rimbert,[36] Shichen Wang,[6] Jesse A. Poland,[25] Ron Knox,[34] Andrzej Kilian,[35] Curtis Pozniak[4]†
Sequence repository: Michael Alaux,[29]† Françoise Alfama,[29] Loïc Couderc,[29] Véronique Jamilloux,[29] Nicolas Guilhot,[14] Claire Viseux,[29] Mikaël Loaec,[29] Hadi Quesneville[29]
Study design: Jane Rogers,[2] Jaroslav Doležel,[3] Kellye Eversole,[2] Catherine Feuillet,[5] Beat Keller,[30] Klaus F. X. Mayer,[1] Odd-Arne Olsen,[20,21] Sebastien Praud[36]

[1]Plant Genome and Systems Biology, Helmholtz Zentrum Munich, Ingolstädter Landstrasse 1, 85764 Neuherberg, Germany. [2]IWGSC, Eversole Associates, 5207 Wyoming Road, Bethesda, MD 20816, USA. [3]Institute of Experimental Botany, Center of Plant Structural and Functional Genomics, Šlechtitelů 31, 783 71 Olomouc, Czech Republic. [4]Crop Development Centre, Department of Plant Sciences, College of Agriculture and Bioresources, University of Saskatchewan, 51 Campus Drive, Saskatoon, SK, Canada. [5]Bayer Crop Science, 3500 Paramount Parkway, Morrisville, NC 27560, USA. [6]Kansas State University, Department of Plant Pathology, Manhattan, KS 66506–5502, USA. [7]College of Natural and Agricultural Sciences, Botany and Plant Sciences, University of California, Riverside, CA 92521, USA. [8]Laboratory of Plant Genetics, Graduate School of Agriculture, Kyoto University, Kyoto 606-8502, Japan. [9]Genomic Sequencing Unit, University of Dundee, Dow Street, Dundee DD1 5EH, UK. [10]Genome Analysis Centre, Norwich Research Park, Norwich, NR4 7UH, UK. [11]School of Agricultural Biotechnology, Punjab Agricultural University, Ludhiana 141 004, India. [12]National Research Centre on Plant Biotechnology, Indian Agricultural Research Institute, New Delhi 110 012, India. [13]Interdisciplinary Centre for Plant Genomics and Department of Plant Molecular Biology, University of Delhi, South Campus, New Delhi 110 021, India. [14]INRA–University Blaise Pascal UMR1095 Genetics, Diversity and Ecophysiology of Cereals, 5 chemin de Beaulieu, 63039 Clermont-Ferrand, France. [15]Commissariat à l'Energie Atomique Genoscope, Centre National de Séquençage, 2 rue Gaston Crémieux, CP5706, 91057 Evry, France. [16]Plant Genome Research Unit, National Institute of Agrobiological Sciences, 2-1-2, Kan-non-dai, Tsukuba 305-8602, Japan. [17]Kihara Institute for Biological Research, Yokohama City University, Maioka-cho 641-12, Totsuka-ku, 244-0813 Yokohama, Japan. [18]National Research Council Canada, 110 Gymnasium Place, Saskatoon, SK, S7N 0W9, Canada. [19]Australian Centre for Plant Functional Genomics, School of Agriculture and Food Sciences, University of Queensland, St. Lucia, QLD 4072, Australia, and School of Plant Biology, University of Western Australia, WA 6009, Australia. [20]Department of Plant Sciences, Center for Integrative Genetics (CIGENE), Norwegian University of Life Sciences, 1432 Ås, Norway. [21]Department of Natural Science and Technology, Hedmark University College, N-2318, Norway. [22]Sainsbury Laboratory, Norwich Research Park, Norwich, NR4 7UH, UK. [23]Bioinformatics and Information Technology, Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), D-06466 Seeland OT Gatersleben, Germany. [24]U.S. Department of Energy Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA. [25]USDA-ARS Hard Winter Wheat Genetics Research Unit and Department of Agronomy, Kansas State University, Manhattan, KS 66506-5502, USA. [26]James Hutton Institute, Invergowrie, Dundee DD2 5DA, UK. [27]Department of Agronomy and Plant Genetics, Department of Plant Biology, University of Minnesota, St. Paul, MN 55108, USA. [28]Genome Diversity, Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), D-06466 Seeland OT Gatersleben, Germany. [29]INRA, UR1164 URGI–Research Unit in Genomics-Info, INRA de Versailles, Route de Saint-Cyr, Versailles, 78026, France. [30]Institute of Plant Biology, University of Zurich, Zollikerstrasse 107, CH-8008 Zurich, Switzerland. [31]Consiglio per la Ricerca e la sperimentazione in Agricoltura–Genomics Research Centre, via San Protaso 302, I-29017 Fiorenzuola d'Arda, Italy. [32]Sabanci University Biological Sciences and Bioengineering Program, 34956 Istanbul, Turkey. [33]Centre for Comparative Genomics, Murdoch University, Perth, WA 6150, Australia. [34]Semiarid Prairie Agricultural Research Centre, Post Office Box 1030, Swift Current, Saskatchewan S9H 3X2, Canada. [35]Diversity Arrays Technology Pty Limited, 1 Wilf Crane Crescent, Yarralumla ACT2600, Australia. [36]Biogemma, Centre de Recherche de Chappes, Route d'Ennezat, 63720 Chappes, France. [37]Department of Animal and Aquicultural Sciences, CIGENE, Norwegian University of Life Sciences, Arboretvelen 6, 1432 Ås, Norway.

Supplementary Materials

www.sciencemag.org/content/345/6194/1251788/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S60
Tables S1 to S48
References (*88*–*160*)

5 February 2014; accepted 2 June 2014
10.1126/science.1251788

EXTENDED PDF FORMAT
SPONSORED BY

Travel Grants Available
Neuroscience 2014
Application Deadline: July 31st
>Apply Now
R&D systems
a bio-techne brand
www.rndsystems.com

Imaging charge transfer in iodomethane upon x-ray photoabsorption

Benjamin Erk *et al.*

Science **345**, 288 (2014);
DOI: 10.1126/science.1253607

This copy is for your personal, non-commercial use only.

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by clicking here.

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines here.

The following resources related to this article are available online at www.sciencemag.org (this information is current as of August 4, 2014):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:
http://www.sciencemag.org/content/345/6194/288.full.html

Supporting Online Material can be found at:
http://www.sciencemag.org/content/suppl/2014/07/16/345.6194.288.DC1.html

A list of selected additional articles on the Science Web sites **related to this article** can be found at:
http://www.sciencemag.org/content/345/6194/288.full.html#related

This article **cites 38 articles**, 3 of which can be accessed free:
http://www.sciencemag.org/content/345/6194/288.full.html#ref-list-1

This article has been **cited by** 1 articles hosted by HighWire Press; see:
http://www.sciencemag.org/content/345/6194/288.full.html#related-urls

This article appears in the following **subject collections:**
Chemistry
http://www.sciencemag.org/cgi/collection/chemistry

Downloaded from www.sciencemag.org on August 4, 2014

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published weekly, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. Copyright 2014 by the American Association for the Advancement of Science; all rights reserved. The title *Science* is a registered trademark of AAAS.

REPORTS

ULTRAFAST DYNAMICS

Imaging charge transfer in iodomethane upon x-ray photoabsorption

Benjamin Erk,[1,2,3]* Rebecca Boll,[1,2,3] Sebastian Trippel,[4] Denis Anielski,[1,2,3] Lutz Foucar,[2,5] Benedikt Rudek,[2,3,6] Sascha W. Epp,[2,3] Ryan Coffee,[7] Sebastian Carron,[7] Sebastian Schorb,[7] Ken R. Ferguson,[7] Michele Swiggers,[7] John D. Bozek,[7] Marc Simon,[8,9] Tatiana Marchenko,[8,9] Jochen Küpper,[4,10,11] Ilme Schlichting,[2,5] Joachim Ullrich,[6,2,3] Christoph Bostedt,[7] Daniel Rolles,[1,2,5] Artem Rudenko[12,2,3]*

Studies of charge transfer are often hampered by difficulties in determining the charge localization at a given time. Here, we used ultrashort x-ray free-electron laser pulses to image charge rearrangement dynamics within gas-phase iodomethane molecules during dissociation induced by a synchronized near-infrared (NIR) laser pulse. Inner-shell photoionization creates positive charge, which is initially localized on the iodine atom. We map the electron transfer between the methyl and iodine fragments as a function of their interatomic separation set by the NIR–x-ray delay. We observe signatures of electron transfer for distances up to 20 angstroms and show that a realistic estimate of its effective spatial range can be obtained from a classical over-the-barrier model. The presented technique is applicable for spatiotemporal imaging of charge transfer dynamics in a wide range of molecular systems.

Charge-transfer and charge-migration processes—in particular, electron transfer—drive numerous important transformations in physics, chemistry, and biology (*1*–*4*). A crucial issue for experimental studies of such phenomena is the need to trace the spatial localization of charge at a given time. Earlier experiments exploited local photoexcitation of donor-acceptor constituents (*5*, *6*) or selective valence photoionization (*2*, *7*) to trigger the electron transfer, which was traced by resonant absorption (*2*, *6*), ion mass spectrometry (*7*) or time-resolved fluorescence (*5*, *6*). Because of their element specificity, x-rays can be efficiently employed to probe charge-transfer dynamics, triggered either by a synchronized optical pulse (*8*), or by the x-ray absorption itself (*9*). When short-pulsed x-rays are used to initiate the dynamics—for example, by inner-shell ionization—the initial charge can be induced with a high degree of spatial and temporal localization. Therefore, recent developments of intense, ultrashort radiation sources in the extreme-ultraviolet and x-ray domain, such as high-harmonic generation and free-electron lasers (FELs), have opened up a promising way to image charge-transfer processes in space and time with atomic resolution. Moreover, ultrafast dynamics after x-ray absorption—i.e., electronic rearrangement coupled to nuclear motion (*10*)—determine mechanisms and time scales of radiation damage (*11*, *12*), which is of particular relevance for imaging applications employing these light sources (*13*–*15*).

Here, we report on the direct experimental mapping of charge-transfer dynamics upon inner-shell ionization of a small polyatomic molecule, iodomethane (CH_3I). As sketched in Fig. 1, we first dissociate the molecule using a femtosecond near-infrared (NIR) laser and then ionize the system with an intense x-ray pulse from the Linac Coherent Light Source (LCLS) (*16*) at different internuclear separations that are defined by the delay between the NIR and the x-ray pulses. At the intensity chosen in our experiment, the NIR dissociation predominantly proceeds as two-body breakup, creating CH_3 and atomic I fragments, both either neutral or singly charged (*17*, *18*). The x-ray pulse, arriving after a well-defined time delay, ionizes almost exclusively the M shell of the iodine atom, because for the photon energy of 1500 eV used here, its x-ray absorption cross section is ~60 times as high as that of the rest of the molecule (*19*). Inner-shell photoionization and subsequent fast (subfemtosecond) Auger decay, which does not involve the valence electrons, induces a positive charge that is initially strongly localized on the iodine site. The charge then spreads over the entire molecule, which undergoes fast breakup. Depending on the internuclear distance at the time of the x-ray absorption, two different regimes of this charge redistribution can be considered. When treating the iodine ion and the methyl group in the dissociating molecule as two independent systems, where the valence

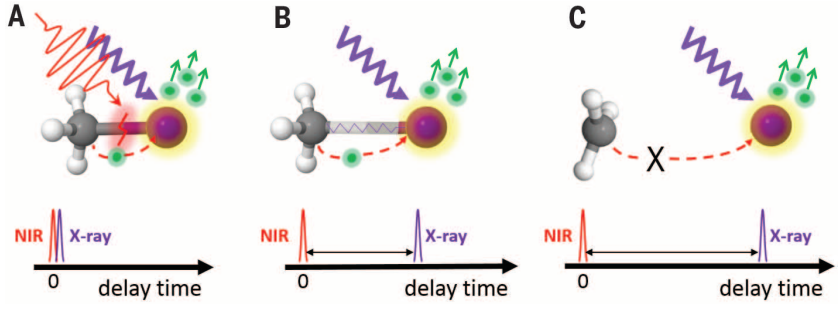


Fig. 1. Schematic of the NIR pump–x-ray probe experiment. The 800-nm NIR pulse (red) dissociates the molecule. The x-ray pulse (purple) arrives after a certain time delay and creates one or a few inner-shell vacancies through absorption of single or multiple photons by the iodine M shell. This is followed by a fast local Auger decay that creates additional charge. For each of the dissociation channels, the delay between the NIR and the x-ray pulses can be mapped onto the resulting internuclear distance R at the time of the x-ray pulse arrival. At small delays (**A**), i.e., in a bound molecule, the delocalized valence electrons can move between the atoms and participate in the decay processes filling the inner-shell holes at the iodine site, such that the resulting charge is shared between the two fragments. In the intermediate regime (**B**), the electrons exhibit a certain degree of localization, and the final charge distribution between the fragments is defined by the R-dependent electron-transfer probability from the methyl group to the iodine. At large delays (and, thus, at large internuclear separations) (**C**), the probability of charge transfer between the two sites becomes negligible, and the final-state charge distribution results from the x-ray interaction with the isolated methyl group and atomic iodine, respectively.

[1]Deutsches Elektronen-Synchrotron (DESY), 22607 Hamburg, Germany. [2]Max Planck Advanced Study Group at CFEL, 22607 Hamburg, Germany. [3]Max-Planck-Institut für Kernphysik, 69117 Heidelberg, Germany. [4]Center for Free-Electron Laser Science (CFEL), DESY, 22607 Hamburg, Germany. [5]Max-Planck-Institut für Medizinische Forschung, 69120 Heidelberg, Germany. [6]Physikalisch-Technische Bundesanstalt, 38116 Braunschweig, Germany. [7]Linac Coherent Light Source, SLAC National Accelerator Laboratory, 94025 Menlo Park, CA, USA. [8]Sorbonne Universités, UPMC Université Paris 06, Laboratoire de Chimie Physique Matière et Rayonnement, F-75005, Paris, France. [9]CNRS, Laboratoire de Chimie Physique Matière et Rayonnement, F-75005, Paris, France. [10]Department of Physics, University of Hamburg, 22761 Hamburg, Germany. [11]Center for Ultrafast Imaging, University of Hamburg, 22761 Hamburg, Germany. [12]J. R. Macdonald Laboratory, Department of Physics, Kansas State University, Manhattan, KS 66506, USA.
*Corresponding author. E-mail: benjamin.erk@desy.de (B.E.), rudenko@phys.ksu.edu (A.R.)

electrons are localized at the two respective sites (atomic regime), charge rearrangement can occur via valence-electron transfer (*20*, *21*) or by nonlocal (interatomic) Auger processes (*22*–*24*). In the case of an intact molecule, which is best described in a molecular orbital picture, this distinction is not well defined, and the relaxation of a localized hole is typically considered as molecular Auger decay, potentially involving any of the delocalized valence electrons (*23*, *24*).

By measuring the charge state and the kinetic energy distributions of the resulting fragment ions as a function of NIR–x-ray delay, we probe the charge redistribution between the fragments as a function of their internuclear separation R. Figure 2A shows the measured kinetic energy distribution for the exemplary I^{6+} fragment as a function of time delay between the NIR and x-ray pulses. The kinetic energy of the iodine ions mainly originates from the Coulomb repulsion between the fragments and thus directly reflects the distance R between them (*24*). Three major features (marked 1, 2, and 3 in Fig. 2) can be clearly distinguished in the time-resolved kinetic energy spectrum. I^{6+} fragments in the broad high-energy band 1 centered at ~7 eV originate from molecules that absorb one or several x-ray photons at or close to the equilibrium geometry of the molecule. These are events in which the molecules either did not interact with the NIR pulse or remained bound after this interaction—for example, as CH_3I^+ or CH_3I^{2+} ions. This broad feature closely resembles the kinetic energy distribution of I^{6+} fragments that was measured with only the x-ray pulse (blue dots in Fig. 2B). The delay-dependent structure 2, which descends from the high-energy band 1 at small time delays to low kinetic energies as the delay increases, asymptotically approaching 0.54 eV, corresponds to events where the molecule is first dissociated by the NIR laser and later multiply ionized by the x-ray pulse. The internuclear distance in the dissociating molecule increases with the pump-probe delay, which is reflected in decreasing fragment kinetic energies. At the NIR intensity used in our measurement (~8×10^{13} W/cm^2), the dominant NIR-induced dissociation channels are $CH_3 + I^+$ (0,1 channel), $CH_3^+ + I$ (1,0), and $CH_3^+ + I^+$ (1,1) (see the supplementary materials for details). Because the x-ray pulse almost exclusively ionizes the iodine fragment, the latter two channels result in a Coulomb explosion after the x-ray absorption. The dashed lines in Fig. 2A show the simulated delay-dependent kinetic energy of I^{6+} ions, which are produced by the NIR dissociation into the (1,1) channel, where both fragments are initially singly charged, and subsequent multiple ionization by the x-ray pulse results in the Coulomb explosion into $I^{6+} + C^+$ (lower-energy line) or $I^{6+} + C^{2+}$ (higher energy line) final states.

For the (0,1) channel, the situation is considerably different. Here, a Coulomb explosion happens only if part of the charge that is induced on the iodine is transferred to the initially neutral methyl group, because x-ray absorption by the CH_3 itself is very unlikely given the corresponding cross sections. If no charge transfer occurs, the final state after the interaction with both the NIR and x-ray pulses is $I^{6+} + CH_3$, which results in I^{6+} fragments with very low kinetic energies. These are visible as the narrow band 3 in Fig. 2A that is centered around the asymptotic NIR dissociation energy at 0.06 eV (see the supplementary materials). Both channels 2 and 3 also appear as two low-energy peaks in the projection in Fig. 2B. They are completely absent in the spectrum recorded with only the x-ray pulse, confirming that they originate from the combination of NIR and x-ray pulses.

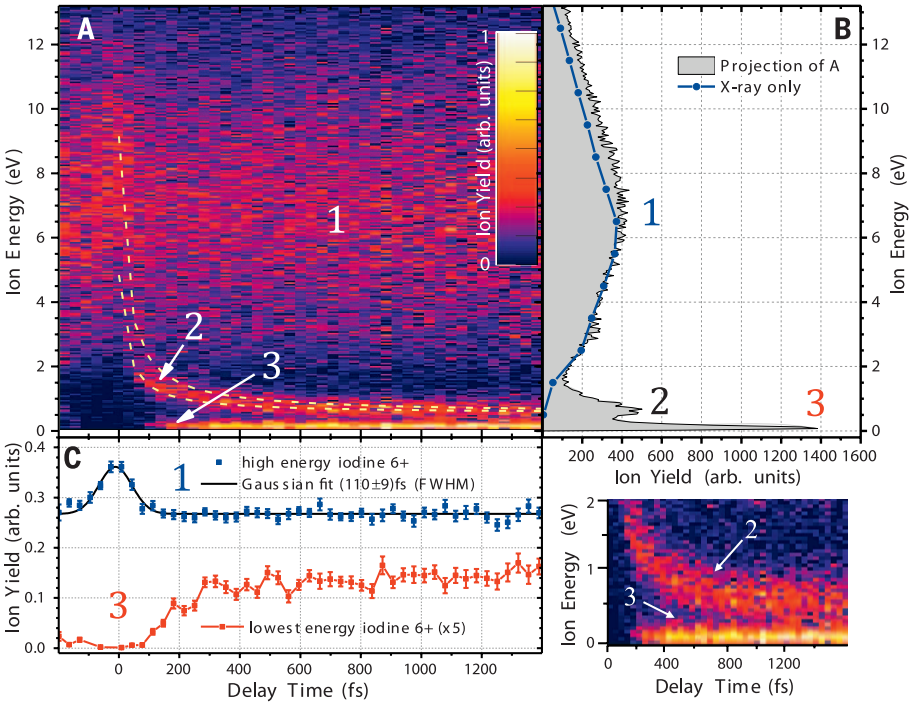


Fig. 2. Map of the delay-dependent kinetic-energy distribution of I^{6+} ions. (**A**) Measured kinetic-energy distribution of all I^{6+} ions as a function of time delay between the NIR and the x-ray pulse. The dashed lines show results of Coulomb explosion simulations for the (1,1) dissociation channel. The inset in the lower right corner displays an enlarged view of the low-energy region of the spectrum. (**B**) Projection of (A) onto the kinetic-energy axis. Channel 1 closely resembles the I^{6+} spectrum measured with only the x-ray pulse (blue dots). Channel 2 peaks at the asymptotic energy of 0.54 eV, and channel 3 exhibits a very sharp peak at 0.06 eV. The latter two features are completely absent in the data recorded with only the x-ray pulse. (**C**) Delay-dependent yields of channel 1 (blue, energies > 6 eV) and channel 3 (red, energies < 0.3 eV). Channel 1 shows a pronounced increase when the NIR and the x-ray pulse overlap in time. This cross-correlation signal can be fitted by a Gaussian of ~110 fs (FWHM), which provides a measure of the experimental time resolution as well as the nominal "zero" delay (*29*).

The low-energy band 3 in Fig. 2A does not appear at time delays smaller than 100 fs, as can be seen in Fig. 2C, where the yield of this channel is plotted as a function of the pump-probe delay. After the first 100 fs, this channel exhibits a fast increase before leveling off above ~300 fs. This indicates that the methyl fragment cannot remain neutral if it is in close proximity to the highly charged iodine ion; that is, charge rearrangement occurs. To investigate the delay dependence, and thus the spatial range of this rearrangement (or electron transfer) in more detail, we measured the delay-dependent yield of the lowest-energy channel 3 (Fig. 3A) for four different charge states of iodine. Assuming a constant dissociation velocity and using the measured asymptotic value of kinetic energy for the NIR-induced (0,1) dissociation channel, we can map the measured NIR–x-ray delay to internuclear distance (see the supplementary materials for details). The initial zero yield of this channel indicates that at rather small distances, at least one electron always leaves the methyl group or, in other words, the electron-transfer probability is saturated. For larger delays, the transfer rate decreases, and there is a finite probability for the CH_3 fragment to remain neutral, which is reflected in the increasing yield of low-energy iodine fragments. Beyond a certain distance, the charge transfer between iodine and carbon becomes prohibitively improbable. As can be clearly seen from the data, this critical internuclear distance R_{cr} depends on the iodine charge state.

This dependence can be understood by considering the classical over-the-barrier charge-transfer model for a two-center system developed in (*20*, *25*). There, the binding energy of the electron in one of the fragments is considered in the Coulomb field of its partner. At the critical distance R_{cr}, the height of the classical potential barrier between the two centers becomes larger than the binding energy of the valence electrons, such that the electrons are localized and cannot be transferred to the other fragment anymore. The critical values of the internuclear distance

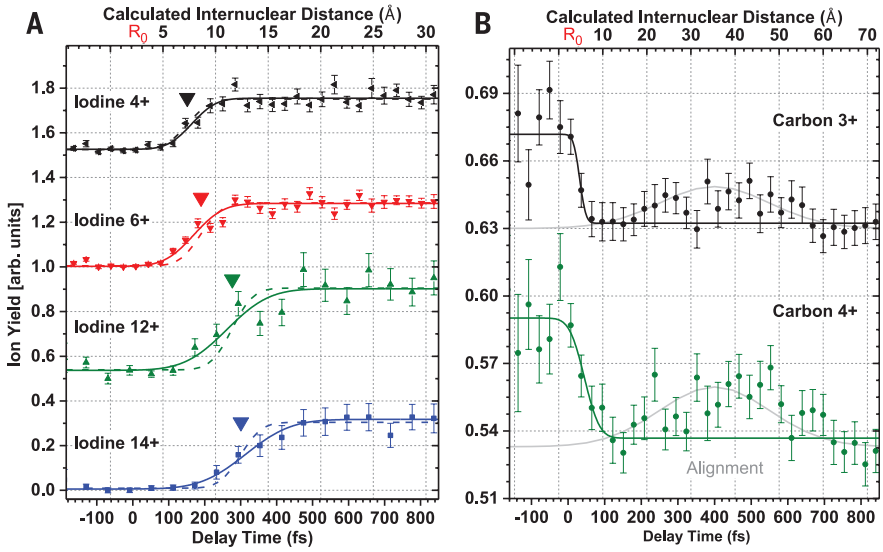


Fig. 3. Delay-dependent yields of iodine and carbon fragments. (**A**) Yields of different iodine charge states in the lowest-energy channel 3 (see Fig. 2) resulting from the iodomethane dissociation into I^{n+} and neutral CH_3. For this channel, the delay time can be mapped into an internuclear distance *R*, as indicated on the top axis. This is possible because these events originate from the unambiguously identified NIR-induced (0,1) dissociation pathway. The solid lines show fits to the data. The triangles indicate critical distances beyond which the charge transfer is forbidden within the classical over-the-barrier model. The dashed lines depict the predictions of this model convolved with the experimental temporal resolution. (**B**) Corresponding yields of C^{3+} and C^{4+} fragment ions, which show a very fast decrease at time delays < 100 fs. Here, the *R* values shown on the top axis are calculated for the fastest of three possible NIR-induced pathways, the (1,1) channel, indicating the upper limit of the internuclear separation at a given delay time. Error bars, mean ± 1 SD. The data result from eight individual delay scans averaging over a total of ~8×10^5 LCLS shots.

for different iodine charge states obtained from this model are shown as triangles in Fig. 3A. According to this classical picture, the charge transfer is forbidden for $R > R_{cr}$ and allowed for $R < R_{cr}$. The dashed lines in Fig. 3A depict the resulting step function convolved with the time resolution of the experiment as estimated from the cross-correlation signal shown in blue in Fig. 2C. For the I^{4+} yield in Fig. 3A, the simulation matches the experimental data well, whereas for higher-charge states, a considerable discrepancy is observed. This discrepancy reflects an *R* dependence of the charge-transfer probability, which goes beyond the simple step function described by the classical model. It can be qualitatively understood by considering that the charge-transfer rate is expected to decrease exponentially with *R*, with the exponent given by a characteristic inverse distance defined by the overlap of the two wave functions localized at the respective sites (*26*). This characteristic distance can be associated with the critical distance R_{cr} obtained from the classical over-the-barrier model, resulting in the longer saturation and the slower rise time of the signal for higher charge states in Fig. 3A.

As a complementary set of observables highlighting the charge-transfer dynamics, we plot the delay-dependent yields of the carbon fragments C^{3+} and C^{4+} in Fig. 3B. These carbon ions mainly acquire their charge not from direct photoabsorption but from electron transfer to the initially ionized iodine. Their highest yield therefore occurs for intact molecules, corresponding to time delays ≤ 0 fs, while their yield decreases for dissociating molecules within the first hundred femtoseconds. The effective time resolution of this experiment is around 110 fs, as can be seen from the x-ray–NIR cross-correlation signal shown in Fig. 2C. Therefore, it is likely that the fall time of the C^{3+} and C^{4+} yield curves in Fig. 3B is defined by this time resolution and that the real decrease in rate due to the suppression of the charge transfer at larger *R* might be even faster. Another pronounced feature in Fig. 3B is a broad hump in the yield of the carbon ions centered at ~400 fs that can be fitted by a Gaussian function with a full width at half maximum (FWHM) of 300 fs. It appears as a consequence of impulsive laser alignment of the CH_3I molecules by the NIR pulse and a limited angular acceptance of our experimental setup for very energetic ions (see the supplementary materials). In brief, the alignment along the NIR laser polarization direction, which is parallel to the spectrometer axis, effectively enhances the energy acceptance, resulting in the increase of the carbon ion signals at the corresponding delays. The position and width of the observed local maximum matches the results of earlier measurements on impulsive alignment of iodomethane (*27*).

Contrary to the lowest-energy iodine fragments discussed above, the "precursor" dissociation channel after the NIR pulse cannot be unambiguously defined for the data of Fig. 3B. In addition, for the carbon ion yields shown in Fig. 3B, the total final charge state of the molecule is not well defined. For these two reasons, we cannot perform the same quantitative analysis for the carbon fragments. Nevertheless, the trend in yields of these carbon ions, produced as a result of the charge transfer, supports our interpretation of electronic rearrangement after x-ray absorption. The very fast initial decrease of the highly charged carbon ion yields indicates that their production is more sensitive to increasing internuclear distance than the yields of the lowest-energy I^{n+} channels.

The experiment described here establishes a link between studies of charge transfer phenomena on the level of individual atoms, as previously realized in atomic collisions with highly charged ion beams (*20*, *21*), and experiments on charge flow in extended systems relevant for chemistry or biology (*2*–*7*). The measured critical internuclear distances supported by the classical over-the-barrier model mark the transition between what can be considered the molecular and the atomic regimes mentioned in the introduction. The electron-transfer rate, which decreases as a function of the internuclear distance, reflects the increasing degree of localization of the valence electrons as the system evolves from a bound molecule to isolated methyl and iodine fragments. Based on the distance dependence measured in this work, we cannot clearly distinguish valence electron transfer after the local (intra-atomic) Auger decay from interatomic Auger-like mechanisms (*22*, *23*) because, for both cases, the *R* dependence is defined by the overlap of the two wave functions, which decreases exponentially with increasing distance *R* (*26*, *28*). This distinction could be resolved by measuring internuclear distance-dependent Auger electron spectra of dissociating molecules.

REFERENCES AND NOTES

1. R. A. Marcus, *Rev. Mod. Phys.* **65**, 599–610 (1993).
2. R. Weinkauf *et al.*, *J. Phys. Chem.* **100**, 18567–18585 (1996).
3. G. L. Closs, J. R. Miller, *Science* **240**, 440–447 (1988).
4. A. de la Lande *et al.*, *Phys. Chem. Chem. Phys.* **14**, 5902–5918 (2012).
5. K. S. Schanze, K. Sauer, *J. Am. Chem. Soc.* **110**, 1180–1186 (1988).
6. M. Maus, W. Rettig, D. Bonafoux, R. Lapouyade, *J. Phys. Chem. A* **103**, 3388–3401 (1999).
7. L. Belshaw *et al.*, *J. Phys. Chem. Lett.* **3**, 3751–3754 (2012).
8. N. Huse *et al.*, *J. Phys. Chem. Lett* **2**, 880–884 (2011).
9. P. A. Brühwiler, O. Karis, N. Mårtensson, *Rev. Mod. Phys.* **74**, 703–740 (2002).
10. B. Erk *et al.*, *Phys. Rev. Lett.* **110**, 053003 (2013).
11. L. Lomb *et al.*, *Phys. Rev. B* **84**, 214111 (2011).
12. A. Barty *et al.*, *Nat. Photonics* **6**, 35–40 (2012).
13. J. Ullrich, A. Rudenko, R. Moshammer, *Annu. Rev. Phys. Chem.* **63**, 635–660 (2012).
14. J. C. H. Spence, U. Weierstall, H. N. Chapman, *Rep. Prog. Phys.* **75**, 102601 (2012).
15. S. Boutet *et al.*, *Science* **337**, 362–364 (2012).
16. P. Emma *et al.*, *Nat. Photonics* **4**, 641–647 (2010).
17. H. Liu, Z. Yang, Z. Gao, Z. Tang, *J. Chem. Phys.* **126**, 044316 (2007).
18. M. E. Corrales *et al.*, *J. Phys. Chem. A* **116**, 2669–2677 (2012).
19. J. J. Yeh, I. Lindau, *At. Data Nucl. Data Tables* **32**, 1–155 (1985).
20. H. Ryufuku, K. Sasaki, T. Watanabe, *Phys. Rev. A* **21**, 745–750 (1980).
21. G. Lubinski, Z. Juhász, R. Morgenstern, R. Hoekstra, *Phys. Rev. Lett.* **86**, 616–619 (2001).

22. J. A. D. Matthew, M. Komninos, *Surf. Sci.* **53**, 716–725 (1975).
23. F. Tarantelli, L. S. Cederbaum, *Phys. Rev. Lett.* **71**, 649–652 (1993).
24. T. D. Thomas *et al.*, *Phys. Rev. Lett.* **89**, 223001 (2002).
25. A. Niehaus, *J. Phys. At. Mol. Opt. Phys.* **19**, 2925–2937 (1986).
26. V. May, O. Kühn, *Charge and Energy Transfer Dynamics in Molecular Systems*, (Wiley-VCH, Weinheim, 2011).
27. E. Hamilton *et al.*, *Phys. Rev. A* **72**, 043402 (2005).
28. V. Averbukh, L. S. Cederbaum, *J. Chem. Phys.* **123**, 204107 (2005).
29. M. Krikunova *et al.*, *New J. Phys.* **11**, 123019 (2009).

ACKNOWLEDGMENTS

Parts of this research were carried out at the Linac Coherent Light Source (LCLS) at the SLAC National Accelerator Laboratory. LCLS is an Office of Science User Facility operated for the U.S. Department of Energy Office of Science by Stanford University. We acknowledge the Max Planck Society for funding the development and operation of the CFEL-ASG-Multi-Purpose (CAMP) instrument within the Advanced Study Group at CFEL. A.R. acknowledges the support from Chemical Sciences, Geosciences, and Biosciences Division, Office of Basic Energy Sciences, Office of Science, U.S. Department of Energy and from Kansas National Science Foundation Experimental Program to Stimulate Competitive Research (NSF EPSCoR) "First Award"; D.R. acknowledges support from the Helmholtz Gemeinschaft through the Young Investigator Program. J. K and S. T. acknowledge support from the excellence cluster "The Hamburg Centre for Ultrafast Imaging - Structure, Dynamics and Control of Matter at the Atomic Scale" of the Deutsche Forschungsgemeinschaft. We are grateful to C. Schmidt for technical support, to M. Cryle for help with sample selection, and to the SLAC staff for their support and hospitality during the beamtime. We thank R. Moshammer, K. Schnorr A. Senftleben, and T. Pfeifer for fruitful discussions and for the idea to adapt the classical over-the-barrier model for the interpretation of the data presented in this manuscript.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/345/6194/288/suppl/DC1
Materials and Methods
Figs. S1 to S4
Tables S1 to S3
References (*30*–*39*)

19 March 2014; accepted 9 June 2014
10.1126/science.1253607

ASYMMETRIC CATALYSIS

High-turnover hypoiodite catalysis for asymmetric synthesis of tocopherols

Muhammet Uyanik,[1] Hiroki Hayashi,[1] Kazuaki Ishihara[1,2]*

The diverse biological activities of tocopherols and their analogs have inspired considerable interest in the development of routes for their efficient asymmetric synthesis. Here, we report that chiral ammonium hypoiodite salts catalyze highly chemo- and enantioselective oxidative cyclization of γ-(2-hydroxyphenyl)ketones to 2-acyl chromans bearing a quaternary stereocenter, which serve as productive synthetic intermediates for tocopherols. Raman spectroscopic analysis of a solution of tetrabutylammonium iodide and *tert*-butyl hydroperoxide revealed the in situ generation of the hypoiodite salt as an unstable catalytic active species and triiodide salt as a stable inert species. A high-performance catalytic oxidation system (turnover number of ~200) has been achieved through reversible equilibration between hypoiodite and triiodide in the presence of potassium carbonate base. We anticipate that these findings will open further prospects for the development of high-turnover redox organocatalysis.

Biologically active compounds containing a chiral chroman skeleton are abundant in nature (*1*). One of the most prominent chiral chromans is α-tocopherol, which is in the same family as vitamin E, together with β-, γ-, and δ-tocopherols and the corresponding α-, β-, γ-, and δ-tocotrienols (Fig. 1A) (*2*). α-Tocopherol acts as a lipid-soluble, chain-breaking antioxidant by capturing free radicals or singlet oxygen formed by oxidative metabolism in tissues (*3*). Furthermore, tocopherols show physiologically diverse properties, including anti-tumor, anti-inflammatory, anti-atherosclerosis, and cell-signaling activities (*4*–*8*). Various non-natural tocopherol analogs have also been developed because of their potency and distinct structural features. For example, trolox and its derivatives have been used as a chiral derivatizing reagent (*9*) and NO-releasing drug candidates (*10*). Dihydrodaedalin A, a synthetic intermediate for the natural product daedalin A (*11*), is a potent tyrosinase inhibitor and has been shown to suppress melanogenesis in human skin without affecting cell viability (*12*). C48 developed by Merck research laboratories (Rahway, New Jersey) is a potent, selective peroxisome proliferator–activated receptor–α/γ dual agonist and exhibits substantial antihyperglycemic and hypolipidemic activities (*13*).

Natural *D*-α-tocopherol is formally (2*R*,4′*R*,8′*R*)-2,5,7,8-tetramethyl-2-(4′,8′,12′-trimethyltridecyl)-6-chromanol. The (2*R*)-configuration at the chroman ring among the three stereocenters of α-tocopherol is critical for its bioactivity; (2*S*)-stereoisomers are not accepted by the tocopherol transfer protein (*5*). Thus, asymmetric construction of the chroman ring has been an important challenge (*1*, *2*). Enantioselective processes have been developed to address this need by using chiral transition metal complexes (Pd or Ru) (*14*–*16*) or organocatalysts (*17*, *18*). However, the low catalytic activities and/or moderate enantioselectivities have limited their utility (*2*). Here, we report a transition metal–free (*19*) approach for the enantioselective synthesis of tocopherols and their analogs by using chiral hypoiodite catalysts (*20*, *21*) generated in situ from the corresponding quaternary ammonium iodide with alkyl hydroperoxides as environmentally benign co-oxidants (Fig. 1B). Chemoselective oxidative cyclization of hydroquinone-derived γ-(2-hydroxyphenyl)ketones

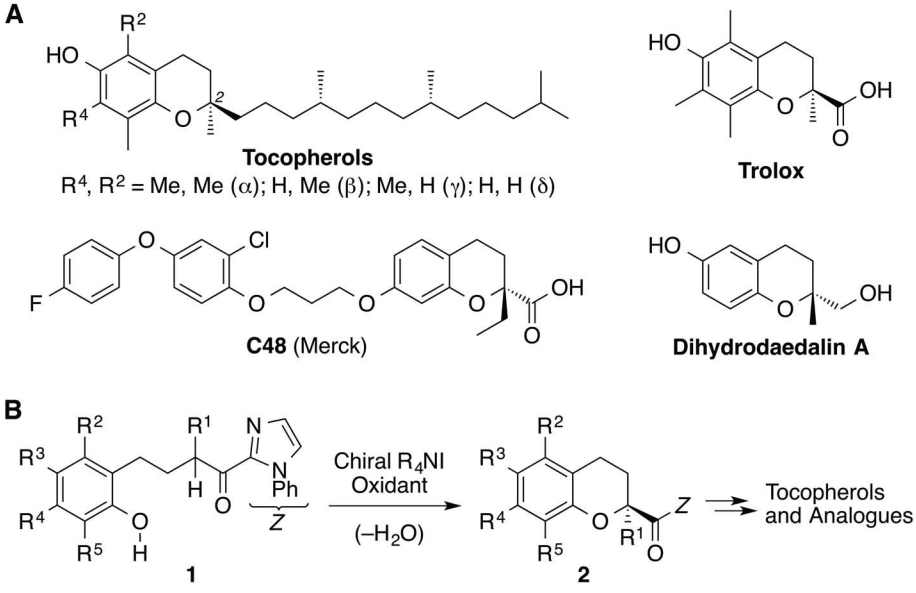
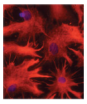


Fig. 1. A catalytic route to the tocopherol core. (**A**) Biologically active tocopherols and other chromans. (**B**) Enantioselective oxidative cycloetherification of γ-(2-hydroxyphenyl)ketones **1** to 2-acyl chromans **2** by in situ–generated chiral hypoiodite catalysts.

[1]Graduate School of Engineering, Nagoya University, Furo-cho, Chikusa, Nagoya 464-8603, Japan. [2]Core Research for Evolutional Science and Technology, Japan Science and Technology Agency, Furo-cho, Chikusa, Nagoya 464-8603, Japan.
***Corresponding author. E-mail: ishihara@cc.nagoya-u.ac.jp**

EXTENDED PDF FORMAT
SPONSORED BY

Travel Grants Available
Neuroscience 2014
Application Deadline: July 31st
>Apply Now
R&D systems
a bietechne brand
www.rndsystems.com

High-turnover hypoiodite catalysis for asymmetric synthesis of tocopherols

Muhammet Uyanik *et al.*

Science **345**, 291 (2014);
DOI: 10.1126/science.1254976

This copy is for your personal, non-commercial use only.

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by clicking here.

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines here.

The following resources related to this article are available online at www.sciencemag.org (this information is current as of August 4, 2014):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:
http://www.sciencemag.org/content/345/6194/291.full.html

Supporting Online Material can be found at:
http://www.sciencemag.org/content/suppl/2014/07/16/345.6194.291.DC1.html

A list of selected additional articles on the Science Web sites **related to this article** can be found at:
http://www.sciencemag.org/content/345/6194/291.full.html#related

This article **cites 46 articles**, 2 of which can be accessed free:
http://www.sciencemag.org/content/345/6194/291.full.html#ref-list-1

This article has been **cited by** 1 articles hosted by HighWire Press; see:
http://www.sciencemag.org/content/345/6194/291.full.html#related-urls

This article appears in the following **subject collections:**
Chemistry
http://www.sciencemag.org/cgi/collection/chemistry

Downloaded from www.sciencemag.org on August 4, 2014

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published weekly, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. Copyright 2014 by the American Association for the Advancement of Science; all rights reserved. The title *Science* is a registered trademark of AAAS.

22. J. A. D. Matthew, M. Komninos, *Surf. Sci.* **53**, 716–725 (1975).
23. F. Tarantelli, L. S. Cederbaum, *Phys. Rev. Lett.* **71**, 649–652 (1993).
24. T. D. Thomas *et al.*, *Phys. Rev. Lett.* **89**, 223001 (2002).
25. A. Niehaus, *J. Phys. At. Mol. Opt. Phys.* **19**, 2925–2937 (1986).
26. V. May, O. Kühn, *Charge and Energy Transfer Dynamics in Molecular Systems*, (Wiley-VCH, Weinheim, 2011).
27. E. Hamilton *et al.*, *Phys. Rev. A* **72**, 043402 (2005).
28. V. Averbukh, L. S. Cederbaum, *J. Chem. Phys.* **123**, 204107 (2005).
29. M. Krikunova *et al.*, *New J. Phys.* **11**, 123019 (2009).

ACKNOWLEDGMENTS

Parts of this research were carried out at the Linac Coherent Light Source (LCLS) at the SLAC National Accelerator Laboratory. LCLS is an Office of Science User Facility operated for the U.S. Department of Energy Office of Science by Stanford University. We acknowledge the Max Planck Society for funding the development and operation of the CFEL-ASG-Multi-Purpose (CAMP) instrument within the Advanced Study Group at CFEL. A.R. acknowledges the support from Chemical Sciences, Geosciences, and Biosciences Division, Office of Basic Energy Sciences, Office of Science, U.S. Department of Energy and from Kansas National Science Foundation Experimental Program to Stimulate Competitive Research (NSF EPSCoR) "First Award"; D.R. acknowledges support from the Helmholtz Gemeinschaft through the Young Investigator Program. J. K and S. T. acknowledge support from the excellence cluster "The Hamburg Centre for Ultrafast Imaging - Structure, Dynamics and Control of Matter at the Atomic Scale" of the Deutsche Forschungsgemeinschaft. We are grateful to C. Schmidt for technical support, to M. Cryle for help with sample selection, and to the SLAC staff for their support and hospitality during the beamtime. We thank R. Moshammer, K. Schnorr A. Senftleben, and T. Pfeifer for fruitful discussions and for the idea to adapt the classical over-the-barrier model for the interpretation of the data presented in this manuscript.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/345/6194/288/suppl/DC1
Materials and Methods
Figs. S1 to S4
Tables S1 to S3
References (*30*–*39*)

19 March 2014; accepted 9 June 2014
10.1126/science.1253607

ASYMMETRIC CATALYSIS

High-turnover hypoiodite catalysis for asymmetric synthesis of tocopherols

Muhammet Uyanik,[1] Hiroki Hayashi,[1] Kazuaki Ishihara[1,2]*

The diverse biological activities of tocopherols and their analogs have inspired considerable interest in the development of routes for their efficient asymmetric synthesis. Here, we report that chiral ammonium hypoiodite salts catalyze highly chemo- and enantioselective oxidative cyclization of γ-(2-hydroxyphenyl)ketones to 2-acyl chromans bearing a quaternary stereocenter, which serve as productive synthetic intermediates for tocopherols. Raman spectroscopic analysis of a solution of tetrabutylammonium iodide and *tert*-butyl hydroperoxide revealed the in situ generation of the hypoiodite salt as an unstable catalytic active species and triiodide salt as a stable inert species. A high-performance catalytic oxidation system (turnover number of ~200) has been achieved through reversible equilibration between hypoiodite and triiodide in the presence of potassium carbonate base. We anticipate that these findings will open further prospects for the development of high-turnover redox organocatalysis.

Biologically active compounds containing a chiral chroman skeleton are abundant in nature (*1*). One of the most prominent chiral chromans is α-tocopherol, which is in the same family as vitamin E, together with β-, γ-, and δ-tocopherols and the corresponding α-, β-, γ-, and δ-tocotrienols (Fig. 1A) (*2*). α-Tocopherol acts as a lipid-soluble, chain-breaking antioxidant by capturing free radicals or singlet oxygen formed by oxidative metabolism in tissues (*3*). Furthermore, tocopherols show physiologically diverse properties, including anti-tumor, anti-inflammatory, anti-atherosclerosis, and cell-signaling activities (*4*–*8*). Various non-natural tocopherol analogs have also been developed because of their potency and distinct structural features. For example, trolox and its derivatives have been used as a chiral derivatizing reagent (*9*) and NO-releasing drug candidates (*10*). Dihydrodaedalin A, a synthetic intermediate for the natural product daedalin A (*11*), is a potent tyrosinase inhibitor and has been shown to suppress melanogenesis in human skin without affecting cell viability (*12*). C48 developed by Merck research laboratories (Rahway, New Jersey) is a potent, selective peroxisome proliferator–activated receptor–α/γ dual agonist and exhibits substantial antihyperglycemic and hypolipidemic activities (*13*).

Natural *D*-α-tocopherol is formally (2*R*,4′*R*,8′*R*)-2,5,7,8-tetramethyl-2-(4′,8′,12′-trimethyltridecyl)-6-chromanol. The (2*R*)-configuration at the chroman ring among the three stereocenters of α-tocopherol is critical for its bioactivity; (2*S*)-stereoisomers are not accepted by the tocopherol transfer protein (*5*). Thus, asymmetric construction of the chroman ring has been an important challenge (*1*, *2*). Enantioselective processes have been developed to address this need by using chiral transition metal complexes (Pd or Ru) (*14*–*16*) or organocatalysts (*17*, *18*). However, the low catalytic activities and/or moderate enantioselectivities have limited their utility (*2*). Here, we report a transition metal–free (*19*) approach for the enantioselective synthesis of tocopherols and their analogs by using chiral hypoiodite catalysts (*20*, *21*) generated in situ from the corresponding quaternary ammonium iodide with alkyl hydroperoxides as environmentally benign co-oxidants (Fig. 1B). Chemoselective oxidative cyclization of hydroquinone-derived γ-(2-hydroxyphenyl)ketones

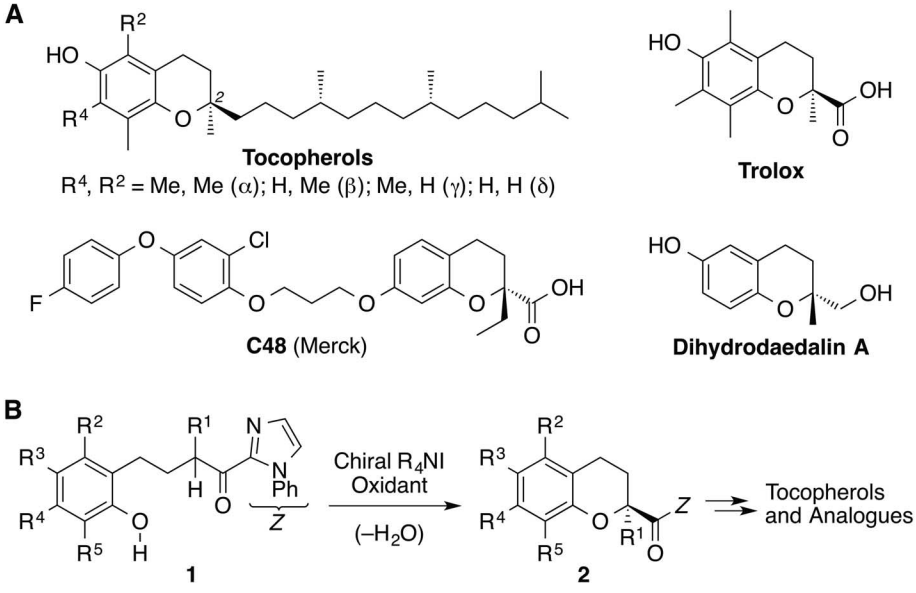


Fig. 1. A catalytic route to the tocopherol core. (**A**) Biologically active tocopherols and other chromans. (**B**) Enantioselective oxidative cycloetherification of γ-(2-hydroxyphenyl)ketones **1** to 2-acyl chromans **2** by in situ–generated chiral hypoiodite catalysts.

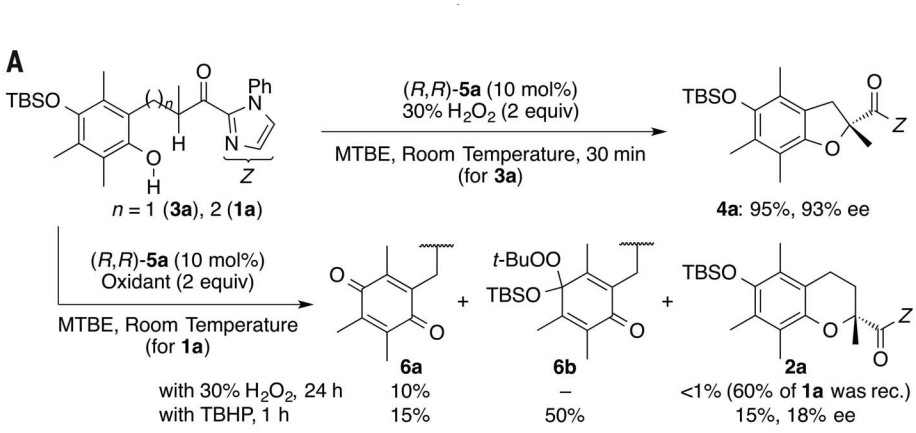
[1]Graduate School of Engineering, Nagoya University, Furo-cho, Chikusa, Nagoya 464-8603, Japan. [2]Core Research for Evolutional Science and Technology, Japan Science and Technology Agency, Furo-cho, Chikusa, Nagoya 464-8603, Japan.
***Corresponding author. E-mail: ishihara@cc.nagoya-u.ac.jp**

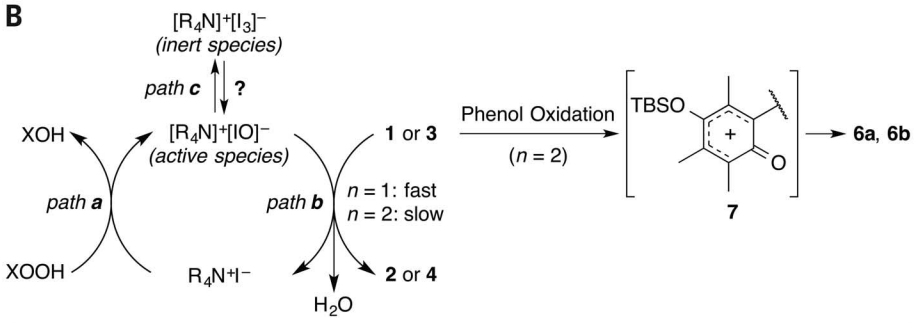
gives >95% yield of the corresponding 2-acyl chromans; with the quaternary stereocenter set in high enantioselectivity, these products are poised for elaboration to a range of tocopherols.

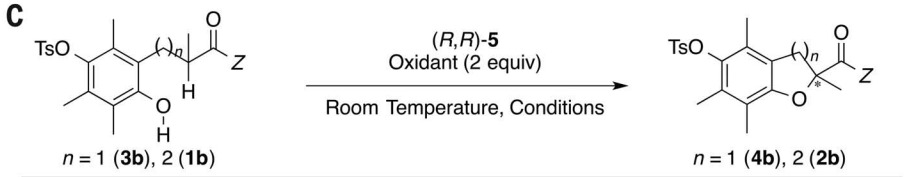
Previously, we developed in situ–generated chiral quaternary ammonium hypoiodite (R_4NOI) catalysis for the enantioselective oxidative cyclization of β-(2-hydroxyphenyl)ketones to the five-membered ring products 2-acyl-2,3-dihydrobenzofurans with hydrogen peroxide or *tert*-butyl hydroperoxide (TBHP) as co-oxidants (*20*). The use of chiral binaphthyl-based quaternary ammonium (*22*) iodide as a precatalyst and an *N*-phenylimidazol-2-yl (*Z*) (*23*) group as an auxiliary of β-(2-hydroxyphenyl) ketones was effective for inducing high enantioselectivities. We envisioned that the enantioselective oxidative cyclization of γ-(2-hydroxyphenyl) ketones (**1**) would give the desired six-membered ring 2-acyl chromans (**2**) (Fig. 1B).

To begin our investigation, we compared the enantioselective oxidative cyclization of (*5-tert*-butyldimethylsilyloxy-2-hydroxyphenyl)ketones (**1a** and ***3a***) derived from trimethyl hydroquinone (Fig. 2A). Oxidative cyclization of ***3a*** under previous conditions by using 10 mole percent (mol %) of ammonium iodide (*R*,*R*)-**5a** and hydrogen peroxide (2 equivalents) as an oxidant in methyl *tert*-butyl ether (MTBE) gave five-membered ring dihydrobenzofuran (*R*)-**4a** in 95% yield with 93% enantiomeric excess (ee). In sharp contrast, oxidative cyclization of **1a** to the six-membered ring under the same conditions was sluggish, and only a trace amount of the desired 2-acyl chroman **2a** was obtained together with a small amount (10% yield) of phenol-oxidation product **6a**.

The posited catalytic mechanism and side-reaction pathway are summarized in Fig. 2B. The above results and competition experiments (table S1) suggested that oxidative cyclization to a six-membered ring was much slower than cyclization to a five-membered ring. Thus, undesired side reactions, such as the dearomatization of **1a**, preferentially proceeded to give side-products **6a** and/or **6b**, presumably via phenoxenium ion **7** (*24*). The cyclization step might be rate-limiting for six-membered ring oxidative cyclization. Consequently, the in situ–generated catalytic active species (hypoiodite) was easily converted to an inert species such as triiodide salts, as confirmed with Raman analysis. For the construction of an efficient catalytic cycle, the oxidative cyclization step should not be rate-limiting. To address this issue, the oxidation of iodide (path **a**) should be decelerated, oxidative cyclization (path **b**) should be accelerated, or the inactivation (path **c**) should be suppressed or reversed. The use of TBHP as a weaker oxidant (*25*) instead of hydrogen peroxide solved this problem partially by deceleration of the generation of active species. Substrate **1a** was consumed within 1 hour; however, the desired product **2a** was obtained in only 15% yield with 18% ee (Fig. 2A). Undesired dearomatization dominated again, and quinone **6a** and peroxy quinol **6b** were obtained in a combined yield of 65%.







Entry	1b *or* 3b	5 (mol%)	Oxidant	Conditions	Product: Yield, ee (%, %)
1	**1b**	**5a** (10)	TBHP	MTBE, 3 h	(*R*)-**2b**: 98, 60
2	**1b**	**5b** (10)	TBHP	MTBE, 24 h	(*S*)-**2b**: 99, 86
3	**1b**	**5c** (10)	TBHP	MTBE, 24 h	(*S*)-**2b**: 94, 82
4	**1b**	**5d** (10)	TBHP	MTBE, 5 h	(*S*)-**2b**: 96, 89
5	**1b**	**5d** (10)	30% H_2O_2	MTBE, 1 h	(*S*)-**2b**: 95, 80
6	**1b**	**5d** (10)	CHP	Et_2O, 45 min.	(*S*)-**2b**: 99, 93
7	**1b**	**5d** (1)	CHP	Et_2O, 24 h	*no reaction*
8*	**1b**	**5d** (1)	CHP	K_2CO_3 (1 equiv), Et_2O, 10 h	(*S*)-**2b**: 98, 93
9	**1b**	**5d** (0.5)	CHP	K_2CO_3 (1 equiv), Et_2O, 48 h	(*S*)-**2b**: 98, 93
10†‡	**3b**	**5a** (0.05)	CHP	K_2CO_3 (1 equiv), Et_2O, 24 h	(*R*)-**4b**: 99, 88

* 1.5 g-scale. † 1 g-scale. ‡ No reaction occurred in the absence of K_2CO_3.

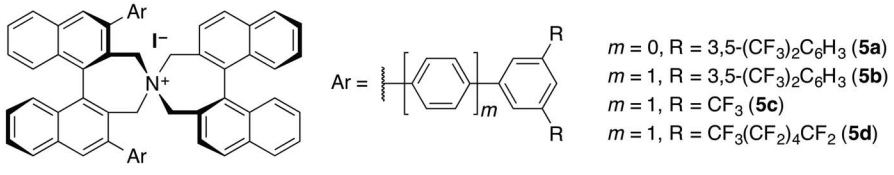


Fig. 2. Investigation of six-membered ring oxidative cyclization. (**A**) Comparison of five- and six-membered enantioselective oxidative cyclizations. Isolated yields of **2a**, **4a**, and **6** are reported. TBS, *tert*-butyldimethylsilyl. (**B**) Proposed catalytic mechanism and side-reaction pathway. "X" (in XOOH), H or alkyl. (**C**) Investigation of reaction parameters toward α-tocopherol. Isolated yields of **2b** and **4b** are reported. ee was determined by means of chiral stationary-phase high-performance liquid chromatography (HPLC). The absolute configuration of **2b** was determined by comparing the optical rotation of **9** (Fig. 3B) derived from **2b** with the literature value (*31*), and all other chromans **2** were assigned through analogy. The absolute configurations of **4a** and **4b** were determined by using their known analogs (*20*).

To prevent undesired dearomatization and accelerate the desired cyclization path, a *para*-toluenesulfonyl (Ts) group was introduced as an electron-withdrawing protective group in place of the silyl group of **1a**. As expected, the oxidative cyclization of **1b** with 10 mol % of (*R*,*R*)-**5a** gave desired (*R*)-**2b** quantitatively, with 60% ee (Fig. 2C, entry 1). The 3,3′-substituents of the binaphthyl moiety of (*R*,*R*)-**5** had a dramatic

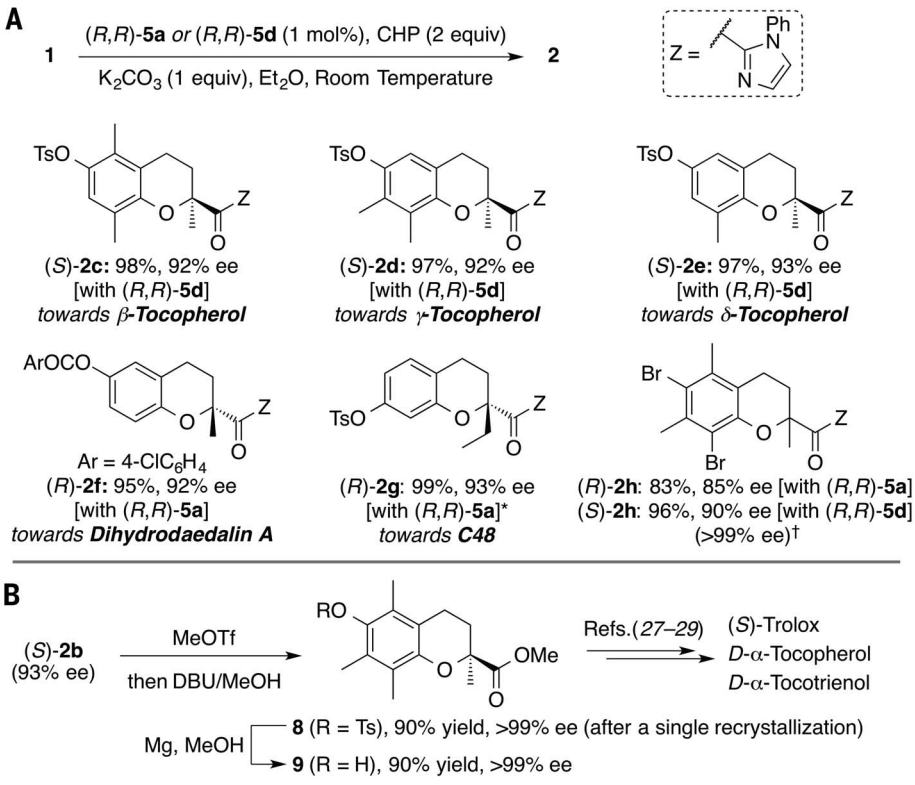


Fig. 3. Application to structurally diverse substrates. (**A**) Enantioselective oxidative cycloetherification of various γ-(2-hydroxyphenyl)ketones **1** to chromans **2**. Unless otherwise noted, 1 mol % of (*R*,*R*)-**5a** or (*R*,*R*)-**5d** was used in the presence of potassium carbonate. Asterisk indicates that the reaction was performed by using 10 mol % (*R*,*R*)-**5a** in the absence of K_2CO_3 with TBHP instead of CHP at 0°C. Dagger symbol indicates that (*R*)-**2h** and (*S*)-**2h** were obtained in optically pure forms after a single re-crystallization. (**B**) Asymmetric formal syntheses of (*S*)-trolox, *D*-α-tocopherol, and *D*-α-tocotrienol. MeOTf, methyl trifluoromethanesulfonate; DBU, 1,8-diazobicyclo[5.4.0]undec-7-ene.

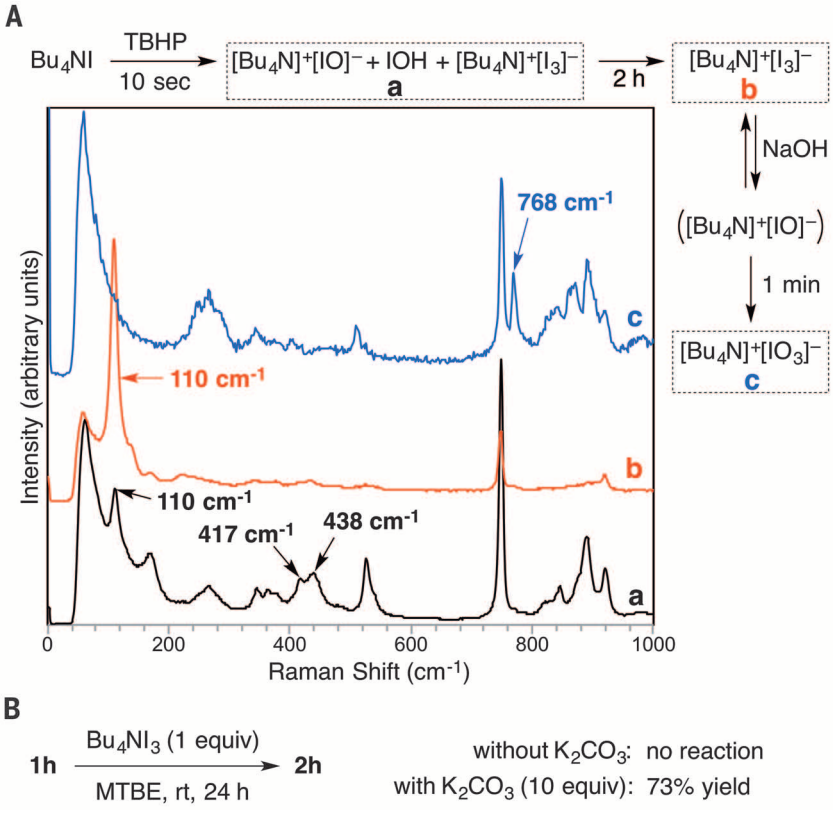


Fig. 4. Probing the active catalyst. (**A**) Detection of catalytic species. A series of Raman spectra (**a** to **c**) obtained upon the addition of sodium hydroxide to a solution of *n*-tetrabutylammonium iodide and TBHP are shown. (**B**) Control experiments with *n*-tetrabutylammonium triiodide in the presence or absence of potassium carbonate.

effect on enantioselectivity and reactivity (Fig. 2C, entries 1 to 4). To our surprise, the use of (*R*,*R*)-**5b** in place of (*R*,*R*)-**5a** gave (*S*)-**2b** as an opposite enantiomer, with higher enantioselectivity (86% ee) (Fig. 2C, entry 2). The opposite absolute stereoselectivity was observed with the use of not only (*R*,*R*)-**5b** but also (*R*,*R*)-**5c** and (*R*,*R*)-**5d**, which have biphenyl groups at the 3,3′ positions (Fig. 2C, entries 2 to 4). The best result (96% yield, 89% ee) was obtained after a shorter reaction time with perfluoroalkyl-substituted ammonium iodide (*R*,*R*)-**5d** (Fig. 2C, entry 4). Because the oxidative cyclization step might not be rate-limiting for **1b**, which was more reactive than **1a**, hydrogen peroxide could be used as an oxidant, albeit with a slightly reduced enantioselectivity (Fig. 2C, entry 5). When cumene hydroperoxide (CHP) was used as an oxidant in diethyl ether, the reaction was complete in 45 min, and (*S*)-**2b** was obtained quantitatively, with 93% ee (Fig. 2C, entry 6).

A reduction in the catalyst loading might cause competition between inactivation (path **c**) and oxidative cyclization (path **b**) in this catalytic system (Fig. 2B). When 1 mol % of **5d** was used, no reaction occurred, and the starting material was recovered fully (Fig. 2C, entry 7). To overcome this problem, suppression or reversible control of the inactivation path was considered. It is known that hypoiodite salts can be prepared with the hydrolysis of triiodide salts in alkaline solutions, and these species are in equilibrium under basic conditions (*26*). We envisioned that the hypoiodite species might be regenerated from triiodide species in the presence of appropriate base additives under our catalytic conditions. After the investigation of various organic and inorganic base additives (table S4), we succeeded in developing a high-performance catalytic oxidation system in the presence of an inorganic base such as potassium carbonate. Thus, the catalyst loading could be reduced to 1 or even 0.5 mol % [turnover number (TON) of the catalyst = 200] for the oxidation of **1b** in the presence of one equivalent of potassium carbonate without reducing the chemical yield or enantioselectivity (Fig. 2C, entries 8 and 9). Furthermore, the TON of the catalyst was 2000 for the five-membered oxidative cyclization of **3b** (Fig. 2C, entry 10). These reaction conditions were compatible with gram-scale synthesis (Fig. 2C, entries 8 and 10).

We examined several γ-(2-hydroxyphenyl) ketones **1** under optimized conditions (Fig. 3A). (*S*)-2-Acylchromans **2c**–**e**, which would be a synthetic intermediate for β-, γ-, and δ-tocopherols, were obtained quantitatively with high enantioselectivities by using 1 mol % of (*R*,*R*)-**5d**. The reactions of γ-[5-(4-chlorobenzoyloxy)-2-hydroxyphenyl]ketone **1f** and γ-(4-tosyloxy-2-hydroxyphenyl)ketone **1g** by using (*R*,*R*)-**5a** gave (*R*)-**2f** and (*R*)-**2g**, respectively. Compounds (*R*)-**2f** and (*R*)-**2g** would potentially offer a different route to dihydrodaedalin A (*11*, *12*) and Merck's compound C48 (*13*), respectively. The oxidative cyclizations of **1h** by using 1 mol % of (*R*,*R*)-**5a** and (*R*,*R*)-**5d** under the same conditions provided both enantiomers of the chroman **2h** with high enantioselectivities. The optically pure enantiomers **2h** could be obtained after a single recrystallization.

The formal syntheses of *D*-α-tocopherol, *D*-α-tocotrienol, and (*S*)-trolox were also achieved (Fig. 3B). The (*N*-phenylimidazol-2-yl)carbonyl group of product **2b** was easily transformed (*23*) to the methyl ester (**8**), which could be obtained in optically pure form after a single recrystallization. Subsequent deprotection of the tosyl group of **8** under mild conditions gave **9** in high chemical yield. The ester **9** is a common synthetic intermediate for *D*-α-tocopherol (*27*), *D*-α-tocotrienol (*28*), and (*S*)-trolox (*29*). Other tocopherols and their biologically active analogs could be easily prepared in a similar manner.

To gain insight into the catalytic mechanism, we performed various control experiments and spectroscopic analysis (Fig. 4, tables S5 and S6, and figs. S1 to S8). A series of Raman spectra obtained upon the addition of sodium hydroxide to a solution of *n*-tetrabutylammonium iodide (Bu_4NI) and TBHP is shown in Fig. 4A. To our delight, we detected unstable hypoiodite $[IO]^-$ and hypoiodous acid [IOH] species (*30*). Spectrum **a**, which includes three main bands at 110, 417, and 438 cm^{-1}, was recorded immediately after the mixing of Bu_4NI with TBHP. The other bands were attributed to the solvents and reagents used. The band at 110 cm^{-1} is characteristic of $[I_3]^-$. The bands at 417 and 438 cm^{-1} were assigned to $[IO]^-$ and [IOH] species, respectively, on the basis of the literature (*30*) and our control experiments (figs. S4 to S7). These two species might be in equilibrium under these conditions and disappeared steadily with time, and only a band of triiodide was observed after 2 hours (spectrum **b**). No other inert species such as iodate and periodate were observed at this time. The band of triiodide decreased immediately under basic conditions (spectrum **c**). A new band at 768 cm^{-1}, which is characteristic of the iodate $[IO_3]^-$ spectrum, was observed (*30*). This indicated the rapid generation and subsequent disproportionation of hypoiodite species (*30*). These results revealed that hypoiodite is an unstable catalytic active species and that triiodide is a stable inert species under our conditions (Fig. 2B). Although iodite species $[IO_2]^-$ could not be detected, we could not completely rule out a catalytic role of $[IO_2]^-$. The regeneration of hypoiodite from triiodide was also confirmed with control experiments (Fig. 4B and table S6). Oxidative cyclization of **1h** did not occur with the use of *n*-tetrabutylammonium triiodide but proceeded in the presence of potassium carbonate under identical conditions.

REFERENCES AND NOTES

1. H. C. Shen, *Tetrahedron* **65**, 3931–3952 (2009) and references therein.
2. T. Netscher, *Vitam. Horm.* **76**, 155–202 (2007) and references therein.
3. J. E. Packer, T. F. Slater, R. L. Willson, *Nature* **278**, 737–738 (1979).
4. The Alpha-Tocopherol Beta Carotene Cancer Prevention Study Group, *N. Engl. J. Med.* **330**, 1029–1035 (1994).
5. D. H. Blatt, W. A. Pryor, J. E. Mata, R. Rodriguez-Proteau, *J. Nutr. Biochem.* **15**, 380–395 (2004).
6. M. Tomasetti, J. Neuzil, *Vitam. Horm.* **76**, 463–491 (2007).
7. A. C. Howard, A. K. McNeil, P. L. McNeil, *Nat. Commun.* **2**, 597 (2011).
8. P.-H. Huang *et al.*, *Sci. Signal.* **6**, ra19 (2013).
9. W. Walther, W. Vetter, T. Netscher, *J. Micro. Sep.* **4**, 45–49 (1992).
10. G. V. López *et al.*, *Bioorg. Med. Chem.* **13**, 5787–5796 (2005).
11. M. Sekimoto, Y. Hattori, K. Morimura, M. Hirota, H. Makabe, *Bioorg. Med. Chem. Lett.* **20**, 1063–1064 (2010).
12. Y. Morimura, M. Hirota, Y. Hattori, H. Makabe, "Chroman compound and method for preparing the same," Japan patent 5212969 (2013).
13. H. Koyama *et al.*, *J. Med. Chem.* **47**, 3255–3263 (2004).
14. B. M. Trost, H. C. Shen, L. Dong, J.-P. Surivet, C. Sylvain, *J. Am. Chem. Soc.* **126**, 11966–11983 (2004).
15. L. F. Tietze, K. M. Sommer, J. Zinngrebe, F. Stecker, *Angew. Chem. Int. Ed. Engl.* **44**, 257–259 (2004).
16. S. Tanaka, T. Seki, M. Kitamura, *Angew. Chem. Int. Ed. Engl.* **48**, 8948–8951 (2009).
17. A. Chougnet, K. Liu, W.-D. Woggon, *Chimia (Aarau)* **64**, 303–308 (2010) and references therein.
18. S. Tokunou, W. Nakanishi, N. Kagawa, T. Kumamoto, T. Ishikawa, *Heterocycles* **84**, 1045–1056 (2012).
19. B. List, K. Maruoka, Eds., *Science of Synthesis: Asymmetric Organocatalysis* (George Thieme Verlag, Stuttgart, 2012).
20. M. Uyanik, H. Okamoto, T. Yasui, K. Ishihara, *Science* **328**, 1376–1379 (2010).
21. M. Uyanik, K. Ishihara, *ChemCatChem* **4**, 177–185 (2012).
22. T. Ooi, K. Maruoka, *Angew. Chem. Int. Ed. Engl.* **46**, 4222–4266 (2007).
23. D. A. Evans, K. R. Fandrick, H.-J. Song, K. A. Scheidt, R. Xu, *J. Am. Chem. Soc.* **129**, 10029–10041 (2007).
24. D. Magdziak, S. J. Meek, T. R. R. Pettus, *Chem. Rev.* **104**, 1383–1430 (2004).
25. R. D. Cadle, H. Huff, *J. Phys. Colloid Chem.* **54**, 1191–1195 (1950).
26. Y.-T. Chia, thesis, University of California, Berkeley (1958).
27. M. Nozawa, K. Takahashi, K. Kato, H. Akita, *Chem. Pharm. Bull. (Tokyo)* **48**, 272–277 (2000).
28. R. Chênevert, G. Courchesne, N. Pelchat, *Bioorg. Med. Chem.* **14**, 5389–5396 (2006).
29. A. Ida, Y. Kyuko, T. Hidaka, "Method of producing (*S*)-(−)-6-hydroxy-2,3,7,8-tetramethylchromane-2-carboxylic acid and product obtained by the method," U.S. Patent 8,080,676 (2011).
30. J. C. Wren, J. Paquette, S. Sunder, B. L. Ford, *Can. J. Chem.* **64**, 2284–2296 (1986).
31. H. Yoda, K. Takabe, *Chem. Lett.* **18**, 465–466 (1989).

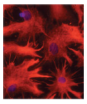
ACKNOWLEDGMENTS

Financial support for this project was partially provided by the Japan Society for the Promotion of Science, Grant-in-Aid for Scientific Research (KAKENHI) (24245020, 22750087, and 25105722), and Program for Leading Graduate Schools "Integrative Graduate Education and Research in Green Natural Sciences," Ministry of Education, Culture, Sports, Science and Technology, Japan. Metrical parameters for the structure of compound **2h** are available free of charge from the Cambridge Crystallographic Data Centre under reference number CCDC 996935. M.U. and K.I. developed the concept and conceived the experiments. H.H. performed the experiments. H.H. and M.U. analyzed the data. M.U. and K.I. prepared the manuscript, with assistance from H.H.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/345/6194/291/suppl/DC1
Materials and Methods
Tables S1 to S7
Figs. S1 to S9
HPLC Traces
NMR Chart
References (*32*–*53*)

18 April 2014; accepted 10 June 2014
10.1126/science.1254976

EXTENDED PDF FORMAT
SPONSORED BY

Travel Grants Available
Neuroscience 2014
Application Deadline: July 31st
>Apply Now
R&D systems
a bio-techne brand
www.rndsystems.com

A hole-conductor–free, fully printable mesoscopic perovskite solar cell with high stability

Anyi Mei *et al.*

Science **345**, 295 (2014);
DOI: 10.1126/science.1254763

This copy is for your personal, non-commercial use only.

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by clicking here.

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines here.

The following resources related to this article are available online at www.sciencemag.org (this information is current as of August 4, 2014):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:
http://www.sciencemag.org/content/345/6194/295.full.html

Supporting Online Material can be found at:
http://www.sciencemag.org/content/suppl/2014/07/16/345.6194.295.DC1.html

This article **cites 22 articles**, 4 of which can be accessed free:
http://www.sciencemag.org/content/345/6194/295.full.html#ref-list-1

This article appears in the following **subject collections:**
Physics, Applied
http://www.sciencemag.org/cgi/collection/app_physics

Downloaded from www.sciencemag.org on August 4, 2014

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published weekly, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. Copyright 2014 by the American Association for the Advancement of Science; all rights reserved. The title *Science* is a registered trademark of AAAS.

PHOTOVOLTAICS

A hole-conductor–free, fully printable mesoscopic perovskite solar cell with high stability

Anyi Mei,[1]* Xiong Li,[1]* Linfeng Liu,[1] Zhiliang Ku,[1] Tongfa Liu,[1] Yaoguang Rong,[1] Mi Xu,[1] Min Hu,[1] Jiangzhao Chen,[1] Ying Yang,[1] Michael Grätzel,[2] Hongwei Han[1]†

We fabricated a perovskite solar cell that uses a double layer of mesoporous TiO_2 and ZrO_2 as a scaffold infiltrated with perovskite and does not require a hole-conducting layer. The perovskite was produced by drop-casting a solution of PbI_2, methylammonium (MA) iodide, and 5-ammoniumvaleric acid (5-AVA) iodide through a porous carbon film. The 5-AVA templating created mixed-cation perovskite $(5\text{-AVA})_x(MA)_{1-x}PbI_3$ crystals with lower defect concentration and better pore filling as well as more complete contact with the TiO_2 scaffold, resulting in a longer exciton lifetime and a higher quantum yield for photoinduced charge separation as compared to $MAPbI_3$. The cell achieved a certified power conversion efficiency of 12.8% and was stable for >1000 hours in ambient air under full sunlight.

Organic-inorganic metal halide perovskite mesoscopic solar cells (*1*–*3*) have rapidly reached conversion efficiencies over 15% (*4*, *5*), and most use a methylammonium lead halide ($MAPbX_3$, X = halogen) and its mixed-halide crystal analog (*6*–*14*). These perovskites have a large absorption coefficient, high carrier mobility, direct band gap, and high stability. They are composed of Earth-abundant materials and can be deposited by low-temperature solution methods. However, perovskite crystallization from solution produces large morphological variations (*14*) and incomplete filling of the mesoporous oxide scaffold, which results in an unwanted spread of the photovoltaic performance of the resulting devices. Most devices use gold as a back contact, in conjunction with organic hole conductors acting as electron-blocking layers. Hole conductors such as the widely used arylamine *spiro*-OMeTAD are not only expensive but can limit the long-term stability of the device. Hole-conductor–free perovskite photovoltaics were initially reported by Etgar *et al.* (*15*), and to date power conversion efficiencies (PCEs) of 10.85% (*16*–*18*) have been achieved, with gold as a counter-electrode.

We fabricated a solar cell using a double layer of mesoporous TiO_2 and ZrO_2 covered by a porous carbon film. The metal halide perovskite was infiltrated into the porous TiO_2/ZrO_2 scaffold by drop-casting a solution through the printed carbon layer that contained PbI_2 in γ-butyrolactone, together with methylammonium (MA) and 5-aminovaleric acid (5-AVA) cations. 5-AVA replaced part of the MA cations in the cuboctahedral site of $MAPbI_3$, forming the new mixed-cation perovskite $(5\text{-AVA})_x(MA)_{1-x}PbI_3$. The orthorhombic perovskite phase formed in the presence of 5-AVA showed greatly increased PCE compared with the neat analog in a hole-conductor–free cell that used a simple mesoscopic $TiO_2/ZrO_2/C$ triple layer as a scaffold to host the perovskite absorber. The role of 5-AVA cations was to template the formation of perovskite crystals in the mesoporous oxide host and induce preferential growth in the normal direction. The resulting mixed-cation perovskite $(5\text{-AVA})_x(MA)_{1-x}PbI_3$ has better surface contact with the TiO_2 surface and lower defect concentration than the single-cation form $MAPbI_3$, resulting in improved conversion efficiency. The triple-layer device also exhibits excellent long-term stability.

A schematic cross section of the triple-layer, perovskite-based, fully printable mesoscopic solar cell (Fig. 1A) shows that the mesoporous layers of TiO_2 and ZrO_2 have thicknesses of ~1 and 2 μm, respectively, that were deposited on a F-doped SnO_2 (FTO)–covered glass sheet. The mesoporous layers were infiltrated with perovskite by drop-casting from solution through a 10-μm-thick carbon layer printed on top of the ZrO_2. The energy band alignment of the device (Fig. 1B) prevents electrons, injected from the photoexcited perovskite into the TiO_2 scaffold, from reaching the back contact because of the 0.6-eV offset between its conduction band and that of ZrO_2. The compact TiO_2 layer deposited on the FTO conducting glass prevents the valence band holes from reaching the FTO-covered front electrode. The $CH_3NH_3PbI_3$ perovskite unit cell (Fig. 1C) contains one Pb^{2+}, one organic ammonium cation, and three I^- anions (*19*).

We introduced the 5-AVA cation into the perovskite lattice with a mixture of MA and 5-AVA in the precursor solution that maintained a 1:1 molar ratio of organic ammonium cations and PbI_2. The optimal molar ratio of ammonium cations in $(5\text{-AVA})_x(MA)_{1-x}PbI_3$ was determined to be between 1:20 and 1:30. The x-ray diffraction (XRD) patterns of neat $MAPbI_3$ and $(5\text{-AVA})_x(MA)_{1-x}PbI_3$ infiltrated into the mesoporous ZrO_2/TiO_2 film are compared in Fig. 2. The newly emerging diffraction peaks arising from the (001) and (111) lattice planes for $(5\text{-AVA})_x(MA)_{1-x}PbI_3$ are much stronger than the corresponding ones for $MAPbI_3$. And even at low 5-AVA/MA molar ratios, 5-AVA substantially increases the *b* and *c* lattice parameters. The large expansion of the *c* axis induced by 5-AVA indicates its preferential alignment along this axis through contact with lead and

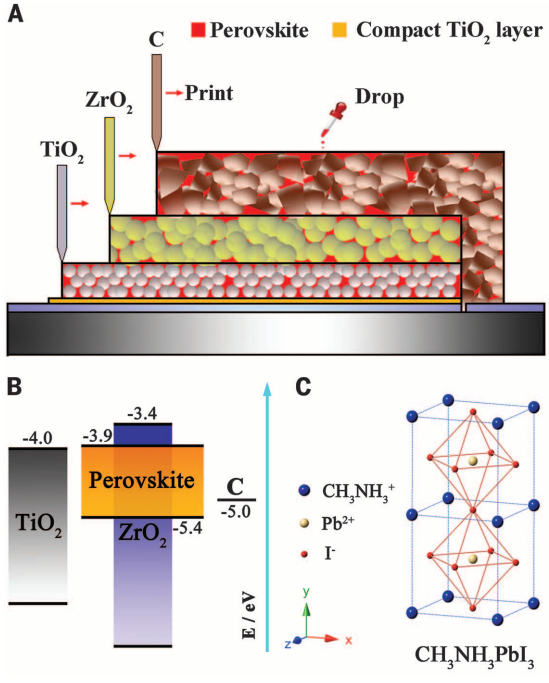


Fig. 1. The triple-layer perovskite junction. (**A**) Schematic drawing showing the cross section of the triple-layer perovskite–based fully printable mesoscopic solar cell. The mesoporous layers of TiO_2 and ZrO_2 have a thickness of ~1 and 2 μm, respectively, and are deposited on a FTO-covered glass sheet shown in blue and gray. They are infiltrated with perovskite by drop-casting from solution. (**B**) Energy band diagram of the triple-layer device. Energies are expressed in electron volts, using the electron energy in vacuum as a reference. The energy levels of the conduction band edges of TiO_2, ZrO_2, and $MAPbI_3$ are at −4.0, −3.4, and −3.9 eV, respectively, whereas the valence band edge of the perovskite is at −5.4 eV and that of the Fermi level of carbon is at −5.0 eV. (**C**) The crystal structure of $MAPbI_3$ perovskite.

[1]Michael Grätzel Center for Mesoscopic Solar Cells, Wuhan National Laboratory for Optoelectronics, School of Optical and Electronic Information, Huazhong University of Science and Technology, Wuhan 430074, Hubei, People's Republic of China. [2]Laboratory for Photonics and Interfaces, Institute of Chemical Sciences and Engineering, School of Basic Sciences, Ecole Polytechnique Federale de Lausanne, CH-1015 Lausanne, Switzerland.
*These authors contributed equally to this work. †Corresponding author. E-mail: hongwei.han@mail.hust.edu.cn

iodide ions, and the *c* axis becomes the dominant orientation during crystal growth. Mercier (*20*) used 4-ammonium-butyric acid (4-ABA) as a template for the engineering of $(4\text{-ABA})_2PbI_4$ and mixed-cation $(4\text{-ABA})_2MAPb_2I_7$ perovskite crystals. The ABA molecules formed linear hydrogen-bonded chains that act as supramolecular synthons for layered plumboiodide perovskite structures. We propose that the 5-AVA affects the crystal growth of $(5\text{-AVA})_x(MA)_{1-x}PbI_3$ in a similar fashion through the formation of hydrogen bonds between its COOH and NH_3^+ groups and I^- ions from the PbI_6 octahedra.

We applied transmission electron microscopy (TEM) and scanning electron microscopy (SEM) to unravel differences in the morphology of the two perovskites. High-resolution TEM images (Fig. 3, A to C) reflecting the morphology of neat TiO_2 particles, as well as $MAPbI_3$ or $(5\text{-AVA})_x(MA)_{1-x}PbI_3$–covered TiO_2 particles that were scraped off of mesoporous TiO_2 films infiltrated with $MAPbI_3$ or $(5\text{-AVA})_x(MA)_{1-x}PbI_3$, respectively. Substantial differences appeared between the morphologies of the samples produced from the two different perovskite precursor solutions. The observed pattern indicates that there is a much denser coverage of the TiO_2 particles with $(5\text{-AVA})_x(MA)_{1-x}PbI_3$ as compared to $MAPbI_3$. The $(5\text{-AVA})_x(MA)_{1-x}PbI_3$ appears to exhibit crystalline features on a much longer length scale than $MAPbI_3$, covering most of the TiO_2 surface in a uniform fashion. In contrast, the single-step solution precipitation of $MAPbI_3$ only partially coated the substrate and left large areas where the perovskite absorber was completely absent. The improvement of the perovskite crystal quality and the higher loading of the mesopores of the oxide scaffold in the presence of 5-AVA cations probably resulted from the templating action of the AVA affecting the perovskite crystal nucleation and growth within the mesoscopic oxide scaffold. The COOH groups of 5-AVA anchor a monolayer of the amino acid to the surface of the mesoporous TiO_2 and ZrO_2 film by coordinative binding to the exposed Ti(IV) or Zr(IV) ions. In the adsorbed state, the terminal $-NH_3^+$ groups of 5-AVA face the perovskite solution and hence serve as nucleation sites. This role we attribute to AVA is confirmed by recent work (*21*) showing that protonated amino acids of the AVA type indeed template the crystal growth of plumbohalide perovskites in mesoscopic TiO_2 films, improving the crystalline network and charge-carrier lifetime of the $CH_3NH_3PbI_3$ inserted into the porous metal oxides.

We performed photoluminescence (PL) decay measurements in order to extract quantitative information on the yield of light-induced charge separation. Excitons generated by light excitation of $MAPbI_3$ dissociate into free charge carriers within 1 ps (*22*). We infiltrated mesoporous TiO_2 and ZrO_2 films with $MAPbI_3$ or $(5\text{-AVA})_x(MA)_{1-x}PbI_3$ and measured the PL decay; ZrO_2 serves as a reference because its conduction band is not accessible for electron injection. The PL decay of the $MAPbI_3$ perovskite contained in ZrO_2 films exhibits a time constant of τ_e = 8.14 ns,

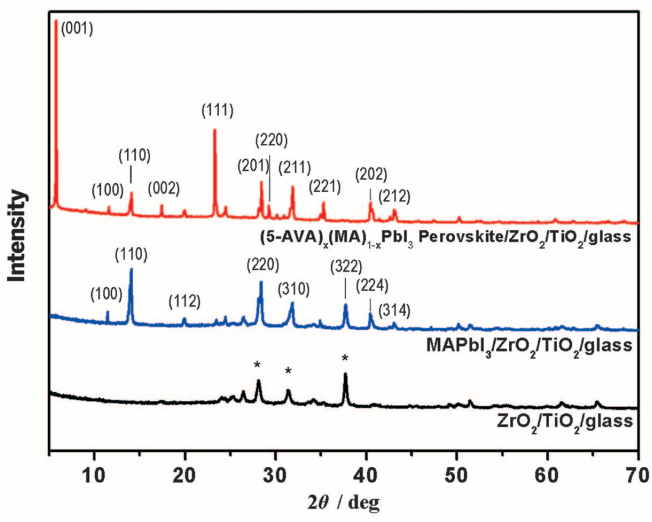


Fig. 2. Diffraction data. XRD patterns of mesoscopic ZrO_2/TiO_2 film on FTO glass infiltrated with the perovskites $(5\text{-AVA})_x(MA)_{1-x}PbI_3$ (red trace) and $MAPbI_3$ (blue trace), as well as a blank ZrO_2/TiO_2 film (black trace).

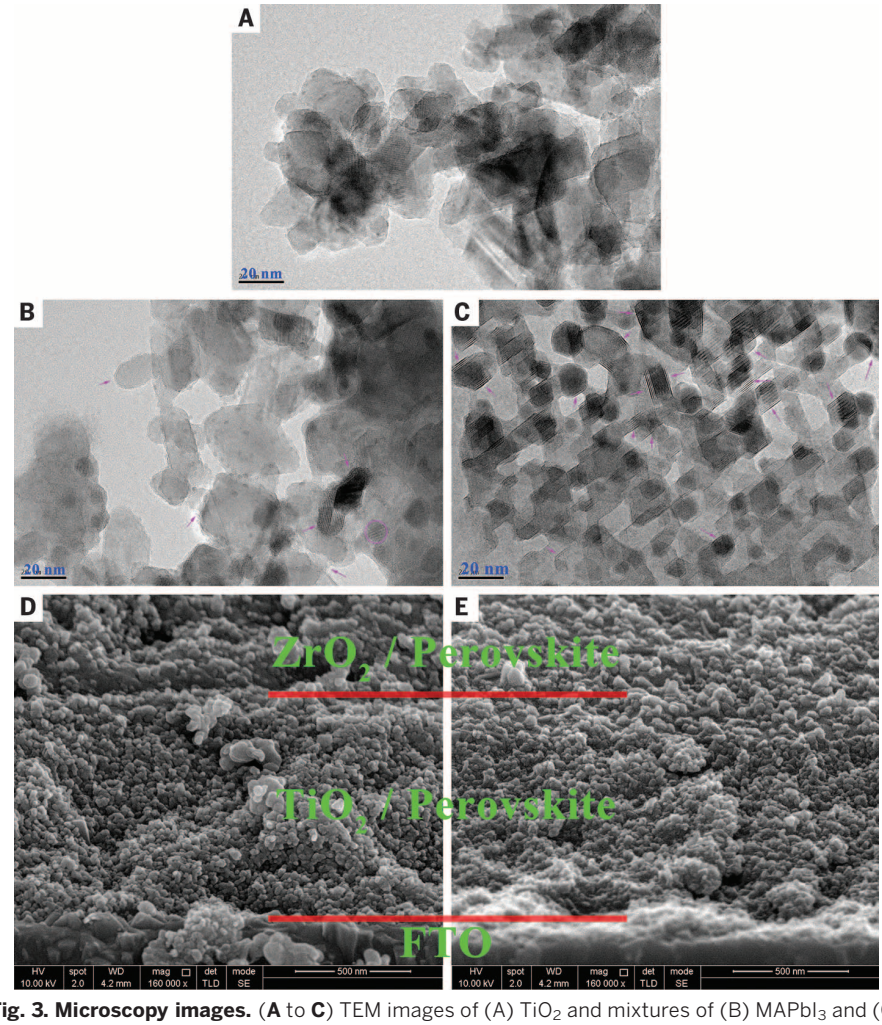


Fig. 3. Microscopy images. (**A** to **C**) TEM images of (A) TiO_2 and mixtures of (B) $MAPbI_3$ and (C) $(5\text{-AVA})_x(MA)_{1-x}PbI_3$, with TiO_2 particles scraped off of a mesoporous film infiltrated with perovskites. (**D** and **E**) SEM images of the cross section of (D) $MAPbI_3$– and (E) $(5\text{-AVA})_x(MA)_{1-x}PbI_3$–based perovskite mesoscopic solar cells.

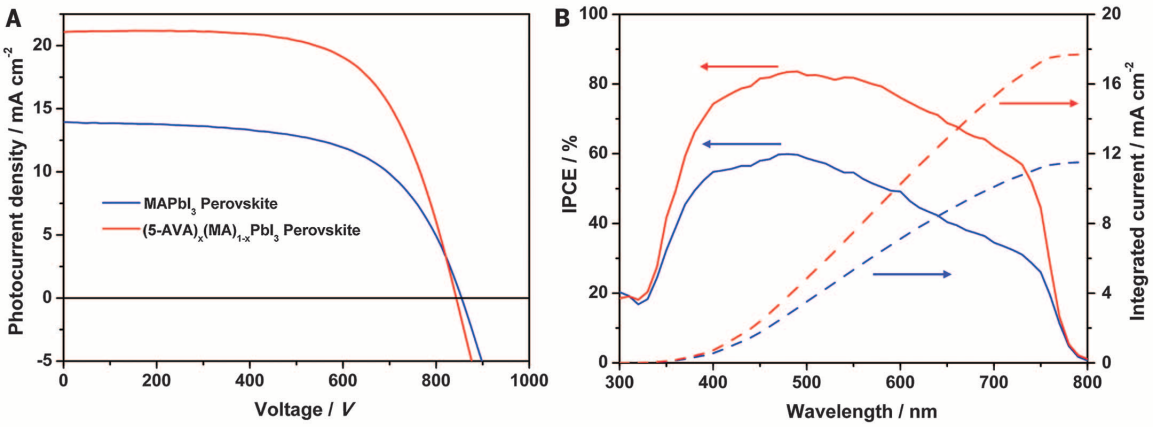


Fig. 4. Device performance. (**A**) *J-V* curves under simulated AM 1.5 solar irradiation at an intensity of 100 mW cm^{-2} measured at room temperature. (**B**) IPCE curves taken with monochromatic light without applied white-light bias for $(5\text{-AVA})_x(MA)_{1-x}PbI_3$ (red curve) or $MAPbI_3$ (blue curve). The integrated photocurrents calculated from the overlap integral of the IPCE spectra with the AM 1.5 solar emission are also shown. The integrated photocurrent of the $(5\text{-AVA})_x(MA)_{(1-x)}PbI_3$–based photovoltaic is 17.8 mA cm^{-2}, which agrees closely with the photocurrent density of 18.4 mA cm^{-2} measured at the beginning of testing, which rose to 21.1 mA cm^{-2} after 3 min of light soaking.

whereas for the $(5\text{-AVA})_x(MA)_{1-x}PbI_3$, τ_e is 23.7 ns [fig. S1 (*23*)]. The longer charge-carrier lifetime observed with $(5\text{-AVA})_x(MA)_{1-x}PbI_3$ indicates a much lower defect concentration. With the TiO_2 scaffold, τ_e for $MAPbI_3$ and $(5\text{-AVA})_x(MA)_{1-x}PbI_3$ were similar, 1.71 and 1.36 ns, respectively; and from the branching ratios, led to calculated quantum efficiencies of 94 and 80%, respectively.

Cross-sectional SEM views of the $(5\text{-AVA})_x(MA)_{1-x}PbI_3$– and $MAPbI_3$–based mesoscopic solar cells, which confirm the different extent of pore filling by the two perovskites are shown in Fig. 3, D and E. Thus, for $(5\text{-AVA})_x(MA)_{1-x}PbI_3$, most of the mesopores of the ZrO_2/TiO_2 double layer are fully loaded with the perovskite, whereas for the device with $MAPbI_3$, the pore filling is less complete, with a substantial amount of voids remaining within the mesoporous ZrO_2/TiO_2 double-layer film. Optical studies also confirmed the higher perovskite loading in the presence of 5-AVA [figs. S2 and S3 (*23*)].

We measured the photocurrent density as a function of forward bias voltage (*J-V* curves) of the $(5\text{-AVA})_x(MA)_{1-x}PbI_3$– and the $MAPbI_3$– based mesoscopic solar cells under standard reporting conditions [air mass 1.5 (AM1.5) global solar light at 100 mW cm^{-2} and room temperature]. Figure S4 (*23*) presents the temporal evolution of the device performance for $(5\text{-AVA})_x(MA)_{1-x}PbI_3$ –based mesoscopic solar cells during the initial phase of illumination. The four key photovoltaic parameters; i.e., the open-circuit voltage (V_{oc}), short-circuit photocurrent (J_{sc}), fill factor (FF), and PCE, increased to a stable value within the first 3 min and then exhibited excellent stability during exposure to full AM 1.5 simulated sunlight over 1008 hours. The final PCE increased slightly during this period, which is more remarkable given that the test was performed with an unsealed device, the perovskite being protected by the 10-μm-thick carbon layer back contact acting as a water-retaining layer.

A hysteresis effect often appears during measurement of the *J-V* curves of perovskite solar cells. For a mesoscopic triple-layer solar cell with perovskite $(5\text{-AVA})_x(MA)_{1-x}PbI_3$ under standard 1 sun irradiation with the scanning rate of 3 mV s^{-1}, there was no substantial effect [fig. S5 (*23*)]. However, a hysteresis effect appeared at higher scan rates unless the cell was subjected to light soaking for a few minutes, which eliminated the hysteresis effects even at scan rates as high as 250 mV s^{-1}. The perovskite systems show mixed ionic and electronic conduction (*24*), and under illumination an electric field is set up, which causes a slow drift of ions through the perovskite device to screen the space charge that is generated by light. Figure 4A presents *J-V* curves collected after 3 min of light soaking with simulated sunlight. In keeping with our previous work (*19*), the optimized $MAPbI_3$-based heterojunction perovskite solar cell produced a J_{sc} of 13.9 mA cm^{-2}, a V_{oc} of 855 mV, and a FF of 0.61, yielding a PCE of 7.2%. The $(5\text{-AVA})_x(MA)_{1-x}PbI_3$–based cells showed greatly improved performance, with J_{sc}, V_{oc}, and FF reaching values of 21.1 mA cm^{-2}, 843 mV, and 0.65, respectively, corresponding to a PCE of 11.6%. We sent one of our cells to an accredited photovoltaic calibration laboratory for certification, which measured a J_{sc} value of 22.8 mA cm^{-2} [fig. S6 (*23*)]. The V_{oc} of the device was 858 mV and the FF was 0.66, corresponding to a PCE of 12.84%. Our certified photocurrent density is superior to values reported so far for any perovskite-based photovoltaic, for which the highest certified J_{sc} is 21.3 mA cm^{-2} (*4*). Also in Fig. 4A, the slope of the *J-V* curve at open circuit voltage is steeper for $(5\text{-AVA})_x(MA)_{1-x}PbI_3$ perovskite than for $MAPbI_3$, suggesting a lower series resistance and higher electric conductivity for the former than for the latter perovskite. This shows the better connectivity of the AVA template perovskite crystals formed in the porous oxide network.

In the absence of ZrO_2, the performance of the perovskite photovoltaic is poor, with J_{sc} = 12.51 mA cm^{-2}, V_{oc} = 592 mV, FF = 0.56, and PCE = 4.18% [fig. S7 (*23*)]. This confirms that ZrO_2 is blocking the flow of photogenerated electrons to the back contact, preventing recombination with the holes from the perovskite at the back contact.

Figure S8 (*23*) reports statistical data concerning the numerical distribution of the key photovoltaic parameters; i.e., V_{oc}, J_{sc}, FF, and PCE for a random selection of 20 photovoltaic devices, which were fully printed on a 10 cm × 10 cm FTO glass as shown by the photograph in fig. S9 (*23*). $(5\text{-AVA})_x(MA)_{1-x}PbI_3$–based mesoscopic solar cells exhibited an average PCE value of 10.3%, J_{sc} of 21.68 mA cm^{-2}, V_{oc} of 740 mV, and FF of 0.64 together with a small standard deviation, indicating that the $(5\text{-AVA})_x(MA)_{1-x}PbI_3$ strongly augments the reproducibility of the perovskite-based cells. Figure 4B shows the incident photon-to-electric current conversion efficiency (IPCE), defined as the number of photogenerated charge carriers contributing to the photocurrent per incident photon. These experiments used monochromatic light without applying a white-light bias. $(5\text{-AVA})_x(MA)_{1-x}PbI_3$ achieved much higher IPCEs than $MAPbI_3$ over the whole spectral range between 300 and 800 nm, matching the difference in photocurrents obtained for the two devices.

REFERENCES AND NOTES

1. G. Hodes, *Science* **342**, 317–318 (2013).
2. G. Xing *et al.*, *Science* **342**, 344–347 (2013).
3. S. D. Stranks *et al.*, *Science* **342**, 341–344 (2013).
4. J. Burschka *et al.*, *Nature* **499**, 316–319 (2013).
5. M. Liu, M. B. Johnston, H. J. Snaith, *Nature* **501**, 395–398 (2013).
6. J. H. Noh, S. H. Im, J. H. Heo, T. N. Mandal, S. I. Seok, *Nano Lett.* **13**, 1764–1769 (2013).
7. H. S. Kim *et al.*, *Nano Lett.* **13**, 2412–2417 (2013).
8. H. S. Kim *et al.*, *Nat. Commun.* **4**, 2242 (2013).
9. H. Cai, J. B. Dunn, Z. Wang, J. Han, M. Q. Wang, *Biotechnol Biofuels* **6**, 141 (2013).
10. J. H. Heo *et al.*, *Nat. Photonics* **7**, 486–491 (2013).
11. A. Abrusci *et al.*, *Adv. Funct. Mater.* **21**, 2571–2579 (2011).
12. M. M. Lee, J. Teuscher, T. Miyasaka, T. N. Murakami, H. J. Snaith, *Science* **338**, 643–647 (2012).
13. H. S. Kim *et al.*, *Sci. Rep.* **2**, 591 (2012).
14. G. E. Eperon, V. M. Burlakov, P. Docampo, A. Goriely, H. J. Snaith, *Adv. Funct. Mater.* **24**, 151–157 (2014).
15. L. Etgar *et al.*, *J. Am. Chem. Soc.* **134**, 17396–17399 (2012).
16. W. A. Laban, L. Etgar, *Energy Environ. Sci.* **6**, 3249–3253 (2013).
17. S. Aharon, S. Gamliel, B. El Cohen, L. Etgar, *Phys. Chem. Chem. Phys.* **16**, 10512–10518 (2014).
18. J. Shi *et al.*, *Appl. Phys. Lett.* **104**, 063901 (2014).
19. Z. Ku, Y. Rong, M. Xu, T. Liu, H. Han, *Sci. Rep.* **3**, 3132 (2013).
20. N. Mercier, *CrystEngComm* **7**, 429 (2005).
21. Y. Ogomi *et al.*, *J. Phys. Chem. C* 10.1021/jp412627n (2014).

22. C. S. Ponseca Jr. *et al.*, *J. Am. Chem. Soc.* **136**, 5189–5192 (2014).
23. See the supplementary materials on *Science* Online.
24. C. C. Stoumpos, C. D. Malliakas, M. G. Kanatzidis, *Inorg. Chem.* **52**, 9019–9038 (2013).

ACKNOWLEDGMENTS

The authors acknowledge financial support from the Ministry of Science and Technology of China (863 project, grant no. SS2013AA50303), the National Natural Science Foundation of China (grant no. 61106056), the Science and Technology Department of Hubei Province (grant no. 2013BAA090), and the Fundamental Research Funds for the Central Universities (grant no. HUSTNY022). We thank the Analytical and Testing Center of Huazhong University of Science and Technology for field emission SEM, TEM, and XRD testing. M.G. thanks the European Research Council for financial support under advanced research grant ARG 247404, "Mesolight."

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/345/6194/295/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S9
References

14 April 2014; accepted 3 June 2014
10.1126/science.1254763

APPLIED OPTICS

Dielectric gradient metasurface optical elements

Dianmin Lin,[1]* Pengyu Fan,[1]* Erez Hasman,[2] Mark L. Brongersma[1]†

Gradient metasurfaces are two-dimensional optical elements capable of manipulating light by imparting local, space-variant phase changes on an incident electromagnetic wave. These surfaces have thus far been constructed from nanometallic optical antennas, and high diffraction efficiencies have been limited to operation in reflection mode. We describe the experimental realization and operation of dielectric gradient metasurface optical elements capable of also achieving high efficiencies in transmission mode in the visible spectrum. Ultrathin gratings, lenses, and axicons have been realized by patterning a 100-nanometer-thick Si layer into a dense arrangement of Si nanobeam antennas. The use of semiconductors can broaden the general applicability of gradient metasurfaces, as they offer facile integration with electronics and can be realized by mature semiconductor fabrication technologies.

The field of nanophotonics has started to facilitate the replacement of bulky optical components by ultrathin, planar elements (*1*–*5*). Such elements can be realized by nanopatterning thin films and can benefit from a more natural integration into electronic and mechanical systems. They can also provide entirely new functions that capitalize on the rapid developments in the field of metamaterials (*6*). The recent realization of metasurfaces, or metamaterials of reduced dimensionality (*7*), is of particular relevance as it is opening up new opportunities to realize virtually flat optics (*8*, *9*). Metasurfaces consist of dense arrangements of resonant optical antennas. The resonant nature of the light-matter interaction with such structures affords substantial control over the local light scattering amplitude and phase. These scattering properties can be manipulated through choices of the antenna material, size, geometry, orientation, and environment. By arranging antennas with distinct scattering properties on a surface, space-variant metasurfaces are created that offer tremendous freedom in manipulating optical wave-fronts. Metasurfaces constructed from high-index antenna arrays are used as antireflection coatings (*10*), as well as to enhance absorption (*11*), create structural color (*12*), and manipulate the state of polarization (*13*, *14*). However, space-variant metasurfaces for beam steering and focusing have thus far been constructed from nanoscale metallic antennas only (*7*). Ohmic losses in the metal and the limited scattering cross sections of the antennas have limited high device efficiencies to reflection mode (*15*, *16*). We leverage recent advances in the realization of high-index dielectric antennas (*12*, *17*, *18*) and Pancharatnam-Berry phase optical elements (PBOEs) (*19*, *20*) to realize easy-to-fabricate and highly efficient transmissive devices. Si-based gradient metasurfaces capable of serving as optical gratings, lenses, and axicons are demonstrated.

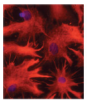
Figure 1 shows a dielectric gradient metasurface optical element (DGMOE) serving as an axicon. Whereas conventional glass axicons are a few millimeters in size (Fig. 1A), the DGMOE features a 100-nm-thick layer of poly-silicon deposited on a quartz substrate (Fig. 1B) generated in a single patterning step (*21*). Figure 1D shows a scanning electron microscopy (SEM) image of the fabricated DGMOE composed of thousands of Si nanobeam antennas.

When illuminated with a collimated Gaussian left circularly polarized (LCP) light beam at a 550-nm wavelength, the DGMOE creates a nondiffracting Bessel beam on the transmission (Fig. 1, C and E, and section S2 of the supplementary materials). The intensity distributions along and across the center of Bessel beam are shown as insets to Fig. 1E. The ratio between the high-intensity central beam spot and the first ring is 6:1, a signature of a Bessel beam. The operation of the lens could be analyzed by understanding all of the coherent far-field and near-field interactions between the Si nanobeam antennas. Here, we describe the behavior of the optical element based on an understanding of the Pancharatnam-Berry (PB) phase.

The PB phase is a geometric phase achieved by space-variant polarization manipulations (*22*, *23*), as opposed to a propagation phase. PBOEs are constructed from wave plate elements for which the orientation of the fast axes depends on the spatial position. Well-defined algorithms exist for realizing specific optical functions. By tiling a surface with half-wave plates with their fast-axes orientations according to a function $\theta(x,y)$, an incident circularly polarized light beam will be fully transformed to a beam of opposite helicity and imprinted with a geometric phase equal to $\varphi_g(x,y) = \pm 2\theta(x,y)$ (*21*, *24*). By controlling the local orientation of the fast axes of the wave plate elements between 0 and π, phase pickups can be achieved that cover the full 0-to-2π range while maintaining equal transmission amplitude for the entire optical component. This provides full control over the wavefront. A continuous desired phase function can be approximated using discrete wave plate orientations. We use eight orientations with which a high theoretical diffraction efficiency of 95% can be achieved (*25*). The realization of wave plate elements represents a crucial step in crafting PBOEs.

In designing ultrathin nanobeam-based wave plates, it is important to first understand the optical resonances of individual semiconductor nanobeams. Top-illuminated single beams support resonances under transverse electric (TE) illumination (with the electric field polarized normal to the length of the structure) and the orthogonal transverse magnetic (TM) polarization (*12*, *18*). The thickness of the resonant structures can be small compared with the free-space wavelength due to the high refractive index of semiconductors. A 120-nm-wide and 100-nm-high Si nanobeam supports a strong resonance and a substantial phase retardation in our target wavelength range of interest. We chose this beam as the basic building block for our wave plate (Fig. 2A) in anticipation that its resonant behavior and the accompanied phase retardation between the orthogonal polarizations would persist in the beam array. Finite element simulations indicate that the wavefront for a TE-polarized light beam at 550 nm is delayed by 0.14π as compared with the wavefront in the absence of the nanobeams. The wavefront for TM-polarized light is delayed even further to 1.15π (Fig. 2B). As a result, the phase retardation between the two orthogonal polarizations

[1]Geballe Laboratory for Advanced Materials, Stanford University, 476 Lomita Mall, Stanford, CA 94305, USA. [2]Micro and Nanooptics Laboratory, Faculty of Mechanical Engineering and Russell Berrie Nanotechnology Institute, Technion–Israel Institute of Technology, Haifa 32000, Israel. *These authors contributed equally to this work. †Corresponding author. E-mail: brongersma@stanford.edu

EXTENDED PDF FORMAT
SPONSORED BY

Travel Grants Available
Neuroscience 2014
Application Deadline: July 31st
>Apply Now
R&D systems
a bio-techne brand
www.rndsystems.com

Dielectric gradient metasurface optical elements

Dianmin Lin *et al.*

Science **345**, 298 (2014);
DOI: 10.1126/science.1253213

This copy is for your personal, non-commercial use only.

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by clicking here.

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines here.

The following resources related to this article are available online at www.sciencemag.org (this information is current as of August 4, 2014):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:
http://www.sciencemag.org/content/345/6194/298.full.html

Supporting Online Material can be found at:
http://www.sciencemag.org/content/suppl/2014/07/16/345.6194.298.DC1.html

This article **cites 28 articles**, 4 of which can be accessed free:
http://www.sciencemag.org/content/345/6194/298.full.html#ref-list-1

This article appears in the following **subject collections:**
Physics, Applied
http://www.sciencemag.org/cgi/collection/app_physics

Downloaded from www.sciencemag.org on August 4, 2014

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published weekly, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. Copyright 2014 by the American Association for the Advancement of Science; all rights reserved. The title *Science* is a registered trademark of AAAS.

22. C. S. Ponseca Jr. *et al.*, *J. Am. Chem. Soc.* **136**, 5189–5192 (2014).
23. See the supplementary materials on *Science* Online.
24. C. C. Stoumpos, C. D. Malliakas, M. G. Kanatzidis, *Inorg. Chem.* **52**, 9019–9038 (2013).

ACKNOWLEDGMENTS

The authors acknowledge financial support from the Ministry of Science and Technology of China (863 project, grant no. SS2013AA50303), the National Natural Science Foundation of China (grant no. 61106056), the Science and Technology Department of Hubei Province (grant no. 2013BAA090), and the Fundamental Research Funds for the Central Universities (grant no. HUSTNY022). We thank the Analytical and Testing Center of Huazhong University of Science and Technology for field emission SEM, TEM, and XRD testing. M.G. thanks the European Research Council for financial support under advanced research grant ARG 247404, "Mesolight."

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/345/6194/295/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S9
References

14 April 2014; accepted 3 June 2014
10.1126/science.1254763

APPLIED OPTICS

Dielectric gradient metasurface optical elements

Dianmin Lin,[1]* Pengyu Fan,[1]* Erez Hasman,[2] Mark L. Brongersma[1]†

Gradient metasurfaces are two-dimensional optical elements capable of manipulating light by imparting local, space-variant phase changes on an incident electromagnetic wave. These surfaces have thus far been constructed from nanometallic optical antennas, and high diffraction efficiencies have been limited to operation in reflection mode. We describe the experimental realization and operation of dielectric gradient metasurface optical elements capable of also achieving high efficiencies in transmission mode in the visible spectrum. Ultrathin gratings, lenses, and axicons have been realized by patterning a 100-nanometer-thick Si layer into a dense arrangement of Si nanobeam antennas. The use of semiconductors can broaden the general applicability of gradient metasurfaces, as they offer facile integration with electronics and can be realized by mature semiconductor fabrication technologies.

The field of nanophotonics has started to facilitate the replacement of bulky optical components by ultrathin, planar elements (*1*–*5*). Such elements can be realized by nanopatterning thin films and can benefit from a more natural integration into electronic and mechanical systems. They can also provide entirely new functions that capitalize on the rapid developments in the field of metamaterials (*6*). The recent realization of metasurfaces, or metamaterials of reduced dimensionality (*7*), is of particular relevance as it is opening up new opportunities to realize virtually flat optics (*8*, *9*). Metasurfaces consist of dense arrangements of resonant optical antennas. The resonant nature of the light-matter interaction with such structures affords substantial control over the local light scattering amplitude and phase. These scattering properties can be manipulated through choices of the antenna material, size, geometry, orientation, and environment. By arranging antennas with distinct scattering properties on a surface, space-variant metasurfaces are created that offer tremendous freedom in manipulating optical wave-fronts. Metasurfaces constructed from high-index antenna arrays are used as antireflection coatings (*10*), as well as to enhance absorption (*11*), create structural color (*12*), and manipulate the state of polarization (*13*, *14*). However, space-variant metasurfaces for beam steering and focusing have thus far been constructed from nanoscale metallic antennas only (*7*). Ohmic losses in the metal and the limited scattering cross sections of the antennas have limited high device efficiencies to reflection mode (*15*, *16*). We leverage recent advances in the realization of high-index dielectric antennas (*12*, *17*, *18*) and Pancharatnam-Berry phase optical elements (PBOEs) (*19*, *20*) to realize easy-to-fabricate and highly efficient transmissive devices. Si-based gradient metasurfaces capable of serving as optical gratings, lenses, and axicons are demonstrated.

Figure 1 shows a dielectric gradient metasurface optical element (DGMOE) serving as an axicon. Whereas conventional glass axicons are a few millimeters in size (Fig. 1A), the DGMOE features a 100-nm-thick layer of poly-silicon deposited on a quartz substrate (Fig. 1B) generated in a single patterning step (*21*). Figure 1D shows a scanning electron microscopy (SEM) image of the fabricated DGMOE composed of thousands of Si nanobeam antennas.

When illuminated with a collimated Gaussian left circularly polarized (LCP) light beam at a 550-nm wavelength, the DGMOE creates a nondiffracting Bessel beam on the transmission (Fig. 1, C and E, and section S2 of the supplementary materials). The intensity distributions along and across the center of Bessel beam are shown as insets to Fig. 1E. The ratio between the high-intensity central beam spot and the first ring is 6:1, a signature of a Bessel beam. The operation of the lens could be analyzed by understanding all of the coherent far-field and near-field interactions between the Si nanobeam antennas. Here, we describe the behavior of the optical element based on an understanding of the Pancharatnam-Berry (PB) phase.

The PB phase is a geometric phase achieved by space-variant polarization manipulations (*22*, *23*), as opposed to a propagation phase. PBOEs are constructed from wave plate elements for which the orientation of the fast axes depends on the spatial position. Well-defined algorithms exist for realizing specific optical functions. By tiling a surface with half-wave plates with their fast-axes orientations according to a function $\theta(x,y)$, an incident circularly polarized light beam will be fully transformed to a beam of opposite helicity and imprinted with a geometric phase equal to $\varphi_g(x,y) = \pm 2\theta(x,y)$ (*21*, *24*). By controlling the local orientation of the fast axes of the wave plate elements between 0 and π, phase pickups can be achieved that cover the full 0-to-2π range while maintaining equal transmission amplitude for the entire optical component. This provides full control over the wavefront. A continuous desired phase function can be approximated using discrete wave plate orientations. We use eight orientations with which a high theoretical diffraction efficiency of 95% can be achieved (*25*). The realization of wave plate elements represents a crucial step in crafting PBOEs.

In designing ultrathin nanobeam-based wave plates, it is important to first understand the optical resonances of individual semiconductor nanobeams. Top-illuminated single beams support resonances under transverse electric (TE) illumination (with the electric field polarized normal to the length of the structure) and the orthogonal transverse magnetic (TM) polarization (*12*, *18*). The thickness of the resonant structures can be small compared with the free-space wavelength due to the high refractive index of semiconductors. A 120-nm-wide and 100-nm-high Si nanobeam supports a strong resonance and a substantial phase retardation in our target wavelength range of interest. We chose this beam as the basic building block for our wave plate (Fig. 2A) in anticipation that its resonant behavior and the accompanied phase retardation between the orthogonal polarizations would persist in the beam array. Finite element simulations indicate that the wavefront for a TE-polarized light beam at 550 nm is delayed by 0.14π as compared with the wavefront in the absence of the nanobeams. The wavefront for TM-polarized light is delayed even further to 1.15π (Fig. 2B). As a result, the phase retardation between the two orthogonal polarizations

[1]Geballe Laboratory for Advanced Materials, Stanford University, 476 Lomita Mall, Stanford, CA 94305, USA. [2]Micro and Nanooptics Laboratory, Faculty of Mechanical Engineering and Russell Berrie Nanotechnology Institute, Technion–Israel Institute of Technology, Haifa 32000, Israel. *These authors contributed equally to this work. †Corresponding author. E-mail: brongersma@stanford.edu

is about π, and the beam array serves as a half-wave plate.

By sweeping the wavelength from 490 to 700 nm in simulations, the phase retardation of the wave plate varies from 0.4π to 1.2π (Fig. 2C). For reference, the figure also shows the comparatively small phase retardation of 0.063π for a 100-nm-thick film of calcite, a naturally birefringent crystal. The substantial swing in the phase retardation is attributed to the fact that the nanobeams support a strong fundamental resonance under TE illumination (as seen in absorption) and only a very weak second-order TM resonance (fig. S5D). The order of the resonance is determined by the number of field maxima inside the nanobeam (Fig. 2D, insets). Figure 2C shows that the array's TE resonance and associated swing in the phase retardation can be translated spectrally by changing the beam width.

To realize ultrathin Si nanobeam wave plates, we fabricated the periodic Si nanobeam array described above (Fig. 2C, inset). These samples were characterized optically through measurements of the four Stokes polarization parameters (supplementary materials section S4). The measured phase retardation of this dielectric metasurface is shown by the red squares in Fig. 2C and is in good agreement with optical simulations (red solid line). The transmission coefficient at 550 nm, where the phase retardation is about π, is ~0.5. Spectral transmission measurements show that transmission coefficients as high as 0.75 can be obtained at longer wavelengths (650 to 700 nm) closer to the Si band gap (*21*).

The most basic DGMOE is one that serves as a blazed diffraction grating capable of steering a light beam into several diffracted orders. The design follows a general strategy that can be used to create more complex optical elements. First, Fourier optics is used to determine the phase profile that will produce a desired field distribution in the far field. Blazed gratings feature a simple phase profile with a linear dependence of the phase on position, which either increases or decreases by 2π across one grating period. The desired phase profile is then discretized into segments that can take on a limited number of discrete phases. The discretized profile $\varphi_d(x,y)$ for a grating with a period Λ of 3.2 μm is shown in Fig. 3A. The phase profile is accomplished by tiling a glass substrate with nanobeam wave plates with space-variant orientations (Fig. 3B). To achieve a desired far-field intensity distribution, it is important to control both the metasurface properties and the polarization state of the incident light. For an incident plane wave $|E_{in}\rangle$ with an arbitrary polarization state, one can show that the output wave $|E_{out}\rangle$ will be composed of three polarization orders (*26*)

$$|E_{out}\rangle = \sqrt{\eta_E}|E_{in}\rangle + \sqrt{\eta_R}e^{i2\theta(x,y)}|R\rangle + \sqrt{\eta_L}e^{-i2\theta(x,y)}|L\rangle \quad (1)$$

Here, $|R\rangle$ and $|L\rangle$ denote the right- and left-handed circularly polarized unit vectors, respectively. The quantities $\eta_E = |\frac{1}{2}(t_x + t_y e^{i\phi})|^2$, $\eta_R = |\frac{1}{2}(t_x - t_y e^{i\phi})\langle L|E_{in}\rangle|^2$, and $\eta_L = |\frac{1}{2}(t_x - t_y e^{i\phi})\langle R|E_{in}\rangle|^2$ provide the magnitude of the coupling efficiencies to the different polarization orders. For convenience, we use Dirac bra-ket notation, where $\langle\alpha|\beta\rangle$ denotes an inner product. The function $\theta(x,y)$ describes the spatially variant distribution of the fast axes of the wave plates. Here, t_x and t_y are the transmission coefficients for light polarized parallel and perpendicular to the fast optical axis, and ϕ is the phase retardation between these linear polarization states. From Eq. 1 it is clear that the optical materials and geometrical properties of the beams, their orientation distribution, and the incident polarization state control the diffraction properties of the array. The operation of DGMOEs can also be explained with simple, intuitive graphics (supplementary materials sections S7 and S8).

The diffraction patterns of the DGMOE blazed grating were observed in transmission with the help of a Bertrand lens (fig. S6). For an incident right circularly polarized (RCP) beam at a wavelength of 550 nm, part of the light changed handedness to LCP and experienced a phase pickup equal to $-2\theta(x,y)$ upon propagation through the DGMOE. As a result, some light steered to the left when viewed from the location of the source (Fig. 3C). A fraction of the light went straight through and emerged again in the incident RCP state due to the fact the transmission magnitudes t_x and t_y are not perfectly equal (fig. S5). For LCP light, part of the light steered to the opposite direction while transforming to RCP light. Again, a fraction of the light went straight through and did not change its state of polarization. For a linearly polarized incident beam, which can be viewed as consisting of equal amounts of RCP and LCP light, light is steered in both directions.

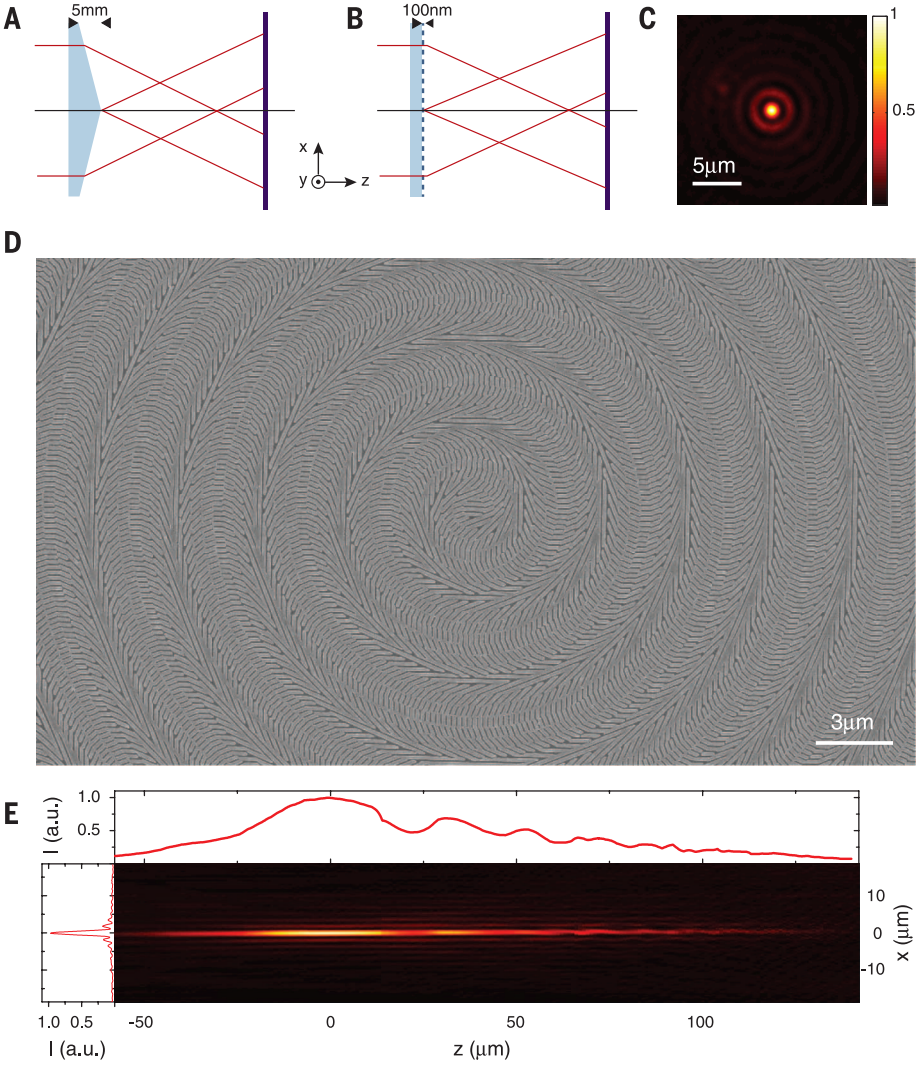


Fig. 1. Example of a DGMOE: An axicon constructed from Si nanoantennas. (**A**) Schematic of a conventional glass axicon focusing light into a Bessel beam. (**B**) Schematic of a DGMOE axicon featuring an ultrathin patterned layer of silicon on a quartz substrate. (**C**) Transversal distribution of Bessel beam generated by the DGMOE. (**D**) SEM image of the fabricated DGMOE. (**E**) Measured intensity profile of the nondiffractive Bessel beam generated behind the DGMOE in the *xz* plane. The intensity (I) along the center of Bessel beam is plotted in the inset along the *z* axis. The inset along the *x* axis shows the cross-sectional intensity profile at the *z* = 0 plane. a.u., arbitrary units.

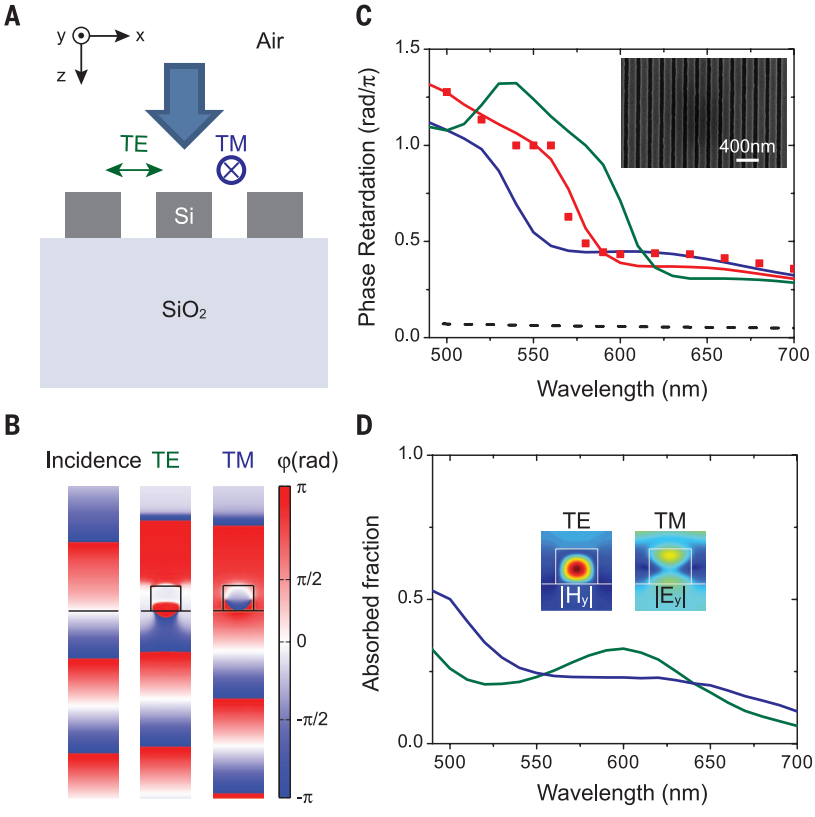


Fig. 2. Ultrathin wave plates constructed from Si nanoantennas. (**A**) Schematic view of a periodic Si nanobeam array. (**B**) Phase wavefronts for TE- and TM-polarized, 550-nm-wavelength light upon transmission through a wave plate with 120-nm-wide Si beams. The incident wave without the Si beams is shown for reference. rad, radians. (**C**) Simulated spectra of the phase retardation of a TM-polarized wave with respect to a TE-polarized wave for nanobeam arrays with beam widths of 100 nm (blue), 120 nm (red), 140 nm (green), with the same thickness of 100 nm and duty cycle of 60%. The dashed black line shows the phase retardation for a 100-nm-thick calcite film. Measurements of the phase retardation (red squares) obtained with an array of 120-nm beams show good agreement with the simulations. (Inset) SEM image of the fabricated array. (**D**) Simulated absorption spectra of the 120-nm-wide nanobeam array in (C) under TE (green) and TM (blue) illumination. (Insets) Magnetic field distribution $|H_y|$ for TE and electrical field distribution $|E_y|$ of TM illumination at a wavelength of 600 nm.

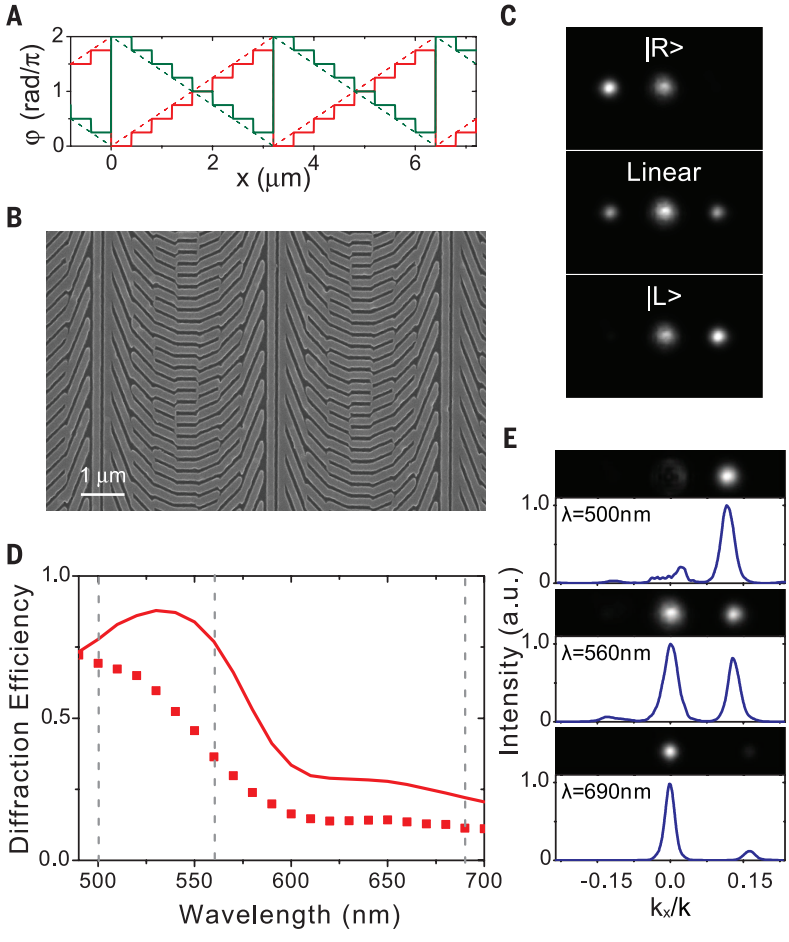


Fig. 3. A DGMOE serving as a blazed grating. (**A**) Discretized (solid line) and continuous (dashed line) phase profile of a DGMOE serving as a blazed grating for illumination with LCP light (red) and RCP light (green). (**B**) SEM image of the fabricated DGMOE. (**C**) Measured diffraction patterns from the DGMOE under illumination with right circular polarization (top), linear polarization (middle), and left circular polarization (bottom) at λ = 550 nm. (**D**) Theoretical (continuous line) and experimental (solid squares) spectra of diffraction efficiency. (**E**) Measured diffraction patterns of the DGMOE blazed grating and associated cross-sectional intensity distributions versus the normalized in-plane momentum k_x/k under illumination with left circular polarization at wavelengths of 500, 560, and 690 nm [vertical dashed lines in (D)].

The expected polarization states of the transmitted light based on Eq. 1 were successfully verified by adding a circular polarizer on the transmission side of the metasurface (supplementary materials section S4 and movie S2). These observations demonstrate the possibility to steer light by changing the incident polarization (movie S1).

DGMOEs can operate over a broadband wavelength range, and it is thus important to understand their wavelength-dependent properties. Because the geometric phase depends only on the orientation function θ, the phase profile experienced by RCP and LCP light will remain the same at each wavelength. This does not imply that the diffraction properties stay the same as well. For example, the steering angle of the blazed grating α increases for longer wavelengths, as this quantity is determined by the ratio of the wavelength and grating period ($\sin\alpha = \lambda/\Lambda$). The diffraction efficiency also varies across the spectrum, as it is linked to the wavelength-dependent transmission and phase-retardation properties of the nanobeam wave plates. The diffraction efficiency η_D is defined by the ratio of the power of circularly polarized light that has been converted to the opposite helicity (i.e., experiencing phase pickup) and the total transmitted power that is transmitted through the metasurface. For a well-designed blazed grating, the intensity in higher diffraction orders is very weak. As such, the total power of transmitted light is approximately equal to the sum power of the zero-order intrinsic beam and the first-order converted beam: $\eta_D \approx I_{1st}/(I_{1st} + I_{0rd})$. The diffraction efficiency for the grating shown in Fig. 3B was determined by measuring the light intensity of the through beam and the diffracted beam with a calibrated charge-coupled device camera (supplementary materials section S6). Figure 3D shows the experimental diffraction efficiency spectra for normally incident LCP light in the wavelength range from 490 to 700 nm (red squares). It also shows theoretical efficiencies calculated using the expressions for η_E and η_R given below Eq. 1. These efficiencies were determined using the measured values of t_x, t_y, and ϕ and by taking into account the expected diffraction efficiency of 95% that is appropriate for an eight-level phase discretization.

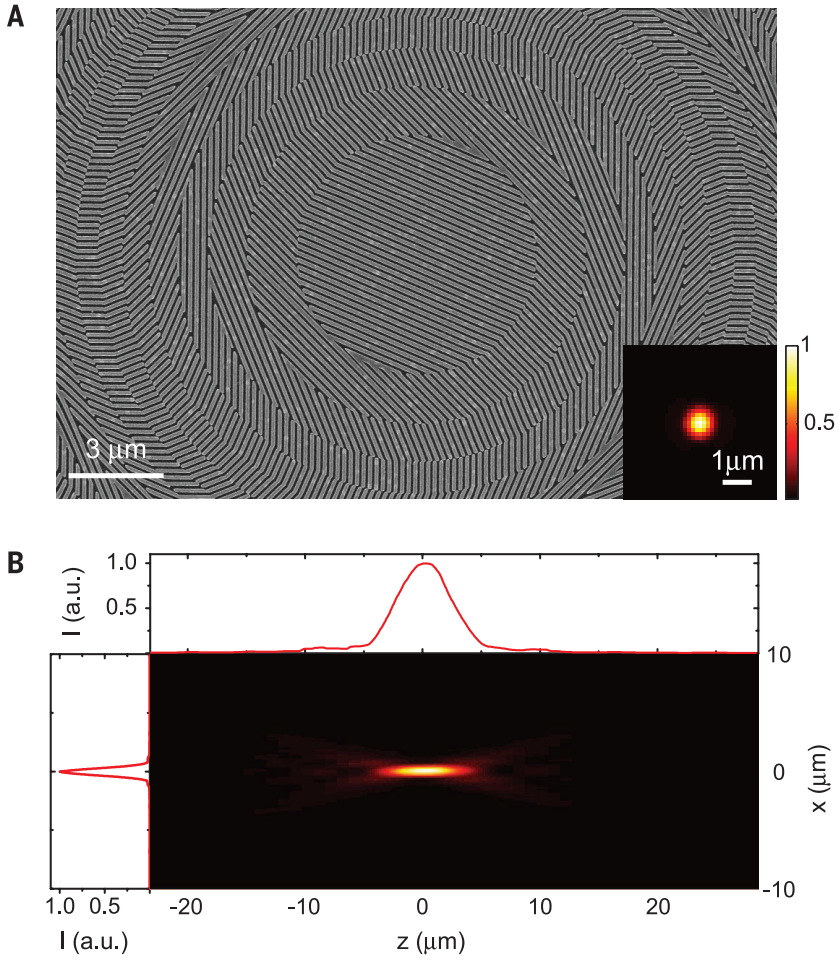


Fig. 4. A DGMOE lens based on Si nanobeams. (**A**) SEM image of a fabricated DGMOE lens with a focal length of 100 μm at λ = 550 nm. (Inset) Two-dimensional intensity profile in the focal plane. (**B**) Measured intensity profile generated behind the DGMOE in the *xz* plane. The intensity distributions along the optical axis and through the focus are shown along the vertical and horizontal axes.

The experimental diffraction patterns at three different wavelengths are plotted in Fig. 3E. At a wavelength of 500 nm, 75% of the light is steered, and the intensity of the zero-order beam (i.e., the straight-through beam) is very low compared with that of the first diffracted order. In contrast, at λ = 690 nm only a small portion of light is steered, and most of the energy remains in the zero-order beam. The pattern also shows the expected increase in the steering angle at this longer wavelength. The overall shape of the measured efficiency spectrum qualitatively agrees with the predicted spectrum, but the measured efficiencies tend to be lower, possibly attributed to imperfections in the fabricated structure. Alternatively, it may be that the birefringent nature of the nanobeam wave plates is size-dependent.

We also realized a DGMOE serving as a conventional lens (Fig. 4A). The phase profile for this DGMOE lens again features eight discrete levels and approximates a hyperboloidal phase profile (supplementary materials section S9). The lens features a focal length of 100 μm and a numerical aperture (NA) of 0.43 at a wavelength of 550 nm. When the DGMOE is uniformly illuminated from the substrate side with RCP light, it concentrates light into a LCP focal spot (Fig. 4B and see supplementary materials section S2 for details). The focal spot size measures 670 nm in full width at half maximum, close to the diffraction-limited spot size.

The presented design principles are generally applicable to any high-index semiconductor or insulator and can be applied at any desired wavelength. The lowest ohmic losses were achieved near the semiconductor band edge. The presented ultrathin DGMOEs address an increased need for low-cost, lightweight, and compact optical elements that can easily be integrated into complex systems requiring assembly of optical, electronic, and mechanical components.

REFERENCES AND NOTES

1. L. Verslegers *et al.*, *Nano Lett.* **9**, 235–238 (2009).
2. D. Fattal, J. Li, Z. Peng, M. Fiorentino, R. G. Beausoleil, *Nat. Photonics* **4**, 466–470 (2010).
3. C. J. Chang-Hasnain, *Semicond. Sci. Technol.* **26**, 014043 (2011).
4. F. Afshinmanesh, J. S. White, W. Cai, M. L. Brongersma, *Nanophotonics* **1**, 125–129 (2012).
5. E. T. F. Rogers *et al.*, *Nat. Mater.* **11**, 432–435 (2012).
6. N. Engheta, *Science* **317**, 1698–1702 (2007).
7. A. V. Kildishev, A. Boltasseva, V. M. Shalaev, *Science* **339**, 1232009 (2013).
8. N. Yu *et al.*, *Science* **334**, 333–337 (2011).
9. X. Ni, N. K. Emani, A. V. Kildishev, A. Boltasseva, V. M. Shalaev, *Science* **335**, 427 (2012).
10. P. Spinelli, M. A. Verschuuren, A. Polman, *Nat. Commun.* **3**, 692 (2012).
11. L. Cao *et al.*, *Nano Lett.* **10**, 439–445 (2010).
12. L. Cao, P. Fan, E. S. Barnard, A. M. Brown, M. L. Brongersma, *Nano Lett.* **10**, 2649–2654 (2010).
13. Y. Yang *et al.*, *Nano Lett.* **14**, 1394–1399 (2014).
14. C. Wu *et al.*, http://arxiv.org/abs/1309.6616 (2013).
15. A. Pors, O. Albrektsen, I. P. Radko, S. I. Bozhevolnyi, *Sci. Rep.* **3**, 2155 (2013).
16. S. Sun *et al.*, *Nano Lett.* **12**, 6223–6229 (2012).
17. R. K. Mongia, A. Ittipiboon, *IEEE Trans. Antenn. Propag.* **45**, 1348–1356 (1997).
18. L. Cao *et al.*, *Nat. Mater.* **8**, 643–647 (2009).

19. Z. Bomzon, G. Biener, V. Kleiner, E. Hasman, *Opt. Lett.* **27**, 285–287 (2002).
20. A. Niv, G. Biener, V. Kleiner, E. Hasman, *Opt. Lett.* **29**, 238–240 (2004).
21. Materials and methods are available as supplementary materials on *Science* Online.
22. S. Pancharatnam, *Proc. Indiana Acad. Sci.* **A44**, 247–262 (1956).
23. M. V. Berry, *J. Mod. Opt.* **34**, 1401–1407 (1987).
24. Z. Bomzon, G. Biener, V. Kleiner, E. Hasman, *Opt. Lett.* **27**, 1141–1143 (2002).
25. A. Niv, G. Biener, V. Kleiner, E. Hasman, *Opt. Commun.* **251**, 306–314 (2005).
26. E. Hasman, V. Kleiner, G. Biener, A. Niv, *Appl. Phys. Lett.* **82**, 328–330 (2003).

ACKNOWLEDGMENTS

We acknowledge funding support from the Global Climate and Energy Project. P.F. acknowledges support from the Stanford Graduate Fellowship. We also thank P. Landreman for his assistance with the optical measurements.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/345/6194/298/suppl/DC1
Materials and Methods
Figs. S1 to S10
References (*27*–*30*)
Movies S1 to S3

11 March 2014; accepted 2 June 2014
10.1126/science.1253213

QUANTUM COMPUTING

Quantum computations on a topologically encoded qubit

D. Nigg,[1]*† M. Müller,[2]*† E. A. Martinez,[1] P. Schindler,[1]‡ M. Hennrich,[1] T. Monz,[1] M. A. Martin-Delgado,[2] R. Blatt[1,3]

The construction of a quantum computer remains a fundamental scientific and technological challenge because of the influence of unavoidable noise. Quantum states and operations can be protected from errors through the use of protocols for quantum computing with faulty components. We present a quantum error-correcting code in which one qubit is encoded in entangled states distributed over seven trapped-ion qubits. The code can detect one bit flip error, one phase flip error, or a combined error of both, regardless on which of the qubits they occur. We applied sequences of gate operations on the encoded qubit to explore its computational capabilities. This seven-qubit code represents a fully functional instance of a topologically encoded qubit, or color code, and opens a route toward fault-tolerant quantum computing.

A fully fledged quantum computer can be used to efficiently solve notoriously difficult problems, such as factoring large numbers or simulating the dynamics of many-body quantum systems (*1*). Technological progress has enabled the implementation of small-scale prototype quantum computing devices on diverse physical platforms (*2*). Sophisticated fault-tolerant quantum computing (FTQC) techniques have been developed for the systematic correction of errors that dynamically occur during storage and manipulation of quantum information (*3*–*5*). For quantum error correction, Calderbank-Shor-Steane (CSS) codes (*4*, *5*) enable independent detection and correction of bit and phase flip errors, as well as combinations thereof. Furthermore, quantum information processing is substantially facilitated in quantum codes in which logical operations on encoded qubits are realized by the bitwise application of the corresponding operations to the underlying physical qubits (i.e., in a transversal way). This property prevents uncontrolled propagation of errors through the quantum hardware, which in turn is essential to enter the FTQC regime (*1*). Ultimately, reliable quantum memories and arbitrarily long quantum computations are predicted to become feasible for appropriately designed quantum codes, once all elementary operations are realized in a fault-tolerant way and with sufficiently low error rates (*6*, *7*).

To date, topological quantum computing represents the most promising and realistic approach toward FTQC. In this method, the encoding of quantum information in global properties of a many-particle system provides protection against noise sources that act locally on individual or small sets of qubits (*8*). Most prominently, topological quantum computing offers highly competitive error thresholds as high as 1% per operation (*9*–*12*), which is within reach of current experimental capabilities (*13*–*15*) and typically about two orders of magnitude larger than in schemes using concatenated quantum codes (*7*).

Within topological quantum computing, topological color codes (*16*, *17*) offer the distinctive feature that the entire group of Clifford gate operations, allowing for arbitrary 90° and 180° qubit rotations, can be implemented transversally (*1*). This versatile set of operations directly enables protocols for quantum distillation of entanglement, quantum teleportation, and dense coding with topological protection (*16*). Moreover, a universal gate set, enabling the implementation of arbitrary quantum algorithms, can be achieved by complementing the Clifford operations with a single non-Clifford gate (*1*). For color codes in two-dimensional (2D) architectures (*16*), such an additional gate can be realized by a technique known as magic-state injection (*18*). Remarkably, this method is not needed in 3D color codes that enable implementation of a universal gate set using exclusively transversal operations (*17*).

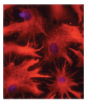
Previous experiments have demonstrated the correction of a single type of error by the three-qubit repetition code (*19*–*21*), correction of bit and phase flip errors by the non–CSS-type five-qubit code in nuclear magnetic resonance systems (*22*, *23*), as well as elements of topological error correction in the framework of measurement-based quantum computation (*24*). Here, we demonstrate a quantum error-correcting seven-qubit CSS code (*5*), which is equivalent to the smallest instance of a 2D topological color code (*16*). The application of multiple operations from the entire, transversally implemented set of logical single-qubit Clifford gates represents the realization of quantum computations on a fully correctable encoded qubit.

Two-dimensional color codes are topological quantum error-correcting codes that are constructed on underlying 2D lattices (*16*) for which three links meet at each vertex and three different colors are sufficient to assign color to all polygons (plaquettes) of the lattice, such that no adjacent plaquettes sharing a link are of the same color. The smallest fully functional 2D color code involves seven qubits (Fig. 1A) and consists of a triangular, planar code structure formed by three adjoined plaquettes with one physical qubit placed at each vertex. Color codes are stabilizer quantum codes (*1*, *25*), which are defined by a set of commuting, so-called stabilizer operators $\{S_i\}$, each having eigenvalues +1 or –1. More precisely, the code space hosting logical or encoded quantum states $|\psi\rangle_L$ is fixed as the simultaneous eigenspace of eigenvalue +1 of all stabilizers, $S_i|\psi\rangle_L = +|\psi\rangle_L$. In color codes, there are two stabilizer operators associated with each plaquette, which for the seven-qubit color code (Fig. 1A) results in the set of four-qubit X and Z-type operators

$$\begin{array}{ll} S_x^{(1)} = X_1X_2X_3X_4 & S_z^{(1)} = Z_1Z_2Z_3Z_4 \\ S_x^{(2)} = X_2X_3X_5X_6 & S_z^{(2)} = Z_2Z_3Z_5Z_6 \\ S_x^{(3)} = X_3X_4X_6X_7 & S_z^{(3)} = Z_3Z_4Z_6Z_7 \end{array} \quad (1)$$

where X_i, Y_i, and Z_i denote the standard Pauli matrices acting on the ith physical qubit with the computational basis states $|0\rangle$ and $|1\rangle$ (*1*). The stabilizers in Eq. 1 impose six independent constraints on the seven physical qubits and thus define a two-dimensional code space, which allows the encoding of one logical qubit. The logical basis states $|0\rangle_L$ and $|1\rangle_L$ spanning the code space are entangled seven-qubit states and

[1]Institut für Experimentalphysik, Universität Innsbruck, A-6020 Innsbruck, Austria. [2]Departamento de Física Teórica I, Universidad Complutense, 28040 Madrid, Spain. [3]Institut für Quantenoptik und Quanteninformation der Österreichischen Akademie der Wissenschaften, A-6020 Innsbruck, Austria. ***Corresponding author: E-mail: daniel.nigg@uibk.ac.at (D.N.); mueller@ucm.es (M.M.)** †These authors contributed equally to this work. ‡Present address: Department of Physics, University of California, Berkeley, CA 94720, USA.

EXTENDED PDF FORMAT
SPONSORED BY

Travel Grants Available
Neuroscience 2014
Application Deadline: July 31st
>Apply Now
R&D systems
a bio-techne brand
www.rndsystems.com

Quantum computations on a topologically encoded qubit

D. Nigg *et al.*

Science **345**, 302 (2014);
DOI: 10.1126/science.1253742

This copy is for your personal, non-commercial use only.

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by clicking here.

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines here.

The following resources related to this article are available online at www.sciencemag.org (this information is current as of August 4, 2014):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:
http://www.sciencemag.org/content/345/6194/302.full.html

Supporting Online Material can be found at:
http://www.sciencemag.org/content/suppl/2014/06/11/science.1253742.DC1.html

This article **cites 23 articles**, 1 of which can be accessed free:
http://www.sciencemag.org/content/345/6194/302.full.html#ref-list-1

This article appears in the following **subject collections:**
Physics
http://www.sciencemag.org/cgi/collection/physics

Downloaded from www.sciencemag.org on August 4, 2014

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published weekly, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. Copyright 2014 by the American Association for the Advancement of Science; all rights reserved. The title *Science* is a registered trademark of AAAS.

19. Z. Bomzon, G. Biener, V. Kleiner, E. Hasman, *Opt. Lett.* **27**, 285–287 (2002).
20. A. Niv, G. Biener, V. Kleiner, E. Hasman, *Opt. Lett.* **29**, 238–240 (2004).
21. Materials and methods are available as supplementary materials on *Science* Online.
22. S. Pancharatnam, *Proc. Indiana Acad. Sci.* **A44**, 247–262 (1956).
23. M. V. Berry, *J. Mod. Opt.* **34**, 1401–1407 (1987).
24. Z. Bomzon, G. Biener, V. Kleiner, E. Hasman, *Opt. Lett.* **27**, 1141–1143 (2002).
25. A. Niv, G. Biener, V. Kleiner, E. Hasman, *Opt. Commun.* **251**, 306–314 (2005).
26. E. Hasman, V. Kleiner, G. Biener, A. Niv, *Appl. Phys. Lett.* **82**, 328–330 (2003).

ACKNOWLEDGMENTS

We acknowledge funding support from the Global Climate and Energy Project. P.F. acknowledges support from the Stanford Graduate Fellowship. We also thank P. Landreman for his assistance with the optical measurements.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/345/6194/298/suppl/DC1
Materials and Methods
Figs. S1 to S10
References (*27*–*30*)
Movies S1 to S3

11 March 2014; accepted 2 June 2014
10.1126/science.1253213

QUANTUM COMPUTING

Quantum computations on a topologically encoded qubit

D. Nigg,[1]*† M. Müller,[2]*† E. A. Martinez,[1] P. Schindler,[1]‡ M. Hennrich,[1] T. Monz,[1] M. A. Martin-Delgado,[2] R. Blatt[1,3]

The construction of a quantum computer remains a fundamental scientific and technological challenge because of the influence of unavoidable noise. Quantum states and operations can be protected from errors through the use of protocols for quantum computing with faulty components. We present a quantum error-correcting code in which one qubit is encoded in entangled states distributed over seven trapped-ion qubits. The code can detect one bit flip error, one phase flip error, or a combined error of both, regardless on which of the qubits they occur. We applied sequences of gate operations on the encoded qubit to explore its computational capabilities. This seven-qubit code represents a fully functional instance of a topologically encoded qubit, or color code, and opens a route toward fault-tolerant quantum computing.

A fully fledged quantum computer can be used to efficiently solve notoriously difficult problems, such as factoring large numbers or simulating the dynamics of many-body quantum systems (*1*). Technological progress has enabled the implementation of small-scale prototype quantum computing devices on diverse physical platforms (*2*). Sophisticated fault-tolerant quantum computing (FTQC) techniques have been developed for the systematic correction of errors that dynamically occur during storage and manipulation of quantum information (*3*–*5*). For quantum error correction, Calderbank-Shor-Steane (CSS) codes (*4*, *5*) enable independent detection and correction of bit and phase flip errors, as well as combinations thereof. Furthermore, quantum information processing is substantially facilitated in quantum codes in which logical operations on encoded qubits are realized by the bitwise application of the corresponding operations to the underlying physical qubits (i.e., in a transversal way). This property prevents uncontrolled propagation of errors through the quantum hardware, which in turn is essential to enter the FTQC regime (*1*). Ultimately, reliable quantum memories and arbitrarily long quantum computations are predicted to become feasible for appropriately designed quantum codes, once all elementary operations are realized in a fault-tolerant way and with sufficiently low error rates (*6*, *7*).

To date, topological quantum computing represents the most promising and realistic approach toward FTQC. In this method, the encoding of quantum information in global properties of a many-particle system provides protection against noise sources that act locally on individual or small sets of qubits (*8*). Most prominently, topological quantum computing offers highly competitive error thresholds as high as 1% per operation (*9*–*12*), which is within reach of current experimental capabilities (*13*–*15*) and typically about two orders of magnitude larger than in schemes using concatenated quantum codes (*7*).

Within topological quantum computing, topological color codes (*16*, *17*) offer the distinctive feature that the entire group of Clifford gate operations, allowing for arbitrary 90° and 180° qubit rotations, can be implemented transversally (*1*). This versatile set of operations directly enables protocols for quantum distillation of entanglement, quantum teleportation, and dense coding with topological protection (*16*). Moreover, a universal gate set, enabling the implementation of arbitrary quantum algorithms, can be achieved by complementing the Clifford operations with a single non-Clifford gate (*1*). For color codes in two-dimensional (2D) architectures (*16*), such an additional gate can be realized by a technique known as magic-state injection (*18*). Remarkably, this method is not needed in 3D color codes that enable implementation of a universal gate set using exclusively transversal operations (*17*).

Previous experiments have demonstrated the correction of a single type of error by the three-qubit repetition code (*19*–*21*), correction of bit and phase flip errors by the non–CSS-type five-qubit code in nuclear magnetic resonance systems (*22*, *23*), as well as elements of topological error correction in the framework of measurement-based quantum computation (*24*). Here, we demonstrate a quantum error-correcting seven-qubit CSS code (*5*), which is equivalent to the smallest instance of a 2D topological color code (*16*). The application of multiple operations from the entire, transversally implemented set of logical single-qubit Clifford gates represents the realization of quantum computations on a fully correctable encoded qubit.

Two-dimensional color codes are topological quantum error-correcting codes that are constructed on underlying 2D lattices (*16*) for which three links meet at each vertex and three different colors are sufficient to assign color to all polygons (plaquettes) of the lattice, such that no adjacent plaquettes sharing a link are of the same color. The smallest fully functional 2D color code involves seven qubits (Fig. 1A) and consists of a triangular, planar code structure formed by three adjoined plaquettes with one physical qubit placed at each vertex. Color codes are stabilizer quantum codes (*1*, *25*), which are defined by a set of commuting, so-called stabilizer operators $\{S_i\}$, each having eigenvalues +1 or –1. More precisely, the code space hosting logical or encoded quantum states $|\psi\rangle_L$ is fixed as the simultaneous eigenspace of eigenvalue +1 of all stabilizers, $S_i|\psi\rangle_L = +|\psi\rangle_L$. In color codes, there are two stabilizer operators associated with each plaquette, which for the seven-qubit color code (Fig. 1A) results in the set of four-qubit X and Z-type operators

$$\begin{array}{ll} S_x^{(1)} = X_1X_2X_3X_4 & S_z^{(1)} = Z_1Z_2Z_3Z_4 \\ S_x^{(2)} = X_2X_3X_5X_6 & S_z^{(2)} = Z_2Z_3Z_5Z_6 \\ S_x^{(3)} = X_3X_4X_6X_7 & S_z^{(3)} = Z_3Z_4Z_6Z_7 \end{array} \quad (1)$$

where X_i, Y_i, and Z_i denote the standard Pauli matrices acting on the ith physical qubit with the computational basis states $|0\rangle$ and $|1\rangle$ (*1*). The stabilizers in Eq. 1 impose six independent constraints on the seven physical qubits and thus define a two-dimensional code space, which allows the encoding of one logical qubit. The logical basis states $|0\rangle_L$ and $|1\rangle_L$ spanning the code space are entangled seven-qubit states and

[1]Institut für Experimentalphysik, Universität Innsbruck, A-6020 Innsbruck, Austria. [2]Departamento de Física Teórica I, Universidad Complutense, 28040 Madrid, Spain. [3]Institut für Quantenoptik und Quanteninformation der Österreichischen Akademie der Wissenschaften, A-6020 Innsbruck, Austria. ***Corresponding author: E-mail: daniel.nigg@uibk.ac.at (D.N.); mueller@ucm.es (M.M.)** †These authors contributed equally to this work. ‡Present address: Department of Physics, University of California, Berkeley, CA 94720, USA.

are given as the eigenstates of the logical operator $Z_L = Z_1Z_2Z_3Z_4Z_5Z_6Z_7$, where $Z_L|0\rangle_L = |0\rangle_L$ and $Z_L|1\rangle_L = -|1\rangle_L$ (*16*, *26*).

For the physical realization of a topologically encoded qubit, we store seven $^{40}Ca^+$ ions in a linear Paul trap. Each ion hosts a physical qubit, which is encoded in (meta)stable electronic states (*27*). Within our setup (Fig. 1B), we realize a high-fidelity universal set of quantum operations consisting of single-ion phase shifts, collective rotations, and a collective entangling gate. Additionally, a single-ion spectroscopic decoupling technique enables the collective entangling operation to act only on subsets of qubits (*26*).

The initial preparation of the logical state $|0\rangle_L$ (encoding) is realized deterministically by the quantum circuit shown in Fig. 1C. First, the seven-ion system is prepared in the product

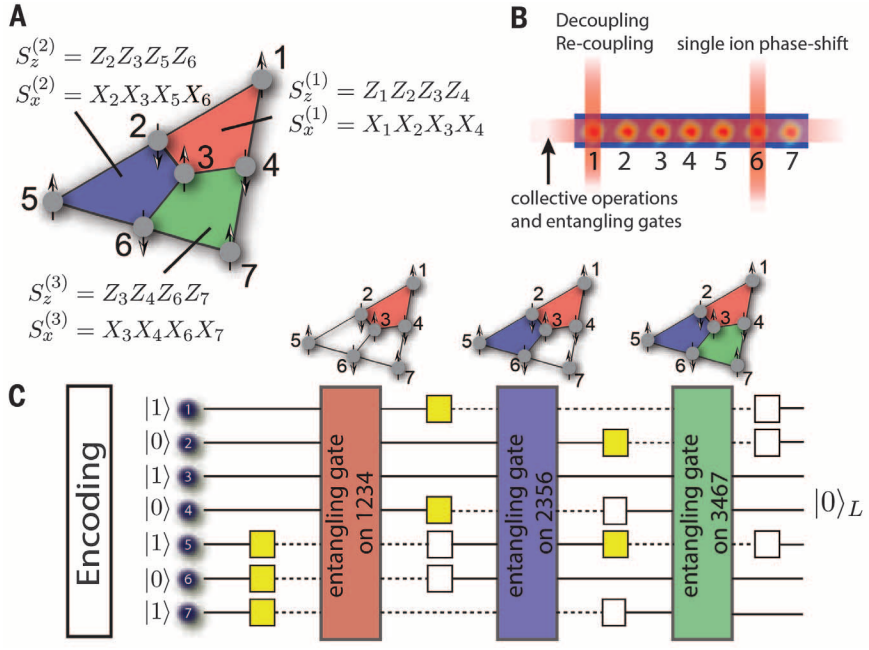


Fig. 1. The topologically encoded qubit: its physical implementation and initialization. (**A**) One logical qubit is embedded in seven physical qubits forming a 2D triangular planar code structure of three plaquettes. The code space is defined via six stabilizer operators $S_X^{(j)}$ and $S_Z^{(j)}$, each acting on a plaquette that involves four physical qubits. (**B**) For the physical realization, qubits are encoded in electronic states of a linear string of ions. These ion qubits can be manipulated via laser interactions that realize a universal gate set consisting of single-ion phase shifts, collective operations, and a collective entangling gate, which is complemented by a single-ion spectroscopic decoupling and recoupling technique (*25*). (**C**) Encoding of the logical qubit is achieved by coherently mapping the input state $|1010101\rangle$ onto the logical state $|0\rangle_L$, using a quantum circuit that combines plaquette-wise entangling operations with decoupling and recoupling pulses (yellow and white squares, respectively) (*26*). Dashed lines denote decoupled or inactive qubits; solid lines denote recoupled or active qubits.

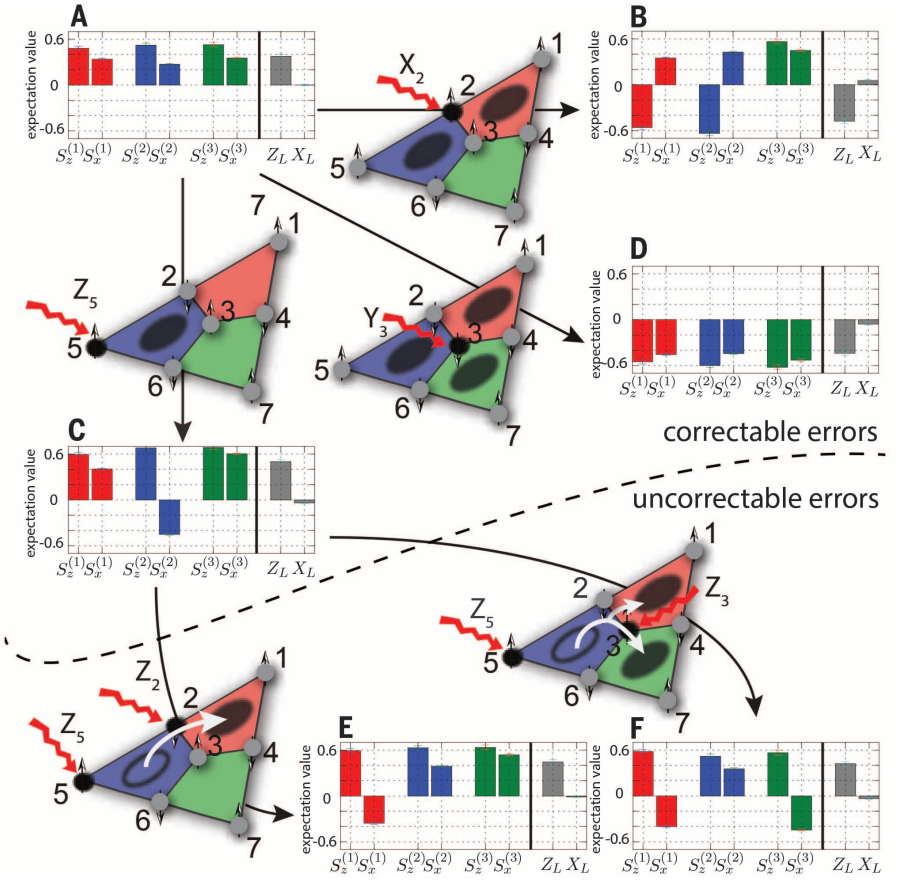


Fig. 2. Effect of arbitrary single-qubit errors on the encoded logical qubit. (**A**) The initial logical state $|0\rangle_L$, prior to the occurrence of single-qubit errors, is reflected (i) by the error syndrome, in which all six $S_X^{(j)}$ and $S_Z^{(j)}$ stabilizers are positive-valued, and (ii) by a positive (vanishing) expectation value of the logical operator Z_L (X_L). (**B**) A bit flip error (red wiggled arrow) on qubit 2 (marked in black) affects the blue and red plaquettes (visualized by gray-shaded circles) and manifests itself by negative $S_Z^{(1)}$ and $S_Z^{(2)}$ expectation values and a Z_L sign flip. (**C**) A Z_5 phase flip error only affects the blue plaquette and results in a sign flip of $S_X^{(2)}$. (**D**) A Y_3 error—equivalent to a combined X_3 and Z_3 error—affects all three plaquettes and induces a sign change in all six stabilizers and Z_L. Double-error events, such as a Z_5 phase flip (Fig. 2C), followed by a Z_2 (**E**) or a Z_3 error (**F**) result in an incorrect assignment of physical errors, as the detected stabilizer patterns are indistinguishable from single-error syndromes—here, the ones induced by a Z_1 (Fig. 2E) or a Z_4 (Fig. 2F) error (*26*). In the correction process, this eventually results in a logical error—here, a Z_L phase flip error. Stabilizer violations can under subsequent errors hop (white non-wiggled arrow) to an adjacent plaquette, as in Fig. 2E, where the violation disappears (open gray circle) from the blue and reappears on the red plaquette. Alternatively (Fig. 2F), they can disappear (from the blue plaquette), split up (white branched arrow), and reappear on two neighboring plaquettes (red and green). This rich dynamical behavior of stabilizer violations is a characteristic signature of the topological order in color codes (*16*, *26*).

state $|1010101\rangle$, which satisfies the required three $\langle S_z^{(i)}\rangle = +1$ and $\langle Z_L\rangle = +1$ conditions. In three subsequent steps, plaquette-wise entangling operations are applied to also satisfy the three $\langle S_x^{(i)}\rangle = +1$ constraints. For each step, three of the seven physical qubits are spectroscopically decoupled prior to the application of the collective entangling gate. Subsequently, Greenberger-Horne-Zeilinger–like entanglement between the four qubits belonging to one plaquette of the code is created with a fidelity of 88.8(5)% (where the value in parentheses is the error in the last significant digit). The entire encoding sequence involves three collective entangling gates and 108 local single-qubit rotations (*26*). The quantum state fidelity of 32.7(8)% of the system in state $|0\rangle_L$ surpasses the threshold value of 25% (by more than 9 standard deviations), above which genuine six-qubit entanglement is witnessed, thereby clearly indicating the mutual entanglement of all three plaquettes of the code (*26*).

Two factors govern the quality of the created logical state $|0\rangle_L$ (Fig. 2A): (i) its overlap with the code space, and (ii) its accordance within the code space with the target encoded state, which is related to the expectation value of the logical operator Z_L. Residual populations outside the code space are indicated by deviations of the six stabilizer expectation values, in our case on average 0.48(2), from the ideal value of +1. A more detailed analysis shows that within the code space, the fidelity between experimental and target state is as high as 95(2)%, whereas the expectation value of $\langle Z_L\rangle = 0.38(3)$ and the overall fidelity between the experimentally realized and the ideal state $|0\rangle_L$ are currently limited by the overlap with the code space of 34(1)% (*26*).

It is a hallmark feature of topologically ordered states that these cannot be characterized by local order parameters, but only reveal their topological quantum order in global system properties (*8*, *26*). We experimentally confirmed this intriguing characteristic for the topologically encoded seven-qubit system in state $|1\rangle_L$ by measuring all subsets of reduced two-qubit density matrices, which yielded an average Uhlmann fidelity of 98.3(2)% with the two-qubit completely mixed state, showing the absence of any single- and two-qubit correlations. In contrast, the presence of global quantum order is signaled for the system size at hand by nonvanishing three-qubit correlations $\langle Z_1 Z_4 Z_7\rangle = -0.46(6)$ (*26*).

We next consider the error correction properties of the encoded qubit. Single-qubit errors lead the system out of the logical code space and manifest themselves as stabilizer eigenstates of eigenvalue −1 (stabilizer violations) associated to one or several plaquettes. Using the available gate set, we coherently induce all single-qubit errors on the encoded state $|0\rangle_L$ and record the induced error syndromes provided by the characteristic pattern of six stabilizer expectation values (*26*) (see Fig. 2 for a selection). The experimental data reveal the CSS character of the quantum code: Starting in $|0\rangle_L$ (Fig. 2A), single-qubit X

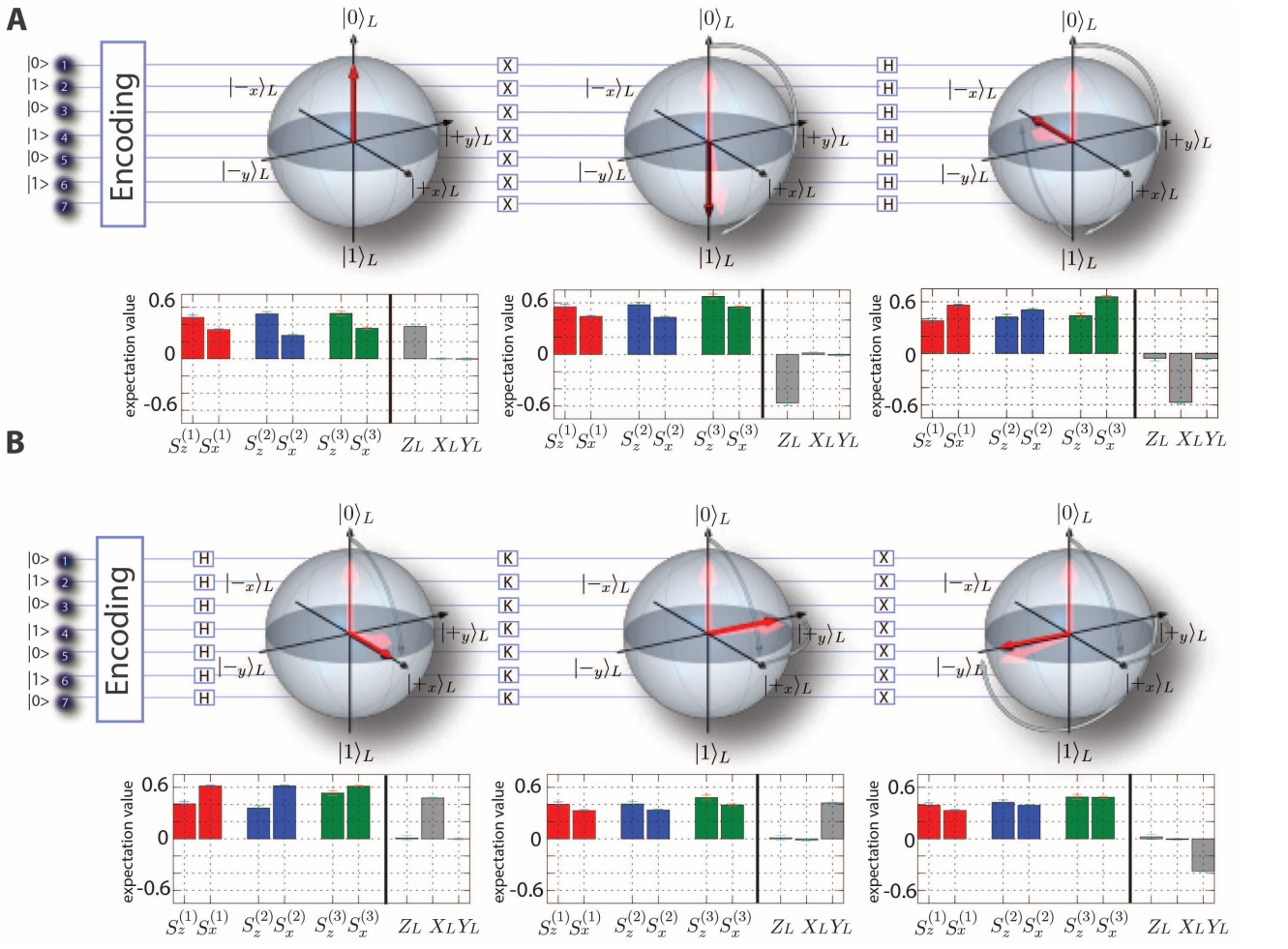


Fig. 3. Single-qubit Clifford gate operations applied on a logical encoded qubit. Starting from the logical $|0\rangle_L$ state, sequences of logical Clifford gate operations $\{X_L, H_L\}$ in (**A**) and $\{H_L, K_L, X_L\}$ in (**B**) are applied consecutively in a transversal way (i.e., bit-wise) to realize all six cardinal states $\{|0\rangle_L, |1\rangle_L, |-x\rangle_L, |+x\rangle_L, |+y\rangle_L, |-y\rangle_L\}$ of the logical space of the topologically encoded qubit. The dynamics under the applied gate operations is illustrated by rotations of the Bloch vector (red arrow) on the logical Bloch sphere as well as by the circuit diagram in the background. Each of the created logical states is characterized by the measured pattern of $S_x^{(j)}$ and $S_z^{(j)}$ stabilizers and the logical Bloch vector, with the three components given by the expectation values of the logical operators X_L, Y_L, and Z_L. The orientation of the logical Bloch vector changes as expected under the logical gate operations.

errors manifest themselves as violations of Z-type stabilizers only, whereas single-qubit Z errors manifest themselves as violations of X-type stabilizers only (Fig. 2, B and C); the effect of single-qubit Y errors is equivalent to a combined X and Z error and is signaled by the simultaneous violation of the corresponding X- and Z-type stabilizers (Fig. 2D). For our experimental statistical uncertainties, the measured characteristic error syndromes can be perfectly assigned to the underlying induced single-qubit errors (*26*). Shown in Fig. 2, E and F, are data in which the code has been exposed to two single-qubit errors, whose correction exceeds the capabilities of the seven-qubit code (*5*, *16*).

In topological color codes, quantum information is processed by logical gate operations acting directly within the code space (*16*, *17*). The entire group of logical Clifford gates generated by the elementary gate operations Z, X, the Hadamard H, and the phase gate K,

$$H = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \quad K = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix} \qquad (2)$$

as well as the C-NOT gate for registers containing several logical qubits, can be realized transversally. We implement the logical Clifford gates $Z_L = Z_1Z_2Z_3Z_4Z_5Z_6Z_7$, $X_L = X_1X_2X_3X_4X_5X_6X_7$, $H_L = H_1H_2H_3H_4H_5H_6H_7$, and $K_L = K_1K_2K_3K_4K_5K_6K_7$ by local operations that are realized by a combination of single-ion and collective rotations (*26*) (Fig. 1B). After initializing the logical qubit in the state $|0\rangle_L$, we prepare all six eigenstates of the logical operators X_L, Y_L, and Z_L, which requires quantum circuits consisting of up to three elementary logical Clifford gate operations (Fig. 3, A and B). The encoded qubit evolves as expected under the logical gate operations, as indicated by the characteristic changes in the pattern of X_L, Y_L, and Z_L, expectation values. This result is corroborated by quantum state fidelities within the code space of {95(2), 85(3), 87(2)}% of the experimental logical states $\{|0\rangle_L, |1\rangle_L, |+_x\rangle_L = (|0\rangle_L + |1\rangle_L)/\sqrt{2}\}$ with the expected ideal states, which indicates a high performance of the transversal (i.e., bitwise) logical Clifford gate operations (*26*).

We further explore the computational capabilities of the encoded qubit by executing a longer encoded quantum computation, which consists of up to 10 logical X_L gate operations applied to the system initially prepared in $|-_y\rangle_L$ (Fig. 4A). Here, the logical qubit flips as expected between the Y_L eigenstates, as witnessed by alternating (vanishing) expectation values of the logical Y_L (X_L, Z_L) operator (Fig. 4B), accompanied by a moderate decay of the average stabilizer expectation values of only 3.8(5)% per logical gate operation, which is in quantitative agreement with the fidelities of our single-ion and collective local operations (*26*). Because the duration of 10 X_L gates is more than one order of magnitude shorter than the measured coherence time of encoded logical states (*26*), the performance of the encoded quantum computation of Fig. 4B is currently not limited by the lifetime of the logical states but is predominantly determined by imperfections of the logical Clifford gate operations.

Future objectives aiming at an extension of the demonstrated quantum information processing capabilities include the use of one additional (unprotected) ancillary qubit for the implementation of a non-Clifford gate, toward universal encoded quantum computation (*1*, *18*). Furthermore, for higher experimental fidelities, repetitive application of complete error correction cycles is achievable (*26*) by incorporating previously demonstrated measurement and feedback techniques (*28*) for quantum nondemolition readout of stabilizers via ancillary qubits; this would represent progress toward the goal of keeping an encoded qubit alive. The demonstrated concepts can be adapted to 2D ion-trap arrays (*29*) as well as other scalable architectures including optical, atomic, molecular, and solid-state systems (*2*). The continuing technological development of these platforms promises fault-tolerant operating of larger numbers of error-resistant, topologically protected qubits, and thereby marks a clear, although technologically challenging, path toward the realization of FTQC.

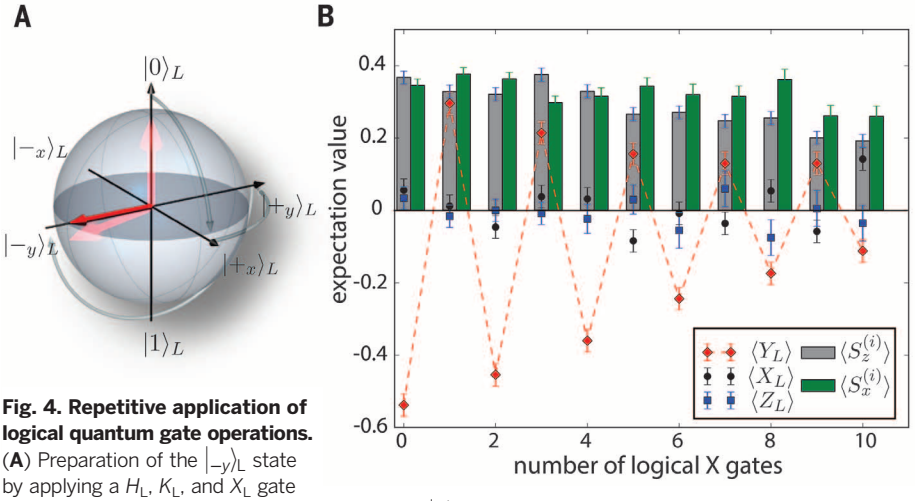


Fig. 4. Repetitive application of logical quantum gate operations. (**A**) Preparation of the $|-_y\rangle_L$ state by applying a H_L, K_L, and X_L gate operation on the qubit initially prepared in the $|0\rangle_L$ state. (**B**) Subsequently, flips between the logical $|+_y\rangle_L$ and $|-_y\rangle$L states are induced by consecutively applying logical X_L gate operations up to 10 times. The sign flip of the Y_L expectation value (red diamonds) after each step clearly signals the induced flips of the logical Bloch vector, whereas the expectation values of Z_L (blue squares) and X_L (black circles) are close to zero, as expected [the average of $\{\langle Z_L\rangle,\langle X_L\rangle\}$ yields {0.01(1), −0.01(1)}] (*26*). Average $S_z^{(i)}$ stabilizer expectation values after each X_L gate are shown as gray bars; average $S_x^{(i)}$ stabilizer expectation values after each X_L gate are shown as green bars.

REFERENCES AND NOTES

1. M. A. Nielsen, I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge Univ. Press, Cambridge, 2000).
2. T. D. Ladd *et al.*, *Nature* **464**, 45–53 (2010).
3. P. W. Shor, *Phys. Rev. A* **52**, R2493–R2496 (1995).
4. A. R. Calderbank, P. W. Shor, *Phys. Rev. A* **54**, 1098–1105 (1996).
5. A. M. Steane, *Phys. Rev. Lett.* **77**, 793–797 (1996).
6. P. W. Shor, Fault-tolerant quantum computation, *IEEE Symposium on Foundations of Computer Science* (1996); http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=548464.
7. J. Preskill, in *Introduction to Quantum Computation and Information*, H.-K. Lo, T. Spiller, S. Popescu, Eds. (World Scientific, Singapore, 1998), pp. 213–269.
8. A. Y. Kitaev, *Ann. Phys.* **303**, 2–30 (2003).
9. E. Dennis, A. Y. Kitaev, A. Landahl, J. Preskill, *J. Math. Phys.* **43**, 4452 (2002).
10. R. Raussendorf, J. Harrington, *Phys. Rev. Lett.* **98**, 190504 (2007).
11. H. G. Katzgraber, H. Bombin, M. A. Martin-Delgado, *Phys. Rev. Lett.* **103**, 090501 (2009).
12. D. S. Wang, A. G. Fowler, C. D. Hill, L. C. L. Hollenberg, *Quant. Info. Comp.* **10**, 780 (2010).
13. J. Benhelm, G. Kirchmair, C. F. Roos, R. Blatt, *Nat. Phys.* **4**, 463–466 (2008).
14. T. P. Harty *et al.*, arXiv:1403.1524 (2014).
15. R. Barends *et al.*, *Nature* **508**, 500–503 (2014).
16. H. Bombin, M. A. Martin-Delgado, *Phys. Rev. Lett.* **97**, 180501 (2006).
17. H. Bombin, M. A. Martin-Delgado, *Phys. Rev. Lett.* **98**, 160502 (2007).
18. S. Bravyi, A. Y. Kitaev, *Phys. Rev. A* **71**, 022316 (2005).
19. J. Chiaverini *et al.*, *Nature* **432**, 602–605 (2004).
20. P. Schindler *et al.*, *Science* **332**, 1059–1061 (2011).
21. M. D. Reed *et al.*, *Nature* **482**, 382–385 (2012).
22. E. Knill, R. Laflamme, R. Martinez, C. Negrevergne, *Phys. Rev. Lett.* **86**, 5811–5814 (2001).
23. J. Zhang, R. Laflamme, D. Suter, *Phys. Rev. Lett.* **109**, 100503 (2012).
24. X. C. Yao *et al.*, *Nature* **482**, 489–494 (2012).
25. D. Gottesman, *Phys. Rev. A* **54**, 1862–1868 (1996).
26. See supplementary materials on *Science* Online.
27. P. Schindler *et al.*, *New J. Phys.* **15**, 123012 (2013).
28. J. T. Barreiro *et al.*, *Nature* **470**, 486–491 (2011).
29. M. Kumph *et al.*, arXiv:1402.0791 (2014).

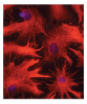
ACKNOWLEDGMENTS

Supported by Spanish MICINN grants FIS2009-10061 and FIS2012-33152; CAM research consortium QUITEMAD grant S2009-ESP-1594; European Commission PICC: FP7 2007-2013, grant 249958; integrated project SIQS (grant 600645); UCM-BS grant GICC-910758; the Austrian Science Fund (FWF) through the SFB FoQus (FWF project no. F4002-N16), and the Institut für Quanteninformation GmbH. This research was partially supported by the U.S. Army Research Office through grant W911NF-14-1-0103. This research was partially funded by the Office of the Director of National Intelligence (ODNI), Intelligence Advanced Research Projects Activity (IARPA), through Army Research Office grant W911NF-10-1-0284. All statements of fact, opinion, or conclusions contained herein are those of the authors and should not be construed as representing the official views or policies of IARPA, the ODNI, or the U.S. Government.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/345/6194/302/suppl/DC1
Supplementary Text
Figs. S1 to S7
Tables S1 and S2
Supplementary References

21 March 2014; accepted 28 May 2014
10.1126/science.1253742

EXTENDED PDF FORMAT
SPONSORED BY

Travel Grants Available
Neuroscience 2014
Application Deadline: July 31st
>Apply Now
R&D systems
a bio-techne brand
www.rndsystems.com

Two-particle quantum interference in tunnel-coupled optical tweezers

A. M. Kaufman *et al.*

Science **345**, 306 (2014);
DOI: 10.1126/science.1250057

This copy is for your personal, non-commercial use only.

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by clicking here.

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines here.

The following resources related to this article are available online at www.sciencemag.org (this information is current as of August 4, 2014):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:
http://www.sciencemag.org/content/345/6194/306.full.html

Supporting Online Material can be found at:
http://www.sciencemag.org/content/suppl/2014/06/25/science.1250057.DC1.html

A list of selected additional articles on the Science Web sites **related to this article** can be found at:
http://www.sciencemag.org/content/345/6194/306.full.html#related

This article **cites 35 articles**, 7 of which can be accessed free:
http://www.sciencemag.org/content/345/6194/306.full.html#ref-list-1

This article has been **cited by** 1 articles hosted by HighWire Press; see:
http://www.sciencemag.org/content/345/6194/306.full.html#related-urls

This article appears in the following **subject collections:**
Physics
http://www.sciencemag.org/cgi/collection/physics

Downloaded from www.sciencemag.org on August 4, 2014

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published weekly, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. Copyright 2014 by the American Association for the Advancement of Science; all rights reserved. The title *Science* is a registered trademark of AAAS.

QUANTUM MECHANICS

Two-particle quantum interference in tunnel-coupled optical tweezers

A. M. Kaufman,[1,2] B. J. Lester,[1,2] C. M. Reynolds,[1,2] M. L. Wall,[1,2] M. Foss-Feig,[3] K. R. A. Hazzard,[1,2] A. M. Rey,[1,2] C. A. Regal[1,2]*

The quantum statistics of atoms is typically observed in the behavior of an ensemble via macroscopic observables. However, quantum statistics modifies the behavior of even two particles. Here, we demonstrate near-complete control over all the internal and external degrees of freedom of two laser-cooled ^{87}Rb atoms trapped in two optical tweezers. This controllability allows us to observe signatures of indistinguishability via two-particle interference. Our work establishes laser-cooled atoms in optical tweezers as a promising route to bottom-up engineering of scalable, low-entropy quantum systems.

Quantum interference between possible detection paths for two indistinguishable particles yields information about quantum statistics and correlations (*1*, *2*). An example is the Hong-Ou-Mandel (HOM) effect, which reveals bosonic quantum statistics through a coalescence effect that causes two indistinguishable photons incident on different ports of a beam splitter to emerge on the same, yet random, output port (*3*). The HOM effect has been observed with photons (*3*–*5*) and in an analogous experiment with electrons (*6*).

Here, we observe two-particle interference akin to the HOM effect with independently prepared bosonic atoms in tunnel-coupled optical tweezers (Fig. 1A). We control the single-atom quantum state by laser cooling each atom to its motional ground state and then directly observe the effects of their quantum indistinguishability. Whereas the role of quantum statistics in macroscopic ensembles of fermionic and bosonic atoms can be observed with Hanbury Brown and Twiss interference experiments (*7*–*13*), our system allows the study of nonclassical few-atom states with single-atom control.

Our results depend on the mobility of two wavelength-scale optical tweezers and single-site imaging, which are realized using the apparatus illustrated in Fig. 1B (*14*–*16*). For laser cooling ^{87}Rb atoms to the three-dimensional (3D) ground state (Fig. 1C) (*16*, *17*) and imaging in position-resolved potentials, our tweezers are positioned far apart compared to the focused spot radius of 710 nm. For tunneling, the tweezers are brought close together such that there is a small, tunable overlap of the single-particle wave functions. Our full experimental sequence consists of the following steps: We image the initial atom positions, laser cool with Raman sideband cooling, perform tunneling experiments, and then image the atoms again. Hence, we can follow the quantum dynamics between initial and final states that are both known with single-site resolution.

In Fig. 2, A to C, we study the single-atom tunneling dynamics by only considering experiments that, after stochastic loading (*14*), yield a single atom in the left or right well in the first image (fig. S1). After imaging and cooling, the atom is in the 3D motional ground state and the $|F = 2, m_F = 2\rangle \equiv |\uparrow\rangle$ spin state, where F and m_F are the total angular momentum quantum number and its projection along a quantization axis, respectively. The tweezer depths and spacing (Fig. 2A) are then decreased rapidly (slowly) with respect to the tunnel coupling (trap frequency) to prepare an initial state ideally localized in the left or right well (fig. S2). The tunnel coupling is given by $J = -\langle \mathrm{R}|H_{\mathrm{sp}}|L\rangle/h$, where $|L\rangle$ ($|R\rangle$) is the lowest-energy, localized state in the left (right) well, and H_{sp} is the single-particle Hamiltonian (*18*). We control the double-well energy bias by varying the relative intensity of each tweezer. On the tunneling resonance, an atom prepared in the left well undergoes the coherent dynamics $|L\rangle \rightarrow \cos(Jt)|L\rangle + i\sin(Jt)|R\rangle$ (*19*–*22*). After an evolution time t in the presence of tunneling, the depth of the traps is rapidly increased to freeze the atom distribution, the traps are pulled apart, and the single-atom location is imaged.

Figure 2B demonstrates resonant coherent tunneling as measured by recording the likelihood of observing the atom in the left well (P_{L}) as a function of time for an atom starting in the left well (blue) or the right well (red). The tunneling can be tuned by varying either the tweezer spacing or the overall potential depth (fig. S3C). In Fig. 2B, a fit to the data reveals $J/2\pi = 262(4)$ Hz; Fig. 2C shows data in which J is increased to 348(4) Hz. One contribution to the subunity oscillation contrast is atom loss due to background collisions; in the duration of our experiments, the loss probability (P_{loss}) ranges from 0.03 to 0.05 and is known precisely for each experiment (fig. S1, table S1, and Fig. 2, B and C, gray regions). We also observe damping of the tunneling oscillations and finite initial contrast that is not accounted for by particle loss ($\tau \approx 10$ ms for $J/2\pi = 262$ Hz) (fig. S3B). These effects are most likely due to experimental fluctuations of the double-well bias, and we experimentally find that the contrast and damping improve with increasing J (*18*).

We now consider the theoretical expectation for an equivalent dynamical experiment starting with two particles, one per well. For perfect cooling and spin preparation of the isolated atoms, the particles will be indistinguishable—that is, all degrees of freedom besides their position (left or right) will have been made the same. We know there is a particle in the left well and there is a particle in the right well, but we cannot associate any additional label to the particles. The bosonic atoms will then, necessarily, occupy the spatially symmetric state $|S\rangle \equiv \frac{1}{\sqrt{2}}(|L\rangle_1|R\rangle_2 + |R\rangle_1|L\rangle_2)$, where the ket subscript is a particle label. For poor cooling or spin preparation, the atoms can be distinguished by a degree of freedom other than their position; hence, the atoms can antisymmetrize in the additional degree of freedom and, in turn, have a projection onto the antisymmetric spatial state $|A\rangle \equiv \frac{1}{\sqrt{2}}(|L\rangle_1|R\rangle_2 - |R\rangle_1|L\rangle_2)$. The bosonic state can then be written as a mixture of the states $|\psi_\pm\rangle = \frac{1}{\sqrt{2}}(|S\rangle|\chi_+\rangle \pm |A\rangle|\chi_-\rangle)$, where $|\chi_\pm\rangle = \frac{1}{\sqrt{2}}(|\chi\rangle_1|\overline{\chi}\rangle_2 \pm |\overline{\chi}\rangle_1|\chi\rangle_2)$ and $\{\chi, \overline{\chi}\}$ describe the other degree of freedom such as motional state $\{n, n'\}$ or spin $\{\uparrow, \downarrow\}$. Two atoms in either of the $|\psi_\pm\rangle$ states are distinguishable because the additional degree of freedom $\{\chi, \overline{\chi}\}$ is uniquely correlated with the atoms' positions: For the $|\psi_+\rangle$ ($|\psi_-\rangle$) state, the atom on the left is in state $|\chi\rangle$ ($|\overline{\chi}\rangle$) and the atom on the right is in state $|\overline{\chi}\rangle$ ($|\chi\rangle$). The ability to measure indistinguishability arises from the different dynamics exhibited in the symmetric and antisymmetric cases. The symmetric spatial state dynamically evolves as $|S\rangle \rightarrow |S\rangle\cos(2Jt) + \frac{i}{\sqrt{2}}(|L\rangle_1|L\rangle_2 + |R\rangle_1|R\rangle_2)\sin(2Jt)$. The antisymmetric state $|A\rangle$ undergoes destructive interference that prevents the two atoms from being in the same well, and hence displays no tunneling dynamics.

In Fig. 2D, we show the ideal dynamics for the distinguishable (purple), symmetric (black), and antisymmetric (brown) cases. We consider the observable $P_{11}(t)$, which is the likelihood to measure the atoms in separate wells as a function of t and is analogous to looking at coincidence counts on a pair of photon detectors. The symmetric $|S\rangle$ state yields unity contrast oscillations of $P_{11}(t)$, whereas $|A\rangle$ yields a time-independent $P_{11}(t) = 1$. Hence, the distinguishable states $|\psi_\pm\rangle$ result in a $P_{11}(t)$ that is the average of these dynamics and, as such, does not attain a value below 0.5. The tunneling between the wells yields an effective atom beam splitter where t varies the reflection and transmission coefficients. Of particular importance are the times $t_b \equiv 2\pi/8J$ and odd-integer multiples thereof (green dashed lines). At $t = t_b$, the tunneling splits each atom equally between the wells and thereby realizes a balanced atom beam splitter according to the transformations $|L\rangle \rightarrow \frac{1}{\sqrt{2}}(|L\rangle + i|R\rangle)$ and $|R\rangle \rightarrow \frac{1}{\sqrt{2}}(|R\rangle + i|L\rangle)$. Accordingly, for the $|S\rangle$ state, $P_{11}(t_b)$ vanishes in analogy to the indistinguishable photons incident on separate ports of a balanced beam splitter in the original HOM experiment (*3*). However, distinguishable atoms, like distinguishable photons, will yield $P_{11} = 0.5$ when equally beam split.

[1]JILA, National Institute of Standards and Technology and University of Colorado, Boulder, CO 80309, USA. [2]Department of Physics, University of Colorado, Boulder, CO 80309, USA. [3]Joint Quantum Institute and the National Institute of Standards and Technology, Gaithersburg, MD 20899, USA.
*Corresponding author. E-mail: regal@colorado.edu

In Fig. 2E, we experimentally investigate the population dynamics observed with two particles. We plot $P_{11}(t)$ for cases in which the stochastic loading results in two atoms, one in each well (black squares); these points are taken in the same experimental sequence as the single-particle data in Fig. 2B. In our atom detection protocol, we image scattered light from the two well-separated traps onto a charge-coupled device (CCD) array. During the 25- to 50-ms imaging time, the atoms are cooled by polarization gradient cooling, during which light-assisted atomic collisions result in a signal corresponding to either zero or one atom (*14*, *22*, *23*). P_{11} is determined by the distinct signature in which the image indicates one atom in each well. If the experiment yields two atoms in one well, P_{20} or P_{02}, this is manifest by final images that yield zero atoms, or in some cases one atom in a single well (fig. S1) (*18*). To accurately interpret P_{11}, we take into account signal depletion caused by the single-particle loss described earlier (P_{loss}). This effect reduces the maximum value that can be achieved by the measured P_{11} to $(1 - P_{\mathrm{loss}})^2$.

In analyzing the results of the two-particle dynamics, our goal is to compare $P_{11}(t)$ from our two-particle measurement to that of a theoretical expectation for uncorrelated, distinguishable atoms, which we refer to as $P_{\mathrm{dist}}(t)$. P_{dist} at any time can be calculated directly from corresponding single-particle data via $P_{\mathrm{dist}} = P_{\mathrm{L}}^1 P_{\mathrm{R}}^2 + P_{\mathrm{R}}^1 P_{\mathrm{L}}^2$ (purple circles in Fig. 2E) (*18*). Here, $P_{\mathrm{L}}^{1(2)}$ corresponds to measuring an atom in the left well when an atom starts in the left (right) well, i.e., the blue (red) data of Fig. 2B, and $P_{\mathrm{R}}^{1(2)}$ is the corresponding information for measuring an atom in the right well. A calculation of P_{dist} directly from the single-particle points inherently contains both loss and finite single-particle contrast. For example, $P_{\mathrm{dist}}(t_b)$ reaches a minimum value consistent with $(1 - P_{\mathrm{loss}})^2/2 \approx 0.5 - P_{\mathrm{loss}}$, and the amplitude of $P_{\mathrm{dist}}(t)$ is consistent with the expectation of one-half the product of the single-particle contrasts (*18*). We can compare the amplitude of oscillation for the distinguishable expectation (purple circles) to our two-particle measurement (black squares). We find that these values differ by 6σ (*18*): $A_{P_{\mathrm{dist}}} = 0.282(12)$ and $A_{P_{11}} = 0.46(2)$.

A full treatment of the observed $P_{11}(t)$ must also consider potential effects of interactions between the atoms. In many experiments with atoms in optical lattices, the on-site interaction energy U is the dominant scale (*19*, *24*); however, we intentionally operate in a regime where U is smaller than J. For the data shown in Fig. 2E, $U = 0.44(4)J$ (*18*). In Fig. 2F, we demonstrate two-particle oscillations for experimental conditions of an even smaller relative interaction $U = 0.22(2)J$, with measurements $A_{P_{11}} = 0.48(2)$ and $A_{P_{\mathrm{dist}}} = 0.306(18)$. The similarity of these results to those in Fig. 2E suggests that interactions are not a relevant scale in either experiment. A careful theoretical analysis also demonstrates that existing interactions between distinguishable atoms could not mimic the observed signal of two-particle interference (*18*) (figs. S5 and S6).

We now study the behavior at the balanced tunneling point t_b, where, in analogy to HOM experiments, the two-particle interference is most readily observed. Using multiple experimental knobs, we can vary the indistinguishability of the atoms and observe a variation in $P_{11}(t_b) = P_b$. We start by studying the dependence of P_b on the relative spin state of the two atoms using two distinct methods. In the first method, after cooling the atoms, we apply a variable-length microwave pulse that couples the $|\uparrow\rangle$ and $|\downarrow\rangle$ spin states in only the right well (fig. S4). This is accomplished by shifting the transition in the left well out of resonance using a circularly polarized, tightly focused laser spot (*18*). Upon π rotation of the right spin, the atoms become distinguishable, and we expect the two-particle interference dip of P_{11} below P_{dist} to disappear. The observed dependence on the microwave pulse area is shown in Fig. 3A; for comparison, we show that P_{dist} (purple circles), calculated from the single-particle measurements, remains constant and near $(1 - P_{\mathrm{loss}})^2/2$. We study multiple spin rotations to show that the interference is recovered after a 2π rotation. We find that the frequency of oscillation is 32.6(6)kHz, which is in agreement with the measured microwave Rabi frequency of 32.05(18)kHz. The displayed fit to the data determines $P_b^{\mathrm{min}} = 0.314(14)$, and the amplitude of the variation in P_b is 0.15(2). Taking into account the spin rotation fidelity, we expect an amplitude of $0.84((1-P_{\mathrm{loss}})^2/2 - P_b^{\mathrm{min}}) = 0.130(13)$ (*18*), which is consistent with the measured value. The agreement further substantiates the irrelevance of interactions, which are nearly constant as the two-particle spin state is varied (*25*), and hence not responsible for the observed variation in P_b.

In the second spin study, we couple the $|\uparrow\rangle$ and $|\downarrow\rangle$ spin states of atoms in both wells using a pair of Raman beams. This global rotation avoids any systematic effects that might be introduced by single-site addressing. During the time (25 ms) between the Raman pulse and the tunneling, the atoms lose their spin coherence, and hence the spin state of each atom is in an incoherent mixture. Ideally, odd-integer multiples of a π/2-pulse yield an equal mixture of all possible two-atom

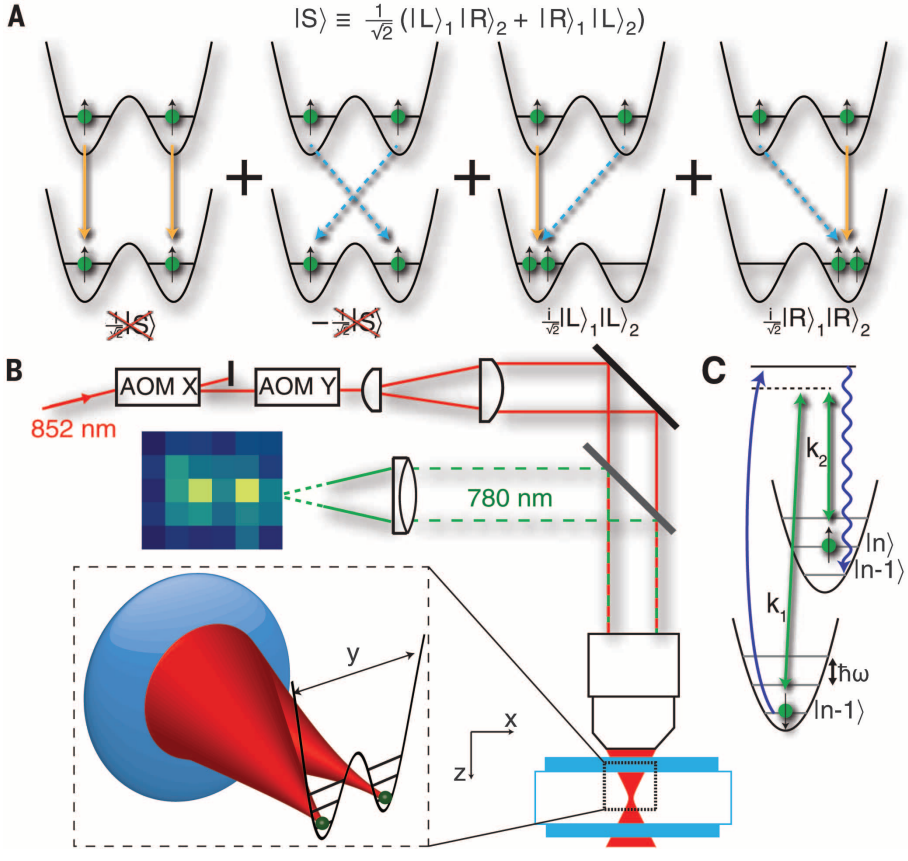


Fig. 1. Quantum interference of two particles. (**A**) The initial two-particle state is a ground state spin-up atom in each optical tweezer, denoted by $|S\rangle \equiv \frac{1}{\sqrt{2}}(|L\rangle_1|R\rangle_2 + |R\rangle_1|L\rangle_2)$, where the ket subscript is a particle label, and the single-particle states $|L\rangle$ ($|R\rangle$) correspond to an atom localized in the left (right) well. For a tunneling time that realizes a balanced atom beam splitter, destructive interference yields the state $\frac{i}{\sqrt{2}}(|L\rangle_1|L\rangle_2 + |R\rangle_1|R\rangle_2)$. (**B**) Experimental setup. The apparatus for realizing tunneling between optical tweezers uses high numerical aperture optics combined with control of the tweezers' positions and depths with acousto-optic modulators (AOMs). The same objective that creates the focused tweezer potentials also collects 780 nm fluorescence from the optically trapped atoms. (**C**) The sideband cooling is accomplished with lasers driving coherent (green) and spontaneous (blue) Raman transitions that couple to the atomic motion and spin states $|F = 1, m_F = 1\rangle \equiv |\downarrow\rangle$ and $|F = 2, m_F = 2\rangle \equiv |\uparrow\rangle$.

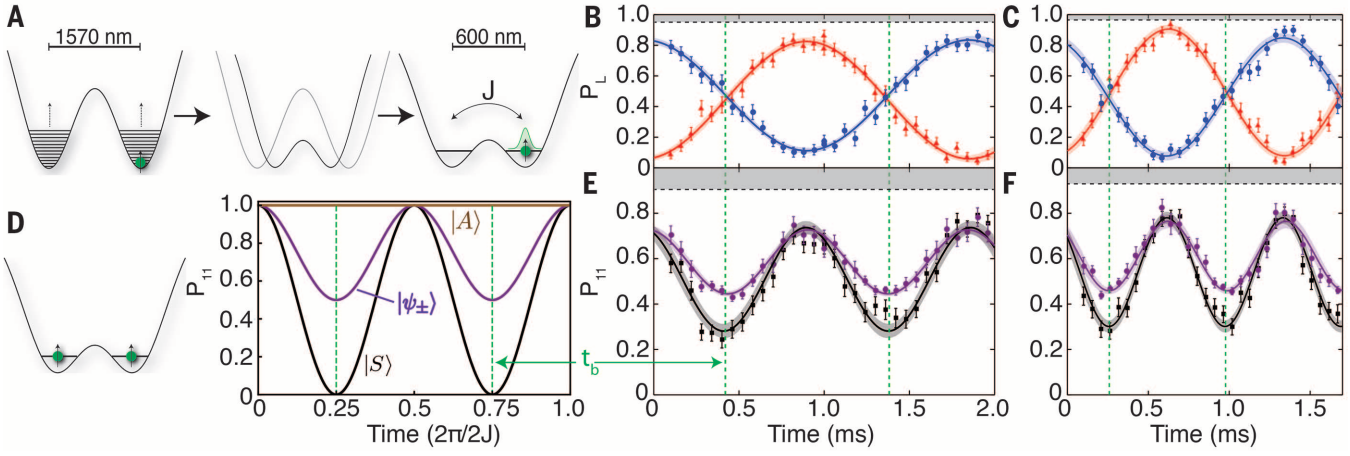


Fig. 2. Single- and two-particle tunneling. (**A**) While the tweezers are 23(1) MHz deep, the atom is imaged, cooled and optically pumped to $|\uparrow\rangle$. For tunneling experiments, the tweezers are swept together such that the two Gaussian functions (double-well minima) are centered ≈800 nm (≈600 nm) apart, and the total trap intensity is decreased to a single-well depth of either 96 kHz or 60 kHz. (**B**) Resonant tunneling oscillations at 2*J* for an 808-nm Gaussian function spacing and a 96-kHz depth. Blue circles (red triangles) are the expectation value $P_L^{1(2)}$ for finding an atom in the left well given an initial single atom in the left (right) well. The gray shaded region indicates atom loss P_{loss}. (**C**) Same as B except with an 805-nm Gaussian function spacing and a depth of 60 kHz. (**D**) Ideal $P_{11}(t)$ for two-particle tunneling dynamics initiated at $t = 0$ and in the symmetric state $|S\rangle$, the distinguishable states $|\psi_\pm\rangle$, and the antisymmetric state $|A\rangle$. The dashed green lines correspond to t_b. (**E**) Measured two-particle dynamics during the same experimental sequence as in (B). Likelihood to measure exactly one atom in each well (P_{11}) for the initial condition of an atom prepared in each well (black squares). Distinguishable expectation P_{dist} calculated from the single-particle data in (B) (purple circles). The gray shaded region above the dashed black line indicates the expected reduction from atom loss. (**F**) Same as (E) except with a larger value of J/U using the double-well parameters of (C). In all plots, the shaded regions are the 95% confidence interval for a sinusoidal fit. The error bars are the standard error in the measurement; each black (red or blue) data point is the mean of ≈140 (100) atom measurements.

spin states, and hence the two-particle interference dip should be reduced by 1/2. As a function of Raman pulse area (Fig. 3B), the frequency of oscillation is 65.5±1.2 kHz, which, as expected, is twice the measured Raman Rabi frequency of 32.3(3) kHz. The displayed fit to the data determines $P_b^{\min} = 0.296(10)$, and the amplitude of the variation in P_b is 0.085(15). Taking into account the relative spin rotation fidelity of 0.90(3), we expect an amplitude of $(0.90^2/2)[(1 - P_{\text{loss}})^2/2 - P_b^{\min}] = 0.069(6)$, which is consistent with the measured amplitude.

Last, we study the dependence of P_b on the motional state of the atoms in Fig. 3C. During the last stage of cooling, we vary the frequency δ_{Cool} of one Raman beam that controls the cooling along the weak axis (z) of both tweezer wells. For a separable potential, motional excitation along this axis would leave the single-particle tunneling unaffected. For our nonseparable tweezer potential, we expect and observe some variation in the tunneling (*18*), but near t_b the single-particle tunneling still results in a relatively constant P_{dist} (purple circles), which is consistent with the distinguishable expectation $(1 - P_{\text{loss}})^2/2 = 0.4660(14)$. For the two-particle measurements, at the primary sideband cooling resonance ($\delta_{\text{Cool}} = 0$), we observe a dip to $P_b^{\min} = 0.28(2)$, which is a value below the distinguishable expectation.

In all of our measurements, the value of $P_b^{\min}$ is observed to be finite, and it is useful to consider the origin of imperfections that could lead to the observed value. If residual atom temperature along the weak z axis were the only contributing factor, the central value of our single-atom ground-state fraction [measured by sideband spectroscopy to be 85% (*16*, *18*)] would correspond to $P_b^{\min} = 0.12$. Hence, it is likely that the finite $P_b^{\min}$ is further enlarged by technical fluctuations similar to those that lead to our finite single-particle contrast.

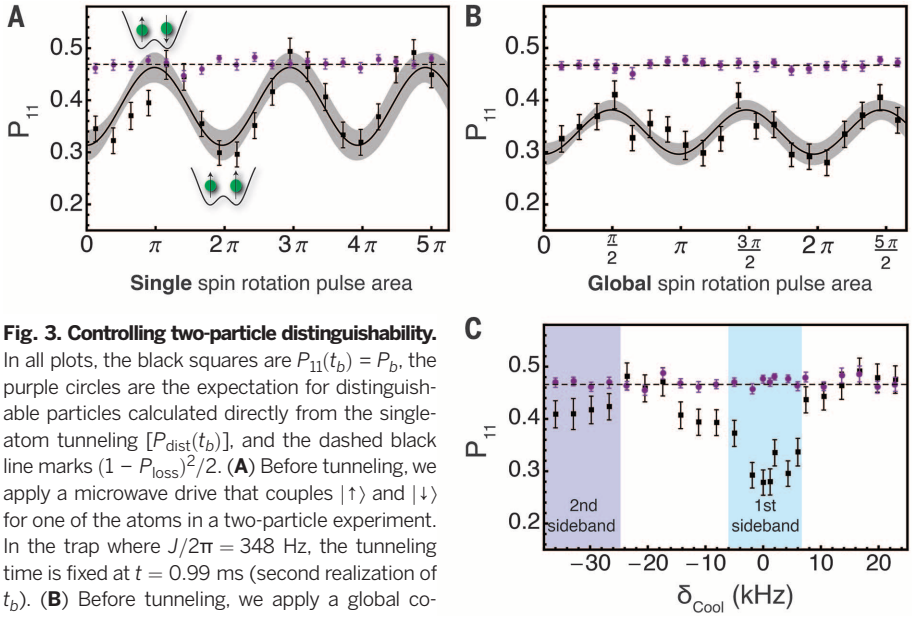


Fig. 3. Controlling two-particle distinguishability. In all plots, the black squares are $P_{11}(t_b) = P_b$, the purple circles are the expectation for distinguishable particles calculated directly from the single-atom tunneling [$P_{\text{dist}}(t_b)$], and the dashed black line marks $(1 - P_{\text{loss}})^2/2$. (**A**) Before tunneling, we apply a microwave drive that couples $|\uparrow\rangle$ and $|\downarrow\rangle$ for one of the atoms in a two-particle experiment. In the trap where $J/2\pi = 348$ Hz, the tunneling time is fixed at $t = 0.99$ ms (second realization of t_b). (**B**) Before tunneling, we apply a global coherent drive of varied pulse area to couple $|\uparrow\rangle$ and $|\downarrow\rangle$ and then allow for decoherence. In the trap where $J/2\pi = 262$ Hz, the tunneling time is fixed at $t = 0.45$ ms. In (A) and (B), the solid line and shaded band are sinusoidal fits and the associated 95% confidence interval. (**C**) We vary the detuning (δ_{Cool}) of the cooling beams of motion along the z axis. In the trap where $J/2\pi = 262$ Hz, the tunneling time is fixed at $t = 0.45$ ms. The two shaded regions correspond to frequency ranges of efficient (1st sideband) and less efficient (2nd sideband) cooling. For all plots, each black data point is the average of ≈360 measurements, and each set of measurements corresponding to a purple point is the average of ≈240 measurements. All error bars are the standard error in the measurement.

We have demonstrated an experimental system with which we achieve quantum control over the motion, position, and spin of single neutral atoms. Through a two-particle interference experiment,

we have shown that it is possible to create a low-entropy bosonic state by individually placing atoms in their motional ground state with laser cooling (*26*–*28*). Our results lay a foundation for linear quantum computing with atoms (*29*), interferometric highly sensitive force detection (*30*), control of neutral atoms in nanoscale optical devices (*31*, *32*), and quantum simulation with laser-cooled atoms in scalable optical tweezer arrays.

REFERENCES AND NOTES

1. R. J. Glauber, *Phys. Rev.* **130**, 2529–2539 (1963).
2. R. Hanbury Brown, R. Q. Twiss, *Nature* **177**, 27 (1956).
3. C. K. Hong, Z. Y. Ou, L. Mandel, *Phys. Rev. Lett.* **59**, 2044–2046 (1987).
4. J. Beugnon *et al.*, *Nature* **440**, 779–782 (2006).
5. C. Lang *et al.*, *Nat. Phys.* **9**, 345–348 (2013).
6. E. Bocquillon *et al.*, *Science* **339**, 1054–1057 (2013).
7. M. Yasuda, F. Shimizu, *Phys. Rev. Lett.* **77**, 3090–3093 (1996).
8. S. Fölling *et al.*, *Nature* **434**, 481–484 (2005).
9. M. Schellekens *et al.*, *Science* **310**, 648–651 (2005).
10. A. Öttl, S. Ritter, M. Köhl, T. Esslinger, *Phys. Rev. Lett.* **95**, 090404 (2005).
11. T. Rom *et al.*, *Nature* **444**, 733–736 (2006).
12. T. Jeltes *et al.*, *Nature* **445**, 402–405 (2007).
13. S. S. Hodgman, R. G. Dall, A. G. Manning, K. G. H. Baldwin, A. G. Truscott, *Science* **331**, 1046–1049 (2011).
14. N. Schlosser, G. Reymond, I. Protsenko, P. Grangier, *Nature* **411**, 1024–1027 (2001).
15. F. Serwane *et al.*, *Science* **332**, 336–338 (2011).
16. A. M. Kaufman, B. J. Lester, C. A. Regal, *Phys. Rev. X* **2**, 041014 (2012).
17. C. Monroe *et al.*, *Phys. Rev. Lett.* **75**, 4011–4014 (1995).
18. Materials and methods are available as supplementary materials on *Science* Online.
19. S. Fölling *et al.*, *Nature* **448**, 1029–1032 (2007).
20. M. Anderlini *et al.*, *Nature* **448**, 452–456 (2007).
21. J. Sebby-Strabley *et al.*, *Phys. Rev. Lett.* **98**, 200405 (2007).
22. C. Weitenberg *et al.*, *Nature* **471**, 319–324 (2011).
23. W. S. Bakr, J. I. Gillen, A. Peng, S. Fölling, M. Greiner, *Nature* **462**, 74–77 (2009).
24. R. Ma *et al.*, *Phys. Rev. Lett.* **107**, 095301 (2011).
25. M. Egorov *et al.*, *Phys. Rev. A* **87**, 053614 (2013).
26. V. Vuletić, C. Chin, A. J. Kerman, S. Chu, *Phys. Rev. Lett.* **81**, 5768–5771 (1998).
27. D. S. Weiss *et al.*, *Phys. Rev. A* **70**, 040302 (2004).
28. A. Steffen *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 9770–9774 (2012).
29. P. Kok, K. Nemoto, T. C. Ralph, J. P. Dowling, G. J. Milburn, *Rev. Mod. Phys.* **79**, 135–174 (2007).
30. V. Giovannetti, S. Lloyd, L. Maccone, *Science* **306**, 1330–1336 (2004).
31. E. Vetsch *et al.*, *Phys. Rev. Lett.* **104**, 203603 (2010).
32. J. D. Thompson *et al.*, *Science* **340**, 1202–1205 (2013).

ACKNOWLEDGMENTS

We thank T. P. Purdy for useful discussions. This work was supported by the David and Lucile Packard Foundation and the National Science Foundation (NSF) under grant no. 1125844. C.A.R. acknowledges support from the Clare Boothe Luce Foundation, A.M.K. and C.M.R. from the National Defense Science and Engineering Graduate Fellowships, and B.J.L. from the NSF Graduate Research Fellowship Program. K.R.A.H., M.L.W., and A.M.R. acknowledge funding from NSF-Physics at the Information Frontier, Army Research Office and Air Force Office of Scientific Research. K.R.A.H. and M.F.-F. acknowledge support from a National Research Council postdoctoral fellowship.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/345/6194/306/suppl/DC1
Materials and Methods
Figs. S1 to S6
Table S1
References (*33*–*36*)

22 December 2013; accepted 6 June 2014
Published online 26 June 2014;
10.1126/science.1250057

OPTICAL METROLOGY

Electro-optical frequency division and stable microwave synthesis

Jiang Li,[1]* Xu Yi,[1]* Hansuek Lee,[1] Scott A. Diddams,[2] Kerry J. Vahala[1]†

Optical frequency division by using frequency combs has revolutionized time keeping and the generation of stable microwave signals. We demonstrate optical frequency division and microwave generation by using a tunable electrical oscillator to create dual combs through phase modulation of two optical signals that have a stable difference frequency. Phase-locked control of the electrical oscillator by means of optical frequency division produces stable microwaves. Our approach transposes the oscillator and frequency reference of a conventional microwave frequency synthesizer. In this way, the oscillator experiences large phase noise reduction relative to the frequency reference. The electro-optical approach additionally relaxes the need for highly linear photodetection of the comb mode spacing. As well as simplicity, the technique is also tunable and scalable to higher division ratios.

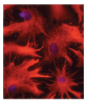
The photomixing of two highly coherent laser signals is a well-known approach to generate a stable radio frequency (RF) or microwave signal (*1*, *2*). Recently, however, a different approach to all-optical signal generation has been demonstrated that may revolutionize applications that require high-stability microwaves. Rather than photomixing stabilized laser signals to directly produce a microwave signal, the approach uses an octave-spanning, self-referenced frequency comb to divide a stable optical reference frequency down to microwave or RF rates (*3*). Frequency dividers are widely used in electronics to generate new frequencies from a single base oscillator or to coherently link different frequency bands. As an ancillary benefit, all frequency dividers reduce the phase noise spectral density of the output signal relative to the input by the square of the division ratio. The new optical frequency dividers perform division by a factor of ~50,000 (the ratio of optical to microwave frequencies), so that phase-noise reduction is greater than 10^9. Moreover, reference-cavity stabilized lasers exhibit a superior, fractional frequency stability in comparison with electrical oscillators (*4*, *5*). Optical dividers applied to such signals thereby generate microwave signals with an exceedingly low phase noise level (*3*).

We present a way to generate high-performance microwave signals through optical frequency division (OFD) by using a cascade of direct phase modulation and self-phase modulation to create an optical comb (*6*–*9*). Because the spectral line spacing is set by the electrical oscillator used to drive the phase modulators (as opposed to an optical resonator), the method of microwave synthesis has similarities to conventional microwave synthesizers while also leveraging the power of OFD so as to reduce phase noise.

In our approach, two laser lines having good relative frequency stability provide an optical reference for the microwave source (Fig. 1A). These laser lines are produced by Brillouin oscillation in a single high-quality-factor (Q) microcavity. However, the lines could also result from any stable optical references, including various types of dual-mode lasers (*10*–*15*), two lasers locked to distinct optical modes of a reference cavity (*16*), or lasers stabilized to atomic transitions (*17*). The laser lines enter the frequency divider portion of the signal generator, where they are phase-modulated by a pair of modulators at a frequency set by a voltage-controlled electrical oscillator (VCO). The sideband spectrum created by the phase modulators is further broadened through pulse-forming and self-phase modulation in an optical fiber (*8*, *18*, *19*). The comb of lines extending from each laser line results in a pair of sidebands near the midpoint of the frequency span. These are optically filtered and detected. The detected beat-note signal contains the phase noise of the VCO, but magnified by the optical division factor. It therefore provides a suitable error signal for phase-lock loop control of the VCO.

We contrast this microwave source that is based on electro-optical frequency division (EOFD) with a conventional microwave source that is based on electrical frequency division of a VCO (*20*). In the conventional approach (Fig. 1B), the VCO provides the highest frequency in the system. It is stabilized through electrical frequency division and phase comparison with a lower-frequency reference oscillator, such as a quartz oscillator. A consequence is that the stabilized VCO has a phase noise level that is always higher than the reference oscillator phase noise by the square of their frequency ratio (the frequency division ratio). In contrast, our optical version reverses the positions of the reference and the VCO in the frequency domain. Specifically, the reference is provided by the frequency difference of the laser lines, and this frequency difference is made much greater than the frequency of the VCO (in the present implementation, this is a nondetectable rate set at ~150 times the VCO frequency). Moreover, this reference frequency is divided down to the VCO frequency, as opposed to dividing the VCO frequency down to

[1]T. J. Watson Laboratory of Applied Physics, California Institute of Technology, Pasadena, CA 91125, USA. [2]Time and Frequency Division, National Institute of Standards and Technology, Boulder, CO 80305, USA.
*These authors contributed equally to this work. **†Corresponding author. E-mail: vahala@caltech.edu**

EXTENDED PDF FORMAT
SPONSORED BY

Travel Grants Available
Neuroscience 2014
Application Deadline: July 31st
>Apply Now
R&D systems
a bio-techne brand
www.rndsystems.com

Electro-optical frequency division and stable microwave synthesis

Jiang Li *et al.*

Science **345**, 309 (2014);
DOI: 10.1126/science.1252909

This copy is for your personal, non-commercial use only.

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by clicking here.

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines here.

The following resources related to this article are available online at www.sciencemag.org (this information is current as of August 4, 2014):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:
http://www.sciencemag.org/content/345/6194/309.full.html

Supporting Online Material can be found at:
http://www.sciencemag.org/content/suppl/2014/06/18/science.1252909.DC1.html

This article appears in the following **subject collections:**
Physics
http://www.sciencemag.org/cgi/collection/physics

Downloaded from www.sciencemag.org on August 4, 2014

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published weekly, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. Copyright 2014 by the American Association for the Advancement of Science; all rights reserved. The title *Science* is a registered trademark of AAAS.

we have shown that it is possible to create a low-entropy bosonic state by individually placing atoms in their motional ground state with laser cooling (*26*–*28*). Our results lay a foundation for linear quantum computing with atoms (*29*), interferometric highly sensitive force detection (*30*), control of neutral atoms in nanoscale optical devices (*31*, *32*), and quantum simulation with laser-cooled atoms in scalable optical tweezer arrays.

REFERENCES AND NOTES

1. R. J. Glauber, *Phys. Rev.* **130**, 2529–2539 (1963).
2. R. Hanbury Brown, R. Q. Twiss, *Nature* **177**, 27 (1956).
3. C. K. Hong, Z. Y. Ou, L. Mandel, *Phys. Rev. Lett.* **59**, 2044–2046 (1987).
4. J. Beugnon *et al.*, *Nature* **440**, 779–782 (2006).
5. C. Lang *et al.*, *Nat. Phys.* **9**, 345–348 (2013).
6. E. Bocquillon *et al.*, *Science* **339**, 1054–1057 (2013).
7. M. Yasuda, F. Shimizu, *Phys. Rev. Lett.* **77**, 3090–3093 (1996).
8. S. Fölling *et al.*, *Nature* **434**, 481–484 (2005).
9. M. Schellekens *et al.*, *Science* **310**, 648–651 (2005).
10. A. Öttl, S. Ritter, M. Köhl, T. Esslinger, *Phys. Rev. Lett.* **95**, 090404 (2005).
11. T. Rom *et al.*, *Nature* **444**, 733–736 (2006).
12. T. Jeltes *et al.*, *Nature* **445**, 402–405 (2007).
13. S. S. Hodgman, R. G. Dall, A. G. Manning, K. G. H. Baldwin, A. G. Truscott, *Science* **331**, 1046–1049 (2011).
14. N. Schlosser, G. Reymond, I. Protsenko, P. Grangier, *Nature* **411**, 1024–1027 (2001).
15. F. Serwane *et al.*, *Science* **332**, 336–338 (2011).
16. A. M. Kaufman, B. J. Lester, C. A. Regal, *Phys. Rev. X* **2**, 041014 (2012).
17. C. Monroe *et al.*, *Phys. Rev. Lett.* **75**, 4011–4014 (1995).
18. Materials and methods are available as supplementary materials on *Science* Online.
19. S. Fölling *et al.*, *Nature* **448**, 1029–1032 (2007).
20. M. Anderlini *et al.*, *Nature* **448**, 452–456 (2007).
21. J. Sebby-Strabley *et al.*, *Phys. Rev. Lett.* **98**, 200405 (2007).
22. C. Weitenberg *et al.*, *Nature* **471**, 319–324 (2011).
23. W. S. Bakr, J. I. Gillen, A. Peng, S. Fölling, M. Greiner, *Nature* **462**, 74–77 (2009).
24. R. Ma *et al.*, *Phys. Rev. Lett.* **107**, 095301 (2011).
25. M. Egorov *et al.*, *Phys. Rev. A* **87**, 053614 (2013).
26. V. Vuletić, C. Chin, A. J. Kerman, S. Chu, *Phys. Rev. Lett.* **81**, 5768–5771 (1998).
27. D. S. Weiss *et al.*, *Phys. Rev. A* **70**, 040302 (2004).
28. A. Steffen *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 9770–9774 (2012).
29. P. Kok, K. Nemoto, T. C. Ralph, J. P. Dowling, G. J. Milburn, *Rev. Mod. Phys.* **79**, 135–174 (2007).
30. V. Giovannetti, S. Lloyd, L. Maccone, *Science* **306**, 1330–1336 (2004).
31. E. Vetsch *et al.*, *Phys. Rev. Lett.* **104**, 203603 (2010).
32. J. D. Thompson *et al.*, *Science* **340**, 1202–1205 (2013).

ACKNOWLEDGMENTS

We thank T. P. Purdy for useful discussions. This work was supported by the David and Lucile Packard Foundation and the National Science Foundation (NSF) under grant no. 1125844. C.A.R. acknowledges support from the Clare Boothe Luce Foundation, A.M.K. and C.M.R. from the National Defense Science and Engineering Graduate Fellowships, and B.J.L. from the NSF Graduate Research Fellowship Program. K.R.A.H., M.L.W., and A.M.R. acknowledge funding from NSF-Physics at the Information Frontier, Army Research Office and Air Force Office of Scientific Research. K.R.A.H. and M.F.-F. acknowledge support from a National Research Council postdoctoral fellowship.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/345/6194/306/suppl/DC1
Materials and Methods
Figs. S1 to S6
Table S1
References (*33*–*36*)

22 December 2013; accepted 6 June 2014
Published online 26 June 2014;
10.1126/science.1250057

OPTICAL METROLOGY

Electro-optical frequency division and stable microwave synthesis

Jiang Li,[1]* Xu Yi,[1]* Hansuek Lee,[1] Scott A. Diddams,[2] Kerry J. Vahala[1]†

Optical frequency division by using frequency combs has revolutionized time keeping and the generation of stable microwave signals. We demonstrate optical frequency division and microwave generation by using a tunable electrical oscillator to create dual combs through phase modulation of two optical signals that have a stable difference frequency. Phase-locked control of the electrical oscillator by means of optical frequency division produces stable microwaves. Our approach transposes the oscillator and frequency reference of a conventional microwave frequency synthesizer. In this way, the oscillator experiences large phase noise reduction relative to the frequency reference. The electro-optical approach additionally relaxes the need for highly linear photodetection of the comb mode spacing. As well as simplicity, the technique is also tunable and scalable to higher division ratios.

The photomixing of two highly coherent laser signals is a well-known approach to generate a stable radio frequency (RF) or microwave signal (*1*, *2*). Recently, however, a different approach to all-optical signal generation has been demonstrated that may revolutionize applications that require high-stability microwaves. Rather than photomixing stabilized laser signals to directly produce a microwave signal, the approach uses an octave-spanning, self-referenced frequency comb to divide a stable optical reference frequency down to microwave or RF rates (*3*). Frequency dividers are widely used in electronics to generate new frequencies from a single base oscillator or to coherently link different frequency bands. As an ancillary benefit, all frequency dividers reduce the phase noise spectral density of the output signal relative to the input by the square of the division ratio. The new optical frequency dividers perform division by a factor of ~50,000 (the ratio of optical to microwave frequencies), so that phase-noise reduction is greater than 10^9. Moreover, reference-cavity stabilized lasers exhibit a superior, fractional frequency stability in comparison with electrical oscillators (*4*, *5*). Optical dividers applied to such signals thereby generate microwave signals with an exceedingly low phase noise level (*3*).

We present a way to generate high-performance microwave signals through optical frequency division (OFD) by using a cascade of direct phase modulation and self-phase modulation to create an optical comb (*6*–*9*). Because the spectral line spacing is set by the electrical oscillator used to drive the phase modulators (as opposed to an optical resonator), the method of microwave synthesis has similarities to conventional microwave synthesizers while also leveraging the power of OFD so as to reduce phase noise.

In our approach, two laser lines having good relative frequency stability provide an optical reference for the microwave source (Fig. 1A). These laser lines are produced by Brillouin oscillation in a single high-quality-factor (Q) microcavity. However, the lines could also result from any stable optical references, including various types of dual-mode lasers (*10*–*15*), two lasers locked to distinct optical modes of a reference cavity (*16*), or lasers stabilized to atomic transitions (*17*). The laser lines enter the frequency divider portion of the signal generator, where they are phase-modulated by a pair of modulators at a frequency set by a voltage-controlled electrical oscillator (VCO). The sideband spectrum created by the phase modulators is further broadened through pulse-forming and self-phase modulation in an optical fiber (*8*, *18*, *19*). The comb of lines extending from each laser line results in a pair of sidebands near the midpoint of the frequency span. These are optically filtered and detected. The detected beat-note signal contains the phase noise of the VCO, but magnified by the optical division factor. It therefore provides a suitable error signal for phase-lock loop control of the VCO.

We contrast this microwave source that is based on electro-optical frequency division (EOFD) with a conventional microwave source that is based on electrical frequency division of a VCO (*20*). In the conventional approach (Fig. 1B), the VCO provides the highest frequency in the system. It is stabilized through electrical frequency division and phase comparison with a lower-frequency reference oscillator, such as a quartz oscillator. A consequence is that the stabilized VCO has a phase noise level that is always higher than the reference oscillator phase noise by the square of their frequency ratio (the frequency division ratio). In contrast, our optical version reverses the positions of the reference and the VCO in the frequency domain. Specifically, the reference is provided by the frequency difference of the laser lines, and this frequency difference is made much greater than the frequency of the VCO (in the present implementation, this is a nondetectable rate set at ~150 times the VCO frequency). Moreover, this reference frequency is divided down to the VCO frequency, as opposed to dividing the VCO frequency down to

[1]T. J. Watson Laboratory of Applied Physics, California Institute of Technology, Pasadena, CA 91125, USA. [2]Time and Frequency Division, National Institute of Standards and Technology, Boulder, CO 80305, USA.
*These authors contributed equally to this work. **†Corresponding author. E-mail: vahala@caltech.edu**

the reference frequency rate. By making this reversal, the present device benefits from the stability of optical oscillators and the power of OFD to quadratically reduce the phase noise of optical oscillators by the division factor. Also, because the divider derives its rate from the electrical VCO, it is continuously tunable. In particular, the frequency output is set by the tuning range of the electrical VCO and the frequency separation of the reference laser lines and is not fixed by a cavity repetition rate. Additionally, by obtaining the output of the electro-optical frequency divider directly from the VCO, this approach greatly relaxes the linearity constraints of high-bandwidth photodetection that have posed challenges for OFD based on mode-locked lasers (*21*, *22*).

The use of a frequency difference as opposed to an absolute frequency to derive a reference for microwave generation has been demonstrated by using mode-locked laser frequency combs. In that case, the mode-locked frequency comb optical divider is implemented by locking a comb at two frequencies of a reference cavity (*16*). This two-point lock approach has also been implemented by using two atomic lines and a frequency microcomb (*17*). Cascaded phase modulation, other than being applied for pulse generation (*6*–*9*), also has been used to stabilize the frequency difference (in the millimeter wave domain) of a dual-frequency laser (*23*) and to create continuously tunable microwave signals through sideband-injection locking (*24*).

To understand phase noise reduction in this divider, one should consider the accumulated phase noise contributions in the side bands as being tracked from the laser sources to the phase difference of the detected sidebands (Fig. 1C). By adjusting the VCO phase to nullify the phase difference, $\Delta\varphi$, the VCO fluctuations are reduced to $\langle\varphi_M^2\rangle = \langle(\varphi_1 - \varphi_2)^2\rangle/(N_1+N_2)^2$, where $\varphi_{1,2}$ and φ_M are the phase fluctuations of the laser field phases and the electrical VCO. Also, N_1 and N_2 are the number of sidebands produced by modulation (sidebands are higher in frequency for laser 1 and lower in frequency for laser 2). As a result, within the servo control bandwidth, the phase noise of the VCO is given by the relative phase noise of the two laser fields reduced by the division factor squared, $(N_1+N_2)^2$. Clearly, to reduce phase noise of the electrical VCO, the laser frequency separation should be made as large as possible. In the present system, this value is set by the span of the dual-pumped Brillouin lasers. It is the relative phase noise of the laser sources that determines the phase noise reference level (before OFD). Because these lasers are co-lasing within the same resonator, the common-mode laser noise (such as pump noise, microphonics, or cavity noise) is largely suppressed, leaving primarily fundamental sources of laser noise, such as Schawlow-Townes noise.

In the experimental setup (Fig. 2), the reference laser signals are provided by Brillouin-laser lines co-lasing within a single silica-on-silicon high-Q disk resonator. The coherence properties of the individual Brillouin laser lines is excellent (*25*, *26*). The silica disk resonators are designed and fabricated with a free spectral range (FSR) of 10.890 GHz that matches the Brillouin shift frequency in silica near 1550 nm. We used a frequency modulation technique to measure the FSR (*27*). In prior work, a single pump configuration was used to demonstrate microwave synthesis up to K-band (22 GHz) rates by the photomixing of cascaded Brillouin laser lines (*28*). Here, the Brillouin lines must be separated by a much larger frequency, and a dual pump configuration is used. Each pump laser is locked to the disk resonator by using the Pound-Drever-Hall (PDH) technique (*29*) and excites its own Brillouin laser in the backward-propagating direction. In the experiment, the stimulated Brillouin scattering (SBS) lines are separated by as much as 1.61 THz by dual pumping on cavity modes separated by 148 FSR. The maximum separation of 1.61 THz is currently set by the tuning range of the pump lasers. The optical spectra of the dual SBS lines at several frequency differences appears in Fig. 3A. Besides inherent stability from co-lasing within the same cavity, the two SBS laser signals share the same fiber optical path, suppressing path-length variation effects.

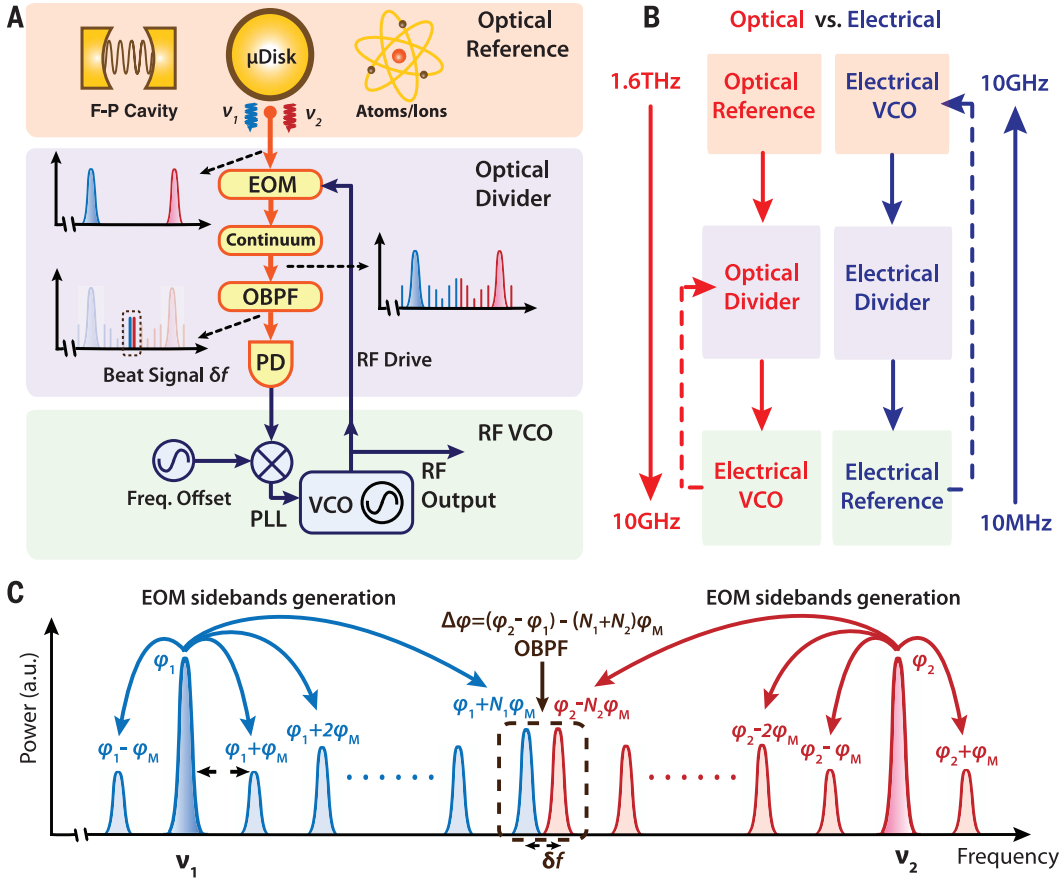


Fig. 1. Conceptual schematics of stable microwave synthesis by use of EOFD. (**A**) Two optical reference laser lines are provided by one of several possible methods, including dual-mode laser action, frequency-locking two lasers to distinct optical modes of a reference cavity, or lasers stabilized to atomic transitions. These laser lines are phase-modulated so as to produce sidebands, two of which nearly overlap at the spectral midpoint of the laser lines. The drive frequency is provided by the electrical VCO that is to be stabilized. Optical filtering and detection of the nearly overlapping sidebands gives a beat note that is used to phase-lock the VCO and lower its phase noise. EOM is the electro-optical modulator, OBPF is the optical band-pass filter, and PD is the photodetector. (**B**) Comparison of conventional, electrical, phase-locked loop control of a VCO with the present method. The conventional method divides the VCO for phase comparison with the low-frequency reference, whereas the present method divides the frequency separation of two lasers so as to stabilize the electrical VCO. (**C**) Analysis of the relative phase of the inner side bands in the optical spectrum. $\varphi_{1,2}$ and φ_M are the phase fluctuations of the laser field phases and the electrical VCO, respectively.

The laser signals are coupled to the optical divider section by use of a circulator. In order to increase the comb bandwidth and so also the division factor, two phase modulators with low $V_\pi \sim 3.9$V (at 12 GHz) are cascaded and phase-synchronized with a RF phase shifter. The phase modulators were driven at 32.5 and 30.7 dBm, corresponding to a total phase-modulation amplitude up to 6π $\{\approx \pi[(V_{drive})/(V_\pi)]\}$, which is also approximately equal to half of the number of sidebands produced by the modulators (*30*). In the experiment, up to 30 electro-optic modulator (EOM) sidebands can be created by using the phase modulators alone, enabling optical division by around a factor of 30. An optical spectrum showing both the original SBS laser lines (18 FSR frequency separation) and the resulting phase modulation sidebands is shown in Fig. 3B. To further enhance the sideband spectral width, nonlinear spectral broadening was performed by introducing an intensity modulator, dispersion compensation, erbium-doped fiber amplifier (EDFA), and nonlinear fiber (*8*, *18*, *19*). The inner sidebands are optically filtered, and their beat note is detected on an amplified photodetector. The beat note frequency was set near 10 MHz and phase-compared with a 10-MHz oven-controlled crystal oscillator (OCXO) in order to generate the error signal used for phase-locking the VCO. The phase noise of the OCXO is very low (compared with the optical frequency reference) and is further divided down in the control loop (supplementary materials). As such, it does not introduce any limitation on the final VCO phase noise. The residual phase noise levels of the amplifier used here [MicroSemi (Aliso Viejo, CA) model L0612-35-T977] are well below the phase noise levels attained through loop control of the VCO and therefore do not present any limitation to the measurement. The sum of the free-running VCO phase noise and the residual phase noise of the RF power amplifier is suppressed by the servo loop (supplementary materials). As a result, by taking the electrical output signal at the output of the amplifier, there is no phase noise limitation set by the RF power amplifier.

The phase noise reference level is set by the relative phase noise of the co-lasing Brillouin laser lines. This phase noise level is expected to be relatively insensitive to the frequency difference (*26*, *28*). Therefore, this noise level was measured by pumping the two lasers on modes that were three cavity FSR apart, enabling direct detection of the beat signal at 32.7 GHz by a fast photodetector. We then analyzed this detected signal using a phase noise analyzer in order to obtain the single sideband (SSB) phase noise spectrum (Fig. 3C, red curve). For offset frequencies higher than 100 Hz, this spectrum is limited by Schawlow-Townes noise, whereas at offset frequencies less than 100 Hz, a technical noise component is present (*28*). The white phase noise floor above 1 MHz offset is due to the thermal noise of the fast photodetector. We measured the SSB phase noise spectrum using a commercial phase noise tester [Rohde Schwartz (Munich, Germany) model FSUP26; sensitivity limit given as the light blue squares in Fig. 3C] based on the phase detector method, with dual-channel cross-correlations so as to further improve the phase noise sensitivity.

To test the optical divider, we locked the VCO to the divided optical reference using the VCO control/tuning port (Fig. 2). The blue and green curves in Fig. 3C are the corresponding phase noise spectra of the VCO at 10.89 GHz when optically dividing by 30 and 148 times relative to the initial frequency separation of the dual SBS lines at 327 GHz and 1.61 THz, respectively. The microwave carrier frequency is held at 10.89 GHz in these measurements by adjusting the frequency separation of the Brillouin laser lines with the tunable pump lasers. Decrease of the phase noise with the increase of the division ratio is clear in the data. For division by 148×, the achieved phase noise level is –104 dBc/Hz at 1 kHz and –121 dBc/Hz at 10 kHz. For comparison, the dashed black curve in Fig. 3C gives the phase noise of the free-running VCO. The servo-locking bandwidth can be seen from the servo bumps in the phase noise spectra, which are located at 820 kHz for the blue curve (adopting path I in Fig. 2) and 300 kHz for the green curve (adopting path II in Fig. 2). As an aside, the locking bandwidth is determined by the loop delay, which includes both optical and electrical path lengths (~20 m for path I and ~50 m for path II) and the FM response of the VCO. For frequencies beyond the servo control bandwidth, the phase noise level is set by the free-running VCO.

In the measured RF power spectra of the VCO at 10.89 GHz (20 kHz span and 30 Hz resolution bandwidth) (Fig. 3D), the black curve is the free-running VCO spectrum, and the red, blue, and green curves are the spectra of the phase-locked VCO when dividing down 18, 30, and 148 times, respectively, from corresponding SBS frequency separations of 196 GHz, 327 GHz, and 1.61 THz. These results are summarized in Fig. 3E by giving the measured phase noise at 1 kHz, 10 kHz, and 100 kHz offset frequencies plotted versus division ratios of 1, 3, 4, 10, 18, 30, and 148. The dashed lines give $1/N^2$ trend lines. For comparison, the phase noise of an Agilent (Santa Clara, CA) MXG microwave synthesizer (carrier 11 GHz, offset 100 kHz, Agilent online data sheet, Literature number 5989-7572EN) and a high-performance Agilent PSG microwave synthesizer (carrier 11 GHz, offset 100 kHz, Agilent online data sheet, Literature

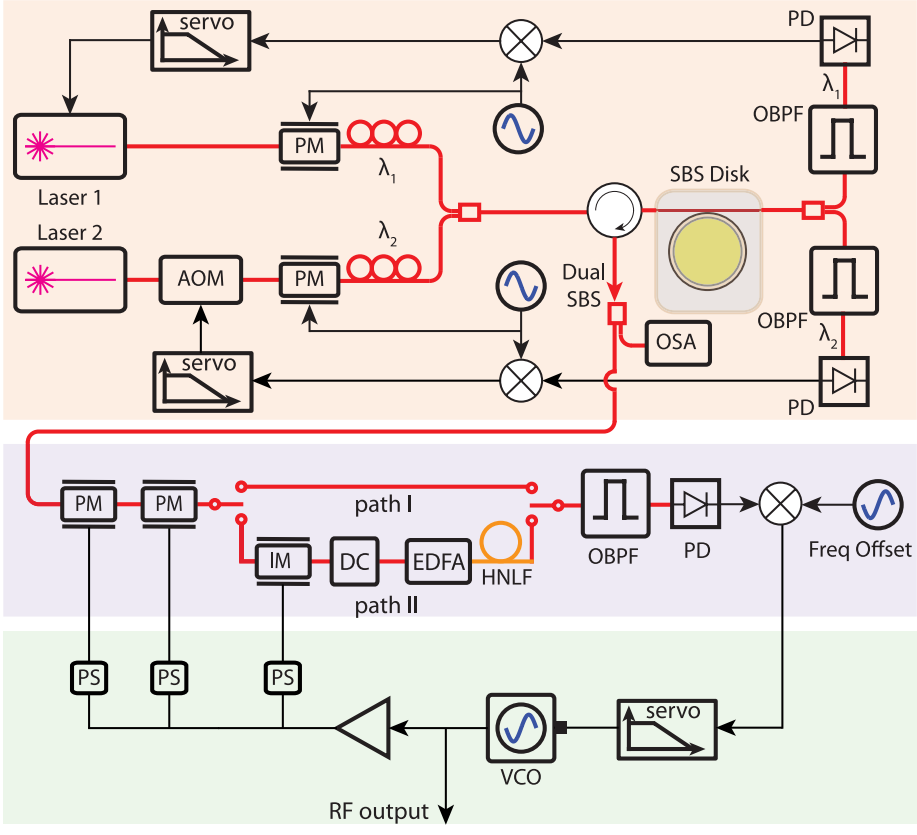


Fig. 2. Experimental setup. (Top) Schematic showing the arrangement used for dual-pumping of the disk-resonator Brillouin laser. Lasers 1 and 2 are locked to distinct modes of a high-Q disk resonator. Servo control involves direct frequency tuning of laser 1 and frequency shifting via an acousto-optic modulator (AOM) for laser 2. The reference laser lines provided by the Brillouin oscillation are input to the optical division function. Depending on the desired division ratio, two pathways are taken (path I and path II). For division up to 30×, a pair of phase modulators are synchronously driven, whereas for larger division ratios, the phase modulators are cascaded with a fiber pulse broadener containing an intensity modulator (IM), dispersion compensation (DC), EDFA, and nonlinear fiber (HNLF). The filtered (OBPF) and detected signal is then processed for servo control of the VCO. Also shown are the optical spectrum analyzer (OSA) and electrical phase shifters (PS).

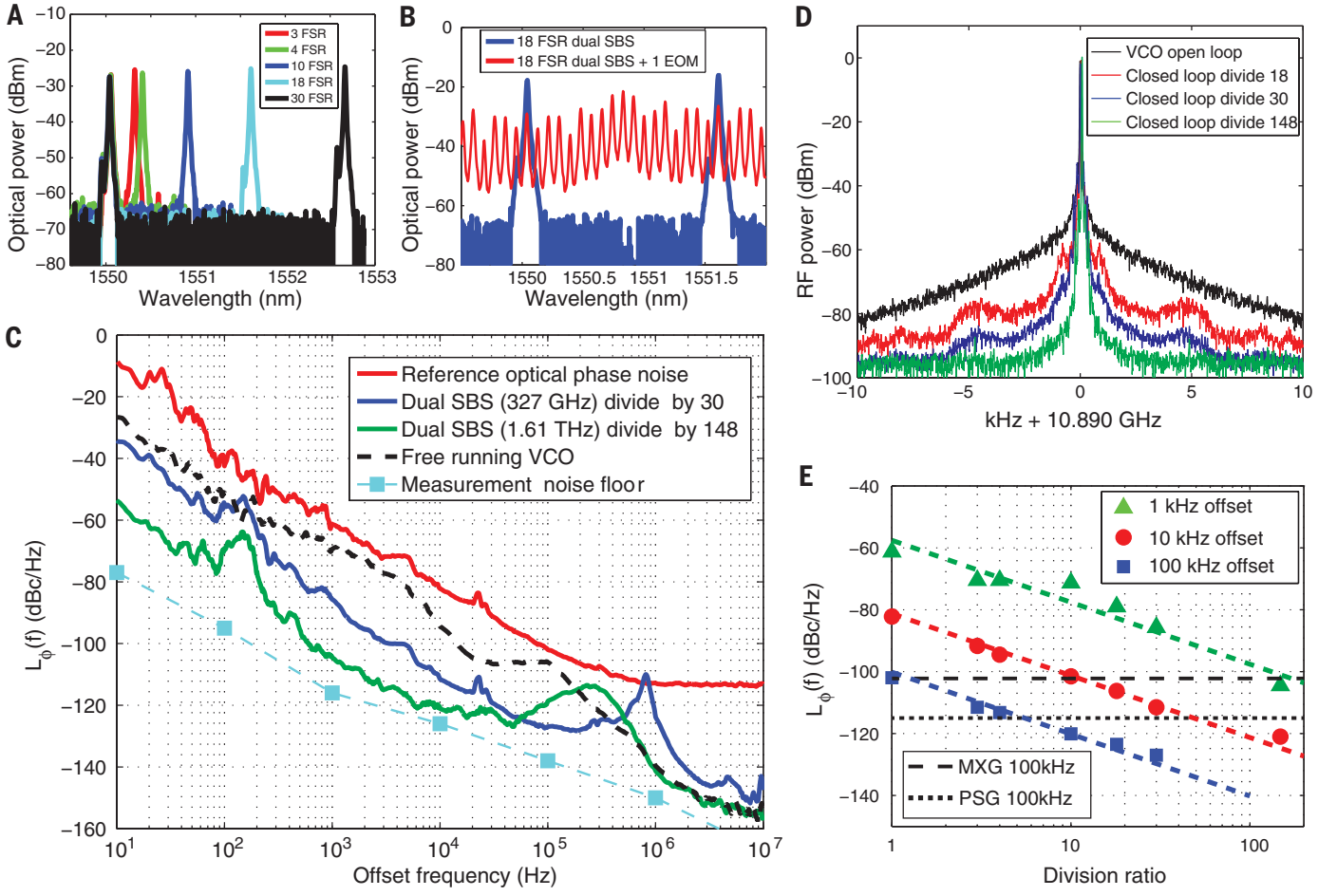
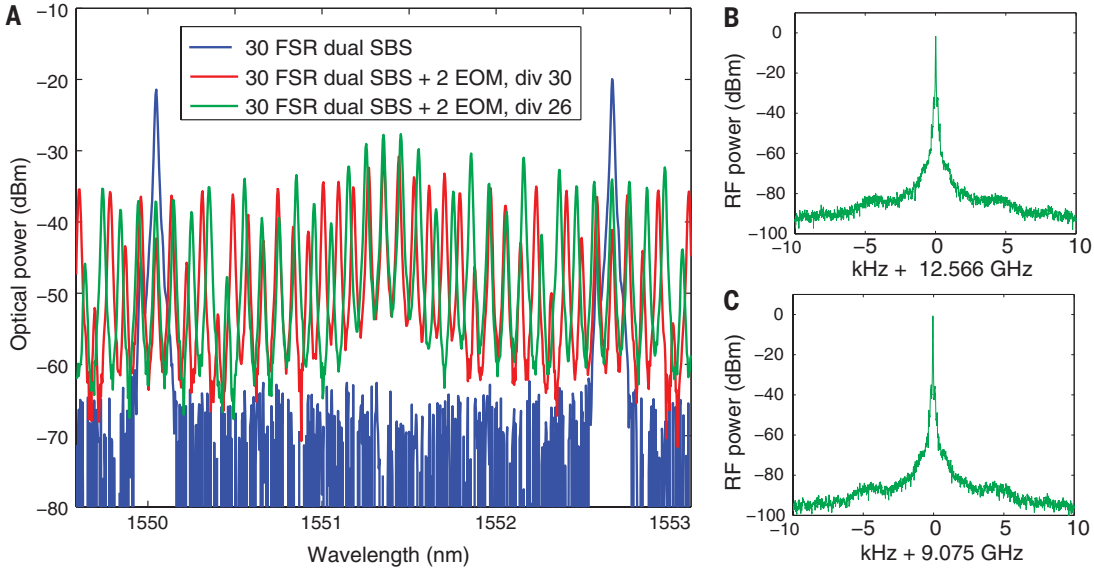


Fig. 3. Summary of experimental data. (**A**) Optical spectra of co-lasing Brillouin lasers at various tuning configurations used in the measurement corresponding to cavity FSR spacings provided in the inset. (**B**) Optical spectrum of Brillouin laser lines with (red) and without (blue) phase modulation. (**C**) SSB phase noise spectrum of the optical reference (red) is shown. For this measurement, the dual SBS lines are tuned to a detectable frequency difference. Also shown are the closed-loop VCO phase noise referenced to dual SBS lines separated by 327 GHz (blue) and 1.61 THz (green), with corresponding division ratios of 30 and 148. The dashed black curve is the phase noise of the free-running VCO. The light blue square markers denote the phase noise sensitivity of the phase noise tester. (**D**) VCO RF power spectrum at 10.89 GHz for open-loop and closed-loop conditions in which the Brillouin frequency separation is divided by 18×, 30×, and 148×. Resolution bandwidth is set at 30 Hz. (**E**) Summary of SSB phase noise levels at 10.89 GHz versus optical division ratio. Cases shown are measured at 1-, 10-, and 100-kHz offset frequencies. The dashed lines give $1/N^2$ trends. Phase noise levels for an Agilent MXG and PSG microwave sources (carrier 11 GHz, 100 kHz offset) are provided for comparison.

Fig. 4. Demonstration of tuning of the phase-locked VCO frequency. (**A**) Optical spectrum showing two Brillouin reference laser lines (blue) and tuning of the optical division by means of variation of the VCO modulation frequency to produce two different combs. (**B**) RF power spectrum corresponding to division by 26× to produce a 12.566-GHz electrical carrier frequency (resolution bandwidth is 30 Hz). (**C**) RF power spectrum corresponding to division by 36× to produce a 9.075-GHz electrical carrier frequency (resolution bandwidth is 30 Hz).



number 5989-0698EN, option UNX) are shown as black, dashed, and dotted lines.

So far, the microwave frequency has been fixed; however, this can be adjusted by varying the division ratio, N, and the VCO frequency for a fixed separation of the dual SBS lines $\nu_1 - \nu_2$ as in Fig. 4A $\{f_M = [(\nu_1 - \nu_2 \pm \delta f)/(N)]\}$. The RF spectrum of a phase-locked 12.566 GHz carrier with a division factor of 26 is shown in Fig. 4B, whereas the RF power spectrum of a microwave signal at 9.075 GHz with a divison factor of 36 is shown in Fig. 4C (in both cases, the dual SBS frequency difference is set at 327 GHz). Moreover, the initial reference frequency difference can also be tuned so that a wide range of final microwave frequencies can be derived from a single divider.

The optical division factor can be scaled to higher values for additional phase noise improvement. Although the frequency reference might be based on some other system, for simplicity we considered the use of one SBS laser operating at 1.55 μm (current system) and another at 1 μm [recently demonstrated by using silica high-Q disks (*31*)]. Matching the Brillouin shift at these wavelengths within the same disk is possible by designing the system to oscillate at a multiple cavity free spectral range. For example, a disk diameter of ~12 mm (FSR of ~5.4 GHz, which is half of the Brillouin shift at 1.55 μm and one third of the Brillouin shift at 1 μm) would enable simultaneous SBS oscillation at 1.55 and 1 μm. To extend the comb over this span would require about one half of an octave. Spectrally broadened EOM combs with more than 500 nm bandwidth have previously been demonstrated (*32*). An electro-optical frequency divider spanning 1- and 1.55-μm SBS lines is therefore feasible. This example system does not necessarily represent a fundamental limit but nonetheless would feature a division factor of ~7000. Assuming that relative phase noise of the co-lasing optical lines does not degrade beyond what has been demonstrated in the present work, this hypothetical system could provide phase noise performance close to levels demonstrated by conventional OFD.

This new photonic architecture for OFD provides a route to improve the phase noise of a common VCO. In the present configuration, dual SBS lasers have been co-generated in a single, chip-based resonator to establish a stable reference as high as 1.61 THz. Although dual-pumped SBS laser lines provide an excellent reference frequency, it should also be possible to lock two lasers to a resonator to establish the reference. Because the current frequency separation is limited by the pump lasers, much larger division ratios and potentially lower phase noise levels should be possible. In comparison with conventional OFD by use of a self-referenced frequency comb, this technique does not presently offer as large a division ratio. However, it is simple and relatively low-cost and also provides tuning of the electrical carrier. Moreover, it relies on a reference signal derived from the relative phase of two lasers as opposed to the absolute phase of a single laser. This can potentially improve the robustness with respect to microphonics and other sources of technical noise.

REFERENCES AND NOTES

1. A. J. Seeds, K. J. Williams, *J. Lightwave Technol.* **24**, 4628–4641 (2006).
2. J. Yao, *J. Lightwave Technol.* **27**, 314–335 (2009).
3. T. Fortier *et al.*, *Nat. Photonics* **5**, 425–429 (2011).
4. B. Young, F. Cruz, W. Itano, J. Bergquist, *Phys. Rev. Lett.* **82**, 3799–3802 (1999).
5. T. Kessler *et al.*, *Nat. Photonics* **6**, 687–692 (2012).
6. H. Murata, A. Morimoto, T. Kobayashi, S. Yamamoto, *IEEE J. Sel. Top. Quantum Electron.* **6**, 1325–1331 (2000).
7. M. Fujiwara, J. Kani, H. Suzuki, K. Araya, M. Teshima, *Electron. Lett.* **37**, 967–968 (2001).
8. C. B. Huang, S. G. Park, D. E. Leaird, A. M. Weiner, *Opt. Express* **16**, 2520–2527 (2008).
9. A. J. Metcalf, V. Torres-Company, D. E. Leaird, A. M. Weiner, *IEEE J. Sel. Top. Quantum Electron.* **19**, 3500306 (2013).
10. G. Pillet *et al.*, *J. Lightwave Technol.* **26**, 2764–2773 (2008).
11. J. Geng, S. Staines, S. Jiang, *Opt. Lett.* **33**, 16–18 (2008).
12. S. Pan, J. Yao, *Opt. Express* **17**, 5414–5419 (2009).
13. M. C. Gross *et al.*, *Opt. Express* **18**, 13321–13330 (2010).
14. P. T. Callahan, M. C. Gross, M. L. Dennis, *IEEE J. Quantum Electron.* **47**, 1142–1150 (2011).
15. G. Carpintero *et al.*, *Opt. Lett.* **37**, 3657–3659 (2012).
16. W. C. Swann, E. Baumann, F. R. Giorgetta, N. R. Newbury, *Opt. Express* **19**, 24387–24395 (2011).
17. S. B. Papp *et al.*, arXiv:1309.3525 (2013).
18. I. Morohashi, T. Sakamoto, H. Sotobayashi, T. Kawanishi, I. Hosako, *Opt. Lett.* **34**, 2297–2299 (2009).
19. A. Ishizawa *et al.*, *Opt. Express* **21**, 29186–29194 (2013).
20. U. L. Rohde, *Microwave and Wireless Synthesizers: Theory and Design* (Wiley, New York, 1997).
21. E. N. Ivanov, S. A. Diddams, L. Hollberg, *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **52**, 1068–1074 (2005).
22. J. Taylor *et al.*, *IEEE Photon. J.* **3**, 140–151 (2011).
23. A. Rolland *et al.*, *Opt. Express* **19**, 17944–17950 (2011).
24. G. J. Schneider, J. A. Murakowski, C. A. Schuetz, S. Shi, D. W. Prather, *Nat. Photonics* **7**, 118–122 (2013).
25. H. Lee *et al.*, *Nat. Photonics* **6**, 369–373 (2012).
26. J. Li, H. Lee, T. Chen, K. J. Vahala, *Opt. Express* **20**, 20170–20180 (2012).
27. J. Li, H. Lee, K. Y. Yang, K. J. Vahala, *Opt. Express* **20**, 26337–26344 (2012).
28. J. Li, H. Lee, K. J. Vahala, *Nat. Commun.* **4**, 2097 (2013).
29. R. Drever *et al.*, *Appl. Phys. B* **31**, 97–105 (1983).
30. T. Sakamoto, T. Kawanishi, M. Izutsu, *Opt. Lett.* **32**, 1515–1517 (2007).
31. J. Li, H. Lee, K. J. Vahala, *Opt. Lett.* **39**, 287–290 (2014).
32. S. Suzuki *et al.*, paper NM3A.3, Nonlinear Optics Conference, Optical Society of America (2013).

ACKNOWLEDGMENTS

This work was sponsored by the Defense Advanced Research Projects Agency ORCHID and PULSE programs and the California Institute of Technology Institute for Quantum Information and Matter, an NSF Physics Frontiers Center with support of the Gordon and Betty Moore Foundation. The authors also acknowledge support from the Kavli NanoScience Institute at the California Institute of Technology. S.A.D. acknowledges the support from the Gordon and Betty Moore Foundation as a Moore Scholar at Caltech.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/345/6194/309/suppl/DC1
Materials and Methods
Supplementary Text
Fig. S1
References (*33*)

4 March 2014; accepted 5 June 2014
Published online 19 June 2014;
10.1126/science.1252909

ASSOCIATIVE LEARNING

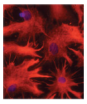
Role of synaptic phosphatidylinositol 3-kinase in a behavioral learning response in *C. elegans*

Hayao Ohno,[1,2] Shinya Kato,[1]* Yasuki Naito,[1,3] Hirofumi Kunitomo,[1] Masahiro Tomioka,[3]† Yuichi Iino[1,2,3]†

The phosphatidylinositol 3-kinase (PI3K) pathway regulates many cellular functions, but its roles in the nervous system are still poorly understood. We found that a newly discovered insulin receptor isoform, DAF-2c, is translocated from the cell body to the synaptic region of the chemosensory neuron in *Caenorhabditis elegans* by a conditioning stimulus that induces taste avoidance learning. This translocation is essential for learning and is dependent on the mitogen-activated protein kinase–regulated interaction of CASY-1 (the calsyntenin ortholog) and kinesin-1. The PI3K pathway is required downstream of the receptor. Light-regulated activation of PI3K in the synaptic region, but not in other parts of the cell, switched taste-attractive behavior to taste avoidance, mimicking the effect of conditioning. Thus, synaptic PI3K is crucial for the behavioral switch caused by learning.

Signaling through the insulin-PI3K pathway is widespread in the mature brain, where it has been implicated in memory performance and neurological disorders (*1*, *2*). However, the regulation and roles of this signaling in the central nervous system are largely unknown. In *Caenorhabditis elegans*, insulin-PI3K signaling acting in the nervous system controls various biological processes (*3*), including associative learning. *C. elegans* are attracted to the previous salt concentration at which they have been fed, whereas they avoid it if they have been starved (Fig. 1A and fig. S1) (*4*, *5*). The avoidance of the salt concentrations associated with starvation (hereafter called taste avoidance

[1]Department of Biological Sciences, Graduate School of Science, University of Tokyo, Tokyo 113-0033, Japan. [2]JST, CREST, 4-1-8 Honcho, Kawaguchi, Saitama 332-0012, Japan. [3]Molecular Genetics Research Laboratory, Graduate School of Science, University of Tokyo, Tokyo 113-0033, Japan. *Present address: Department of Mathematical Informatics, Graduate School of Information Science and Technology, University of Tokyo, Tokyo 113-0033, Japan. **†Corresponding author. E-mail: seidai@gen.s.u-tokyo.ac.jp (M.T.); iino@bs.s.u-tokyo.ac.jp (Y.I.)**

EXTENDED PDF FORMAT
SPONSORED BY

Travel Grants Available
Neuroscience 2014
Application Deadline: July 31st
>Apply Now
R&D systems
a bietechne brand
www.rndsystems.com

Role of synaptic phosphatidylinositol 3-kinase in a behavioral learning response in *C. elegans*

Hayao Ohno *et al.*

Science **345**, 313 (2014);
DOI: 10.1126/science.1250709

This copy is for your personal, non-commercial use only.

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by clicking here.

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines here.

The following resources related to this article are available online at www.sciencemag.org (this information is current as of August 4, 2014):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:
http://www.sciencemag.org/content/345/6194/313.full.html

Supporting Online Material can be found at:
http://www.sciencemag.org/content/suppl/2014/07/16/345.6194.313.DC1.html

A list of selected additional articles on the Science Web sites **related to this article** can be found at:
http://www.sciencemag.org/content/345/6194/313.full.html#related

This article **cites 47 articles**, 24 of which can be accessed free:
http://www.sciencemag.org/content/345/6194/313.full.html#ref-list-1

This article appears in the following **subject collections:**
Neuroscience
http://www.sciencemag.org/cgi/collection/neuroscience

Downloaded from www.sciencemag.org on August 4, 2014

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published weekly, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. Copyright 2014 by the American Association for the Advancement of Science; all rights reserved. The title *Science* is a registered trademark of AAAS.

number 5989-0698EN, option UNX) are shown as black, dashed, and dotted lines.

So far, the microwave frequency has been fixed; however, this can be adjusted by varying the division ratio, N, and the VCO frequency for a fixed separation of the dual SBS lines $\nu_1 - \nu_2$ as in Fig. 4A $\{f_M = [(\nu_1 - \nu_2 \pm \delta f)/(N)]\}$. The RF spectrum of a phase-locked 12.566 GHz carrier with a division factor of 26 is shown in Fig. 4B, whereas the RF power spectrum of a microwave signal at 9.075 GHz with a divison factor of 36 is shown in Fig. 4C (in both cases, the dual SBS frequency difference is set at 327 GHz). Moreover, the initial reference frequency difference can also be tuned so that a wide range of final microwave frequencies can be derived from a single divider.

The optical division factor can be scaled to higher values for additional phase noise improvement. Although the frequency reference might be based on some other system, for simplicity we considered the use of one SBS laser operating at 1.55 μm (current system) and another at 1 μm [recently demonstrated by using silica high-Q disks (*31*)]. Matching the Brillouin shift at these wavelengths within the same disk is possible by designing the system to oscillate at a multiple cavity free spectral range. For example, a disk diameter of ~12 mm (FSR of ~5.4 GHz, which is half of the Brillouin shift at 1.55 μm and one third of the Brillouin shift at 1 μm) would enable simultaneous SBS oscillation at 1.55 and 1 μm. To extend the comb over this span would require about one half of an octave. Spectrally broadened EOM combs with more than 500 nm bandwidth have previously been demonstrated (*32*). An electro-optical frequency divider spanning 1- and 1.55-μm SBS lines is therefore feasible. This example system does not necessarily represent a fundamental limit but nonetheless would feature a division factor of ~7000. Assuming that relative phase noise of the co-lasing optical lines does not degrade beyond what has been demonstrated in the present work, this hypothetical system could provide phase noise performance close to levels demonstrated by conventional OFD.

This new photonic architecture for OFD provides a route to improve the phase noise of a common VCO. In the present configuration, dual SBS lasers have been co-generated in a single, chip-based resonator to establish a stable reference as high as 1.61 THz. Although dual-pumped SBS laser lines provide an excellent reference frequency, it should also be possible to lock two lasers to a resonator to establish the reference. Because the current frequency separation is limited by the pump lasers, much larger division ratios and potentially lower phase noise levels should be possible. In comparison with conventional OFD by use of a self-referenced frequency comb, this technique does not presently offer as large a division ratio. However, it is simple and relatively low-cost and also provides tuning of the electrical carrier. Moreover, it relies on a reference signal derived from the relative phase of two lasers as opposed to the absolute phase of a single laser. This can potentially improve the robustness with respect to microphonics and other sources of technical noise.

REFERENCES AND NOTES

1. A. J. Seeds, K. J. Williams, *J. Lightwave Technol.* **24**, 4628–4641 (2006).
2. J. Yao, *J. Lightwave Technol.* **27**, 314–335 (2009).
3. T. Fortier *et al.*, *Nat. Photonics* **5**, 425–429 (2011).
4. B. Young, F. Cruz, W. Itano, J. Bergquist, *Phys. Rev. Lett.* **82**, 3799–3802 (1999).
5. T. Kessler *et al.*, *Nat. Photonics* **6**, 687–692 (2012).
6. H. Murata, A. Morimoto, T. Kobayashi, S. Yamamoto, *IEEE J. Sel. Top. Quantum Electron.* **6**, 1325–1331 (2000).
7. M. Fujiwara, J. Kani, H. Suzuki, K. Araya, M. Teshima, *Electron. Lett.* **37**, 967–968 (2001).
8. C. B. Huang, S. G. Park, D. E. Leaird, A. M. Weiner, *Opt. Express* **16**, 2520–2527 (2008).
9. A. J. Metcalf, V. Torres-Company, D. E. Leaird, A. M. Weiner, *IEEE J. Sel. Top. Quantum Electron.* **19**, 3500306 (2013).
10. G. Pillet *et al.*, *J. Lightwave Technol.* **26**, 2764–2773 (2008).
11. J. Geng, S. Staines, S. Jiang, *Opt. Lett.* **33**, 16–18 (2008).
12. S. Pan, J. Yao, *Opt. Express* **17**, 5414–5419 (2009).
13. M. C. Gross *et al.*, *Opt. Express* **18**, 13321–13330 (2010).
14. P. T. Callahan, M. C. Gross, M. L. Dennis, *IEEE J. Quantum Electron.* **47**, 1142–1150 (2011).
15. G. Carpintero *et al.*, *Opt. Lett.* **37**, 3657–3659 (2012).
16. W. C. Swann, E. Baumann, F. R. Giorgetta, N. R. Newbury, *Opt. Express* **19**, 24387–24395 (2011).
17. S. B. Papp *et al.*, arXiv:1309.3525 (2013).
18. I. Morohashi, T. Sakamoto, H. Sotobayashi, T. Kawanishi, I. Hosako, *Opt. Lett.* **34**, 2297–2299 (2009).
19. A. Ishizawa *et al.*, *Opt. Express* **21**, 29186–29194 (2013).
20. U. L. Rohde, *Microwave and Wireless Synthesizers: Theory and Design* (Wiley, New York, 1997).
21. E. N. Ivanov, S. A. Diddams, L. Hollberg, *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **52**, 1068–1074 (2005).
22. J. Taylor *et al.*, *IEEE Photon. J.* **3**, 140–151 (2011).
23. A. Rolland *et al.*, *Opt. Express* **19**, 17944–17950 (2011).
24. G. J. Schneider, J. A. Murakowski, C. A. Schuetz, S. Shi, D. W. Prather, *Nat. Photonics* **7**, 118–122 (2013).
25. H. Lee *et al.*, *Nat. Photonics* **6**, 369–373 (2012).
26. J. Li, H. Lee, T. Chen, K. J. Vahala, *Opt. Express* **20**, 20170–20180 (2012).
27. J. Li, H. Lee, K. Y. Yang, K. J. Vahala, *Opt. Express* **20**, 26337–26344 (2012).
28. J. Li, H. Lee, K. J. Vahala, *Nat. Commun.* **4**, 2097 (2013).
29. R. Drever *et al.*, *Appl. Phys. B* **31**, 97–105 (1983).
30. T. Sakamoto, T. Kawanishi, M. Izutsu, *Opt. Lett.* **32**, 1515–1517 (2007).
31. J. Li, H. Lee, K. J. Vahala, *Opt. Lett.* **39**, 287–290 (2014).
32. S. Suzuki *et al.*, paper NM3A.3, Nonlinear Optics Conference, Optical Society of America (2013).

ACKNOWLEDGMENTS

This work was sponsored by the Defense Advanced Research Projects Agency ORCHID and PULSE programs and the California Institute of Technology Institute for Quantum Information and Matter, an NSF Physics Frontiers Center with support of the Gordon and Betty Moore Foundation. The authors also acknowledge support from the Kavli NanoScience Institute at the California Institute of Technology. S.A.D. acknowledges the support from the Gordon and Betty Moore Foundation as a Moore Scholar at Caltech.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/345/6194/309/suppl/DC1
Materials and Methods
Supplementary Text
Fig. S1
References (*33*)

4 March 2014; accepted 5 June 2014
Published online 19 June 2014;
10.1126/science.1252909

ASSOCIATIVE LEARNING

Role of synaptic phosphatidylinositol 3-kinase in a behavioral learning response in *C. elegans*

Hayao Ohno,[1,2] Shinya Kato,[1]* Yasuki Naito,[1,3] Hirofumi Kunitomo,[1] Masahiro Tomioka,[3]† Yuichi Iino[1,2,3]†

The phosphatidylinositol 3-kinase (PI3K) pathway regulates many cellular functions, but its roles in the nervous system are still poorly understood. We found that a newly discovered insulin receptor isoform, DAF-2c, is translocated from the cell body to the synaptic region of the chemosensory neuron in *Caenorhabditis elegans* by a conditioning stimulus that induces taste avoidance learning. This translocation is essential for learning and is dependent on the mitogen-activated protein kinase–regulated interaction of CASY-1 (the calsyntenin ortholog) and kinesin-1. The PI3K pathway is required downstream of the receptor. Light-regulated activation of PI3K in the synaptic region, but not in other parts of the cell, switched taste-attractive behavior to taste avoidance, mimicking the effect of conditioning. Thus, synaptic PI3K is crucial for the behavioral switch caused by learning.

Signaling through the insulin-PI3K pathway is widespread in the mature brain, where it has been implicated in memory performance and neurological disorders (*1*, *2*). However, the regulation and roles of this signaling in the central nervous system are largely unknown. In *Caenorhabditis elegans*, insulin-PI3K signaling acting in the nervous system controls various biological processes (*3*), including associative learning. *C. elegans* are attracted to the previous salt concentration at which they have been fed, whereas they avoid it if they have been starved (Fig. 1A and fig. S1) (*4*, *5*). The avoidance of the salt concentrations associated with starvation (hereafter called taste avoidance

[1]Department of Biological Sciences, Graduate School of Science, University of Tokyo, Tokyo 113-0033, Japan. [2]JST, CREST, 4-1-8 Honcho, Kawaguchi, Saitama 332-0012, Japan. [3]Molecular Genetics Research Laboratory, Graduate School of Science, University of Tokyo, Tokyo 113-0033, Japan. *Present address: Department of Mathematical Informatics, Graduate School of Information Science and Technology, University of Tokyo, Tokyo 113-0033, Japan. **†Corresponding author. E-mail: seidai@gen.s.u-tokyo.ac.jp (M.T.); iino@bs.s.u-tokyo.ac.jp (Y.I.)**

learning) was impaired in mutants of the components of the insulin-PI3K pathway consisting of the insulin receptor DAF-2, the PI3K AGE-1, and downstream kinases (Fig. 1A and fig. S2) (*4*, *5*). DAF-2, which is the only *C. elegans* insulin receptor, relays the signals of various ligands (*6*) and controls multiple downstream effectors (*7*).

To understand how the *daf-2* gene can exert diverse biological functions, we searched for possible splicing isoforms by screening *daf-2* transcripts with extensive use of reverse transcription polymerase chain reaction. We discovered an isoform (hereafter called *daf-2c*) that has an insertion of a new alternative exon (hereafter called exon 11.5) located between exons 11 and 12 of the *daf-2a* isoform (fig. S3A). This splicing pattern is analogous to that of the B isoform of the mammalian insulin receptor (IR-B), which has an additional cassette exon (exon 11) not present in the A isoform (*8*). Both the alternative segments of DAF-2c and IR-B are inserted near the C terminus of the α chain that has been predicted to be exposed extracellularly (Fig. 1B and fig. S3B)

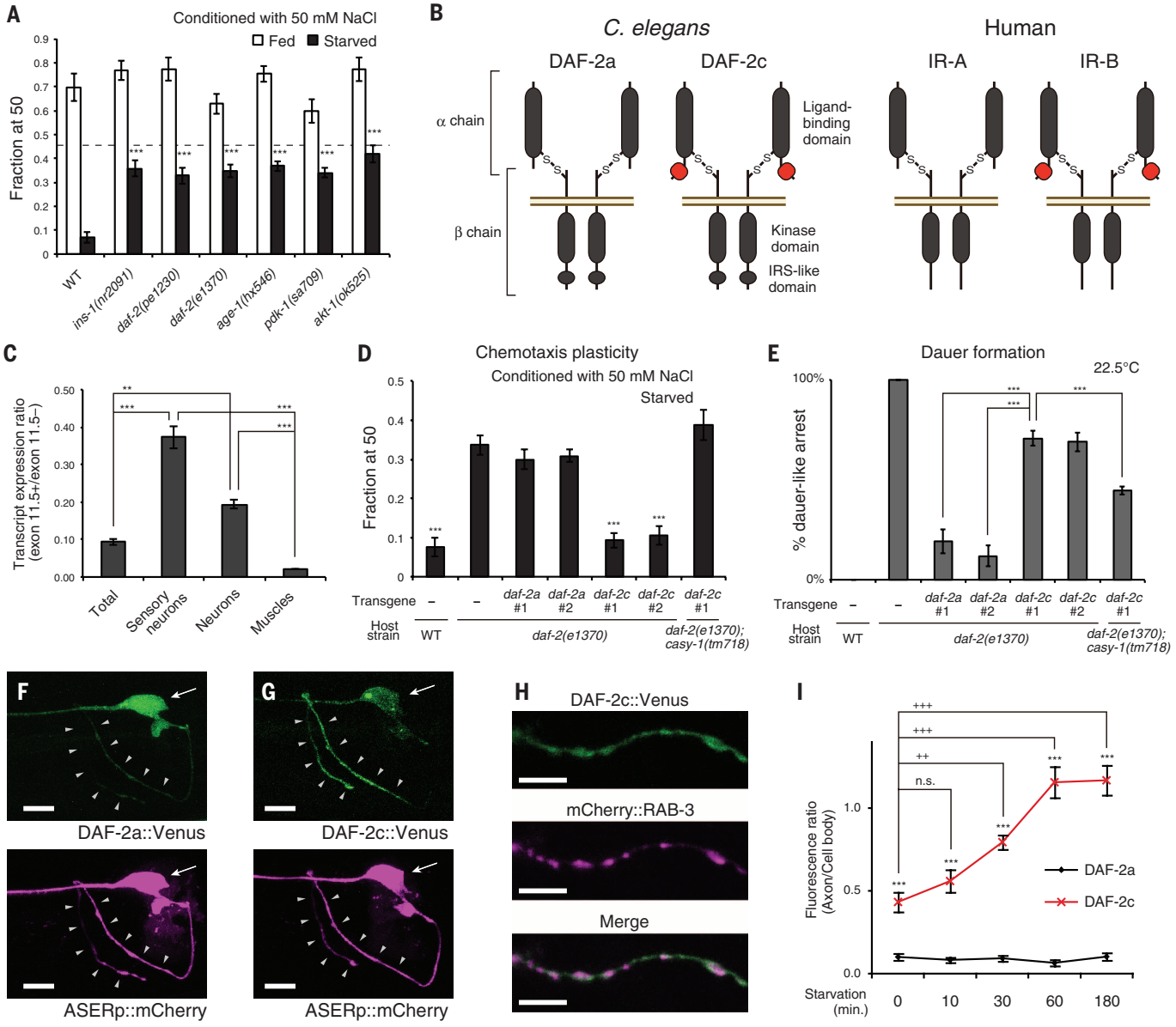


Fig. 1. Alternative splicing yields *daf-2* isoforms with different intracellular localization and biological functions. (**A**) Preference for the ~50 mM NaCl area was examined in insulin-signaling mutants that were conditioned with 50 mM NaCl. The dashed line shows the control score of the wild type in the absence of a NaCl gradient (0.45 ± 0.03, n = 14); $n \geq 14$. ***$P < 0.001$ (Dunnett's test). (**B**) Schematic of the predicted structural difference (red) between the two insulin receptor isoforms in *C. elegans* and humans. IRS-like domain, a domain similar to the central part of insulin receptor substrate (IRS). (**C**) Expression ratios of *daf-2* (exon 11.5+) to *daf-2* (exon 11.5–) in mRNA that was isolated from the indicated tissues by mRNA tagging; n = 4. **$P < 0.01$, ***$P < 0.001$ (Tukey's test). (**D**) Salt concentration preferences of animals ubiquitously expressing *daf-2a* or *daf-2c*. #1 and #2 indicate two independent transgenic lines; n = 14. ***$P < 0.001$ (Dunnett's test). (**E**) Dauer formation of animals expressing *daf-2a* or *daf-2c*; $n \geq 5$. ***$P < 0.001$ (Tukey's test). (**F** and **G**) Localization of DAF-2a::Venus (F) and DAF-2c::Venus (G) in ASER. The lower panels show ASER labeled with mCherry. Arrowheads and arrows indicate the axon (synaptic region) and the cell body, respectively. Scale bars, 10 μm. (**H**) DAF-2c::Venus and a presynaptic marker, mCherry::RAB-3, in the distal axon of ASER. Scale bars, 5 μm. (**I**) Axonal localization of DAF-2a/c::Venus, expressed as ratios of axonal to cell body fluorescence, was quantified in ASER of animals that had been starved for 0 to 180 min; $n \geq 16$ for each condition. Difference between two isoforms: ***$P < 0.001$ (Bonferroni's test). Difference between time points: ++$P < 0.01$, +++$P < 0.001$ (Dunnett's test). Error bars represent SEM in all figures.

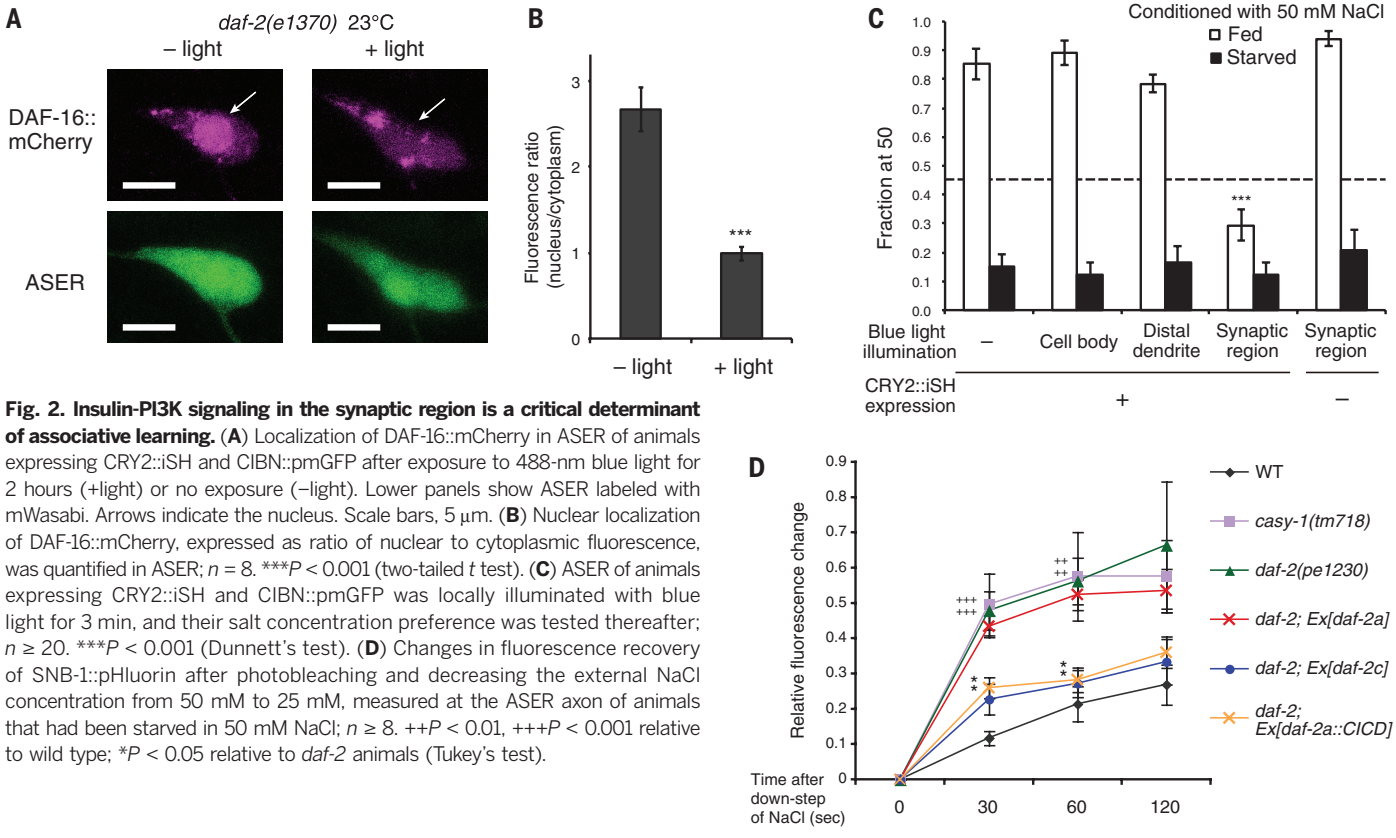


Fig. 2. Insulin-PI3K signaling in the synaptic region is a critical determinant of associative learning. (**A**) Localization of DAF-16::mCherry in ASER of animals expressing CRY2::iSH and CIBN::pmGFP after exposure to 488-nm blue light for 2 hours (+light) or no exposure (–light). Lower panels show ASER labeled with mWasabi. Arrows indicate the nucleus. Scale bars, 5 μm. (**B**) Nuclear localization of DAF-16::mCherry, expressed as ratio of nuclear to cytoplasmic fluorescence, was quantified in ASER; $n = 8$. ***$P < 0.001$ (two-tailed t test). (**C**) ASER of animals expressing CRY2::iSH and CIBN::pmGFP was locally illuminated with blue light for 3 min, and their salt concentration preference was tested thereafter; $n \geq 20$. ***$P < 0.001$ (Dunnett's test). (**D**) Changes in fluorescence recovery of SNB-1::pHluorin after photobleaching and decreasing the external NaCl concentration from 50 mM to 25 mM, measured at the ASER axon of animals that had been starved in 50 mM NaCl; $n \geq 8$. ++$P < 0.01$, +++$P < 0.001$ relative to wild type; *$P < 0.05$ relative to *daf-2* animals (Tukey's test).

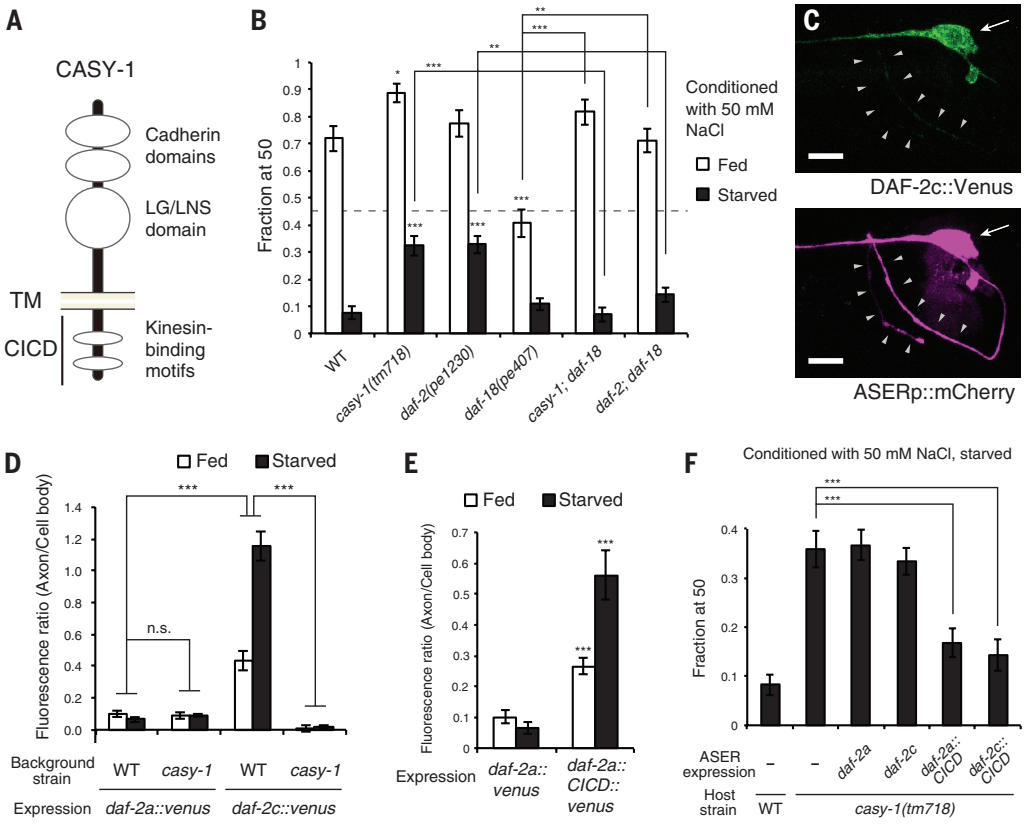


Fig. 3. Calsyntenin-dependent translocation to the synaptic region confers functional specificity to DAF-2c. (**A**) Predicted structure of CASY-1. (**B**) Salt concentration preference in *casy-1*, *daf-2*, *daf-18*, *casy-1;daf-18*, and *daf-2;daf-18* mutants; $n \geq 12$. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$ (Tukey's test). (**C**) Localization of DAF-2c in ASER of *casy-1(tm718)* mutants. Arrowheads and arrows indicate the axon and the cell body, respectively. Scale bars, 10 μm. (**D**) Axonal localization of DAF-2a/c::Venus, expressed as ratios of axonal to cell body fluorescence, was quantified in ASER of wild type and *casy-1(tm718)* mutants; $n \geq 12$. ***$P < 0.001$ (Tukey's test); n.s., not significant. (**E**) Axonal localization of DAF-2a fused with the intracellular domain (amino acids 872 to 984) of CASY-1 (DAF-2a::CICD); $n \geq 20$. ***$P < 0.001$ (two-tailed t test). (**F**) Salt concentration preference in *casy-1* animals that express DAF-2a::CICD or DAF-2c::CICD chimeric proteins in ASER; $n \geq 20$. ***$P < 0.001$ (Dunnett's test).

(*5*, *8*, *9*), which suggests that the roles of the additional exon might be conserved between *C. elegans* and mammals. In *C. elegans*, tissue-specific mRNA pull-down showed that the inclusion of exon 11.5 occurs frequently in the nervous system (Fig. 1C) (*5*).

We expressed the *daf-2a* or *daf-2c* isoform from single-copy transgenes that were integrated into the genome with *Mos1*-mediated single-copy insertion (MosSCI) (*10*). For taste avoidance learning, the expression of *daf-2c*, but not of *daf-2a*, was sufficient to rescue the defect of *daf-2* mutants (Fig. 1D and fig. S5, A and B). In contrast, for three other *daf-2*–regulated processes—the regulation of dauer formation, life span, and intrinsic thermotolerance—the expression of *daf-2c* was less effective than that of *daf-2a* (Fig. 1E, fig. S5C, and table S1).

We attempted to elucidate the basis of the functional difference between DAF-2a and DAF-2c and discovered that these two forms show different intracellular localizations. In the salt-sensing neuron ASER, which is essential for taste avoidance learning (fig. S6) (*4*, *5*), DAF-2a fused with Venus [a modified green fluorescent protein (GFP)] was mainly localized to the cell body (Fig. 1F). DAF-2c was strongly localized to the synapse-rich neurite that extends to the neuropil (called axon; Fig. 1, G and H) (*5*, *11*). The axonal localization of DAF-2c increased in response to starvation (Fig. 1I and fig. S7), whereas that of DAF-2a did not (Fig. 1I).

To explore whether axonal insulin-PI3K signaling plays a directive role in taste avoidance learning, we developed a worm version of the PI3K photoactivation system (fig. S8A) (*5*, *12*). When the PI3K photoactivation constructs (CRY2::iSH and CIBN::pmGFP) were expressed in the ASER neuron of *daf-2* mutants and blue light was applied, nuclear accumulation of DAF-16 [which is negatively regulated by the insulin-PI3K pathway (*13*)] was reduced (Fig. 2, A and B). When the axon of ASER was illuminated with blue light in animals expressing CRY2::iSH and CIBN::pmGFP, attraction to the salt concentration experienced with food was reduced, and avoidance was observed instead (Fig. 2C and fig. S8, B to E). This behavioral change was not observed when other parts of ASER, the cell body, or the dendritic neurite were illuminated, nor when other chemosensory neurons were illuminated after expression of CRY2::iSH and CIBN::pmGFP (Fig. 2C and fig. S8, B to G).

Considering these results, we monitored the effects of DAF-2a and DAF-2c upon the release

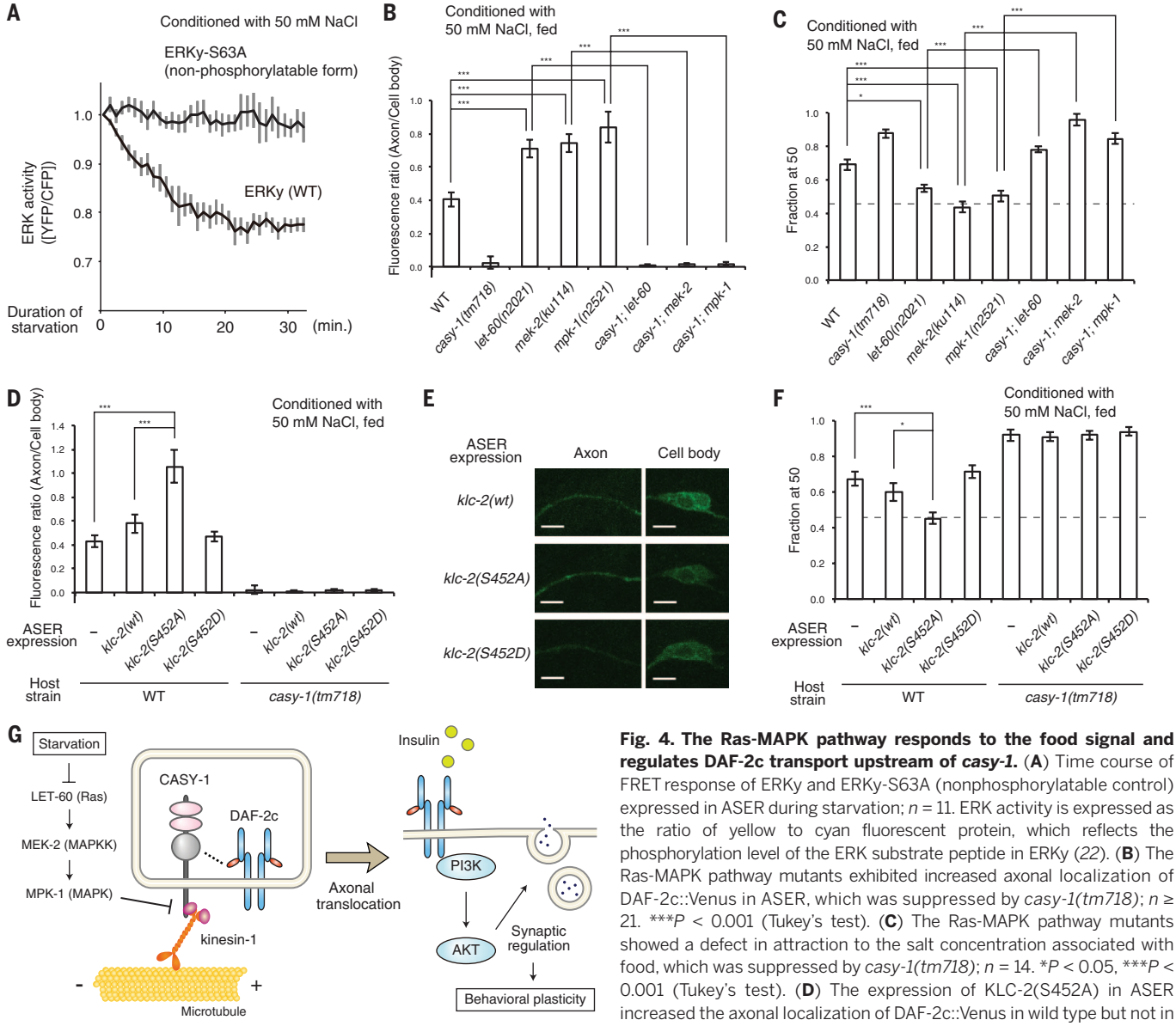


Fig. 4. The Ras-MAPK pathway responds to the food signal and regulates DAF-2c transport upstream of *casy-1*. (**A**) Time course of FRET response of ERKy and ERKy-S63A (nonphosphorylatable control) expressed in ASER during starvation; $n = 11$. ERK activity is expressed as the ratio of yellow to cyan fluorescent protein, which reflects the phosphorylation level of the ERK substrate peptide in ERKy (*22*). (**B**) The Ras-MAPK pathway mutants exhibited increased axonal localization of DAF-2c::Venus in ASER, which was suppressed by *casy-1(tm718)*; $n \geq 21$. $^{***}P < 0.001$ (Tukey's test). (**C**) The Ras-MAPK pathway mutants showed a defect in attraction to the salt concentration associated with food, which was suppressed by *casy-1(tm718)*; $n = 14$. $^{*}P < 0.05$, $^{***}P < 0.001$ (Tukey's test). (**D**) The expression of KLC-2(S452A) in ASER increased the axonal localization of DAF-2c::Venus in wild type but not in *casy-1(tm718)*; $n \geq 21$. $^{***}P < 0.001$ (Tukey's test). (**E**) Localization of DAF-2c::Venus in ASER of animals that express KLC-2, KLC-2(S452A), or KLC-2(S452D). Scale bars, 5 μm. (**F**) The expression of KLC-2(S452A) in ASER caused a defect in the attraction to the salt concentration associated with food in wild type but not in *casy-1(tm718)*; $n \geq 20$. $^{*}P < 0.05$, $^{***}P < 0.001$ (Tukey's test). (**G**) A model for the mechanism of the behavioral switch induced by starvation.

of synaptic vesicles from ASER using synaptopHluorin, which is a pH-sensitive derivative of GFP fused to SNB-1 (a *C. elegans* ortholog of synaptobrevin) (*14*–*16*). *daf-2* mutants exhibited increased synaptic release from ASER upon a stepwise decrease in the external NaCl concentration under starvation (Fig. 2D) (*16*). The expression of DAF-2c, but not that of DAF-2a, rescued the increased synaptic release of *daf-2* mutants (Fig. 2D).

Calsyntenins, also called alcadeins, have a *C. elegans* ortholog called CASY-1 (Fig. 3A). Calsyntenins are type I transmembrane proteins of the cadherin superfamily that have been implicated in several types of learning in humans and *C. elegans* (*17*–*19*), but their physiological roles are not well understood. *casy-1* mutants showed a defect in avoidance of previously experienced salt concentrations after starvation, similar to that seen in the insulin-PI3K pathway mutants (Fig. 3B and fig. S9, B to D) (*19*). This defect of *casy-1* in taste avoidance learning was suppressed by new mutant alleles of *daf-18*/PTEN (the PI3K-antagonizing phosphatase) that were identified in an unbiased genetic screen (Fig. 3B and fig. S9) (*5*). This result prompted us to examine the relationship between DAF-2 and CASY-1 and led us to find that DAF-2c was not localized to the axon but accumulated in the cell body in *casy-1* mutants (Fig. 3, C and D; compare with wild type in Fig. 1G). The localization of DAF-2a was unaffected (Fig. 3D), which suggests that CASY-1 is involved in the isoform-specific axonal localization of DAF-2. Consistent with these results, *casy-1* mutants showed increased synaptic transmission from the ASER axon, like that seen in *daf-2* mutants (Fig. 2D) (*5*). *casy-1* deletion increased the rescue effect of *daf-2c* in the regulation of dauer formation, life span, and intrinsic thermotolerance (Fig. 1E, fig. S5C, and table S1).

How does CASY-1 control the axonal localization of DAF-2c? Coimmunoprecipitation experiments showed that CASY-1 physically interacts in vivo with DAF-2c but not with DAF-2a (fig. S10A). When the exon 11.5–encoded segment, which distinguishes DAF-2c from DAF-2a, was fused with a signal peptide for secretion and then expressed, it was coprecipitated with CASY-1 (fig. S10B). On the other hand, the CASY-1 intracellular C-terminal domain (CICD) binds to the kinesin-1 light chain KLC-2 (*20*) in a yeast two-hybrid assay, like its mammalian orthologs (fig. S11A) (*5*, *21*). CASY-1 can thus function as a molecular linker for DAF-2c and the kinesin-1 motor. Consistent with this idea, the hypomorphic mutations or ASER-specific RNA interference (RNAi) knockdown of kinesin-1 motor components (*unc-116* and *klc-2*) caused a decrease in the axonal DAF-2c localization (fig. S11, B and C).

To determine whether the functional difference between DAF-2a and DAF-2c can be attributed to the difference in localization, we attempted to ectopically localize DAF-2a to the axons. CICD, which binds to KLC-2 and is expected to be transported to the axon, was C-terminally fused with DAF-2a (DAF-2a::CICD). DAF-2a::CICD was more localized to the axon than was wild-type DAF-2a (Fig. 3E). The expression of this chimera protein restored the learning defect of *casy-1* and *daf-2* mutants (Fig. 3F and fig. S12, A to C). The rescue effect of DAF-2c::CICD was comparable to that of DAF-2a::CICD (Fig. 3F and fig. S12, A and B). These results support the following two concepts: (i) DAF-2a contributes to taste avoidance learning if it is forcibly localized to the axon (*5*); therefore, the functional difference between DAF-2a and DAF-2c is mainly due to the difference in axonal localization. (ii) The forced localization of DAF-2 to the axon rescues the learning defect of *casy-1*; therefore, the learning defect of *casy-1* is caused by the deficit in the axonal localization of DAF-2c. The expression of DAF-2a::CICD also rescued the increased synaptic release of *daf-2* mutants, further suggesting that axonal DAF-2 regulates synaptic transmission (Fig. 2D). Quantitative relationships between axonally localized DAF-2 protein and salt preference behavior supported the key roles of axonal localization of this receptor (fig. S12E) (*5*).

We monitored the activity of extracellular signal–regulated kinase (ERK)/mitogen-activated protein kinase (MAPK) in ASER during starvation with a FRET-based probe, ERKy (*22*), and found that the ERK activity in ASER gradually decreases after food deprivation at a constant NaCl concentration of 50 mM (Fig. 4A). A biochemical study has shown that ERK phosphorylates the mammalian kinesin light chain KLC1 and inhibits the binding of KLC1 to the intracellular domain of calsyntenins (*23*). The sequence context of this phosphorylation site is almost perfectly conserved in *C. elegans* KLC-2 (*5*). In fact, ERK phosphorylated recombinant KLC-2, but not KLC-2(S452A), a mutant form in which the putative phosphorylation site Ser^{452} was mutated to alanine (fig. S13).

We therefore investigated whether ERK regulates the CASY-1–dependent axonal localization of DAF-2c and salt chemotaxis in vivo. Hypomorphic mutations or ASER-specific RNAi knockdown of each of the Ras-MAPK pathway genes *let-60*, *mek-2*, and *mpk-1* (encoding Ras, ERK kinase, and ERK, respectively) increased axonal localization of DAF-2c and reduced worm attraction to the salt concentration associated with food (Fig. 4, B and C, and fig. S14, A to G). The deletion of *casy-1* suppressed these phenotypes (Fig. 4, B and C, and fig. S14, A, B, and H to J), which suggests that CASY-1 is essential for axonal transport of DAF-2c even in the absence of MAPK. The forced ASER expression of KLC-2 (S452A) significantly increased the axonal localization of DAF-2c (Fig. 4, D and E) and decreased the attraction of the previous salt concentration that was experienced with food (Fig. 4F and fig. S14, K and L), whereas the expression of a phosphomimic form, KLC-2(S452D), had no effect (Fig. 4, D and E, and fig. S14, K and L). In the *casy-1* mutant background, the expression of KLC-2(S452A) did not alter DAF-2c localization or salt chemotaxis (Fig. 4, D and F, and fig. S14, K and L). Taken together, these findings indicate that the Ras-MAPK pathway is a candidate for transmitting a food signal and for negatively regulating axonal localization of DAF-2c upstream of CASY-1 and KLC-2.

Our data demonstrate that the CASY-1 and kinesin-1 complex that is regulated by the Ras-MAPK pathway confers functional specificity to the insulin receptor long isoform through rapid translocation to the synaptic region (Fig. 4G). The mammalian insulin receptor also localizes to the synaptic region of neurons (*24*), and both insulin signaling and calsyntenins have been considered relevant to memory performance and Alzheimer's disease in humans (*1*, *2*, *17*, *25*). Our findings may lead to further understanding of the mechanisms underlying human memory formation and dementia.

REFERENCES AND NOTES

1. S.-L. Chiu, H. T. Cline, *Neural Dev.* **5**, 7 (2010).
2. H. B. Schiöth, S. Craft, S. J. Brooks, W. H. Frey 2nd, C. Benedict, *Mol. Neurobiol.* **46**, 4–10 (2012).
3. C. A. Wolkow, K. D. Kimura, M.-S. Lee, G. Ruvkun, *Science* **290**, 147–150 (2000).
4. H. Kunitomo *et al*., *Nat. Commun.* **4**, 2210 (2013).
5. See supplementary materials on *Science* Online.
6. A. Cornils, M. Gloeck, Z. Chen, Y. Zhang, J. Alcedo, *Development* **138**, 1183–1193 (2011).
7. J. M. A. Tullet *et al*., *Cell* **132**, 1025–1038 (2008).
8. A. Belfiore, F. Frasca, G. Pandini, L. Sciacca, R. Vigneri, *Endocr. Rev.* **30**, 586–623 (2009).
9. D. S. Patel *et al*., *Genetics* **178**, 931–946 (2008).
10. C. Frøkjær-Jensen, M. W. Davis, M. Ailion, E. M. Jorgensen, *Nat. Methods* **9**, 117–118 (2012).
11. J. G. White, E. Southgate, J. N. Thomson, S. Brenner, *Philos. Trans. R. Soc. London Ser. B* **314**, 1–340 (1986).
12. O. Idevall-Hagren, E. J. Dickson, B. Hille, D. K. Toomre, P. De Camilli, *Proc. Natl. Acad. Sci. U.S.A.* **109**, E2316–E2323 (2012).
13. K. Lin, H. Hsin, N. Libina, C. Kenyon, *Nat. Genet.* **28**, 139–145 (2001).
14. S. Sankaranarayanan, D. De Angelis, J. E. Rothman, T. A. Ryan, *Biophys. J.* **79**, 2199–2208 (2000).
15. A. D. Samuel, R. A. Silva, V. N. Murthy, *J. Neurosci.* **23**, 373–376 (2003).
16. S. Oda, M. Tomioka, Y. Iino, *J. Neurophysiol.* **106**, 301–308 (2011).
17. C. Preuschhof *et al*., *Neuropsychologia* **48**, 402–408 (2010).
18. F. J. Hoerndli *et al*., *PLOS ONE* **4**, e4880 (2009).
19. D. D. Ikeda *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **105**, 5260–5265 (2008).
20. R. Sakamoto *et al*., *Mol. Biol. Cell* **16**, 483–496 (2005).
21. Y. Araki *et al*., *EMBO J.* **26**, 1475–1486 (2007).
22. T. Tomida, S. Oda, M. Takekawa, Y. Iino, H. Saito, *Sci. Signal.* **5**, ra76 (2012).
23. A. Vagnoni, L. Rodriguez, C. Manser, K. J. De Vos, C. C. J. Miller, *J. Cell Sci.* **124**, 1032–1042 (2011).
24. M. A. Abbott, D. G. Wells, J. R. Fallon, *J. Neurosci.* **19**, 7300–7308 (1999).
25. A. Vagnoni *et al*., *Hum. Mol. Genet.* **21**, 2845–2854 (2012).

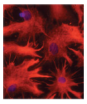
ACKNOWLEDGMENTS

We thank the Caenorhabditis Genetics Center (CGC), the National Bioresource Project, R. Iwata, O. Hobert, K. Matsumoto, M. Han, G. Ruvkun, M. Doi, and S. Oda for strains; N. Chronis for the design of the olfactory chip; K. D. Kimura and G. Ruvkun for *daf-2a* cDNA; and K. Shen, C. Bargmann, K. Takeo, and members of the Iino lab for thoughtful comments and discussions. The primary data of this study are archived in the Iino laboratory in the Department of Biological Sciences, University of Tokyo, and are available upon request. Y.I. was supported by a Grant-in-Aid for Innovative Areas "Systems molecular ethology" (20115002) and "Memory dynamism" (25115010) and for Scientific Research (B) (21370002), and M.T. was supported by a Grant-in-Aid for Innovative Areas "RNA regulation" (23112704), from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/345/6194/313/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S21
Tables S1 to S3
References (*26*–*48*)

10 January 2014; accepted 13 June 2014
10.1126/science.1250709

EXTENDED PDF FORMAT
SPONSORED BY

Travel Grants Available
Neuroscience 2014
Application Deadline: July 31st
>Apply Now
R&D systems
a bio-techne brand
www.rndsystems.com

Thermohaline circulation crisis and impacts during the mid-Pleistocene transition

Leopoldo D. Pena and Steven L. Goldstein

Science **345**, 318 (2014);
DOI: 10.1126/science.1249770

This copy is for your personal, non-commercial use only.

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by clicking here.

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines here.

The following resources related to this article are available online at www.sciencemag.org (this information is current as of August 4, 2014):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:
http://www.sciencemag.org/content/345/6194/318.full.html

Supporting Online Material can be found at:
http://www.sciencemag.org/content/suppl/2014/06/25/science.1249770.DC1.html

This article **cites 46 articles**, 9 of which can be accessed free:
http://www.sciencemag.org/content/345/6194/318.full.html#ref-list-1

This article appears in the following **subject collections:**
Oceanography
http://www.sciencemag.org/cgi/collection/oceans

Downloaded from www.sciencemag.org on August 4, 2014

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published weekly, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. Copyright 2014 by the American Association for the Advancement of Science; all rights reserved. The title *Science* is a registered trademark of AAAS.

PALEOCEANOGRAPHY

Thermohaline circulation crisis and impacts during the mid-Pleistocene transition

Leopoldo D. Pena[1]* and Steven L. Goldstein[1,2]

The mid-Pleistocene transition (MPT) marked a fundamental change in glacial-interglacial periodicity, when it increased from ~41-thousand-year to 100-thousand-year cycles and developed higher-amplitude climate variability without substantial changes in the Milankovitch forcing. Here, we document, by using Nd isotopes, a major disruption of the ocean thermohaline circulation (THC) system during the MPT between marine isotope stages (MISs) 25 and 21 at ~950 to 860 thousand years ago, which effectively marks the first 100-thousand-year cycle, including an exceptional weakening through a critical interglacial (MIS 23) at ~900 thousand years ago. Its recovery into the post-MPT 100-thousand-year world is characterized by continued weak glacial THC. The MPT ocean circulation crisis facilitated the coeval drawdown of atmospheric CO_2 and high-latitude ice sheet growth, generating the conditions that stabilized 100-thousand-year cycles.

The causes of the mid-Pleistocene transition (MPT) from 41- to 100-thousand-year (ky) glacial-interglacial periodicity and more intense glacials (*1*, *2*) are not well understood. Because the astronomical drivers of climate variability remained the same (*3*), there must have been primary changes in the ocean-atmosphere-continental ice system (*4*–*8*). We used neodymium (Nd) isotopes in South Atlantic deep-sea cores to trace North Atlantic Deep Water (NADW) export to the Southern Ocean, which document a major shift in the ocean thermohaline circulation (THC) system (*9*) between marine isotope stages (MISs) 25 and 21 at ~950 to 860 thousand years ago (ka) (*6*). During the 41-ky world covered by this study (MISs 31 to 25 at 1100 to 950 ka), differences in glacial-interglacial THC vigor were much smaller than in the 100-ky world (MISs 21 to 15 at 860 to 600 ka), which is characterized by much weaker glacial THC. The important interval includes the first major THC weakening during MISs 24 to 22 and MIS 23 as the only known interglacial during which the THC system did not strengthen. The weak THC stabilized this first 100-ky glacial cycle, which helped to drive the climate system across the threshold to continued 100-ky cycles.

Hypotheses to explain what caused the MPT [e.g., (*7*, *10*, *11*) and references therein] include nonlinear climate responses to Milankovitch forcing [e.g., (*12*, *13*)], cooling of sea-surface temperatures and increased sea-ice cover (*11*, *14*, *15*), changes in THC vigor inferred from benthic foraminiferal carbon isotopes (*10*), long-term or sudden decrease of atmospheric partial pressure of CO_2 (PCO_2) (*16*), and stabilization of ice sheets from loss of underlying regolith (*5*, *7*). A point of consensus is that decreased PCO_2 was important for enhancing cooling and facilitating ice-sheet growth, and modeling studies conclude that PCO_2 reduction during glacial periods may be the result of decreased ocean ventilation, increased carbonate alkalinity, terrestrial carbon storage, and increased marine productivity and that the deep ocean's storage capacity would be enhanced by a weaker THC (*17*–*21*) (supplementary materials).

The utility of Nd isotopes as a paleo-ocean circulation tracer reflects the short Nd ocean residence time [500 to 1000 years, e.g., (*22*)], which is less than the mixing time. As a result, Nd isotope ratios (presented as ε_{Nd}, defined in the tables S1 and S2 caption) vary in space and time. ε_{Nd} values in seawater and marine precipitates are not substantially fractionated by biological and physical processes (unlike light stable isotope $\delta^{13}C$ and $\delta^{18}O$) and reflect radioactive decay of ^{147}Sm over Earth's history. In seawater, ε_{Nd} values behave quasi-conservatively, that is, approximating the mixture of end-member deep-water masses even far from the source areas (*23*). Over the past decade or so, studies have shown that ε_{Nd} in Fe-Mn oxide precipitates in deep-sea cores can be used to monitor THC variability [e.g., (*24*)]. A key observation for applying ε_{Nd} as a paleo-THC tracer is that ε_{Nd} values have remained close to ~–4 and ~–13 in the deep North Pacific (NPW) and North Atlantic end-members, respectively, at least over the past ~2 million years [e.g., (*25*–*27*) and figs. S1 and S2]. These differences reflect erosion of old continent into the North Atlantic and a substantial young volcanic component into the Pacific.

We report ε_{Nd} on Fe-Mn oxide encrusted foraminifera samples from Ocean Drilling Project (ODP) sites 1088 (41°8.16′S, 13°33.77′E, 2082 m) and 1090 (42°54.82′S, 8°53.98E′, 3702 m) in the Cape Basin, southeast Atlantic, through the MPT (tables S1 and S2). We sampled glacial and interglacial peaks between MISs 31 and 15 (~1100 to 600 ka) in order to capture first-order changes in the THC system across the MPT; the stratigraphy is based on benthic foraminiferal oxygen isotopes ($\delta^{18}O_b$) in these cores (*21*, *28*) (Fig. 2). Both sites are in the sub-Antarctic zone, near the upper and lower boundaries between NADW and Circumpolar Deep Water (CDW) (Fig. 1). At this latitude in the southeast Atlantic, the high-salinity NADW core today has ε_{Nd} ~–10.5, and at the ODP core site depths ε_{Nd} values are ~–10 (*29*). As NADW moves southward, it mixes with CDW, which today has ε_{Nd} ~ –8.5 (*29*). By using the ε_{Nd} and Nd concentrations in NADW and deep NPW (Fig. 3 caption), the mixing percentages are ~75:25

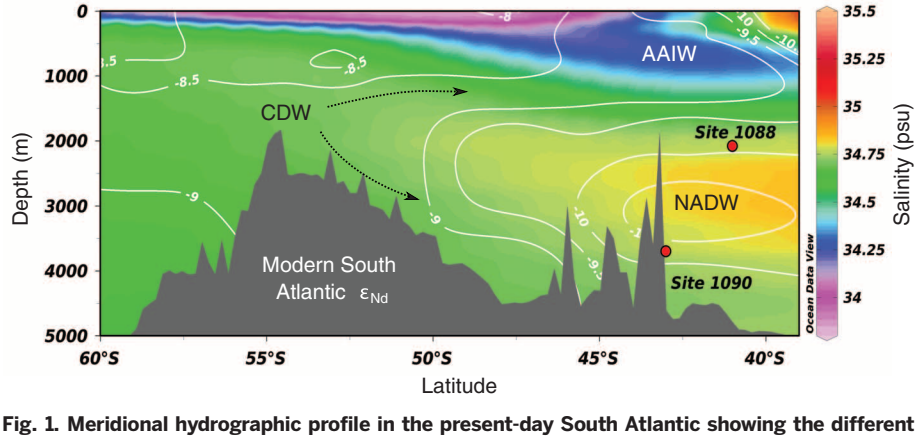


Fig. 1. Meridional hydrographic profile in the present-day South Atlantic showing the different water masses based on the salinity distribution (World Ocean Atlas). ODP sites 1088 and 1090 are located at the upper and lower boundaries of the southward-flowing NADW wedge, which is partly mixed with northward-flowing CDW and Antarctic Intermediate Water (AAIW). White isolines correspond to the modern South Atlantic seawater ε_{Nd} (*29*) and show that NADW has more-negative ε_{Nd} (–10 to –10.5), whereas present-day CDW and AAIW have higher ε_{Nd} (–8.5 to –9), which reflects the mixing of Atlantic and Pacific source waters (*23*). psu, practical salinity unit.

[1]Lamont-Doherty Earth Observatory of Columbia University, 61 Route 9W, Palisades, NY 10964, USA. [2]Department of Earth and Environmental Sciences, Columbia University, 61 Route 9W, Palisades, NY 10964, USA.
*Corresponding author. E-mail: leopoldo@ldeo.columbia.edu

NADW:NPW in Cape Basin NADW and ~60:40 NADW:NPW in CDW. Given the stability of ε_{Nd} values in deep North Atlantic and North Pacific end-members through the MPT interval (fig. S1), we interpret our sample ε_{Nd} values to reflect changes in the end-member mixing fractions. Reductions in the southward penetration of NADW appear as more-positive ε_{Nd} values. Because Nd isotopes cannot distinguish between different North Pacific– or North Atlantic–sourced waters having the same end-member values [e.g., NADW today and glacial North Atlantic intermediate water (GNAIW) during recent glacials (*30*)], we refer to these water masses as North Pacific– or North Atlantic–sourced waters (NPSW and NASW, respectively).

During the studied pre-MPT interval (MISs 31 to 25, 1100 to 950 ka) at the mid-depth site, the differences in glacial and interglacial ε_{Nd} values are small (Figs. 2A and 3A). During glacials, ε_{Nd} values are only slightly more positive than during interglacials (−8.96 ± 0.19 versus −9.40 ± 0.38, respectively, 1 SD, n = 3). The NASW mixing percentages (Fig. 3A and table S1) during interglacials are similar to those of the modern Cape Basin (~72 ± 3%) and are only slightly lower during glacials (~68 ± 2%), consistent with strong glacial export of NASW. The deep-ocean site (Figs. 2B and 3B) displays a similar pattern but with pre-MPT glacials and interglacials showing larger ε_{Nd} differences than the mid-depth site (−8.45 ± 0.15 versus −9.55 ± 0.33, respectively, 1 SD, n = 3 for glacials, n = 4 for interglacials), indicating slightly higher interglacial and lower glacial NASW mixing percentages (Fig. 3C and table S2) than the mid-depth site (~73 ± 3% versus 62 ± 2%, respectively). Thus, ε_{Nd} values in both sites indicate notable NASW presence in the deep South Atlantic during pre-MPT interglacials and glacials, and in particular the mid-depth site documents sustained export of NASW and a vigorous THC even during glacials (Fig. 4, A and B).

The post-MPT regime (MISs 21 to 15, 860 to 600 ka) shows major differences in the THC mode during glacials compared with the 41-ky world (Figs. 2 to 4). Both mid-depth and deep sites record large shifts toward more-positive glacial ε_{Nd} values [−7.77 ± 0.29 and −7.39 ± 0.05, respectively, 1 SD, n = 4; versus pre-MPT values of ~−9.0 and −8.5, respectively (Fig. 2)], indicating much weaker glacial NASW export in the 100-ky world [~53 ± 4% and ~49 ± 1%, respectively, versus ~68 ± 2% and ~62 ± 2%, respectively, during pre-MPT glacials (Fig. 3A)]. In contrast to glacials, the ε_{Nd} values during interglacials are essentially the same before and after the MPT at both mid-depth and deep sites (remaining at ~−9.4; Figs. 2 and 3). The data indicate similar NASW export to the Southern Ocean during pre- and post-MPT interglacials, whereas it has been substantially reduced during post-MPT glacials (Figs. 3 and 4).

The Nd data underscore the importance of MISs 25 to 21 as an interval of major transition in the THC system (Figs. 2 and 3). At the mid-depth site, the change from strong pre-MPT glacial THC to weak post-MPT THC occurs suddenly in MIS 24, with a large ε_{Nd} increase (from ~−9 during MIS 26 to ~−7 during MIS 24), whereas in the deep site the change is more gradual, reaching ε_{Nd} ~−7 at MIS 22 (Fig. 2C). During this period, MIS 23 at ~900 ka (*6*) stands out as an exceptional interglacial (Figs. 2 and 3) during which ε_{Nd} values did not shift to more-negative values, indicating a stronger THC. Rather, the mid-depth and deep sites show similarly high ε_{Nd} values like the preceding glacial MIS 24 (Fig. 2). This observation at both sites precludes any sampling artifact (supplementary materials). MIS 23 is thus unique as the only known interglacial in which the THC did not strengthen, not even to pre-MPT glacial levels. Rather, it operated in the same weak mode as post-MPT glacials and thus can be considered a transglacial (a glacial-like interglacial), with ε_{Nd} values of −7.1 and −8.0 at the mid-depth and the deep sites, respectively, corresponding to NASW end-member mixing percentages of only ~46 ± 2% and ~58 ± 3%, respectively (Fig. 3, A and C). By using the present day as an analog, we can estimate the equivalent flux

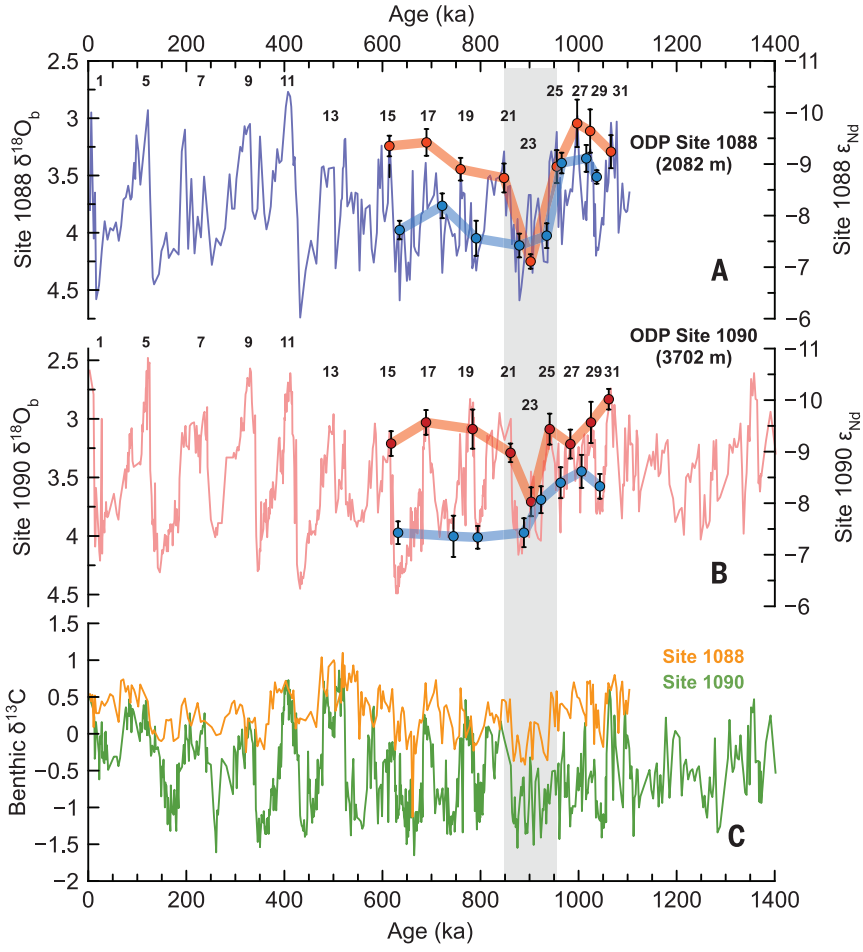


Fig. 2. Glacial-interglacial ε_{Nd} in the South Atlantic (Cape Basin) at mid-depth (site 1088) and deep (site 1090) sites through the MPT (~1100 to 600 ka). Red and blue lines connect data for samples from interglacial and glacial maximum, respectively. Error bars indicate the propagated error (2σ) of the ε_{Nd} measurements. (**A** and **B**) ODP site 1088 (purple) and site 1090 (pink) benthic foraminiferal $\delta^{18}O$ records (*21*, *28*). MIS numbers of interglacials are listed. (**C**) Sites 1088 (yellow) and 1090 (green) benthic $\delta^{13}C$ isotope records plotted in the same scale, showing the small $\delta^{13}C$ amplitude at site 1088 compared with 1090. The vertical gray bar indicates the time interval between MISs 24 and 22, when a drastic change in the ε_{Nd} values took place at both sites, linked to a major weakening of NASW advection into the South Atlantic.

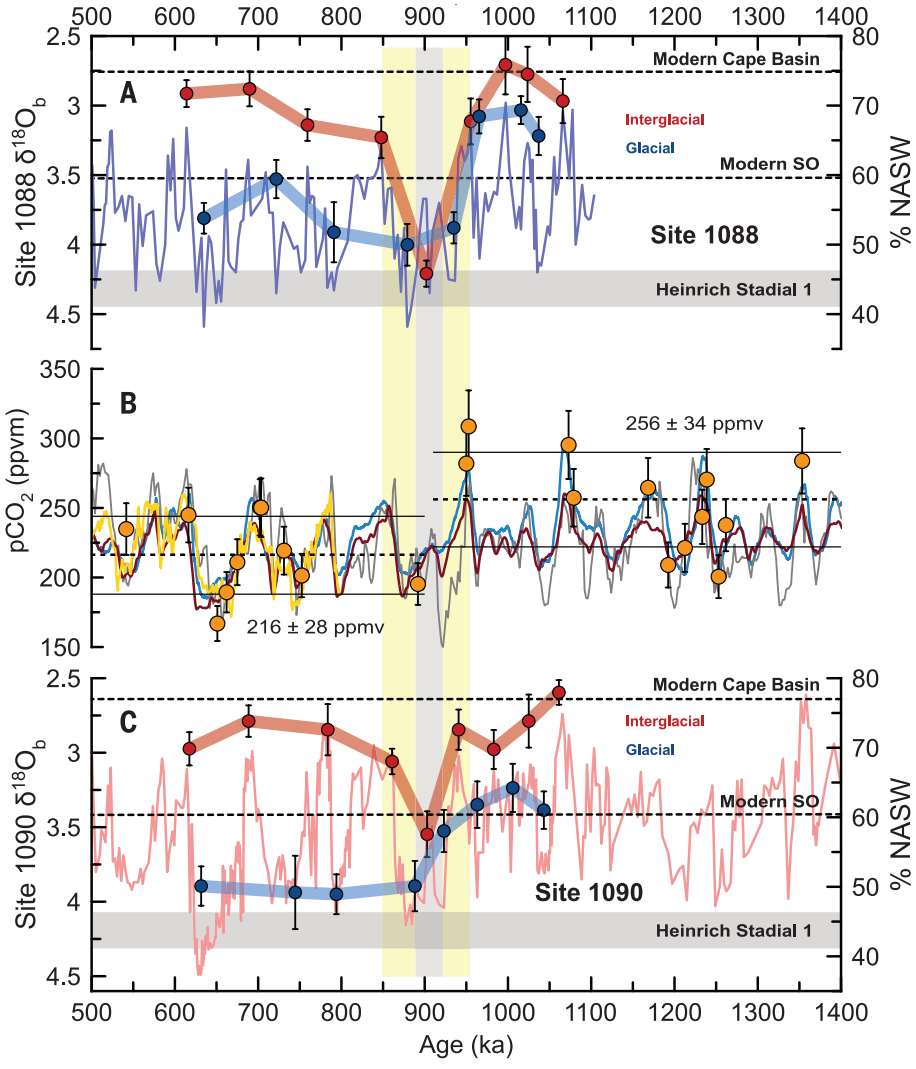


Fig. 3. Estimation of the mixing proportions of NASW and NPSW water masses during glacial (blue) and interglacial (red) time periods using ε_{Nd} and Nd concentrations and assuming binary mixing for (A) sites 1088 and (C) 1090. The Atlantic-Pacific mixing curve is based on the following parameters: for NADW, ε_{Nd} = –13 and [Nd] = 22 pmol/kg; for NPW, ε_{Nd} = –4.5 and [Nd] = 42 pmol/kg (*23*). The horizontal dashed lines and gray bar correspond to modern mixing ratios for the Cape Basin (ε_{Nd} = –10.3), Southern Ocean (ε_{Nd} = –9), and the very high ε_{Nd} values in the South Atlantic during Heinrich events (ε_{Nd} = –6.9 to –7.3) (*24*). (**B**) Yellow circles correspond to atmospheric P_{CO_2} estimates based on B-isotope reconstructions (*16*) and compared with ice-core atmospheric P_{CO_2} (*35*). Dashed lines indicate pre- and post-MPT P_{CO_2} averages (2σ deviations), indicating a net drop of 40 parts per million by volume (ppmv), mostly driven by glacial P_{CO_2} drop (*16*). Also shown are P_{CO_2} reconstructions based on benthic foraminiferal $\delta^{13}C$ (gray) (*36*), and two model outputs based on the LR04 stack [light blue and magenta (*19*, *32*)]. Error bars indicate the propagated ε_{Nd} error (2σ) into the mixing calculation (supplementary materials).

drop of NASW (details in supplementary materials) on the basis of the mixing percentages and the present-day NADW flux of 15 sverdrups (*31*). The flux drop is equivalent to reductions in NASW southward export of ~11 ± 2 and ~8 ± 2 sverdrups for the mid-depth and deep sites, respectively, and is similar to low NASW mixing percentages in the South Atlantic during recent Heinrich events (~42 to 46%) (*24*), which cause especially weak THC (Fig. 3).

The ε_{Nd} data point to a major role for the THC in facilitating the MPT during the critical period between MISs 25 and 21 and document a major reduction in NASW export during MIS 24. Such a THC change would most likely originate in the North Atlantic and thus indicates that an important threshold was reached in the North Atlantic region that resulted in an unprecedented weakening of the THC compared with earlier glacials, likely associated with increased volume and/or stability of the continental ice sheets [e.g. (*5*, *7*)]. The MIS 24 THC weakening predated the main phase of ice sheet growth during MISs 24 to 22 that was recently documented in a $\delta^{18}O_b$-based MPT sea-level reconstruction by Elderfield *et al.* (*8*). That study noted that Southern Hemisphere summer insolation during MIS 23 was anomalously low (fig. S3) and concluded that the increase in continental ice volume (corresponding to a ~120-m sea level decrease over the entire MIS 24 to 22 interval) took place primarily in Antarctica (fig. S3). Our Nd data indicate even lower NASW export during MIS 23 and 22 than MIS 24 (Fig. 3A), indicating substantial impacts also on North Atlantic ice sheets during this interval. The timing of the decrease in the glacial THC vigor also corresponds to the drawdown in glacial atmospheric P_{CO_2}, shown to have occurred abruptly also during between MISs 25 and 23 in recent proxy data and modeling studies, which further indicates similar interglacial P_{CO_2} levels during 41- and 100-ky glacial cycles and much lower glacial P_{CO_2} in the 100-ky world (*16*, *19*, *32*) (Fig. 3). The observation that post-MPT glacials are characterized by decreased THC vigor and reduced atmospheric P_{CO_2} compared with pre-MPT glacials, whereas the differences are much smaller for pre- and post-MPT interglacials, strongly indicates a linkage. During the critical interval, the weak THC facilitated drawdown of atmospheric CO_2 by reducing exchange between Antarctic surface and deep waters [e.g., (*21*)], which in turn resulted in further cooling. As a corollary effect, in the 100-ky world, glacial periods have been characterized by weak THC (Fig. 4, A and C), which facilitates reduced atmospheric P_{CO_2} and more intense glacials.

Studies of foraminiferal $\delta^{13}C_b$ have been fundamental for understanding past ocean circulation [e.g., (*30*)] and used as evidence for THC variability through the MPT [e.g., (*10*)]. At the mid-depth site, $\delta^{13}C_b$ values are always heavier (more NASW-like) than the deep site, whereas the deep site shows greater glacial-interglacial variability (*21*) (Fig. 2). The $\delta^{13}C_b$ data contrast with $\delta^{18}O_b$, which displays similar amplitudes in both sites in response to global changes in deep-ocean temperature and ice volume. Moreover, $\delta^{13}C_b$ in the deep site cannot be used to quantify THC changes because glacial values overshoot the Pacific end-member during glacial periods (*21*), indicating disruption of the deep $\delta^{13}C_b$ THC signal by processes such as air-sea exchange and biological productivity [e.g., (*33*, *34*)]. The high glacial $\delta^{13}C_b$ gradient (Fig. 2) is consistent with greater glacial water column stratification and provides a means to store CO_2 in the deep ocean, which would enhance glacial cooling (*21*).

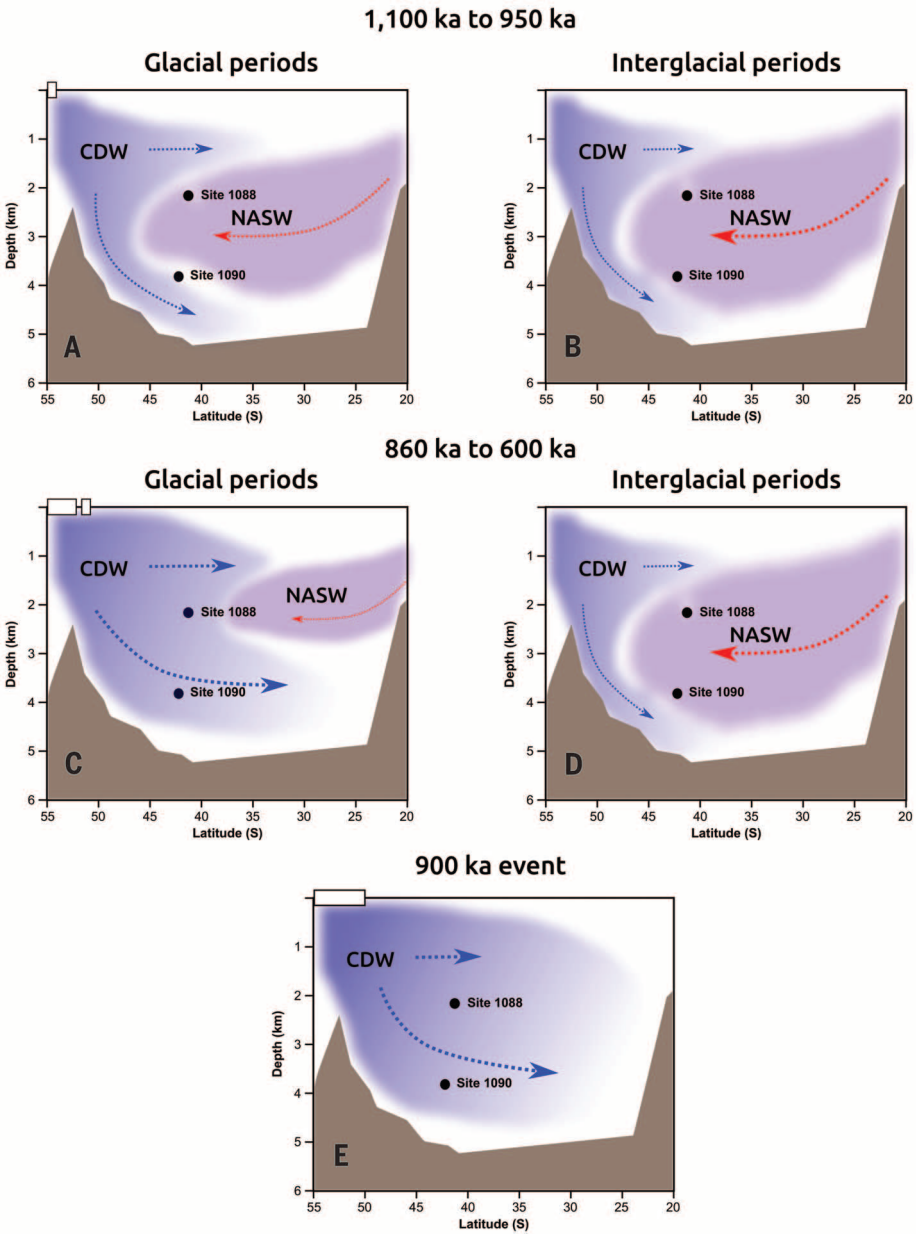


Fig. 4. Schematic illustrations of the changes in South Atlantic THC across the MPT. The pre-MPT time period of MISs 31 to 25 (~1100 to 950 ka) is characterized by substantial export of NASW to the South Atlantic during both (**A**) glacials and (**B**) interglacials. ε_{Nd} values indicate little glacial-interglacial variability at mid-depth (site 1088) and important NASW contributions during glacials. In the deep ocean (site 1090), the difference between glacials and interglacials is larger, but pre-MPT glacials show substantially more NASW than post-MPT glacials. The post-MPT time period since MISs 21 to 15 (~860 to 600 ka) is characterized by (**C**) strong THC during interglacials, like the pre-MPT period, but (**D**) weak THC during glacials. (**E**) The MIS 24 to 22 interval, centered at ~900 ka, shows an unprecedented weakening of NASW, similar to post-MPT glacials. MIS 23 is unique as an interglacial with THC like post-MPT glacial periods.

In contrast to $\delta^{13}C_b$, the ε_{Nd} values at both sites remain between the NASW and NPSW end-members and may clarify processes affecting $\delta^{13}C_b$. For example, GNAIW has been proposed to explain higher $\delta^{13}C_b$ values at mid-depths (*21*, *30*), and this is in agreement with our ε_{Nd} data, where the mid-depth site systematically shows more NASW-like values than the deep site during glacials (Figs. 2 and 3, excepting MIS 24 when the THC crisis began), thus supporting the presence of partially ventilated NASW-derived intermediate waters. During interglacials, the deep site shows stronger NASW-like character (tables S1 and S2), likely reflecting deeper interglacial convection of NADW.

We conclude that the MPT was marked by a drastic change in the deep-ocean THC during glacials, which shifted from an active pre-MPT glacial-THC mode to a reduced post-MPT glacial-THC mode. In contrast, the vigor of the THC pre- and post-MPT has remained similar during interglacials. A variety of observations document a series of events that fed back upon each other during the first 100-ky glacial cycle between MISs 25 and 21 (~950 to 850 ka) and led to permanent increases in the length of glacial-interglacial cycles and the intensity of glacials. The North Atlantic ice sheets reached a milestone in size and/or stability that resulted in substantial reduction of THC vigor during MIS 24 (Fig. 2). The weak THC in turn facilitated atmospheric CO_2 drawdown (*16*) (Fig. 3) by reducing exchange between Antarctic surface and deep waters, which in turn resulted in further cooling. The lower temperatures facilitated ice sheet expansion in Antarctica (*8*) as well as the North Atlantic region, because the weak export of NASW, typical of post-MPT glacials, continued through the critical transglacial MIS 23. These impacts were amplified by anomalously low Southern Hemisphere summer insolation during MIS 23, resulting in suppressed interglacial melting of Antarctic continental ice, likely amplified by sustained sea-ice cover around the Southern Ocean (*14*). The cold glacial conditions and weak THC continued through MIS 22. Together, these effects facilitated major ice-sheet expansion through the MIS 24 to 22 interval and generated conditions for crossing the threshold to further 100-ky glacial-interglacial cycles and colder glacials.

REFERENCES AND NOTES

1. J. D. Hays, J. Imbrie, N. J. Shackleton, *Science* **194**, 1121–1132 (1976).
2. N. J. Shackleton, *Science* **289**, 1897–1902 (2000).
3. N. G. Pisias, T. C. Moore Jr., *Earth Planet. Sci. Lett.* **52**, 450–458 (1981).
4. W. H. Berger, E. Jansen, in *The Polar Oceans and Their Role in Shaping the Global Environment*, O. M. Johannessen,

R. D. Muench, J. E. Overland, Eds. (Geophysical Monograph Series, American Geophysical Union, Washington, DC, 1994), vol. 85, pp. 295–311.
5. P. U. Clark, D. Pollard, *Paleoceanography* **13**, 1–9 (1998).
6. L. Lisiecki, M. E. Raymo, *Paleoceanography* **20**, PA1003 (2005).
7. P. U. Clark *et al*., *Quat. Sci. Rev.* **25**, 3150–3184 (2006).
8. H. Elderfield *et al*., *Science* **337**, 704–709 (2012).
9. W. S. Broecker, *Oceanography* **4**, 79–89 (1991).
10. M. E. Raymo, W. F. Ruddiman, N. J. Shackleton, D. W. Oppo, *Earth Planet. Sci. Lett.* **97**, 353–368 (1990).
11. E. L. McClymont, S. M. Sosdian, A. Rosell-Melé, Y. Rosenthal, *Earth Sci. Rev.* **123**, 173–193 (2013).
12. M. E. Raymo, *Paleoceanography* **12**, 577–585 (1997).
13. P. Huybers, *Quat. Sci. Rev.* **26**, 37–55 (2007).
14. H. Gildor, E. Tziperman, *Paleoceanography* **15**, 605–615 (2000).
15. A. Martínez-Garcia, A. Rosell-Melé, E. L. McClymont, R. Gersonde, G. H. Haug, *Science* **328**, 1550–1553 (2010).
16. B. Hönisch, N. G. Hemming, D. Archer, M. Siddall, J. F. McManus, *Science* **324**, 1551–1554 (2009).
17. J. R. Toggweiler, *Paleoceanography* **14**, 571–588 (1999).
18. J. Schmitt *et al*., *Science* **336**, 711–714 (2012).
19. P. Köhler, R. Bintanja, *Clim. Past* **4**, 311–332 (2008).
20. P. Köhler, H. Fischer, J. Schmitt, *Paleoceanography* **25**, PA1213 (2010).
21. D. A. Hodell, K. A. Venz, C. D. Charles, U. S. Ninnemann, *Geochem. Geophys. Geosyst.* **4**, 1–19 (2003).
22. K. Tachikawa, C. Jeandel, M. Roy-Barman, *Earth Planet. Sci. Lett.* **170**, 433–446 (1999).
23. S. L. Goldstein, S. R. Hemming, "Long-lived isotopic tracers in oceanography, paleoceanography and ice sheet dynamics," in *Treatise on Geochemistry* (Elsevier, New York, 2003), vol. 6, pp. 453–498.
24. A. M. Piotrowski, S. L. Goldstein, S. R. Hemming, R. G. Fairbanks, *Science* **307**, 1933–1938 (2005).
25. W. Abouchami, S. L. Goldstein, S. J. G. Gazer, A. Eisenhauer, A. Mangini, *Geochim. Cosmochim. Acta* **61**, 3957–3974 (1997).
26. K. W. Burton, D.-C. Lee, J. N. Christensen, A. N. Halliday, J. R. Hein, *Earth Planet. Sci. Lett.* **171**, 149–156 (1999).
27. H.-F. Ling *et al*., *Earth Planet. Sci. Lett.* **232**, 345–361 (2005).
28. K. A. Venz, D. A. Hodell, *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **182**, 197–220 (2002).
29. T. Stichel, M. Frank, J. Rickli, B. A. Haley, *Earth Planet. Sci. Lett.* **317-318**, 282–294 (2012).
30. W. B. Curry, D. W. Oppo, *Paleoceanography* **20**, PA1017 (2005).
31. A. Ganachaud, C. Wunsch, *Nature* **408**, 453–457 (2000).
32. R. Bintanja, R. S. W. van de Wal, *Nature* **454**, 869–872 (2008).
33. W. S. Broecker, E. Maier-Reimer, *Global Biogeochem. Cycles* **6**, 315–320 (1992).
34. A. Mackensen, H. W. Hubberten, T. Bickert, G. Fischer, D. K. Fütterer, *Paleoceanography* **8**, 587–610 (1993).
35. D. Lüthi *et al*., *Nature* **453**, 379–382 (2008).
36. L. E. Lisiecki, *Geophys. Res. Lett.* **37**, L21708 (2010).

ACKNOWLEDGMENTS

We acknowledge support from NSF grant OCE-10-31198. We also thank two anonymous reviewers for valuable comments.This study was partly supported by the Storke Endowment of the Department of Earth and Environmental Sciences, Columbia University. Samples were provided by the ODP, which is sponsored by the NSF and participating countries under management of Joint Oceanographic Institutions. R. Lupien helped with sample processing. The data in this paper will be available in the supplementary materials and deposited with National Oceanic and Atmospheric Administration–National Climate Data Center at www.ncdc.noaa.gov. This is LDEO contribution number 7809.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/345/6194/318/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S3
Tables S1 and S2
References (*37*–*50*)

16 December 2013; accepted 16 June 2014
Published online 26 June 2014;
10.1126/science.1249770

OCEANOGRAPHY

Oceanic mass transport by mesoscale eddies

Zhengguang Zhang,[1] Wei Wang,[1]* Bo Qiu[2]

Oceanic transports of heat, salt, fresh water, dissolved CO_2, and other tracers regulate global climate change and the distribution of natural marine resources. The time-mean ocean circulation transports fluid as a conveyor belt, but fluid parcels can also be trapped and transported discretely by migrating mesoscale eddies. By combining available satellite altimetry and Argo profiling float data, we showed that the eddy-induced zonal mass transport can reach a total meridionally integrated value of up to 30 to 40 sverdrups (Sv) (1 Sv = 10^6 cubic meters per second), and it occurs mainly in subtropical regions, where the background flows are weak. This transport is comparable in magnitude to that of the large-scale wind- and thermohaline-driven circulation.

Heat and material transports in oceans play a dominant role in regulating Earth's climate and in controlling the oceanic absorption of greenhouse gases that are responsible for global warming. Large-scale wind- and thermohaline-driven ocean circulation has traditionally been regarded to constitute the major part of the oceanic transport, and its spatiotemporal variations to have profound climate and biogeochemical impacts. However, by tracking water masses (*1*) and radiocarbons (*2*), observational studies have revealed that the large-scale circulation alone cannot explain the oceanic transports. Since the detection of their ubiquitous presence in the oceans more than three decades ago, oceanic mesoscale eddies with length scales about 50 to 300 km have been recognized as key contributors in transporting heat, dissolved carbon, and other biogeochemical tracers (*3*–*6*).

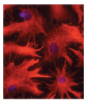
Unlike the large-scale circulation that transports fluids and their properties continuously, mesoscale eddies can trap fluid parcels within the eddy core and transport them discretely (*7*, *8*). Our theoretical understanding of the processes involving mesoscale eddies remains, however, incomplete. For example, many existing studies have treated transport by eddies as a simple diffusion process that tends to obscure the redistribution processes of dynamical and biogeochemical tracers (*9*). In addition to the theoretical understanding, it is equally important to quantify the eddy-induced fluid volume transport on the basis of available observational data. There exist, however, two major challenges for such quantification: First, there is no widely accepted criterion that defines the volume of fluid trapped by an eddy; and, second, accurate vertical structures of the mesoscale eddies are often observationally inaccessible.

In defining the fluid trapped by eddies, all existing criteria are kinematic in nature; they are inherently descriptive and depend on the choice of reference frame (*6*–*8*, *10*). From a dynamical point of view, the movement of a rotating fluid can be depicted by a dynamically conserved quantity: potential vorticity (PV), its tendency to spin. PV contours have long been used to identify the trajectory of fluid particles in the ocean (*11*), and fluid tends to be trapped within the closed PV contours (*12*). A classical demonstration of this constraint is the Taylor Column in the homogeneous-fluid rotating tank experiment (*13*). For an adiabatic stratified ocean, fluid motion is constrained on isopycnal surfaces. If a PV contour on an isopycnal is closed because of eddy perturbations, the fluid trapped inside this closed contour will move with the eddy. In the present work, the outmost closed PV contours on isopycnals were used as the criterion to determine the boundary of fluid trapped by eddies.

Regarding the second challenge of capturing the three-dimensional (3D) structures of mesoscale eddies, our recent study has demonstrated that they can be reconstructed by combining the high-resolution satellite altimeter data and concurrent Argo profiling float temperature/salinity data (*14*). Once the 3D potential density (called density hereafter) field is reconstructed from the altimeter sea surface height (SSH) measurements, we are able to evaluate the eddy-perturbed PV field and finally quantify the fluid volume trapped by the moving eddies (*15*). To verify our method of reconstructing the eddy density field, we used independent mooring observations from different regions of the world oceans: the Labrador Sea, the Arabian Sea, and the Kuroshio Extension (Fig. 1). The reconstructed eddy density structures are found to compare favorably with the mooring results for both the warm- and cold-core eddies. The relative errors of the reconstructed isopycnal surface displacement are estimated to be 20% of the variance (*15*), an uncertainty level that is acceptable for estimating the fluid volume trapped by mesoscale eddies.

For illustration, three isopycnal surfaces of an observed warm-core eddy in the subtropical North Pacific are presented in Fig. 2A. The isopycnal surfaces of this warm-core eddy exhibit a concave

[1]Physical Oceanography Lab, Qingdao Collaborative Innovation Center of Marine Science and Technology, Ocean University of China, Qingdao, People's Republic of China. [2]Department of Oceanography, University of Hawaii at Manoa, Honolulu, HI, USA.
*Corresponding author. E-mail: wei@ouc.edu.cn

EXTENDED PDF FORMAT
SPONSORED BY

Travel Grants Available
Neuroscience 2014
Application Deadline: July 31st
>Apply Now
R&D systems
a bio-techne brand
www.rndsystems.com

Oceanic mass transport by mesoscale eddies

Zhengguang Zhang *et al.*

Science **345**, 322 (2014);
DOI: 10.1126/science.1252418

This copy is for your personal, non-commercial use only.

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by clicking here.

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines here.

The following resources related to this article are available online at www.sciencemag.org (this information is current as of August 4, 2014):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:
http://www.sciencemag.org/content/345/6194/322.full.html

Supporting Online Material can be found at:
http://www.sciencemag.org/content/suppl/2014/06/25/science.1252418.DC1.html

This article **cites 22 articles**, 4 of which can be accessed free:
http://www.sciencemag.org/content/345/6194/322.full.html#ref-list-1

This article appears in the following **subject collections:**
Oceanography
http://www.sciencemag.org/cgi/collection/oceans

Downloaded from www.sciencemag.org on August 4, 2014

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published weekly, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. Copyright 2014 by the American Association for the Advancement of Science; all rights reserved. The title *Science* is a registered trademark of AAAS.

R. D. Muench, J. E. Overland, Eds. (Geophysical Monograph Series, American Geophysical Union, Washington, DC, 1994), vol. 85, pp. 295–311.
5. P. U. Clark, D. Pollard, *Paleoceanography* **13**, 1–9 (1998).
6. L. Lisiecki, M. E. Raymo, *Paleoceanography* **20**, PA1003 (2005).
7. P. U. Clark *et al*., *Quat. Sci. Rev.* **25**, 3150–3184 (2006).
8. H. Elderfield *et al*., *Science* **337**, 704–709 (2012).
9. W. S. Broecker, *Oceanography* **4**, 79–89 (1991).
10. M. E. Raymo, W. F. Ruddiman, N. J. Shackleton, D. W. Oppo, *Earth Planet. Sci. Lett.* **97**, 353–368 (1990).
11. E. L. McClymont, S. M. Sosdian, A. Rosell-Melé, Y. Rosenthal, *Earth Sci. Rev.* **123**, 173–193 (2013).
12. M. E. Raymo, *Paleoceanography* **12**, 577–585 (1997).
13. P. Huybers, *Quat. Sci. Rev.* **26**, 37–55 (2007).
14. H. Gildor, E. Tziperman, *Paleoceanography* **15**, 605–615 (2000).
15. A. Martínez-Garcia, A. Rosell-Melé, E. L. McClymont, R. Gersonde, G. H. Haug, *Science* **328**, 1550–1553 (2010).
16. B. Hönisch, N. G. Hemming, D. Archer, M. Siddall, J. F. McManus, *Science* **324**, 1551–1554 (2009).
17. J. R. Toggweiler, *Paleoceanography* **14**, 571–588 (1999).
18. J. Schmitt *et al*., *Science* **336**, 711–714 (2012).
19. P. Köhler, R. Bintanja, *Clim. Past* **4**, 311–332 (2008).
20. P. Köhler, H. Fischer, J. Schmitt, *Paleoceanography* **25**, PA1213 (2010).
21. D. A. Hodell, K. A. Venz, C. D. Charles, U. S. Ninnemann, *Geochem. Geophys. Geosyst.* **4**, 1–19 (2003).
22. K. Tachikawa, C. Jeandel, M. Roy-Barman, *Earth Planet. Sci. Lett.* **170**, 433–446 (1999).
23. S. L. Goldstein, S. R. Hemming, "Long-lived isotopic tracers in oceanography, paleoceanography and ice sheet dynamics," in *Treatise on Geochemistry* (Elsevier, New York, 2003), vol. 6, pp. 453–498.
24. A. M. Piotrowski, S. L. Goldstein, S. R. Hemming, R. G. Fairbanks, *Science* **307**, 1933–1938 (2005).
25. W. Abouchami, S. L. Goldstein, S. J. G. Gazer, A. Eisenhauer, A. Mangini, *Geochim. Cosmochim. Acta* **61**, 3957–3974 (1997).
26. K. W. Burton, D.-C. Lee, J. N. Christensen, A. N. Halliday, J. R. Hein, *Earth Planet. Sci. Lett.* **171**, 149–156 (1999).
27. H.-F. Ling *et al*., *Earth Planet. Sci. Lett.* **232**, 345–361 (2005).
28. K. A. Venz, D. A. Hodell, *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **182**, 197–220 (2002).
29. T. Stichel, M. Frank, J. Rickli, B. A. Haley, *Earth Planet. Sci. Lett.* **317-318**, 282–294 (2012).
30. W. B. Curry, D. W. Oppo, *Paleoceanography* **20**, PA1017 (2005).
31. A. Ganachaud, C. Wunsch, *Nature* **408**, 453–457 (2000).
32. R. Bintanja, R. S. W. van de Wal, *Nature* **454**, 869–872 (2008).
33. W. S. Broecker, E. Maier-Reimer, *Global Biogeochem. Cycles* **6**, 315–320 (1992).
34. A. Mackensen, H. W. Hubberten, T. Bickert, G. Fischer, D. K. Fütterer, *Paleoceanography* **8**, 587–610 (1993).
35. D. Lüthi *et al*., *Nature* **453**, 379–382 (2008).
36. L. E. Lisiecki, *Geophys. Res. Lett.* **37**, L21708 (2010).

ACKNOWLEDGMENTS

We acknowledge support from NSF grant OCE-10-31198. We also thank two anonymous reviewers for valuable comments.This study was partly supported by the Storke Endowment of the Department of Earth and Environmental Sciences, Columbia University. Samples were provided by the ODP, which is sponsored by the NSF and participating countries under management of Joint Oceanographic Institutions. R. Lupien helped with sample processing. The data in this paper will be available in the supplementary materials and deposited with National Oceanic and Atmospheric Administration–National Climate Data Center at www.ncdc.noaa.gov. This is LDEO contribution number 7809.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/345/6194/318/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S3
Tables S1 and S2
References (*37*–*50*)

16 December 2013; accepted 16 June 2014
Published online 26 June 2014;
10.1126/science.1249770

OCEANOGRAPHY

Oceanic mass transport by mesoscale eddies

Zhengguang Zhang,[1] Wei Wang,[1]* Bo Qiu[2]

Oceanic transports of heat, salt, fresh water, dissolved CO_2, and other tracers regulate global climate change and the distribution of natural marine resources. The time-mean ocean circulation transports fluid as a conveyor belt, but fluid parcels can also be trapped and transported discretely by migrating mesoscale eddies. By combining available satellite altimetry and Argo profiling float data, we showed that the eddy-induced zonal mass transport can reach a total meridionally integrated value of up to 30 to 40 sverdrups (Sv) (1 Sv = 10^6 cubic meters per second), and it occurs mainly in subtropical regions, where the background flows are weak. This transport is comparable in magnitude to that of the large-scale wind- and thermohaline-driven circulation.

Heat and material transports in oceans play a dominant role in regulating Earth's climate and in controlling the oceanic absorption of greenhouse gases that are responsible for global warming. Large-scale wind- and thermohaline-driven ocean circulation has traditionally been regarded to constitute the major part of the oceanic transport, and its spatiotemporal variations to have profound climate and biogeochemical impacts. However, by tracking water masses (*1*) and radiocarbons (*2*), observational studies have revealed that the large-scale circulation alone cannot explain the oceanic transports. Since the detection of their ubiquitous presence in the oceans more than three decades ago, oceanic mesoscale eddies with length scales about 50 to 300 km have been recognized as key contributors in transporting heat, dissolved carbon, and other biogeochemical tracers (*3*–*6*).

Unlike the large-scale circulation that transports fluids and their properties continuously, mesoscale eddies can trap fluid parcels within the eddy core and transport them discretely (*7*, *8*). Our theoretical understanding of the processes involving mesoscale eddies remains, however, incomplete. For example, many existing studies have treated transport by eddies as a simple diffusion process that tends to obscure the redistribution processes of dynamical and biogeochemical tracers (*9*). In addition to the theoretical understanding, it is equally important to quantify the eddy-induced fluid volume transport on the basis of available observational data. There exist, however, two major challenges for such quantification: First, there is no widely accepted criterion that defines the volume of fluid trapped by an eddy; and, second, accurate vertical structures of the mesoscale eddies are often observationally inaccessible.

In defining the fluid trapped by eddies, all existing criteria are kinematic in nature; they are inherently descriptive and depend on the choice of reference frame (*6*–*8*, *10*). From a dynamical point of view, the movement of a rotating fluid can be depicted by a dynamically conserved quantity: potential vorticity (PV), its tendency to spin. PV contours have long been used to identify the trajectory of fluid particles in the ocean (*11*), and fluid tends to be trapped within the closed PV contours (*12*). A classical demonstration of this constraint is the Taylor Column in the homogeneous-fluid rotating tank experiment (*13*). For an adiabatic stratified ocean, fluid motion is constrained on isopycnal surfaces. If a PV contour on an isopycnal is closed because of eddy perturbations, the fluid trapped inside this closed contour will move with the eddy. In the present work, the outmost closed PV contours on isopycnals were used as the criterion to determine the boundary of fluid trapped by eddies.

Regarding the second challenge of capturing the three-dimensional (3D) structures of mesoscale eddies, our recent study has demonstrated that they can be reconstructed by combining the high-resolution satellite altimeter data and concurrent Argo profiling float temperature/salinity data (*14*). Once the 3D potential density (called density hereafter) field is reconstructed from the altimeter sea surface height (SSH) measurements, we are able to evaluate the eddy-perturbed PV field and finally quantify the fluid volume trapped by the moving eddies (*15*). To verify our method of reconstructing the eddy density field, we used independent mooring observations from different regions of the world oceans: the Labrador Sea, the Arabian Sea, and the Kuroshio Extension (Fig. 1). The reconstructed eddy density structures are found to compare favorably with the mooring results for both the warm- and cold-core eddies. The relative errors of the reconstructed isopycnal surface displacement are estimated to be 20% of the variance (*15*), an uncertainty level that is acceptable for estimating the fluid volume trapped by mesoscale eddies.

For illustration, three isopycnal surfaces of an observed warm-core eddy in the subtropical North Pacific are presented in Fig. 2A. The isopycnal surfaces of this warm-core eddy exhibit a concave

[1]Physical Oceanography Lab, Qingdao Collaborative Innovation Center of Marine Science and Technology, Ocean University of China, Qingdao, People's Republic of China. [2]Department of Oceanography, University of Hawaii at Manoa, Honolulu, HI, USA.
*Corresponding author. E-mail: wei@ouc.edu.cn

shape. The upper isopycnal surface outcrops at the sea surface and envelops a bulk of warm water. All PV contours on this isopycnal are closed, and the boundary of the trapped water coincides with the outcrop line. On the middle isopycnal of the eddy (at ~500 m depth), the PV contours near the eddy core are closed, whereas those outside the eddy core are open; in this case, the eddy traps only the fluid within the outermost closed PV contour. As the eddy perturbation decays with depth, the area enclosed by the closed PV contours diminishes on denser isopycnals, and the eddy loses its ability to trap water in these deep layers. The outmost closed PV contour on each isopycnal forms a 3D conic surface (the transparent black surface in Fig. 2A), and it is this surface that delineates the boundary of the fluid trapped by this warm-core eddy. For comparison, a cold-core eddy detected in the same geographical region is presented in Fig. 2B, and the fluid trapped by this cold-core eddy exhibits a somewhat different vertical enveloping shape.

From the sequential SSH maps from the satellite altimeter measurements, we evaluated the propagation speed of individual eddies by adopting a Lagrangian tracking method (*16*, *17*). With the eddy propagation speed and its volume of trapped fluid at hand, the eddy-induced mass transport is readily quantifiable (*15*). Figure 3A shows the global distribution of the eddy-induced zonal transport based on the altimeter SSH data of 1992–2010. A remarkable feature is that the regions of large westward eddy transport are located in all latitude bands between 20° to 40°. Also noteworthy is that an enhanced eddy-induced eastward transport occurs along the Antarctic Circumpolar Current path south of 40°S. When integrated over the entire latitude range from 80°S to 80°N, the total zonal eddy-induced transports add up to 30 to 40 Sv westward and 5 to 10 Sv eastward (Fig. 3B). The global distribution of eddy-induced meridional transport is shown in Fig. 4A. Geographically, the meridional eddy transport has a poleward tendency within the tropical regions (20°S to 20°N) and an equatorward tendency in the subtropical regions (20°N to 60°N and 20°S to 60°S). This meridional eddy transport in many parts of the world oceans moves in the same direction as the meridional wind-forced Ekman transport. The convergent latitudes for the meridional eddy transport, however, appear equatorward of the convergent latitudes for the meridional Ekman transport. Because of the predominance of westward propagations by the mesoscale eddies (*17*), the zonally integrated eddy-induced meridional transport is on the order of several sverdrups, which is much smaller than the zonal eddy-induced mass transport (Fig. 4B). The relative errors in these transport values are estimated at about 20% (*15*).

As a comparison, the transports of the wind-driven gyres in the Pacific, Atlantic, and Indian Oceans are 20 to 60 Sv, 20 to 50 Sv, and 10 to 25 Sv, respectively, and the transports of the thermohaline circulation in these three oceans are 10 to 20 Sv, 9 to 20 Sv, and 10 to 15 Sv, respectively (*18*, *19*). The 30- to 40-Sv value of the eddy-induced zonal transport is comparable in magnitude with the large-scale wind- and thermohaline-driven circulation. Apart from the Southern Ocean, these zonal transports work to accumulate water onto the western side of the ocean basins, implying that the mesoscale eddies can exert a strong impact on the transports of the western boundary currents. Indeed, observations of the Kuroshio transport east of Taiwan reveal that its interannual variation has no obvious correlation with the large-scale wind fluctuations but exhibits a good correspondence with the eddy kinetic energy level east of the island of Taiwan (*20*). As indicated in Fig. 3A, the region east of Taiwan is the strongest hot spot of eddy-induced tranport in the North Pacific Ocean. With the amount of mesoscale eddies modulating on the interannual and decadal time scales (*21*), the eddy-induced zonal transport is likely to fluctuate as well. By altering the poleward mass and heat transport carried by the western boundary currents, the eddy-induced transport fluctuations can contribute substantially to regional and basin-scale climate variability (*22*–*24*).

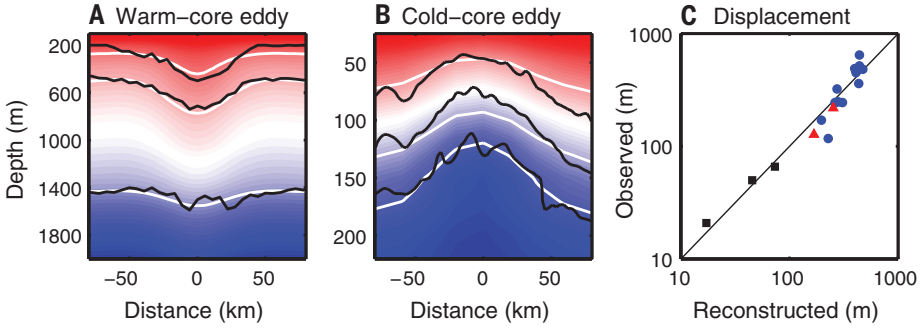


Fig. 1. Comparison between observed and reconstructed eddy density fields. (**A**) Potential density section across a warm-core eddy at 60.6°N, 52.4°W in the Labrador Sea on 11 October 2008. The horizontal ordinate is the distance to the eddy center, and the vertical ordinate is depth. Colored shades denote the reconstructed density field, with white curves representing the isopycnals at 27.65 kg m^{-3}, 27.70 kg m^{-3}, and 27.75 kg m^{-3}. The corresponding isopycnal surfaces from the mooring observations are indicated by black curves (*25*). (**B**) Potential density section across a cold-core eddy at 15.5°N, 61.5°E in the Arabian Sea on 30 November 1994. White and black contours denote the isopycnals at 24.5 kg m^{-3}, 25.5 kg m^{-3}, and 26.0 kg m^{-3}. (**C**) Comparison of the reconstructed isopycnal displacement with the observations at the eddy centers. Blue circles represent the 27.7 kg m^{-3} isopycnal displacement of 12 eddies observed in the Labrador Sea. Black squares represent the 26.0 kg m^{-3} isopycnal displacement of three eddies observed in the Arabian Sea. Red triangles represent the 27.4 kg m^{-3} isopycnal displacement of two eddies observed in the Kuroshio Extension (*15*).

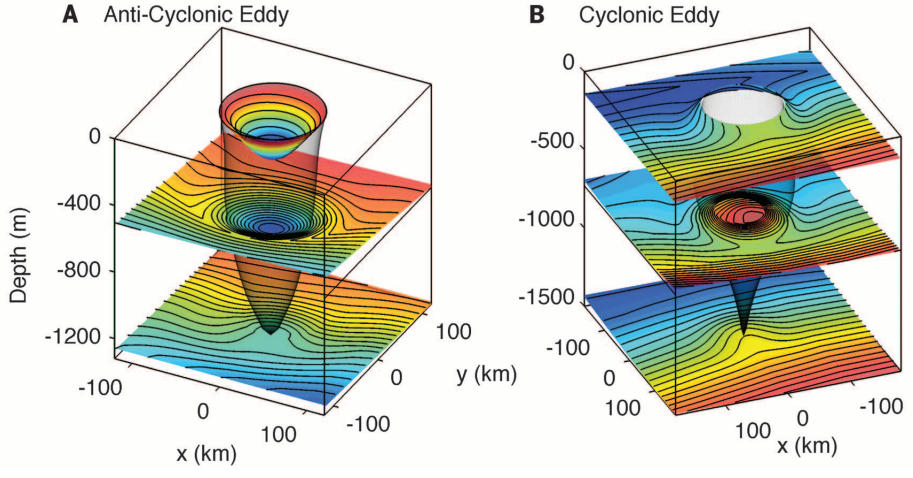


Fig. 2. 3D structures of trapped fluid by mesoscale eddies and PV distributions on representative isopycnals. (**A**) PV contours on three isopycnal surfaces of a warm-core eddy identified by the altimetry data at 30.1°N, 150.1°E in the subtropical North Pacific on 25 May 2005. The sea-level anomaly at the eddy center is 12 cm. The shape and depth of isopycnals are represented by the 3D surfaces. The potential densities of the upper, middle, and lower isopycnals are 23.35 kg m^{-3}, 26.59 kg m^{-3}, and 27.48 kg m^{-3}, respectively. PV distribution on the isopycnal is depicted by colored contours. The transparent black surface, defined by the outmost closed PV contours, signifies the boundary of fluid trapped by this warm-core eddy. (**B**) Same as (A) except for a cold-core eddy identified by the altimetry data at 30.7°N, 151.0°E in the subtropical North Pacific on 25 May 2005. The sea-level anomaly at the eddy center is –14 cm. The depicted isopycnal surfaces are 23.99 kg m^{-3}, 26.75 kg m^{-3}, and 27.56 kg m^{-3}, respectively.

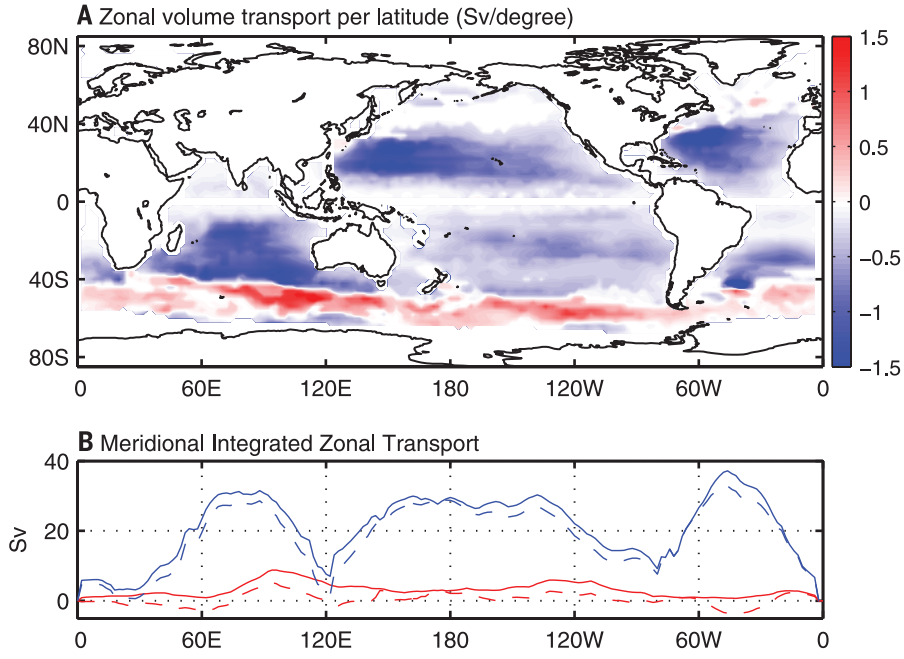


Fig. 3. Global distribution of the zonal transport of fluid trapped by mesoscale eddies. (**A**) Eddy-induced zonal transport through meridional cross section per degree of latitude. The figure is based on the statistical average of eddy propagation speed and trapped fluid volume (*15*). (**B**) Distribution of the total meridionally integrated zonal transport induced by eddies as a function of longitude. The solid blue curve represents the westward transport and the solid red curve represents the eastward transport, computed by integrating the zonal transports in (A) from 80°S to 80°N. The dashed blue curve represents integrated total transport north of 40°S, and the dashed red curve represents integrated total transport south of 40°S.

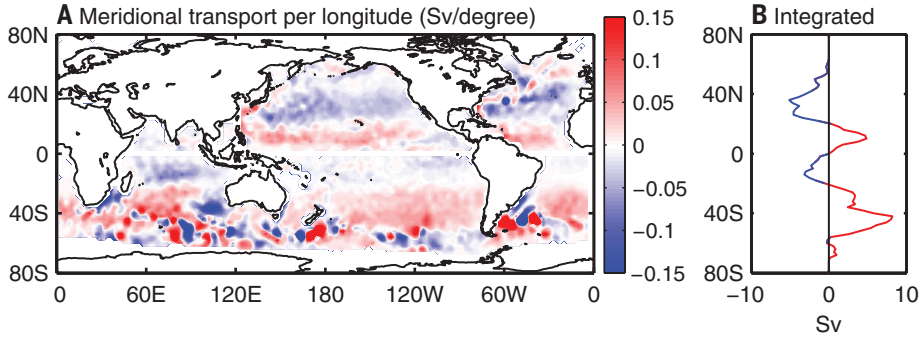


Fig. 4. Global distribution of the meridional transport of fluid trapped by mesoscale eddies. (**A**) Eddy-induced meridional transport through a zonal cross-section per degree of longitude. (**B**) Distribution of the total zonal-integrated meridional transport induced by eddies as a function of latitude.

It is worth noting that there exist no strong currents in many regions with large westward eddy-induced transport identified in Fig. 3A. Equally important, the eddy-induced transports can penetrate to the depth below the main thermocline, where the wind-driven circulation is generally weak. In these regions and depth ranges without strong background circulation, the mesoscale eddies can dominate the transport of heat, fresh water, and dissolved carbons and other biogeochemical tracers, which are important for understanding long-term climate change.

In climate models that simulate global warming scenarios, the transport by oceanic mesoscale eddies is generally underestimated and/or improperly resolved because of their coarse grid resolution. As the model resolution increases with enhanced computing power, mesoscale eddies will be progressively resolved by future climate models, and the impact of the eddy-induced transport will probably emerge. This makes the task of establishing proper theories about eddy-induced transport particularly important and urgent.

REFERENCES AND NOTES

1. G. Danabasoglu, J. C. McWilliams, P. R. Gent, *Science* **264**, 1123–1126 (1994).
2. S. Peacock, *Global Biogeochem. Cycles* **18**, GB2022 (2004).
3. H. L. Bryden, E. C. Brady, *J. Mar. Res.* **47**, 55–79 (1989).
4. D. B. Chelton, P. Gaube, M. G. Schlax, J. J. Early, R. M. Samelson, *Science* **334**, 328–332 (2011).
5. D. J. McGillicuddy Jr. *et al.*, *Science* **316**, 1021–1026 (2007).
6. C. Dong, J. C. McWilliams, Y. Liu, D. Chen, *Nat. Commun.* **5**, 3294 (2014).
7. G. R. Flierl, *Geophys. Astrophys. Fluid Dyn.* **18**, 39–74 (1981).
8. J. J. Early, R. M. Samelson, D. B. Chelton, *J. Phys. Oceanogr.* **41**, 1535–1555 (2011).
9. M. M. Lee, D. P. Marshall, R. G. Williams, *J. Mar. Res.* **55**, 483–505 (1997).
10. F. J. Beron-Vera, Y. Wang, M. J. Olascoaga, G. J. Goni, G. Haller, *J. Phys. Oceanogr.* **43**, 1426–1438 (2013).
11. J. R. Luyten, J. Pedlosky, H. M. Stommel, *J. Phys. Oceanogr.* **13**, 292–309 (1983).
12. P. B. Rhines, W. R. Young, *J. Fluid Mech.* **122**, 347–367 (1982).
13. G. I. Taylor, *Proc. R. Soc. London A Contain. Pap. Math. Phys. Character* **93**, 99–113 (1917).
14. Z. G. Zhang, Y. Zhang, W. Wang, R. X. Huang, *Geophys. Res. Lett.* **40**, 3677–3681 (2013).
15. Information on the observational data and analysis methods is available in the supplementary materials on *Science* Online.
16. D. B. Chelton, M. G. Schlax, R. M. Samelson, R. A. de Szoeke, *Geophys. Res. Lett.* **34**, L15606 (2007).
17. D. B. Chelton, M. G. Schlax, R. M. Samelson, *Prog. Oceanogr.* **91**, 167–216 (2011).
18. W. J. Schmitz Jr., *On the World Ocean Circulation: Volume I: Some Global Features/North Atlantic Circulation* (Technical Report WHOI-96-03, Woods Hole Oceanographic Institution, Woods Hole, MA, 1996).
19. W. J. Schmitz Jr., *On the World Ocean Circulation. Volume II: The Pacific and Indian Oceans/A Global Update* (Technical Report WHOI-96-08, Woods Hole Oceanographic Institution, Woods Hole, MA, 1996).
20. Y. L. Chang, L. Y. Oey, *Geophys. Res. Lett.* **38**, L08603 (2011).
21. B. Qiu, S. Chen, *J. Phys. Oceanogr.* **43**, 344–358 (2013).
22. S. Minobe, A. Kuwano-Yoshida, N. Komori, S. P. Xie, R. J. Small, *Nature* **452**, 206–209 (2008).
23. B. Qiu, *J. Phys. Oceanogr.* **30**, 1486–1502 (2000).
24. B. Qiu, *J. Phys. Oceanogr.* **33**, 2465–2482 (2003).
25. M. F. de Jong, A. S. Bower, H. H. Furey, *J. Phys. Oceanogr.* **44**, 427–444 (2014).

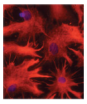
ACKNOWLEDGMENTS

We thank the reviewers for their constructive and detailed comments. This research was supported by the National Basic Research Priorities Program of China through grant 2013CB430303, the National Natural Science Foundation of China under grant 41276014, and the National Global Change Major Research Project of China through grant 2013CB956201. The sources of the data used in this paper are contained in the supplementary materials.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/345/6194/322/suppl/DC1
Materials and Methods
Figs. S1 to S4
References

19 February 2014; accepted 16 June 2014
Published online 26 June 2014;
10.1126/science.1252418

EXTENDED PDF FORMAT
SPONSORED BY

Travel Grants Available
Neuroscience 2014
Application Deadline: July 31st
>Apply Now
R&D systems
a bio-techne brand
www.rndsystems.com

Leverage points for improving global food security and the environment

Paul C. West *et al.*

Science **345**, 325 (2014);
DOI: 10.1126/science.1246067

This copy is for your personal, non-commercial use only.

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by clicking here.

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines here.

The following resources related to this article are available online at www.sciencemag.org (this information is current as of August 4, 2014):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:
http://www.sciencemag.org/content/345/6194/325.full.html

Supporting Online Material can be found at:
http://www.sciencemag.org/content/suppl/2014/07/16/345.6194.325.DC1.html

This article **cites 42 articles**, 14 of which can be accessed free:
http://www.sciencemag.org/content/345/6194/325.full.html#ref-list-1

This article appears in the following **subject collections:**
Ecology
http://www.sciencemag.org/cgi/collection/ecology

Downloaded from www.sciencemag.org on August 4, 2014

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published weekly, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. Copyright 2014 by the American Association for the Advancement of Science; all rights reserved. The title *Science* is a registered trademark of AAAS.

FOOD SECURITY

Leverage points for improving global food security and the environment

Paul C. West,[1]* James S. Gerber,[1] Peder M. Engstrom,[1] Nathaniel D. Mueller,[2] Kate A. Brauman,[1] Kimberly M. Carlson,[1] Emily S. Cassidy,[3] Matt Johnston,[1] Graham K. MacDonald,[1] Deepak K. Ray,[1] Stefan Siebert[4]

Achieving sustainable global food security is one of humanity's contemporary challenges. Here we present an analysis identifying key "global leverage points" that offer the best opportunities to improve both global food security and environmental sustainability. We find that a relatively small set of places and actions could provide enough new calories to meet the basic needs for more than 3 billion people, address many environmental impacts with global consequences, and focus food waste reduction on the commodities with the greatest impact on food security. These leverage points in the global food system can help guide how nongovernmental organizations, foundations, governments, citizens' groups, and businesses prioritize actions.

Sustainably meeting global food demands is one of humanity's grand challenges and has attracted considerable attention in the past few years (*1*–*7*). However, there have been few attempts to systematically test the effectiveness of potential strategies and identify where to target different proposed solutions. In particular, finding whether proposed solutions can truly scale to the global scope of the problem remains a challenge.

In this study, we aimed to identify a small set of regions, crops, and actions that provide strategic global opportunities to increase yields, reduce the environmental impact of agriculture, and deliver food more efficiently from what is already grown. In order to identify these "leverage points" where interventions would have a high global impact in one or more of the three categories, we used recently published geospatial data and models to analyze how specific issues at the nexus of agriculture and the environment vary across the globe.

We focused our analysis on 17 key global crops (table S1), composing the 16 highest-calorie–producing crops consumed as food, as well as cotton, because of its intensive water and nutrient use. These 17 crops cover 58% of the global cropland area harvested and produce 86% of the world's crop calories. They also account for most resource use on croplands: 95% of irrigated area, 92% of irrigation water consumption, and ~70% of all nitrogen and phosphorus application. The percentages in the main text are all relative to the 17 major crops. Additional details for individual crops and methods are provided in the supplementary materials (*8*).

One goal for achieving future food security is to grow more food on the existing land base in ways that limit additional pressure on natural ecosystems (*9*–*11*). Current yields are 50% below realistically attainable potentials in many regions of the world (*3*, *12*). This difference between the current and attainable yields is commonly referred to as the "yield gap" (*3*, *9*). To assess possible production increases on existing croplands, we calculated the yield gap using previously developed methods (*3*, *13*). Increasing yields in low-performing areas by closing the yield gap to 50% of attainable yields (Fig. 1) could increase total production by 358 megatons per year [8.5×10^{14} kilocalories (kcal)] (table S2, A and B), which is enough calories to meet the basic needs of ~850 million people.

These potential gains from closing the yield gap to 50% are not evenly distributed. For example, half of these potential gains are concentrated in only ~5% of the total harvested area for these crops. Approximately 92% of these gains (4.3×10^{14} kcal) are in Africa (43%), Asia (29%), and Europe (mainly eastern) (20%) (table S2C). Maximizing calorie gains is essential for improved food security, but nutrition, access, and cultural preferences must also be addressed. Further, the yield gap estimates are based on current yield and climate conditions. Climate change will generally decrease mean yields (*14*) and thus probably also change the magnitude and possibly the location of leverage points.

Although humanity needs to produce more food to meet future demand, agriculture is the major driving force on the planet for greenhouse gas (GHG) emissions (*7*, *15*), water quality degradation from soil loss and nutrient runoff (*16*), and water use (*17*). Prioritizing a small set of leverage points could greatly increase the efficiency and sustainability of food production.

Agriculture accounts for 20 to 35% of global GHG emissions (*7*, *15*). The largest sources of agricultural GHGs are carbon dioxide (CO_2) from tropical deforestation, methane (CH_4) from livestock and rice production, and nitrous oxide (N_2O) from nutrient additions to croplands (*7*, *15*). The bulk of GHG emissions from croplands are concentrated in just a few countries. Approximately 51% of 2000–2012 tropical deforestation (in areas with >50% tree canopy cover) occurred in Brazil (34%) and Indonesia (17%) (*18*). Deforestation in these countries is linked mainly to timber, cattle, and soy production in Brazil (*19*) and oil palm and wood plantations in Indonesia (*20*). Similarly, methane emissions from rice cultivation are produced mainly in China (29%) and India (24%) (*21*).

For N_2O, we estimated the distribution of emissions from major crops using the Intergovernmental Panel on Climate Change Tier 1 approach

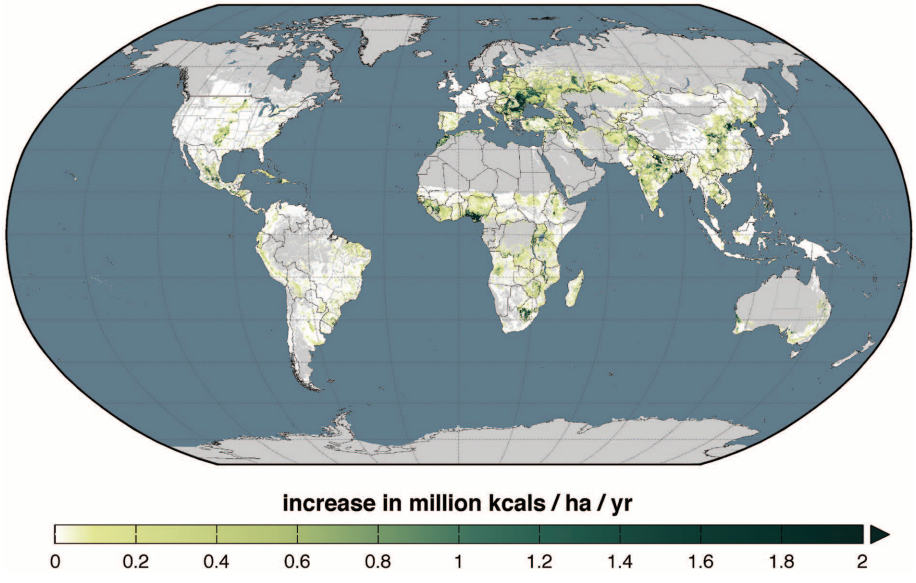


Fig. 1. Increasing crop production from closing the yield gaps to 50% of potential yields. Increasing yields to 50% of the potential yield in all low-performing areas could increase annual production by 8.46×10^{14} kcal, which is enough to meet the basic caloric requirements of ~850 million people. Further, half of these potential gains (4.2×10^{14} kcal) are concentrated in only 5% of the harvested area for the 17 crops analyzed here.

[1]Institute on the Environment, University of Minnesota, St. Paul, MN 55108, USA. [2]Center for the Environment, Harvard University, Cambridge, MA 02138, USA. [3]Environmental Working Group, Washington, DC 20009, USA. [4]Institute of Crop Science and Resource Conservation, University of Bonn, Katzenburgweg 5, D-53115 Bonn, Germany.
*Corresponding author. E-mail: pcwest@umn.edu

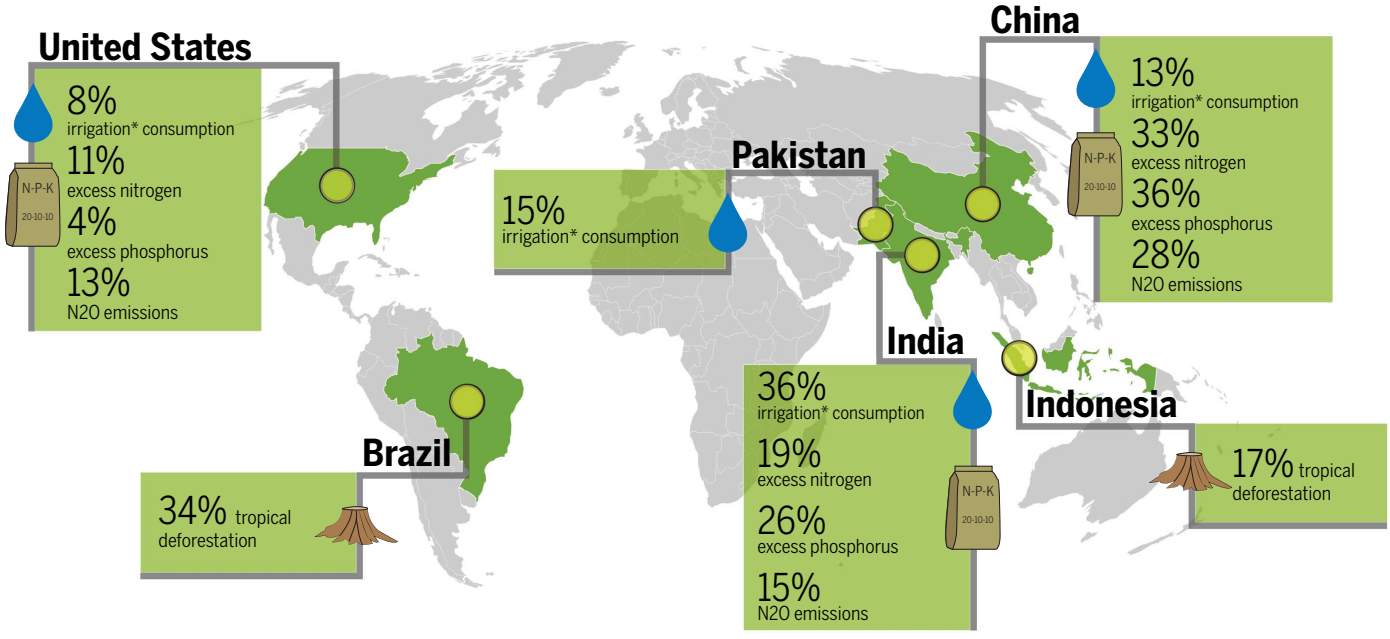


Fig. 2. Leverage points to reduce agriculture's effect on climate, water quality, and water consumption. The majority of global environmental effects of agriculture are in a few countries, driven by a few commodities. All nutrient and irrigation values are relative to the 17 major crops in this study. Figures S1 to S3 provide maps of N_2O emissions, nutrient input and excess, and water consumption, respectively. Irrigation consumption is relative only to precipitation-limited areas.

(*22*), with data sets of global fertilizer and manure use, atmospheric nitrogen deposition, and crop nitrogen fixation (*8*). We find that China, India, and the United States contribute 56% of all N_2O emissions from croplands, with 28% coming from China alone. Three crops (wheat, maize, and rice) account for 68% of global N_2O emissions from the 17 major crops analyzed here (table S3A). These results highlight the fact that agriculture's biogeochemical influence on global climate is not evenly distributed and that there are several key leverage points for reducing GHG emissions. Fortunately, regional efforts have been effective at reducing CO_2 from deforestation (*23*), and similar strategies are proposed for reducing N_2O and CH_4 in other regions (*15*, *24*).

Although nutrient inputs on farmland are needed to grow crops, nutrient loss from agricultural areas is a major source of pollution for freshwater and coastal ecosystems (*16*). We estimated the amount of excess nutrient application (defined as the difference between rates of nutrient input versus nutrient removal from plant harvesting) for the 17 major crops using a simple nutrient mass balance model and spatial data sets for each of the nutrient inputs and outputs (*8*). Across the 17 crops, we find that ~60% of nitrogen and ~48% of phosphorus inputs are in excess (table S3, A and B, and fig. S2).

Excess nutrients are not evenly distributed across regions or crops. We find that ~50% of the excess nitrogen and phosphorus is concentrated in only 24% and 21% of the world's cropland area, respectively. China, India, and the United States together account for ~64 to 66% of excess nitrogen and phosphorus (Fig. 2 and fig. S2). Furthermore, rice, wheat, and maize alone are responsible for ~58 to 60% of the excess nitrogen and phosphorous of the 17 crops we analyzed globally. Using a yield-response model (*13*), we estimate that nitrogen and phosphorus applied to wheat, rice, and maize could be reduced by 14 to 29% and 13 to 22%, respectively, while maintaining current yields (table S3). Targeting reductions in fertilizer use to a small set of crops and countries could therefore have a large effect on global nitrogen and phosphorus pollution. Further efficiency gains are possible by altering the timing, placement, and type of fertilizer (*25*).

Irrigation accounts for ~70% of global water withdrawals and ~90% of water consumption and is frequently a driver of water stress (*17*). To identify hot spots of irrigation use, we used spatially explicit, crop-specific irrigation data (*6*). We restricted the analysis to precipitation-limited areas, defined here as places where potential evapotranspiration exceeds precipitation.

We find that India, Pakistan, China, and the United States account for 72% of all irrigation water used in precipitation-limited areas, with India alone accounting for 36% (fig. S3) (*8*). However, water productivity varies widely among countries. For example, although China consumes 13% of irrigation in precipitation-limited areas, it is responsible for just 2% of irrigation consumption in systems with very low water productivity (below the 20th percentile). Raising very low water productivity in precipitation-limited regions up to the 20th percentile has the potential to decrease water consumption by about 8 to 15% while keeping food production constant (*26*).

Relative to the 17 major crops, rice and wheat together cover 63% of the total irrigated area and consume 59% of irrigation water globally. Cotton, maize, and sugarcane use an additional 30% of total irrigation water. Despite covering only 3% of cropland area globally, sugarcane and cotton are the most water-intensive crops, with mean irrigation water consumption that is 2.4 and 1.6 times, respectively, the amount used for each hectare of wheat (table S4).

Meat and dairy consumption is increasing globally and generally increases with wealth (*11*, *27*). Between 1961 and 2009, the percentage of calories from crop production consumed as food decreased from 57 to 51% (fig. S4). Although crops used for animal feed ultimately produce human food in the form of meat and dairy products, they do so with a substantial loss of caloric efficiency. If current crop production used for animal feed and other nonfood uses (including biofuels) were targeted for direct consumption, ~70% more calories would become available, potentially providing enough calories to meet the basic needs of an additional 4 billion people (*28*). The human-edible crop calories that do not end up in the food system are referred to as the "diet gap."

Using methods described in Cassidy *et al.* (*28*), Food and Agriculture Organization (FAO) Food Balance Sheets (*29*), and crop production data (*30*), we estimate that the United States, China, Western Europe, and Brazil account for 26, 17, 11, and 6% of the global diet gap, respectively (fig. S5). Changing crop allocation to directly feed people in these four regions alone could provide enough calories to meet the basic needs of 2.4 billion people. Maize represents the largest potential gain, accounting for 41% of the global diet gap. Maize in the United States accounts for 19%

Table 1. Calories lost to consumer food waste and associated land requirements for key food commodities in the United States, India, and China.

Food	Calorie supply (*29*) (kcal per capita day^{-1})	Consumer food waste (*31*) (% and kcal per capita day^{-1})	Embodied feed in calorie waste (kcal per capita day^{-1})	Harvested area required to support waste (m^2 per capita year^{-1})
United States	**3688***			
Wheat and rice	693	27% (187)	–	65
Vegetables	76	28% (21)	–	8
Beef	110	11% (12)	278	43
Pork	132	11% (15)	136	21
Poultry	193	11% (21)	199	31
China	**3036†**			
Wheat and rice	1380	20% (276)	–	59
Vegetables	205	15% (31)	–	10
Beef	27	8% (2)	33	9
Pork	355	8% (28)	178	47
Poultry	59	8% (5)	30	8
India	**2321‡**			
Wheat and rice	1168	3% (35)	–	14
Vegetables	46	7% (3)	–	3
Beef	5	4% (<1)	3	2
Pork	3	4% (<1)	1	1
Poultry	6	4% (<1)	2	2

*22% from cereals, with the remainder mostly a mix of vegetable oils, sweeteners, meat, and dairy. †48% from cereals, with the remainder mostly a mix of meat, vegetables, oils, and starchy roots. ‡57% from cereals, with the remainder mostly a mix of sugar, vegetable oils, dairy, and pulses.

of the global diet gap, which is enough calories for 760 million people; maize in China represents 9% of the diet gap. Although the diet gap presents opportunities to improve food security, cultural preferences and political obstacles create large challenges to reducing meat as well as overconsumption. However, adapting how these crops are used could provide short-term relief and serve as a safety net or release valve in years with high production losses from weather or pests.

Another major opportunity for increasing crop availability is to reduce food waste. Approximately ~30 to 50% of food production is wasted (*31*). Several broad strategies are recommended for reducing waste along different points in the supply chain, but they provide little insight into the relative impact of waste among food groups. For example, we estimate that wasting 1 kg of boneless beef has ~24 times the effect on available calories as wasting 1 kg of wheat (~98,000 kcal versus ~3800 to 4125 kcal), because of inefficiencies in converting feed to animal calories and protein (*8*). Wasting 1 kg of other animal products, such as poultry and pork, has a less dramatic effect on edible calorie losses. This analysis applies only to animals raised in feedlot systems.

We compared consumer food waste in the United States, China, and India to illustrate the connections among waste, diet, and crop yields. Per-capita food losses range from <3 kcal per person day^{-1} for pork or vegetables in India to as much as 290 kcal per person day^{-1} for beef in the United States (Table 1). Consequently, the land base required to support this waste is ~7 to 8 times greater in the United States than in India. Curbing consumer waste of major food crops (i.e., wheat, rice, and vegetables) and meats (i.e., beef, pork, and poultry) in these three countries alone could feed ~413 million people year^{-1} if the feed calories embodied in meat are included (*8*). This illustration demonstrates that small changes in the consumption and waste of animal products could have a large effect on available calories.

The challenges and opportunities for improving global food security and the environment are not evenly distributed across the globe. The analyses here illustrate, in principle, what is possible and where to focus such possible actions. Local and regional successes in any of these leverage points can have national or global impact, and provide examples of increasing food production, reducing deforestation, and managing water and nutrients more efficiently. Additional analysis is needed to estimate how the leverage points identified here may change in response to changes in climate, diet, technology, and markets. Although this large-scale analysis is useful for identifying the major leverage points to create a more sustainable food system, there are many other regional and global factors that would need to be included to address food security and environmental sustainability—and the tradeoffs among the many ecosystem services in agricultural landscapes—more holistically.

More work is needed for the widespread adoption of best practices by adapting current successful approaches to local socioeconomic and environmental conditions, within the context of a changing climate. Fortunately, many institutions, including NGOs, foundations, governments, citizens' groups, and businesses, are committed to improving global and local food security while reducing environmental harm. The leverage points presented here represent a relatively small set of regions, crops, and actions that could provide additional food to a growing world, as well as address a sizeable fraction of agriculture's global environmental impact.

REFERENCES AND NOTES

1. International Assessment of Agricultural Knowledge, Science, and Technology for Development, *Agriculture at a Crossroads* (Island Press, Washington, DC, 2009).
2. The Royal Society, *Reaping the Benefits: Science and the Sustainable Intensification of Global Agriculture* (The Royal Society, London, 2009).
3. J. A. Foley *et al.*, *Nature* **478**, 337–342 (2011).
4. H. C. J. Godfray *et al.*, *Science* **327**, 812–818 (2010).
5. J. Kearney, *Philos. Trans. R. Soc. London Ser. B* **365**, 2793–2807 (2010).
6. S. Siebert, P. Döll, *J. Hydrol. (Amst.)* **384**, 198–217 (2010).
7. S. J. Vermeulen, B. M. Campbell, J. S. I. Ingram, *Annu. Rev. Environ. Resour.* **37**, 195–222 (2012).
8. Materials and methods are available as supplementary materials on *Science* Online.
9. K. G. Cassman, A. Dobermann, D. T. Walters, H. Yang, *Annu. Rev. Environ. Resour.* **28**, 315–358 (2003).
10. P. C. West *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 19645–19648 (2010).
11. D. Tilman, C. Balzer, J. Hill, B. L. Befort, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 20260–20264 (2011).
12. D. B. Lobell, K. G. Cassman, C. B. Field, *Annu. Rev. Environ. Resour.* **34**, 179–204 (2009).
13. N. D. Mueller *et al.*, *Nature* **490**, 254–257 (2012).
14. A. J. Challinor *et al.*, *Nat. Clim. Change* **4**, 287–291 (2014).
15. P. Smith *et al.*, *Philos. Trans. R. Soc. London Ser. B* **363**, 789–813 (2008).
16. S. R. Carpenter *et al.*, *Ecol. Appl.* **8**, 559–568 (1998).
17. P. H. Gleick, H. Cooley, M. Morikawa, *The World's Water 2008-2009: The Biennial Report on Freshwater Resources* (Island Press, Washington, DC, 2009).
18. M. C. Hansen *et al.*, *Science* **342**, 850–853 (2013).
19. D. C. Morton *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 14637–14641 (2006).
20. K. M. Carlson *et al.*, *Nat. Clim. Change* **3**, 283–287 (2013).
21. X. Yan, H. Akiyama, K. Yagi, H. Akimoto, *Global Biogeochem. Cycles* **23**, n/a (2009).

22. C. De Klein *et al.*, in *2006 IPCC Guidelines for National Greenhouse Gas Inventories* (Intergovernmental Panel on Climate Change, IGES, Hayama, Japan, 2006), vol. Technical, chap. 11, p. V4.
23. M. N. Macedo *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 1341–1346 (2012).
24. W.-F. Zhang *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 8375–8380 (2013).
25. R. T. Venterea *et al.*, *Front. Ecol. Environ* **10**, 562–570 (2012).
26. K. A. Brauman, S. Siebert, J. A. Foley, *Environ. Res. Lett.* **8**, 24030 (2013).
27. T. Kastner, M. J. I. Rivas, W. Koch, S. Nonhebel, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 6868–6872 (2012).
28. E. S. Cassidy, P. C. West, J. S. Gerber, J. A. Foley, *Environ. Res. Lett.* **8**, 034015 (2013).
29. FAO, *Food Balance Sheets* (FAO, Rome, 2011); http://faostat.fao.org/site/368/default.aspx#ancor.
30. C. Monfreda, N. Ramankutty, J. Foley, *Global Biogeochem. Cycles* **22**, GB1022 (2008).
31. J. Gustavsson, C. Cederberg, U. Sonesson, R. van Otterdijk, A. Meybeck, *Global Food Losses and Food Waste: Extent, Causes and Prevention* (FAO, Rome, 2011).

ACKNOWLEDGMENTS

This research greatly benefited from discussions with J. Foley. S. Carpenter, and E. Bennett, and three anonymous reviewers provided critical feedback that strengthened this paper. Research support was primarily provided by the Gordon and Betty Moore Foundation. Additional research support was from the Institute on the Environment, NASA's Interdisciplinary Earth Science program and Laurel Salton Clark Memorial Fellowship, the McKnight Foundation, the Grantham Foundation, the World Wildlife Fund, the Nature Conservancy, and Stanford University. NSF provided support for N.D.M. The Natural Sciences and Engineering Research Council of Canada provided support for G.K.M. Contributions by General Mills, Mosaic, Cargill, Pentair, Google, Kellogg's, Mars, and PepsiCo supported stakeholder outreach and public engagement. The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. Additional data and method details are provided in the supplementary materials.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/345/6194/325/suppl/DC1
Materials and Methods
Figs. S1 to S5
Tables S1 to S4
References (*32*–*53*)

16 September 2013; accepted 17 June 2014
10.1126/science.1246067

ALTERNATIVE SPLICING

Human tRNA synthetase catalytic nulls with diverse functions

Wing-Sze Lo,[1,2] Elisabeth Gardiner,[3,4] Zhiwen Xu,[1,2] Ching-Fun Lau,[1,2] Feng Wang,[1,2] Jie J. Zhou,[1,2] John D. Mendlein,[4] Leslie A. Nangle,[4] Kyle P. Chiang,[4] Xiang-Lei Yang,[1,3] Kin-Fai Au,[5] Wing Hung Wong,[6] Min Guo,[7] Mingjie Zhang,[1,8] Paul Schimmel[1,3,7]*

Genetic efficiency in higher organisms depends on mechanisms to create multiple functions from single genes. To investigate this question for an enzyme family, we chose aminoacyl tRNA synthetases (AARSs). They are exceptional in their progressive and accretive proliferation of noncatalytic domains as the Tree of Life is ascended. Here we report discovery of a large number of natural catalytic nulls (CNs) for each human AARS. Splicing events retain noncatalytic domains while ablating the catalytic domain to create CNs with diverse functions. Each synthetase is converted into several new signaling proteins with biological activities "orthogonal" to that of the catalytic parent. We suggest that splice variants with nonenzymatic functions may be more general, as evidenced by recent findings of other catalytically inactive splice-variant enzymes.

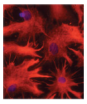
Aminoacyl tRNA synthetases (AARSs) establish the genetic code by esterifying specific amino acids to the 3′ ends of their cognate tRNAs (*1*–*5*) and have adaptations of this reaction for specific physiological responses (*6*). A few literature examples show that natural proteolysis or alternative splicing of AARS can reveal previously unknown AARS proteins (*7*, *8*) with new functions (*9*–*11*). With this in mind, we investigated potential mechanisms for achieving genetic efficiency through functional expansions. The enzymes are divided into two classes of 10 proteins each, with each class being defined by the architecture of the highly conserved catalytic domain (CD) that is retained through evolution (*12*–*14*). As the Tree of Life is ascended, 13 new domains, which have no obvious association with aminoacylation or editing, have collectively been added to AARSs and maintained over the course of evolution, with no appreciable benefit or detriment to primary function (*15*–*17*). The extent of these domain additions appears to be particular to AARSs (*15*). Some of these new domains are appended to each of several synthetases, whereas others are specific to a single synthetase. Notably, these novel domain additions are accretive and progressive; and while their persistence provides no major benefit to aminoacylation, the strong evolutionary pressure for their retention suggests they are not random functionless stochastic fusions, but may be conserved for a specific biological purpose, perhaps distinct from the canonical enzymatic function.

We made a comprehensive search for alternative splice variants of AARSs to understand how splicing changes the domain organization and underlying architecture of each synthetase. We selectively targeted the AARS family of genes by enriching the AARS transcriptome in six distinct human samples [human fetal and adult brain, primary human leukocytes, and three cultured leukocyte cell types (Raji B cells, Jurkat T cells, and THP1 monocytes)]. A polymerase chain reaction (PCR)–based gene-capture and enrichment method was integrated with high-throughput deep sequencing to increase sequencing depth for each AARS transcript (materials and methods and Fig. 1A). This methodology allowed for high enrichment of AARS mRNAs and mainly targeted exon-exon junctions for discovery of exon-skipping events. We defined the AARS transcriptome as the transcripts of 37 AARS genes, including those for 17 cytoplasmic synthetases, 17 mitochondrial synthetases, and 3 that encode both cytoplasmic and mitochondrial forms. For efficient capture, transcripts were amplified by multiplex PCR using AARS gene–specific primers and optimized PCR conditions (see materials and methods). Sensitive detection of low-abundance splice variants was achieved with an optimized multiplex PCR that amplified gene regions close to exon-exon junctions of AARS transcripts and produced short PCR fragments (Fig. 1A). Fragments were assembled into cDNA libraries and sequenced by high-throughput deep sequencing (*18*, *19*).

Approximately 42 million 50-base reads were obtained and analyzed, using established methods (*19*). About 70% (30.4 million) mapped to the 37 AARS genes, and about two-thirds of the AARS-specific reads (21.4 million) covered AARS exon-exon junctions. When compared to previously published whole-transcriptome studies (*20*, *21*), the AARS transcriptome enrichment method employed here successfully improved sequencing depth so that we could detect all of the 61 previously reported exon-exon junctions for AARS transcripts, as well as identify 248 previously unreported junctions (Fig. 1B and table S1). These new splice forms allowed for the ablation of specific coding regions and simultaneous creation of new exon-exon junctions.

In addition, the tissue origin and the overlap of AARS splice variants in different tissues were examined. Although there was obvious tissue

[1]IAS HKUST–Scripps R&D Laboratory, Institute for Advanced Study, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, China. [2]Pangu Biopharma, Edinburgh Tower, The Landmark, 15 Queen's Road Central, Hong Kong, China. [3]The Scripps Laboratories for tRNA Synthetase Research, The Scripps Research Institute, 10650 North Torrey Pines Road, La Jolla, CA 92037, USA. [4]aTyr Pharma, 3545 John Hopkins Court, Suite 250, San Diego, CA 92121, USA. [5]Department of Internal Medicine, University of Iowa, Iowa City, IA 52242, USA. [6]Department of Statistics, Stanford University, Stanford, CA 94305, USA. [7]The Scripps Laboratories for tRNA Synthetase Research, Scripps Florida, 130 Scripps Way, Jupiter, FL 33458, USA. [8]Division of Life Science, State Key Laboratory of Molecular Neuroscience, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, China.
***Corresponding author. E-mail: schimmel@scripps.edu**

EXTENDED PDF FORMAT
SPONSORED BY

Travel Grants Available
Neuroscience 2014
Application Deadline: July 31st
>Apply Now
R&D systems
a bio-techne brand
www.rndsystems.com

Human tRNA synthetase catalytic nulls with diverse functions

Wing-Sze Lo *et al.*

Science **345**, 328 (2014);
DOI: 10.1126/science.1252943

This copy is for your personal, non-commercial use only.

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by clicking here.

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines here.

The following resources related to this article are available online at www.sciencemag.org (this information is current as of August 4, 2014):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:
http://www.sciencemag.org/content/345/6194/328.full.html

Supporting Online Material can be found at:
http://www.sciencemag.org/content/suppl/2014/07/16/345.6194.328.DC1.html

A list of selected additional articles on the Science Web sites **related to this article** can be found at:
http://www.sciencemag.org/content/345/6194/328.full.html#related

This article **cites 34 articles**, 13 of which can be accessed free:
http://www.sciencemag.org/content/345/6194/328.full.html#ref-list-1

This article appears in the following **subject collections:**
Biochemistry
http://www.sciencemag.org/cgi/collection/biochem

Downloaded from www.sciencemag.org on August 4, 2014

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published weekly, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. Copyright 2014 by the American Association for the Advancement of Science; all rights reserved. The title *Science* is a registered trademark of AAAS.

22. C. De Klein *et al.*, in *2006 IPCC Guidelines for National Greenhouse Gas Inventories* (Intergovernmental Panel on Climate Change, IGES, Hayama, Japan, 2006), vol. Technical, chap. 11, p. V4.
23. M. N. Macedo *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 1341–1346 (2012).
24. W.-F. Zhang *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 8375–8380 (2013).
25. R. T. Venterea *et al.*, *Front. Ecol. Environ* **10**, 562–570 (2012).
26. K. A. Brauman, S. Siebert, J. A. Foley, *Environ. Res. Lett.* **8**, 24030 (2013).
27. T. Kastner, M. J. I. Rivas, W. Koch, S. Nonhebel, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 6868–6872 (2012).
28. E. S. Cassidy, P. C. West, J. S. Gerber, J. A. Foley, *Environ. Res. Lett.* **8**, 034015 (2013).
29. FAO, *Food Balance Sheets* (FAO, Rome, 2011); http://faostat.fao.org/site/368/default.aspx#ancor.
30. C. Monfreda, N. Ramankutty, J. Foley, *Global Biogeochem. Cycles* **22**, GB1022 (2008).
31. J. Gustavsson, C. Cederberg, U. Sonesson, R. van Otterdijk, A. Meybeck, *Global Food Losses and Food Waste: Extent, Causes and Prevention* (FAO, Rome, 2011).

ACKNOWLEDGMENTS

This research greatly benefited from discussions with J. Foley. S. Carpenter, and E. Bennett, and three anonymous reviewers provided critical feedback that strengthened this paper. Research support was primarily provided by the Gordon and Betty Moore Foundation. Additional research support was from the Institute on the Environment, NASA's Interdisciplinary Earth Science program and Laurel Salton Clark Memorial Fellowship, the McKnight Foundation, the Grantham Foundation, the World Wildlife Fund, the Nature Conservancy, and Stanford University. NSF provided support for N.D.M. The Natural Sciences and Engineering Research Council of Canada provided support for G.K.M. Contributions by General Mills, Mosaic, Cargill, Pentair, Google, Kellogg's, Mars, and PepsiCo supported stakeholder outreach and public engagement. The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. Additional data and method details are provided in the supplementary materials.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/345/6194/325/suppl/DC1
Materials and Methods
Figs. S1 to S5
Tables S1 to S4
References (*32*–*53*)

16 September 2013; accepted 17 June 2014
10.1126/science.1246067

ALTERNATIVE SPLICING

Human tRNA synthetase catalytic nulls with diverse functions

Wing-Sze Lo,[1,2] Elisabeth Gardiner,[3,4] Zhiwen Xu,[1,2] Ching-Fun Lau,[1,2] Feng Wang,[1,2] Jie J. Zhou,[1,2] John D. Mendlein,[4] Leslie A. Nangle,[4] Kyle P. Chiang,[4] Xiang-Lei Yang,[1,3] Kin-Fai Au,[5] Wing Hung Wong,[6] Min Guo,[7] Mingjie Zhang,[1,8] Paul Schimmel[1,3,7]*

Genetic efficiency in higher organisms depends on mechanisms to create multiple functions from single genes. To investigate this question for an enzyme family, we chose aminoacyl tRNA synthetases (AARSs). They are exceptional in their progressive and accretive proliferation of noncatalytic domains as the Tree of Life is ascended. Here we report discovery of a large number of natural catalytic nulls (CNs) for each human AARS. Splicing events retain noncatalytic domains while ablating the catalytic domain to create CNs with diverse functions. Each synthetase is converted into several new signaling proteins with biological activities "orthogonal" to that of the catalytic parent. We suggest that splice variants with nonenzymatic functions may be more general, as evidenced by recent findings of other catalytically inactive splice-variant enzymes.

Aminoacyl tRNA synthetases (AARSs) establish the genetic code by esterifying specific amino acids to the 3′ ends of their cognate tRNAs (*1*–*5*) and have adaptations of this reaction for specific physiological responses (*6*). A few literature examples show that natural proteolysis or alternative splicing of AARS can reveal previously unknown AARS proteins (*7*, *8*) with new functions (*9*–*11*). With this in mind, we investigated potential mechanisms for achieving genetic efficiency through functional expansions. The enzymes are divided into two classes of 10 proteins each, with each class being defined by the architecture of the highly conserved catalytic domain (CD) that is retained through evolution (*12*–*14*). As the Tree of Life is ascended, 13 new domains, which have no obvious association with aminoacylation or editing, have collectively been added to AARSs and maintained over the course of evolution, with no appreciable benefit or detriment to primary function (*15*–*17*). The extent of these domain additions appears to be particular to AARSs (*15*). Some of these new domains are appended to each of several synthetases, whereas others are specific to a single synthetase. Notably, these novel domain additions are accretive and progressive; and while their persistence provides no major benefit to aminoacylation, the strong evolutionary pressure for their retention suggests they are not random functionless stochastic fusions, but may be conserved for a specific biological purpose, perhaps distinct from the canonical enzymatic function.

We made a comprehensive search for alternative splice variants of AARSs to understand how splicing changes the domain organization and underlying architecture of each synthetase. We selectively targeted the AARS family of genes by enriching the AARS transcriptome in six distinct human samples [human fetal and adult brain, primary human leukocytes, and three cultured leukocyte cell types (Raji B cells, Jurkat T cells, and THP1 monocytes)]. A polymerase chain reaction (PCR)–based gene-capture and enrichment method was integrated with high-throughput deep sequencing to increase sequencing depth for each AARS transcript (materials and methods and Fig. 1A). This methodology allowed for high enrichment of AARS mRNAs and mainly targeted exon-exon junctions for discovery of exon-skipping events. We defined the AARS transcriptome as the transcripts of 37 AARS genes, including those for 17 cytoplasmic synthetases, 17 mitochondrial synthetases, and 3 that encode both cytoplasmic and mitochondrial forms. For efficient capture, transcripts were amplified by multiplex PCR using AARS gene–specific primers and optimized PCR conditions (see materials and methods). Sensitive detection of low-abundance splice variants was achieved with an optimized multiplex PCR that amplified gene regions close to exon-exon junctions of AARS transcripts and produced short PCR fragments (Fig. 1A). Fragments were assembled into cDNA libraries and sequenced by high-throughput deep sequencing (*18*, *19*).

Approximately 42 million 50-base reads were obtained and analyzed, using established methods (*19*). About 70% (30.4 million) mapped to the 37 AARS genes, and about two-thirds of the AARS-specific reads (21.4 million) covered AARS exon-exon junctions. When compared to previously published whole-transcriptome studies (*20*, *21*), the AARS transcriptome enrichment method employed here successfully improved sequencing depth so that we could detect all of the 61 previously reported exon-exon junctions for AARS transcripts, as well as identify 248 previously unreported junctions (Fig. 1B and table S1). These new splice forms allowed for the ablation of specific coding regions and simultaneous creation of new exon-exon junctions.

In addition, the tissue origin and the overlap of AARS splice variants in different tissues were examined. Although there was obvious tissue

[1]IAS HKUST–Scripps R&D Laboratory, Institute for Advanced Study, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, China. [2]Pangu Biopharma, Edinburgh Tower, The Landmark, 15 Queen's Road Central, Hong Kong, China. [3]The Scripps Laboratories for tRNA Synthetase Research, The Scripps Research Institute, 10650 North Torrey Pines Road, La Jolla, CA 92037, USA. [4]aTyr Pharma, 3545 John Hopkins Court, Suite 250, San Diego, CA 92121, USA. [5]Department of Internal Medicine, University of Iowa, Iowa City, IA 52242, USA. [6]Department of Statistics, Stanford University, Stanford, CA 94305, USA. [7]The Scripps Laboratories for tRNA Synthetase Research, Scripps Florida, 130 Scripps Way, Jupiter, FL 33458, USA. [8]Division of Life Science, State Key Laboratory of Molecular Neuroscience, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, China.
***Corresponding author. E-mail: schimmel@scripps.edu**

specificity for certain transcripts, many of the same splice variant transcripts were found across distinct tissue pools (Fig. 1B). Surprisingly, most of the splice variants of both class I and class II family members abrogated the CD (Fig. 2A and fig. S1). These included both truncations of N- or C-terminal coding regions, as well as in-frame internal deletions (Fig. 2B). For instance, 79% of the 66 discovered in-frame splice variants (Fig. 2C and table S2) had a disrupted or ablated canonical CD and thereby created a catalytic null (CN) (Fig. 2B and fig. S1). Because three-dimensional (3D) structures are available for many human AARSs and their orthologs, events that removed entire specific exons could be diagrammatically portrayed as linear arrangements of domain structure elements (fig. S1). These virtual structures suggest that the new domain-domain interactions created by internal deletions might engender new structural conformations [compare (*22*)] and thereby might lead to new interactions.

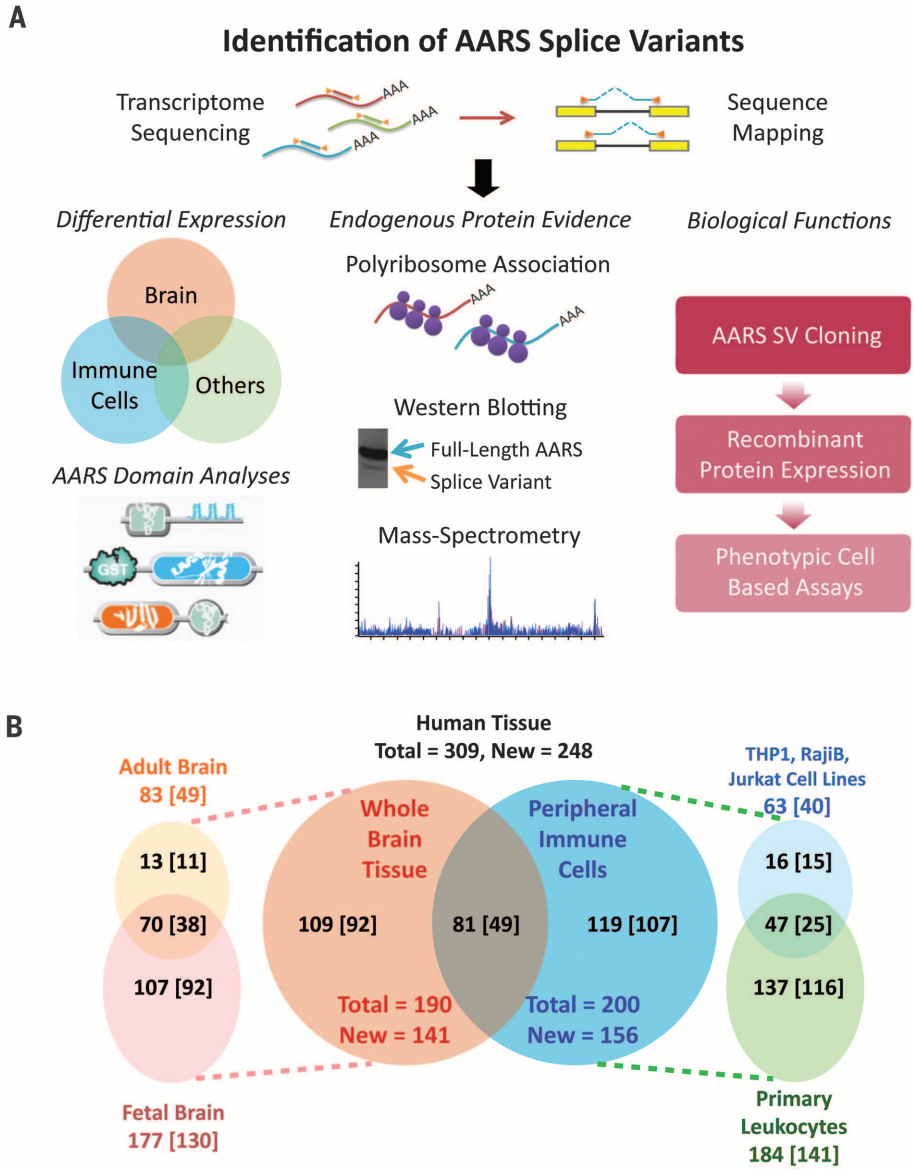


Fig. 1. Identification of AARS splice variants and assays for their functions. (**A**) AARS transcriptome discovery and analysis funnel. The AARS transcriptome was amplified by gene-specific multiplex PCR and sequenced by high-throughput deep sequencing. Sequencing reads were then mapped to the AARS transcriptome for identification of exon-skipping events and alternative splice sites. Products of this effort were submitted to various analyses. Regulation and functions of the identified AARS splice variants were studied in the context of their distribution across various human tissues, through identification of the endogenously expressed protein products, and the effects of expressed protein products on cells in a diverse set of in vitro cell-based assays. (**B**) The Venn diagram in the center shows the total number of exon-skipping junctions that were identified by RNA sequencing and by "exon-flanking" PCR in fetal and adult brain tissues (orange circle) and in immune cells that include primary total leukocytes and three different cultured leukocytic cell types (blue circle). In the diagram on the left, the light pink circle annotates the number of exon-skipping junctions identified in fetal brain, while the light orange circle shows those found in adult brain. The diagram on the right gives the number of identified exon-skipping junctions in the three cultured leukocytic cell types (Raji B cells, Jurkat T cells, and THP1 monocytes) (light blue) and in primary leukocytes (pale green). Overlapping areas give the number of splice junctions that were common between tissues or cells. The numbers in square brackets give the total new alternative splice junctions identified in each respective tissue.

The tissue-specific association of transcripts suggested that AARS mRNA splice variants encode endogenously expressed proteins. To explore this possibility, we examined polysome association of the splice variant–encoding mRNAs [materials and methods as described in (*24*)]. Of the 48 CN mRNAs tested, all were associated with polysomes in naïve Jurkat cells (fig. S2 and table S3). AARS-specific antibodies probed the same Jurkat cell lysates that were used for detecting polysome association of the mRNAs. To detect endogenous translation products of the CN splice variant transcripts, we performed Western blot analysis with antibodies specific for AlaRS, CysRS, LysRS, TyrRS, and ValRS (Fig. 3A and table S4). These synthetase fragments were chosen based on the availability of suitable antibodies for immunoprecipitation. By Western blot analysis, we detected the expected endogenous AARS splice forms that lacked CDs but retained appended domains (Fig. 3A). Mass spectrometry identified specific GlnRS, ValRS, and TyrRS CN-sized fragments, and multiple peptides were identified for all of these CNs (fig. S3A). In addition to finding representative peptides from these CNs, we found no support for the possibility of

As specific examples, all eight in-frame splice variants of HisRS showed an ablated CD, and only one of the six in-frame splice variants of TyrRS retained the CD (Fig. 2B and table S2). In contrast to the consistent abrogation of the canonical CD, 60 of the 70 in-frame splice variants (85%) are CNs that retain at least one of the 13 added domains appearing in the AARSs of higher eukaryotes (Fig. 2B and table S2). Of particular interest are the UNE domains, which are specific to AARSs and have, like the other appended domains, no notable aminoacylation function. The UNE domains are almost universally retained in the CNs (Fig. 2B and fig. S1). An interesting case, suggesting a noncanonical role for the CNs, is the retention of the UNE-S domain of SerRS. Recent work established a nuclear activity for SerRS that is dependent on the UNE-S domain and showed that the addition of UNE-S to SerRS was essential for development of the closed circulatory systems of vertebrates (*23*). Motifs found in other proteins of higher eukaryotes, such as the glutathione S-transferase domain (GST)-, single-helix–, WHEP [named for discovery in TrpRS (W), HisRS(H), GluProRS (EP)]-, and endothelial monocyte-activating polypeptide II (EMAPII)-like-domain, also remain intact in many of the AARS CN splice variants (fig. S1 and table S2).

proteolytic cleavage of full-length TyrRS giving rise to its assigned CN peptides (legend to fig. S3A). In a separate experiment, we identified a 23-kD protein as HisRS1-C9 in the public PROTOMAP mass spectrometry (MS) database (*25*). We aligned MS-scored peptides on both sides of the sequence encompassing the splice junction reported here for HisRS1-C9 (fig. S3B). Finally, in vitro translation of a copy of the mRNA encoding an endogenously expressed TyrRS1-C7 splice variant (identified by Western blot analysis of whole-cell lysates as shown in Fig. 3) confirmed that the transcripts could be stably translated into proteins (fig. S3C). MS confirmed peptides on both sides of the internal splice junction.

We observed tissue-specific expression of specific CNs. Across 19 human adult tissues or cells, 38 of 48 CN transcripts (79%) were differentially expressed with gene up-regulation (by five times or more of median) in at least one of the tissues, whereas the full-length parent AARS genes were evenly distributed (table S3). We also found that some CN transcripts were expressed differentially in one developmental stage over another. For example, six specific CNs were highly expressed (by 10-fold or more) in adult versus fetal lung tissue. These included ArgRS1-AS01, CysRS1-AS04, MetRS1-AS13, SerRS1-AS02, ThrRS1-AS05, and TyrRS1-AS10 (Fig. 3B and table S3).

Because the splice-variant mRNAs prominently ablate the CD-encoding portion, we sought to investigate the potential for these fragments to exert biological activities distinct from the canonical aminoacylation function. To this end, recombinant human AARS fragments, including CNs, were expressed as soluble proteins and purified to >95% homogeneity. Phenotypic cell-based assays were performed largely in primary human cells to monitor potential biological activities (fig. S4 and table S5). The assay types were clustered into assay groups (Fig. 4), including proliferation (different cell types were profiled for effects of splice variants on proliferation or cell death), cytoprotection, immunomodulation, acute inflammatory response, transcriptional regulation (four assays in two cell types at two distinct time points across a set of 88 genes), "regenerative responses," cell differentiation in primary human cell types and, finally, cholesterol transport. All assays were run at minimum in duplicate for each protein, and many proteins

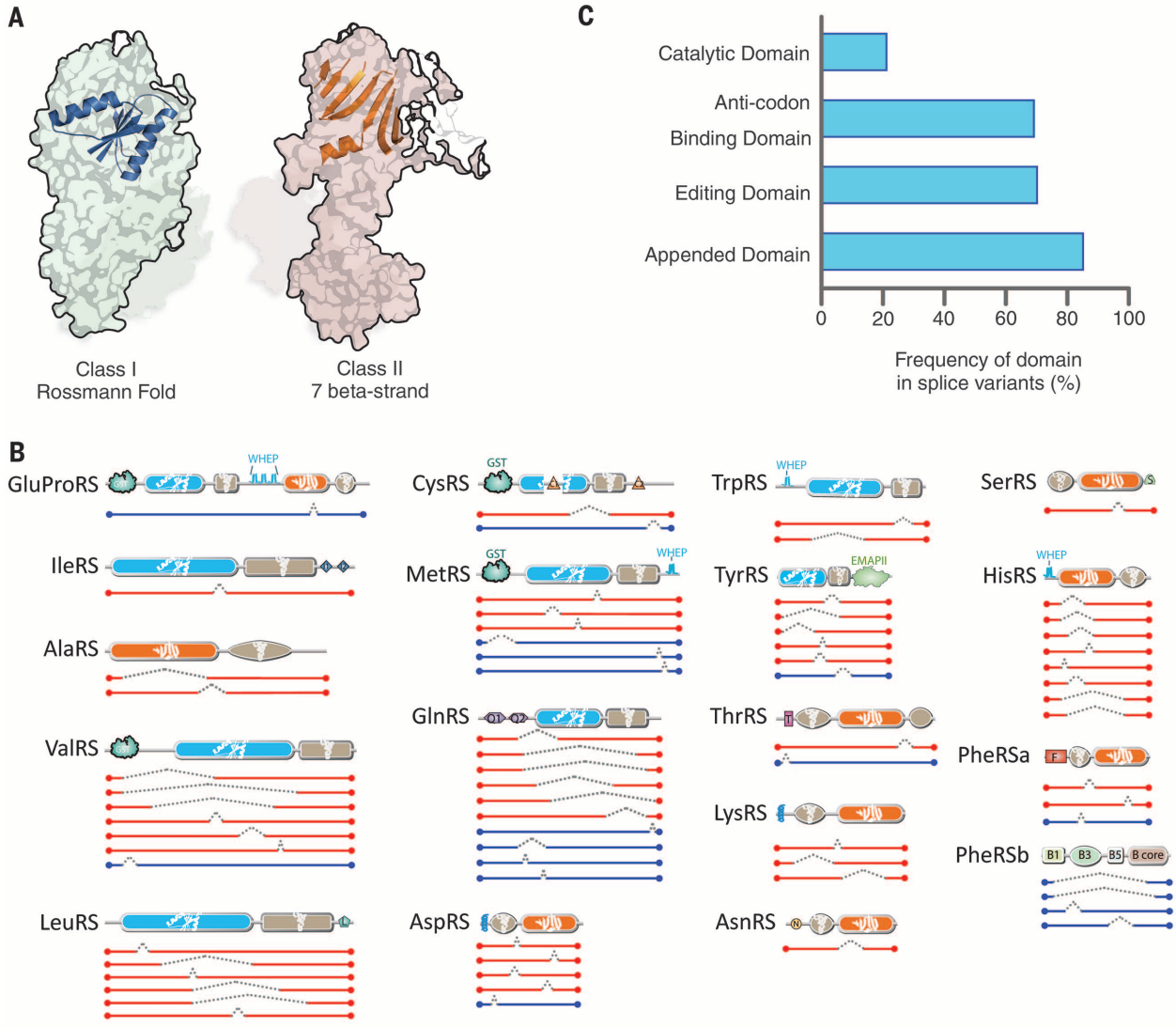


Fig. 2. Most AARS splice variants are catalytic nulls. (**A**) Architectures of the CDs for aminoacylation—class I versus class II. The conserved core Rossmann fold is represented on the structure of MetRS [Protein Data Base (PDB) code: 2CT8] (*33*) in class I, and the conserved core 7 β strand with motif-3 helix is represented on the structure of LysRS (PDB: 4DPG) (*34*) in class II. (**B**) In-frame splice variants of cytoplasmic AARS are illustrated. Splice variants with CDs deleted (catalytic nulls) are highlighted in red whereas those with CDs retained are represented in blue. (**C**) The CD is abrogated in most AARS splice variants. By contrast, domains that have functions distinct from aminoacylation are predominantly retained. Of note, the UNE domains [such as UNE-S, and UNE-L; abbreviated as S and L, respectively, in (B)], which constitute part of the "Appended Domain" category and are idiosyncratic to specific synthetases, are retained in the AARS splice variants identified here.

were run in multiple batches, and at a range of concentrations, to confirm activity. All proteins were generated as His-tagged recombinant forms, with either the N or C termini, or both, having the tag (table S6). Full-length forms of AspRS, TyrRS, HisRS, and AsnRS synthetases were expressed in parallel and run in assays as controls for the expressed synthetase fragments. In all cases, the full-length parental form was either inactive across all assays or had a single activity that was not the same as any of its splice variants.

More than 100,000 data points were evaluated across the cell-based assay panel (fig. S4). Of the 94 AARS-derived proteins analyzed here, 88% tested positive for one or more biological activities. The cell-based activities associated with each recombinant protein were specific and idiosyncratic to the variant. This observation provided a system-wide "internal control," largely ruling out the potential for nonspecific readouts of cell signaling by the various proteins. MetRS1-C5 is presented as a specific example. This CN strongly stimulated skeletal muscle fiber formation in vitro (fig. S5). After exposure to the recombinant MetRS1-C5 for 2 days, quantitative PCR assessment of primary human skeletal myoblasts showed up-regulation of key genes for

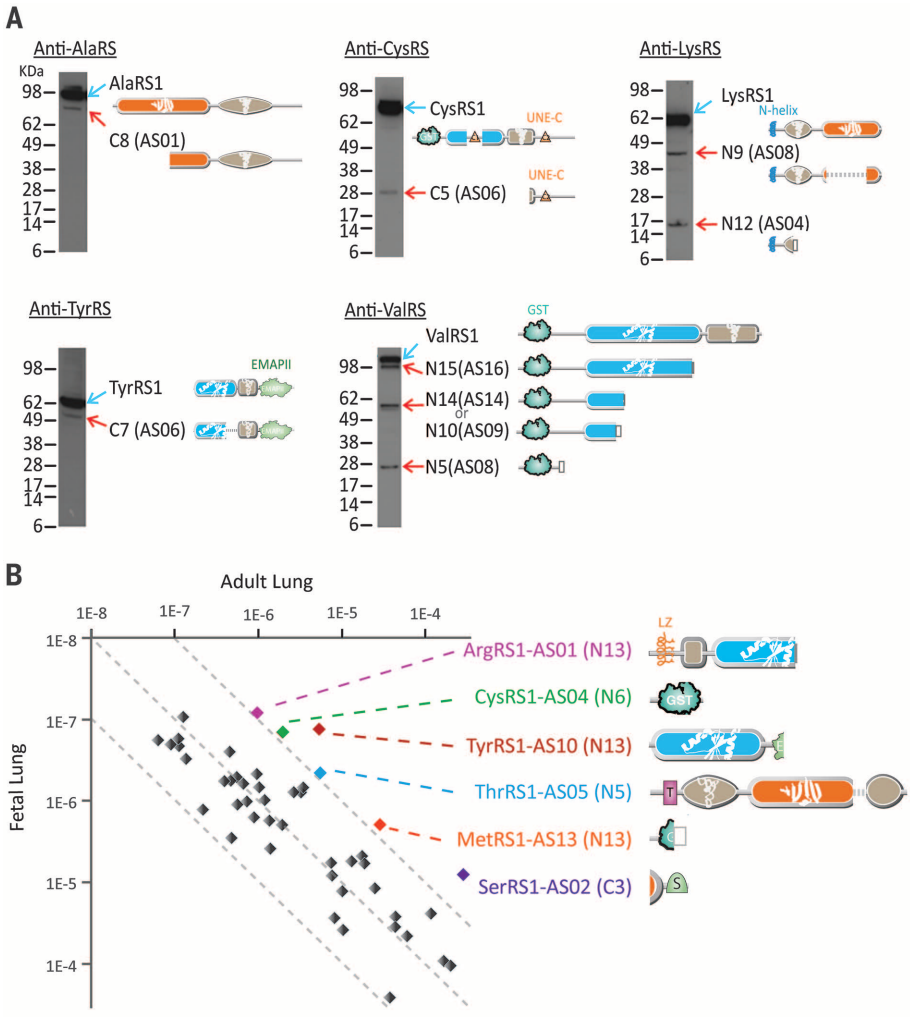


Fig. 3. Detection of endogenous AARS splice variants. (**A**) Western blot detection of AARS splice variants in Jurkat cell lysates. Detailed information for these splice variants is shown in table S4. (**B**) Tissue-specific expression of selected AARS splice variants in adult and fetal lung tissues. Gene expression of the target gene was normalized by the gene expression of house-keeping genes (see materials and methods) in the same sample.

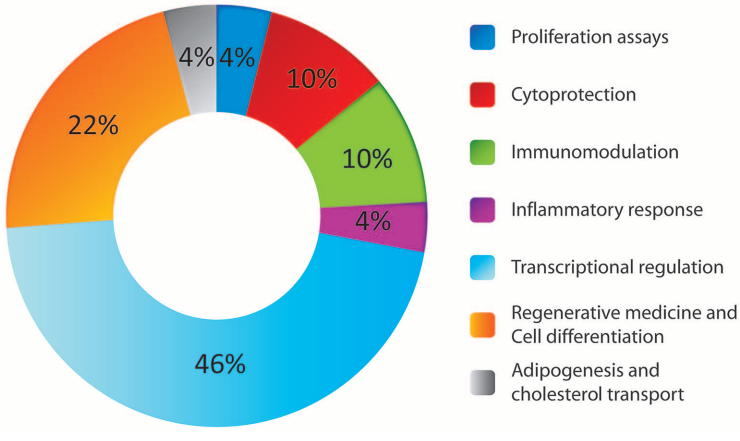


Fig. 4. Recombinant AARS variants have specific biological activities across a spectrum of cell-based assays. Proteins were expressed in *Escherichia coli* and purified for use in cell-based assays. Most of the proteins were soluble and highly expressed. The AARS variants (table S6) were tested in a variety of different cell-based assays (fig. S4) spanning a range of biological activities, largely using primary human cells (table S5).

muscle cell differentiation and metabolism, including insulin growth factor 1 (*IGFI*) and lipoprotein lipase (*LPL*) (fig. S6).

While deliberately ablating the canonical catalytic function, alternative splicing of the AARS family of genes has created a large ensemble of CNs that specifically retain the domain expansions. The successful expression of more than 100 recombinant forms as soluble proteins suggests that splice-site selection has been tailored to create stable folded structures. The canonical function and structure of the ancient aminoacyl tRNA synthetase CD are strongly preserved throughout all taxa, which makes the ablation of this essential domain (for aminoacylation) especially provocative. The paradox of strongly conserved noncatalytic domains progressively added to AARSs protein structure over the course of evolution appears to be at least in part an evolutionary reshaping of tRNA synthetases for other functions.

Although splice variants of other proteins also exist, it is the extent of these novel domain additions specifically to AARSs, and their retention by the CNs, that make the AARSs splice variants distinct. Possibly, functional expansion of AARSs was to link translation at the first step of protein synthesis to a variety of cell signaling pathways. Recent studies have demonstrated roles for specific AARSs in pathways associated with angiogenesis (*9*, *26*–*28*), inflammation (*29*, *30*), the immune response, mammalian target of rapamycin (mTOR) signaling, apoptosis, tumorigenesis, and interferon-γ (IFN-γ) and p53 signaling (*15*). The work detailed here suggests that the universe of AARS-derived entities, which are active for nontranslational functions, may be far greater than anticipated. The mechanism of erasing the canonical function, while adding noncatalytic domains, engenders a clear implementation of orthogonal functions. Members of other enzyme families, though perhaps to a lesser extent, likely also gain new functions through splice variants. The recently reported catalytically impaired natural splice variants of several oncogenic kinases (*31*) and of the sirtuin-2 (SIRT2) histone deacetylase (*32*) suggest that other enzyme families have undergone similar, though perhaps less extensive, variation.

REFERENCES AND NOTES

1. M. T. Boniecki, M. T. Vu, A. K. Betha, S. A. Martinis, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 19223–19228 (2008).
2. S. S. Yadavalli, K. Musier-Forsyth, M. Ibba, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 19031–19032 (2008).
3. R. Giegé, *Nat. Struct. Biol.* **10**, 414–416 (2003).
4. C. W. Carter Jr., *Annu. Rev. Biochem.* **62**, 715–748 (1993).
5. M. Ibba, D. Söll, *Genes Dev.* **18**, 731–738 (2004).
6. N. Netzer *et al.*, *Nature* **462**, 522–526 (2009).
7. H. Schwenzer *et al.*, *Biochimie* **100**, 18–26 (2014).
8. M. Miyanokoshi, T. Tanaka, M. Tamai, Y. Tagawa, K. Wakasugi, *Sci. Rep.* **3**, 3477 (2013).
9. K. Wakasugi *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 173–177 (2002).
10. K. Wakasugi, P. Schimmel, *Science* **284**, 147–151 (1999).
11. L. F. Lareau, R. E. Green, R. S. Bhatnagar, S. E. Brenner, *Curr. Opin. Struct. Biol.* **14**, 273–282 (2004).
12. S. W. Ludmerer, P. Schimmel, *J. Biol. Chem.* **262**, 10807–10813 (1987).
13. G. Eriani, M. Delarue, O. Poch, J. Gangloff, D. Moras, *Nature* **347**, 203–206 (1990).
14. S. Cusack, *Curr. Opin. Struct. Biol.* **7**, 881–889 (1997).
15. M. Guo, P. Schimmel, *Nat. Chem. Biol.* **9**, 145–153 (2013).
16. M. Guo, X. L. Yang, P. Schimmel, *Nat. Rev. Mol. Cell Biol.* **11**, 668–674 (2010).
17. J. Jia, A. Arif, P. S. Ray, P. L. Fox, *Mol. Cell* **29**, 679–690 (2008).
18. J. Takeda *et al.*, *Nucleic Acids Res.* **34**, 3917–3928 (2006).
19. H. Jiang, W. H. Wong, *Bioinformatics* **25**, 1026–1032 (2009).
20. J. M. Toung, M. Morley, M. Li, V. G. Cheung, *Genome Res.* **21**, 991–998 (2011).
21. D. Thierry-Mieg, J. Thierry-Mieg, *Genome Biol.* **7** (suppl. 1), S12.1–14 (2006).
22. Z. Xu *et al.*, *Structure* **20**, 1470–1477 (2012).
23. X. Xu *et al.*, *Nat. Commun.* **3**, 681 (2012).
24. F. Wang *et al.*, *J. Biol. Chem.* **288**, 29223–29228 (2013).
25. M. M. Dix, G. M. Simon, B. F. Cravatt, *Cell* **134**, 679–691 (2008).
26. E. Tzima *et al.*, *J. Biol. Chem.* **280**, 2405–2408 (2005).
27. A. Kawahara, D. Y. Stainier, *Trends Cardiovasc. Med.* **19**, 179–182 (2009).
28. Q. Zhou *et al.*, *Nat. Struct. Mol. Biol.* **17**, 57–61 (2010).
29. S. G. Park *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 6356–6361 (2005).
30. A. Arif, J. Jia, R. A. Moodt, P. E. DiCorleto, P. L. Fox, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 1415–1420 (2011).
31. K. Anamika, N. Garnier, N. Srinivasan, *BMC Genomics* **10**, 622 (2009).
32. J. G. Rack, M. R. VanLinden, T. Lutter, R. Aasland, M. Ziegler, *J. Mol. Biol.* **426**, 1677–1691 (2014).
33. K. Nakanishi, Y. Ogiso, T. Nakama, S. Fukai, O. Nureki, *Nat. Struct. Mol. Biol.* **12**, 931–932 (2005).
34. Y. Ofir-Birin *et al.*, *Mol. Cell* **49**, 30–42 (2013).

ACKNOWLEDGMENTS

This work was supported by the Innovation and Technology Fund from the Hong Kong Government (UIM181, UIM192, and UIM199); by a fellowship from the National Foundation for Cancer Research; and by NIH grants R01CA92577, R01GM088278, R01NS085092, R01HG005717, and R01GM100136. We thank K. Piehl and J. Li (aTyr Pharma) for help with splice variant cloning and protein expression. We also thank A. Cubitt (aTyr Pharma) for assistance on the transcriptional profiling, V. Trinh and J. Zhao (both formerly of aTyr Pharma) for assistance with cell profiling of recombinant tRNA synthetase fragments, and B. Cravatt (Scripps Research Institute) for help with interpretation of the PROTOMAP data. P.S. is a cofounder and member of the Board of Directors of aTyr Pharma, W.H.W. is a member of the aTyr Pharma Scientific Advisory Board, and X.-L.Y is a scientific cofounder and a member of the Scientific Advisory Board of aTyr Pharma. E.G. has no financial interest with aTyr Pharma. Clones of catalytic nulls are available for research purposes under a Material Transfer Agreement with aTyr Pharma. W.-S.L, L.A.N, and K.P.C. are inventors on patent applications filed by aTyr Pharma Inc. and Pangu Biopharma Ltd. that cover the splice variants described in this work.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/345/6194/328/suppl/DC1
Materials and Methods
Figs. S1 to S6
Tables S1 to S6

5 March 2014; accepted 18 June 2014
10.1126/science.1252943

CELL DIVISION

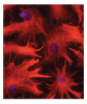
Feedback control of chromosome separation by a midzone Aurora B gradient

Olga Afonso,[1]* Irina Matos,[1]*† António J. Pereira,[1] Paulo Aguiar,[1,2] Michael A. Lampson,[3] Helder Maiato[1,4]‡

Accurate chromosome segregation during mitosis requires the physical separation of sister chromatids before nuclear envelope reassembly (NER). However, how these two processes are coordinated remains unknown. Here, we identified a conserved feedback control mechanism that delays chromosome decondensation and NER in response to incomplete chromosome separation during anaphase. A midzone-associated Aurora B gradient was found to monitor chromosome position along the division axis and to prevent premature chromosome decondensation by retaining Condensin I. PP1/PP2A phosphatases counteracted this gradient and promoted chromosome decondensation and NER. Thus, an Aurora B gradient appears to mediate a surveillance mechanism that prevents chromosome decondensation and NER until effective separation of sister chromatids is achieved. This allows the correction and reintegration of lagging chromosomes in the main nuclei before completion of NER.

The formation of a nuclear envelope that compartmentalizes genomic DNA involves the recruitment of membranes around the decondensing chromatin and insertion of nuclear pore complexes (NPCs) at the anaphase-telophase transition of mitosis (*1*). However, it is unknown how cells coordinate nuclear envelope reassembly (NER) with the spatial separation of chromosomes during anaphase. Here, we found that the spindle elongation velocity and the respective duration of anaphase in *Drosophila* S2 cells were inversely correlated, which suggested that incomplete chromosome separation in spindles that elongate more slowly is compensated by increasing anaphase duration (fig. S1, A to C). Pharmacological or RNA interference (RNAi)–based attenuation of spindle elongation velocity also correlated with an increase in

[1]Chromosome Instability and Dynamics Laboratory, Instituto de Biologia Molecular e Celular, Universidade do Porto, Rua do Campo Alegre 823, 4150-180 Porto, Portugal. [2]Center for Mathematics, Universidade do Porto, Rua do Campo Alegre 687, 4169-007 Porto, Portugal. [3]Department of Biology, University of Pennsylvania, Philadelphia, PA 19104, USA. [4]Cell Division Unit, Department of Experimental Biology, Faculdade de Medicina, Universidade do Porto, Alameda Prof. Hernâni Monteiro, 4200-319 Porto, Portugal.
*These authors contributed equally to this work. †Present address: Laboratory of Mammalian Cell Biology and Development, Howard Hughes Medical Institute, The Rockefeller University, New York, NY 10065, USA. ‡**Corresponding author. E-mail: maiato@ibmc.up.pt**

EXTENDED PDF FORMAT
SPONSORED BY

Travel Grants Available
Neuroscience 2014
Application Deadline: July 31st
>Apply Now
R&D systems
a bio-techne brand
www.rndsystems.com

Feedback control of chromosome separation by a midzone Aurora B gradient

Olga Afonso *et al.*

Science **345**, 332 (2014);
DOI: 10.1126/science.1251121

This copy is for your personal, non-commercial use only.

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by clicking here.

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines here.

The following resources related to this article are available online at www.sciencemag.org (this information is current as of August 4, 2014):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:
http://www.sciencemag.org/content/345/6194/332.full.html

Supporting Online Material can be found at:
http://www.sciencemag.org/content/suppl/2014/06/11/science.1251121.DC1.html

A list of selected additional articles on the Science Web sites **related to this article** can be found at:
http://www.sciencemag.org/content/345/6194/332.full.html#related

This article **cites 29 articles**, 14 of which can be accessed free:
http://www.sciencemag.org/content/345/6194/332.full.html#ref-list-1

This article has been **cited by** 1 articles hosted by HighWire Press; see:
http://www.sciencemag.org/content/345/6194/332.full.html#related-urls

This article appears in the following **subject collections:**
Cell Biology
http://www.sciencemag.org/cgi/collection/cell_biol

Downloaded from www.sciencemag.org on August 4, 2014

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published weekly, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. Copyright 2014 by the American Association for the Advancement of Science; all rights reserved. The title *Science* is a registered trademark of AAAS.

muscle cell differentiation and metabolism, including insulin growth factor 1 (*IGFI*) and lipoprotein lipase (*LPL*) (fig. S6).

While deliberately ablating the canonical catalytic function, alternative splicing of the AARS family of genes has created a large ensemble of CNs that specifically retain the domain expansions. The successful expression of more than 100 recombinant forms as soluble proteins suggests that splice-site selection has been tailored to create stable folded structures. The canonical function and structure of the ancient aminoacyl tRNA synthetase CD are strongly preserved throughout all taxa, which makes the ablation of this essential domain (for aminoacylation) especially provocative. The paradox of strongly conserved noncatalytic domains progressively added to AARSs protein structure over the course of evolution appears to be at least in part an evolutionary reshaping of tRNA synthetases for other functions.

Although splice variants of other proteins also exist, it is the extent of these novel domain additions specifically to AARSs, and their retention by the CNs, that make the AARSs splice variants distinct. Possibly, functional expansion of AARSs was to link translation at the first step of protein synthesis to a variety of cell signaling pathways. Recent studies have demonstrated roles for specific AARSs in pathways associated with angiogenesis (*9*, *26*–*28*), inflammation (*29*, *30*), the immune response, mammalian target of rapamycin (mTOR) signaling, apoptosis, tumorigenesis, and interferon-γ (IFN-γ) and p53 signaling (*15*). The work detailed here suggests that the universe of AARS-derived entities, which are active for nontranslational functions, may be far greater than anticipated. The mechanism of erasing the canonical function, while adding noncatalytic domains, engenders a clear implementation of orthogonal functions. Members of other enzyme families, though perhaps to a lesser extent, likely also gain new functions through splice variants. The recently reported catalytically impaired natural splice variants of several oncogenic kinases (*31*) and of the sirtuin-2 (SIRT2) histone deacetylase (*32*) suggest that other enzyme families have undergone similar, though perhaps less extensive, variation.

REFERENCES AND NOTES

1. M. T. Boniecki, M. T. Vu, A. K. Betha, S. A. Martinis, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 19223–19228 (2008).
2. S. S. Yadavalli, K. Musier-Forsyth, M. Ibba, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 19031–19032 (2008).
3. R. Giegé, *Nat. Struct. Biol.* **10**, 414–416 (2003).
4. C. W. Carter Jr., *Annu. Rev. Biochem.* **62**, 715–748 (1993).
5. M. Ibba, D. Söll, *Genes Dev.* **18**, 731–738 (2004).
6. N. Netzer *et al.*, *Nature* **462**, 522–526 (2009).
7. H. Schwenzer *et al.*, *Biochimie* **100**, 18–26 (2014).
8. M. Miyanokoshi, T. Tanaka, M. Tamai, Y. Tagawa, K. Wakasugi, *Sci. Rep.* **3**, 3477 (2013).
9. K. Wakasugi *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 173–177 (2002).
10. K. Wakasugi, P. Schimmel, *Science* **284**, 147–151 (1999).
11. L. F. Lareau, R. E. Green, R. S. Bhatnagar, S. E. Brenner, *Curr. Opin. Struct. Biol.* **14**, 273–282 (2004).
12. S. W. Ludmerer, P. Schimmel, *J. Biol. Chem.* **262**, 10807–10813 (1987).
13. G. Eriani, M. Delarue, O. Poch, J. Gangloff, D. Moras, *Nature* **347**, 203–206 (1990).
14. S. Cusack, *Curr. Opin. Struct. Biol.* **7**, 881–889 (1997).
15. M. Guo, P. Schimmel, *Nat. Chem. Biol.* **9**, 145–153 (2013).
16. M. Guo, X. L. Yang, P. Schimmel, *Nat. Rev. Mol. Cell Biol.* **11**, 668–674 (2010).
17. J. Jia, A. Arif, P. S. Ray, P. L. Fox, *Mol. Cell* **29**, 679–690 (2008).
18. J. Takeda *et al.*, *Nucleic Acids Res.* **34**, 3917–3928 (2006).
19. H. Jiang, W. H. Wong, *Bioinformatics* **25**, 1026–1032 (2009).
20. J. M. Toung, M. Morley, M. Li, V. G. Cheung, *Genome Res.* **21**, 991–998 (2011).
21. D. Thierry-Mieg, J. Thierry-Mieg, *Genome Biol.* **7** (suppl. 1), S12.1–14 (2006).
22. Z. Xu *et al.*, *Structure* **20**, 1470–1477 (2012).
23. X. Xu *et al.*, *Nat. Commun.* **3**, 681 (2012).
24. F. Wang *et al.*, *J. Biol. Chem.* **288**, 29223–29228 (2013).
25. M. M. Dix, G. M. Simon, B. F. Cravatt, *Cell* **134**, 679–691 (2008).
26. E. Tzima *et al.*, *J. Biol. Chem.* **280**, 2405–2408 (2005).
27. A. Kawahara, D. Y. Stainier, *Trends Cardiovasc. Med.* **19**, 179–182 (2009).
28. Q. Zhou *et al.*, *Nat. Struct. Mol. Biol.* **17**, 57–61 (2010).
29. S. G. Park *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 6356–6361 (2005).
30. A. Arif, J. Jia, R. A. Moodt, P. E. DiCorleto, P. L. Fox, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 1415–1420 (2011).
31. K. Anamika, N. Garnier, N. Srinivasan, *BMC Genomics* **10**, 622 (2009).
32. J. G. Rack, M. R. VanLinden, T. Lutter, R. Aasland, M. Ziegler, *J. Mol. Biol.* **426**, 1677–1691 (2014).
33. K. Nakanishi, Y. Ogiso, T. Nakama, S. Fukai, O. Nureki, *Nat. Struct. Mol. Biol.* **12**, 931–932 (2005).
34. Y. Ofir-Birin *et al.*, *Mol. Cell* **49**, 30–42 (2013).

ACKNOWLEDGMENTS

This work was supported by the Innovation and Technology Fund from the Hong Kong Government (UIM181, UIM192, and UIM199); by a fellowship from the National Foundation for Cancer Research; and by NIH grants R01CA92577, R01GM088278, R01NS085092, R01HG005717, and R01GM100136. We thank K. Piehl and J. Li (aTyr Pharma) for help with splice variant cloning and protein expression. We also thank A. Cubitt (aTyr Pharma) for assistance on the transcriptional profiling, V. Trinh and J. Zhao (both formerly of aTyr Pharma) for assistance with cell profiling of recombinant tRNA synthetase fragments, and B. Cravatt (Scripps Research Institute) for help with interpretation of the PROTOMAP data. P.S. is a cofounder and member of the Board of Directors of aTyr Pharma, W.H.W. is a member of the aTyr Pharma Scientific Advisory Board, and X.-L.Y is a scientific cofounder and a member of the Scientific Advisory Board of aTyr Pharma. E.G. has no financial interest with aTyr Pharma. Clones of catalytic nulls are available for research purposes under a Material Transfer Agreement with aTyr Pharma. W.-S.L, L.A.N, and K.P.C. are inventors on patent applications filed by aTyr Pharma Inc. and Pangu Biopharma Ltd. that cover the splice variants described in this work.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/345/6194/328/suppl/DC1
Materials and Methods
Figs. S1 to S6
Tables S1 to S6

5 March 2014; accepted 18 June 2014
10.1126/science.1252943

CELL DIVISION

Feedback control of chromosome separation by a midzone Aurora B gradient

Olga Afonso,[1]* Irina Matos,[1]*† António J. Pereira,[1] Paulo Aguiar,[1,2] Michael A. Lampson,[3] Helder Maiato[1,4]‡

Accurate chromosome segregation during mitosis requires the physical separation of sister chromatids before nuclear envelope reassembly (NER). However, how these two processes are coordinated remains unknown. Here, we identified a conserved feedback control mechanism that delays chromosome decondensation and NER in response to incomplete chromosome separation during anaphase. A midzone-associated Aurora B gradient was found to monitor chromosome position along the division axis and to prevent premature chromosome decondensation by retaining Condensin I. PP1/PP2A phosphatases counteracted this gradient and promoted chromosome decondensation and NER. Thus, an Aurora B gradient appears to mediate a surveillance mechanism that prevents chromosome decondensation and NER until effective separation of sister chromatids is achieved. This allows the correction and reintegration of lagging chromosomes in the main nuclei before completion of NER.

The formation of a nuclear envelope that compartmentalizes genomic DNA involves the recruitment of membranes around the decondensing chromatin and insertion of nuclear pore complexes (NPCs) at the anaphase-telophase transition of mitosis (*1*). However, it is unknown how cells coordinate nuclear envelope reassembly (NER) with the spatial separation of chromosomes during anaphase. Here, we found that the spindle elongation velocity and the respective duration of anaphase in *Drosophila* S2 cells were inversely correlated, which suggested that incomplete chromosome separation in spindles that elongate more slowly is compensated by increasing anaphase duration (fig. S1, A to C). Pharmacological or RNA interference (RNAi)–based attenuation of spindle elongation velocity also correlated with an increase in

[1]Chromosome Instability and Dynamics Laboratory, Instituto de Biologia Molecular e Celular, Universidade do Porto, Rua do Campo Alegre 823, 4150-180 Porto, Portugal. [2]Center for Mathematics, Universidade do Porto, Rua do Campo Alegre 687, 4169-007 Porto, Portugal. [3]Department of Biology, University of Pennsylvania, Philadelphia, PA 19104, USA. [4]Cell Division Unit, Department of Experimental Biology, Faculdade de Medicina, Universidade do Porto, Alameda Prof. Hernâni Monteiro, 4200-319 Porto, Portugal.
*These authors contributed equally to this work. †Present address: Laboratory of Mammalian Cell Biology and Development, Howard Hughes Medical Institute, The Rockefeller University, New York, NY 10065, USA. ‡**Corresponding author. E-mail: maiato@ibmc.up.pt**

anaphase duration (Fig. 1, A to D, and fig. S1, D to G). A similar response was observed in human cells (fig. S2, A to C and F to H, and fig. S3, A to C). NER was also delayed on lagging chromosomes and DNA bridges relative to the main nuclei (Fig. 2, A to C). In most cases (~50%), this promoted the correction and reintegration of lagging chromosomes into the main nuclei, and only a smaller fraction (~20%) formed micronuclei. Thus, the anaphase-telophase transition in metazoans cannot be explained by a "clock" model that is set at the onset of anaphase but must take into account the effective separation of sister chromatids before triggering NER. This spatial control of NER appears to be important for the fidelity of chromosome segregation.

The chromosomal passenger protein Aurora B relocates from centromeres to the spindle midzone in anaphase, which produces a phosphorylation gradient (*2*). Aurora B inactivation on chromatin is also required for chromosome decondensation and NER (*3*). We noticed that NER was inversely correlated with Aurora B activity on chromosomes and to the proximity of the spindle midzone (Fig. 2, D and D′; fig. S3, D and D′; and fig. S4, A and A′). Measurement of Aurora B activity in S2 cells using a chromatin-targeted Förster resonance energy transfer (FRET) sensor (*2*) revealed a lower Aurora B activity (FRET increase) as chromosomes separated during anaphase (fig. S4, B to G). Thus, chromosome position along the division axis appears to be monitored by a midzone-associated Aurora B activity gradient that spatially controls NER.

To test this idea, we performed laser microsurgery in metaphase S2 cells to generate acentric chromosome fragments (i.e., devoid of kinetochores and centromeric Aurora B) and found that NER was significantly delayed or inhibited on the lagging acentric fragments (9 out of 11 cells) (Fig. 3, A and B). Furthermore, Aurora B accumulation at the spindle midzone and the respective duration of anaphase were correlated (fig. S5, A and C). Prevention of Aurora B localization at the spindle midzone by depletion of the conserved kinesin-6 Subito (Mklp2) (*4*, *5*) extended spindle elongation and anaphase duration (Fig. 3C and fig. S5, A, B, D, and E). Given that global Aurora B inhibition at anaphase onset with the *Drosophila*-specific Aurora B inhibitor Binucleine-2 (Bi-2) (*6*) did not perturb anaphase duration (fig. S6A and Fig. 4, A and B), we attributed the anaphase delay after Subito RNAi to Aurora B retention on chromatin (*3*, *4*). NER occurred simultaneously on all chromosomes after Subito RNAi (7 out of 8 cells) or global Aurora B inhibition (12 out of 12 cells), independently of their position along the cell division axis (Fig. 3D and fig. S6A). Thus, Aurora B localization at the spindle midzone is required for spatial regulation of NER. Global Aurora B inhibition at anaphase onset compromised the separation of sister chromatids and led to the formation of polyploid

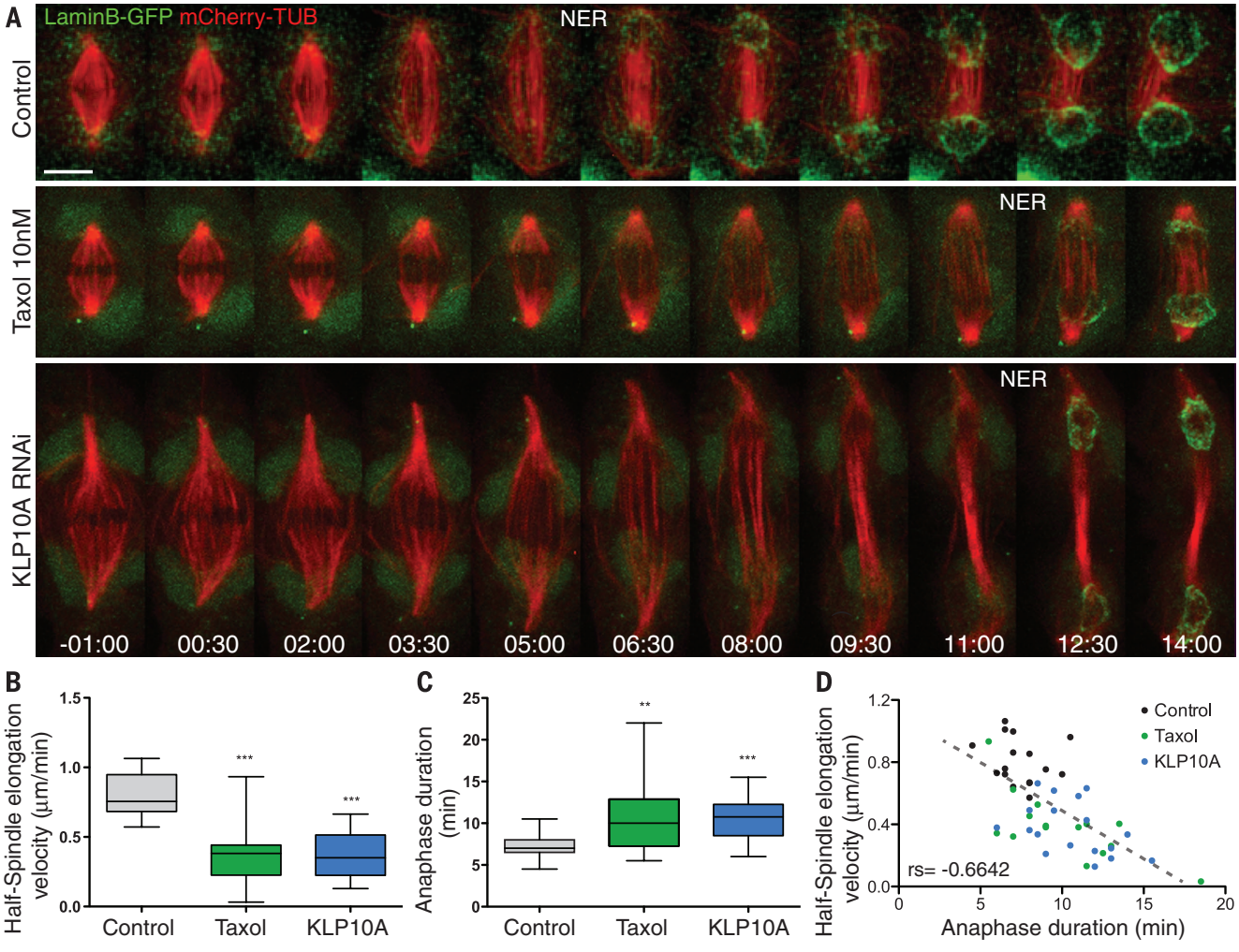


Fig. 1. Incomplete chromosome separation delays the anaphase-telophase transition. (**A**) Control, Taxol-treated, and kinesin-like protein KLP10A RNAi S2 cells stably expressing Lamin B–GFP and mCherry–α-tubulin (mCherry-TUB). Time shown as minutes:seconds. (**B** and **C**) Quantification (box plots) of anaphase duration in response to experimental attenuation of half-spindle elongation velocity. Whiskers define the maximum and the minimum value. Statistics: *t* test or Mann-Whitney U test depending on the normality of the samples (***$P < 0.001$; ** $P < 0.01$). (**D**) Linear regression and distribution of anaphase duration versus half-spindle elongation velocity in the different conditions. rs, Spearman correlation coefficient, $P < 0.001$. Note that there is a minimal anaphase duration time (~5 min).

cells with a single nucleus (Fig. 4, A to C, and fig. S7), which indicated that Aurora B activity is required to delay NER in response to incomplete chromosome separation. Similar findings were observed in human cells (fig. S2, D to H, and fig. S3, A to C).

PP1 and PP2A phosphatases are implicated in NER and counteract Aurora B activity on chromosomes (*7*–*9*). Inhibition of PP1/PP2A with 300 nM okadaic acid (OA) at anaphase onset prevented chromosome decondensation and NER and only slightly attenuated chromosome segregation velocity relative to controls (Fig. 4, A and C, and fig. S7). Specific inhibition of PP1-87B and PP2A-C by RNAi confirmed that both phosphatases were independently required for timely NER (Fig. 4, D and E, and fig. S8, A to F).

Aurora B regulates chromosome condensation by recruiting its substrate Condensin I (Barren/Cap-H in *Drosophila*) to chromosomes (*10*–*14*), and its localization at the spindle midzone is required for the coordination of chromosome compaction with anaphase spindle length (*15*). Chromosome localization of Barren fused with green fluorescent protein (Barren-GFP) correlated with chromosome condensation and disappeared as sister chromatids separated (fig. S9, A and B). Moreover, Barren-GFP was enriched on lagging chromosomes, and Aurora B inhibition at anaphase onset caused the removal of Barren-GFP from chromatin, including lagging chromosomes (fig. S9, A to C′). In contrast, PP1/PP2A inhibition with OA at anaphase onset prevented Barren-GFP removal from chromatin, and chromosomes remained condensed (fig. S9, D and D′). Finally, RNAi-mediated depletion of Barren was reminiscent of Aurora B inhibition, with cells completing NER before effective chromosome separation (fig. S9, E to H). Histone H3 phosphorylation on S10 and S28 did not affect the anaphase-telophase transition in S2 cells (fig. S10, A to E). Condensin I recruitment to chromosomes during anaphase is thus spatially regulated by the counteracting activities of Aurora B and PP1/PP2A to control chromosome decondensation and NER.

To investigate whether Aurora B inhibition is sufficient to trigger chromosome decondensation and NER, we kept Cdk1 activity constitutively high by the transient expression of nondegradable Δ90Cyclin B–GFP (fig. S11, A, B, and D), combined or not, with Aurora B inhibition by Bi-2 at anaphase onset. Cells arrested in anaphase for several hours under both conditions (fig. S11, C and G). Thus, in addition to Aurora B, Cdk1 inhibition is required for chromosome decondensation and NER, as shown previously (*16*–*18*). Expression of Δ90Cyclin B–GFP further impaired spindle elongation and Aurora B relocalization from centromeres to the spindle midzone during anaphase (fig. S11, E, F, H, and I), as reported in mammalian cells (*18*–*20*). In contrast, Cdk1 inhibition at anaphase onset with RO-3306

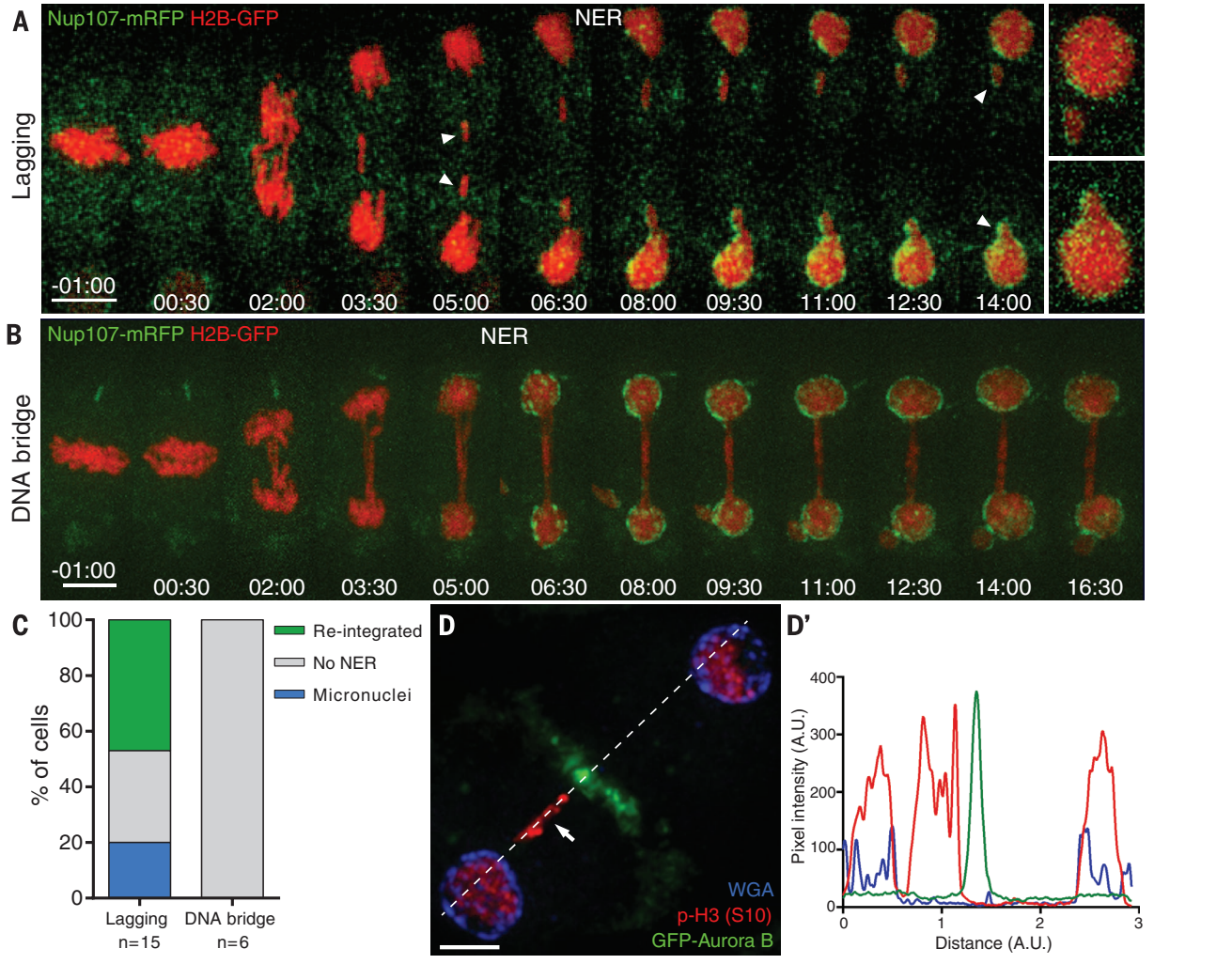


Fig. 2. NER is delayed on lagging chromosomes and DNA bridges. (**A** and **B**) S2 cells stably expressing Nup107-mRFP and H2B-GFP showing a delay in NER on lagging chromosomes [white arrowheads in (A)] and (B) DNA bridges. Time shown as minutes:seconds. (**C**) Many lagging chromosomes are reintegrated into the main nuclei or form micronuclei. Those cases in which the envelope did not form 30 min after anaphase onset were classified as "No NER." (**D** and **D′**) NER correlates with lower Aurora B activity [p-H3 (S10)] and chromosome separation. Wheat germ agglutinin (WGA) staining was used to reveal the nuclear envelope. A.U., arbitrary units. Scale bars, 5 μm.

accelerated NER in the main nuclei relative to controls and Bi-2–treated cells, independently of chromosome separation, and also decreased spindle elongation velocity (fig. S6, A and B; fig. S12, A to D; and fig. S13, A to D). Thus, Cdk1 appears to work as a "clock" that temporally regulates the formation of a midzone Aurora B–based "ruler," which explains the observed minimal anaphase duration time (~5 min) (Fig. 1D and fig. S1C). Indeed, Cdk1 inhibition at anaphase onset led to a fast accumulation of Aurora B at the spindle midzone (fig. S12, E and F) and, contrary to Aurora B inhibition, NER never took place on lagging chromosomes (10 out of 10 cells) (fig. S6B). Simultaneous inhibition of Aurora B and PP1/PP2A or Cdk1 and PP1/PP2A (with or without Aurora B inhibition) at anaphase onset prevented timely chromosome decondensation (fig. S14). Thus, PP1/PP2A activities are required to dephosphorylate Aurora B and Cdk1 substrates during anaphase and for the spatiotemporal regulation of chromosome decondensation and NER.

Here, we have described a conserved surveillance mechanism that coordinates chromosome separation with chromosome decondensation and NER involving three essential components: (i) a sensor (Aurora B gradient) that monitors incompletely separated chromosomes, (ii) an effector (PP1/PP2A) that promotes decondensation and NER on fully separated chromosomes, and (iii) a target (Condensin I) that maintains chromosome condensation in response to incomplete chromosome separation (fig. S15). This surveillance mechanism operates after spindle assembly checkpoint (SAC) satisfaction and fulfills the criteria to qualify as a "chromosome separation checkpoint." Accordingly, "checkpoints" delay or arrest a late cell-cycle event until completion of an early event and, typically, are only detected in the event of errors (*21*). One of the strongest pieces of evidence for a checkpoint is the "relief of dependence" when one finds conditions that permit a late event to occur even when an early, normally a prerequisite, event is prevented (*21*). Here, we showed that chromosome decondensation and NER depend on the previous separation of chromosomes during anaphase. This is monitored by a midzone-associated Aurora B phosphorylation gradient, perturbation of which (either by inactivating global or localized kinase activity) permits chromosome decondensation and NER without completing separation. Thus, the dependence of chromosome decondensation and NER on chromosome separation can be relieved by inactivating global or localized Aurora B kinase activity. This could give rise to polyploidy or micronuclei associated with chromosome aberrations (*22*, *23*). Because NER starts at the distal regions of chromatin during late anaphase (*1*), a chromosome separation checkpoint would provide an opportunity to correct errors "invisible" to the SAC (e.g., merotelic attachments or catenated DNA) and allow reintegration of lagging chromosomes in the main nuclei before completion of NER. Lagging chromosomes sustain Aurora B activity to inhibit abscission during cytokinesis—the so-called NoCut checkpoint that monitors clearance of chromatin

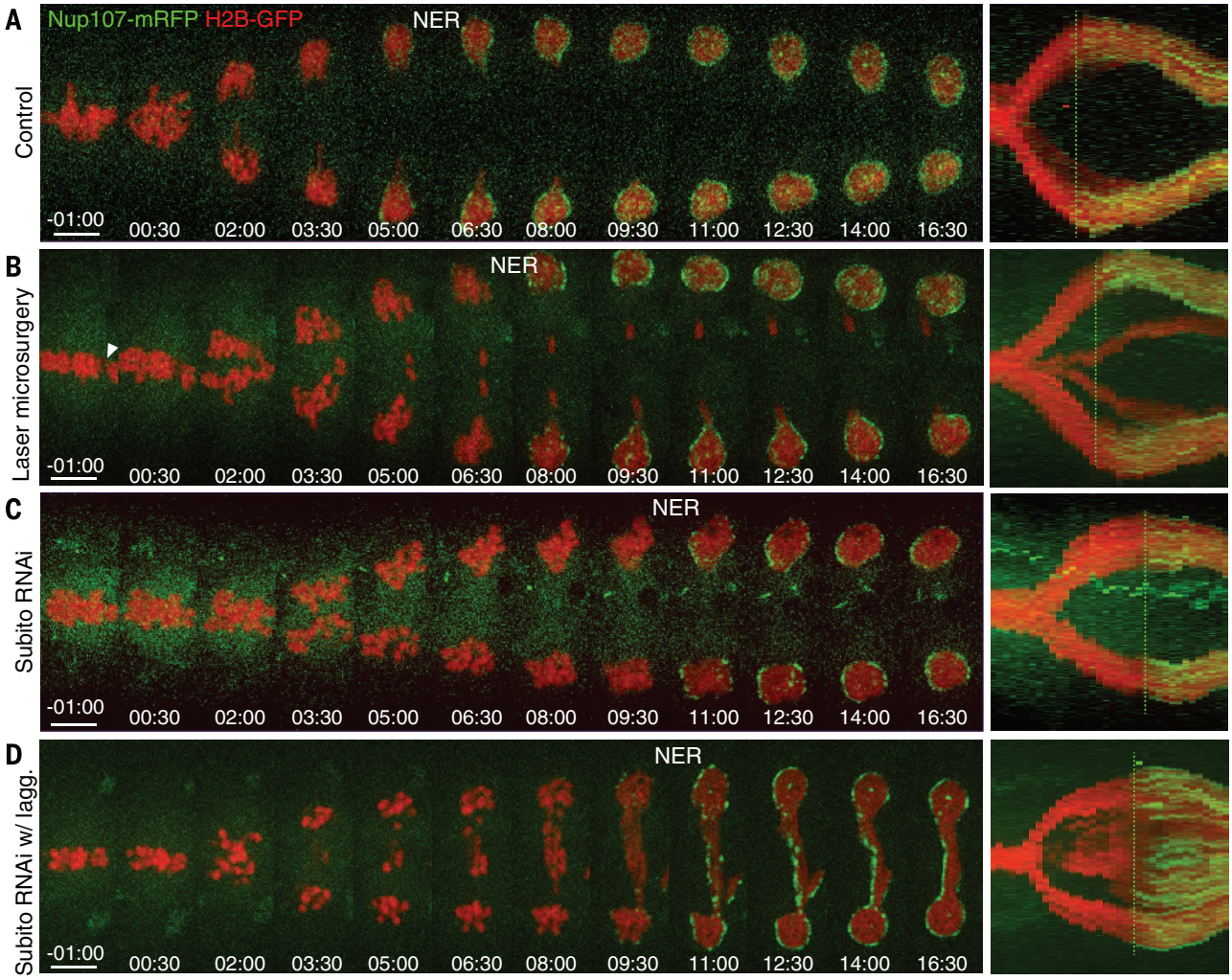


Fig. 3. Aurora B at the spindle midzone is required for spatial control of NER. (**A**) S2 cell stably expressing Nup107-mRFP and H2B-GFP. Time shown as minutes:seconds. Scale bars, 5 μm. (**B**) Induction of acentric chromosomes with laser microsurgery (arrowhead) generates lagging fragments that delay NER. (**C**) Subito/Mklp2 RNAi delays NER. (**D**) Simultaneous NER on lagging chromosomes and main nuclei after Subito/Mklp2 depletion.

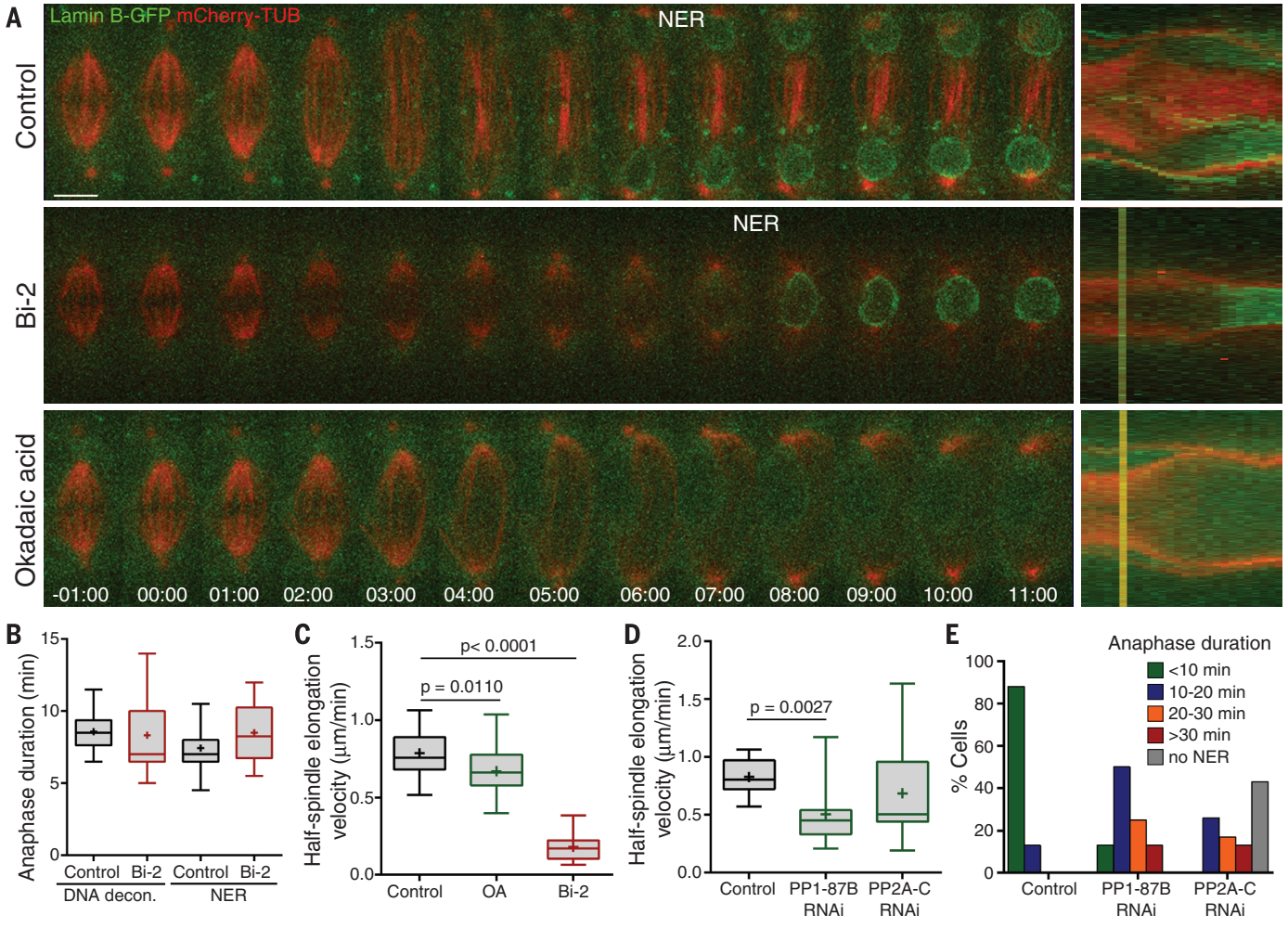


Fig. 4. PP1/PP2A phosphatases counteract Aurora B and are required to promote chromosome decondensation and NER. (**A**) Acute Aurora B (Bi-2) or PP1/PP2A inhibition (OA) at anaphase onset in S2 cells stably expressing Lamin B–GFP and mCherry–α-tubulin. (**B**) Quantification of anaphase duration until DNA decondensation or NER in control cells (n = 16 cells per condition) and Bi-2–treated cells (n = 10 cells per condition). Time shown as minutes:seconds. (**C**) Quantification of half-spindle elongation velocity. (**D**) RNAi against specific PP1 or PP2A subunits (n = 8 cells per condition) in the Lamin B–GFP/mCherry–α-tubulin cell line (Mann-Whitney U test). The whiskers define the maximum and minimum values. (**E**) Distribution of anaphase duration in control (n = 16 cells), PP1-87B (n = 8 cells), and PP2A-C RNAi (n = 23 cells). Scale bars, 5 µm.

from the spindle midzone (*24*–*26*). The critical differences between these potential checkpoints are that the spatial regulation of NER involves a default mechanism mediated by a midzone Aurora B gradient that is active during a normal mitosis, whereas NoCut is activated only in the presence of chromatin at the cleavage furrow for successful completion of cytokinesis.

REFERENCES AND NOTES

1. S. Güttinger, E. Laurell, U. Kutay, *Nat. Rev. Mol. Cell Biol.* **10**, 178–191 (2009).
2. B. G. Fuller *et al.*, *Nature* **453**, 1132–1136 (2008).
3. K. Ramadan *et al.*, *Nature* **450**, 1258–1262 (2007).
4. J. M. Cesario *et al.*, *J. Cell Sci.* **119**, 4770–4780 (2006).
5. U. Gruneberg, R. Neef, R. Honda, E. A. Nigg, F. A. Barr, *J. Cell Biol.* **166**, 167–172 (2004).
6. Y. Smurnyy, A. V. Toms, G. R. Hickson, M. J. Eck, U. S. Eggert, *ACS Chem. Biol.* **5**, 1015–1020 (2010).
7. P. Vagnarelli *et al.*, *Dev. Cell* **21**, 328–342 (2011).
8. M. H. Schmitz *et al.*, *Nat. Cell Biol.* **12**, 886–893 (2010).
9. J. Qian, M. Beullens, B. Lesage, M. Bollen, *Curr. Biol.* **23**, 1136–1143 (2013).
10. N. Nakazawa, R. Mehrotra, M. Ebe, M. Yanagida, *J. Cell Sci.* **124**, 1795–1807 (2011).
11. R. Giet, D. M. Glover, *J. Cell Biol.* **152**, 669–682 (2001).
12. J. J. Lipp, T. Hirota, I. Poser, J. M. Peters, *J. Cell Sci.* **120**, 1245–1255 (2007).
13. A. Takemoto *et al.*, *Nucleic Acids Res.* **35**, 2403–2412 (2007).
14. K. Tada, H. Susumu, T. Sakuno, Y. Watanabe, *Nature* **474**, 477–483 (2011).
15. G. Neurohr *et al.*, *Science* **332**, 465–468 (2011).
16. S. Sigrist, H. Jacobs, R. Stratmann, C. F. Lehner, *EMBO J.* **14**, 4827–4838 (1995).
17. D. H. Parry, P. H. O'Farrell, *Curr. Biol.* **11**, 671–683 (2001).
18. S. P. Wheatley *et al.*, *J. Cell Biol.* **138**, 385–393 (1997).
19. M. Murata-Hori, M. Tatsuka, Y. L. Wang, *Mol. Biol. Cell* **13**, 1099–1108 (2002).
20. S. Hümmer, T. U. Mayer, *Curr. Biol.* **19**, 607–612 (2009).
21. L. H. Hartwell, T. A. Weinert, *Science* **246**, 629–634 (1989).
22. A. Janssen, M. van der Burg, K. Szuhai, G. J. Kops, R. H. Medema, *Science* **333**, 1895–1898 (2011).
23. K. Crasta *et al.*, *Nature* **482**, 53–58 (2012).
24. C. Norden *et al.*, *Cell* **125**, 85–98 (2006).
25. P. Steigemann *et al.*, *Cell* **136**, 473–484 (2009).
26. M. Mendoza *et al.*, *Nat. Cell Biol.* **11**, 477–483 (2009).

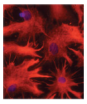
ACKNOWLEDGMENTS

We thank all colleagues who provided reagents and P. Sampaio for technical help with FRET. H.M. is funded by FCOMP-01-0124-FEDER-015941 (PTDC/SAU-ONC/112917/2009) through COMPETE and Fundação para a Ciência e a Tecnologia of Portugal, the Human Frontier Science Program, and PRECISE grant from the European Research Council. Data described can be found in the main figures and supplementary materials. The authors declare no conflict of interests.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/345/6194/332/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S15
References (*27*–*29*)

21 January 2014; accepted 2 June 2014
Published online 12 June 2014;
10.1126/science.1251121

EXTENDED PDF FORMAT
SPONSORED BY

Travel Grants Available
Neuroscience 2014
Application Deadline: July 31st
>Apply Now
R&D systems
a bietechne brand
www.rndsystems.com

Activation of cytoplasmic dynein motility by dynactin-cargo adapter complexes

Richard J. McKenney *et al.*

Science **345**, 337 (2014);
DOI: 10.1126/science.1254198

This copy is for your personal, non-commercial use only.

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by clicking here.

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines here.

The following resources related to this article are available online at www.sciencemag.org (this information is current as of August 4, 2014):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:
http://www.sciencemag.org/content/345/6194/337.full.html

Supporting Online Material can be found at:
http://www.sciencemag.org/content/suppl/2014/06/18/science.1254198.DC1.html

A list of selected additional articles on the Science Web sites **related to this article** can be found at:
http://www.sciencemag.org/content/345/6194/337.full.html#related

This article **cites 36 articles**, 14 of which can be accessed free:
http://www.sciencemag.org/content/345/6194/337.full.html#ref-list-1

This article has been **cited by** 2 articles hosted by HighWire Press; see:
http://www.sciencemag.org/content/345/6194/337.full.html#related-urls

This article appears in the following **subject collections:**
Cell Biology
http://www.sciencemag.org/cgi/collection/cell_biol

Downloaded from www.sciencemag.org on August 4, 2014

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published weekly, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. Copyright 2014 by the American Association for the Advancement of Science; all rights reserved. The title *Science* is a registered trademark of AAAS.

MOLECULAR MOTORS

Activation of cytoplasmic dynein motility by dynactin-cargo adapter complexes

Richard J. McKenney, Walter Huynh, Marvin E. Tanenbaum, Gira Bhabha, Ronald D. Vale*

Cytoplasmic dynein is a molecular motor that transports a large variety of cargoes (e.g., organelles, messenger RNAs, and viruses) along microtubules over long intracellular distances. The dynactin protein complex is important for dynein activity in vivo, but its precise role has been unclear. Here, we found that purified mammalian dynein did not move processively on microtubules in vitro. However, when dynein formed a complex with dynactin and one of four different cargo-specific adapter proteins, the motor became ultraprocessive, moving for distances similar to those of native cargoes in living cells. Thus, we propose that dynein is largely inactive in the cytoplasm and that a variety of adapter proteins activate processive motility by linking dynactin to dynein only when the motor is bound to its proper cargo.

Cytoplasmic dynein 1 (dynein), a member of the AAA adenosine triphosphatase (ATPase) family, is the major minus-end–directed microtubule (MT) motor in most eukaryotic cells (*1*). Several adapter proteins control the recruitment of a soluble pool of cytoplasmic dynein to transport cargos at the right place and time in the cell (*2*–*4*). Dynein motor activity also appears to be governed by two general regulatory factors—the Lis1-NudEL complex and the dynactin complex (*5*). Lis1 may act as a "clutch" that suppresses dynein motility and causes it to form a tight binding complex with the MT (*6*, *7*). Whether dynein motility requires just the detachment of Lis1 or needs an additional activation process has been unclear.

Yeast cytoplasmic dynein, the best characterized dynein in terms of its motility, moves processively on its own (run length of 1 to 2 μm) (*8*), and dynactin only increases its run length by ~twofold (*9*). Mammalian cytoplasmic dynein is also generally thought to be constitutively active for motility, as it produces movement when attached to solid surfaces such as glass slides (*10*), plastic beads (*11*), or quantum dots (*12*). However, surface binding of kinesin activates this normally autoinhibited motor (*13*). Without direct visualization of the motor itself, it also can be difficult to determine whether one or multiple motors are contributing to movement. Prior studies of fluorescently labeled mammalian dynactin, but with unlabeled dynein, have reported processive run lengths of <2 μm in both directions on the MT (*14*).

Here, we examined the motility of purified rat brain cytoplasmic dynein (termed "brain dynein") by single-molecule fluorescence without attachment to surfaces. Brain dynein, which exhibited a characteristic two-headed shape by electron microscopy (EM) (Fig. 1A), produces fast motility (~0.6 μm/s) of MTs in a multiple motor gliding assay (*15*). However, individual Cy3-labeled native dynein molecules, examined by total internal reflection (TIRF) microscopy in the presence of 1 mM adenosine 5′-triphosphate (ATP), mostly either bound statically to MTs or exhibited short back-and-forth movements (Fig. 1A, fig. S1A, and movie S1), which are likely due to thermal-driven diffusion, as they persist after addition of the ATPase inhibitor vanadate (fig. S1B) (*16*). Directional movements were only occasionally observed (<1% of MT-bound dynein); those movements were very slow (~90 nm/s, fig. S1A) and inhibited by vanadate (fig. S1B). A recombinant, glutathione *S*-transferase (GST)–dimerized human motor domain construct (fig. S1C) also did not show fast, processive motion (Fig. 1B and movie S2), consistent with recent studies of a recombinant human cytoplasmic dynein holoenzyme (*17*). Thus, in contrast to yeast dynein, purified mammalian dynein rarely displays processive motility.

The poor single-molecule motility by mammalian dynein might be due to the absence of an activator in our purified preparations. BicD2 is a conserved, dimeric adapter protein that links dynein to Rab6 GTPase on membrane organelles (*2*, *18*); the BicD N-terminal coiled coil (*19*) facilitates an interaction between dynein and dynactin, forming a stable dynein-dynactin-BicD2 ternary complex (DDB) that can be purified (*20*). Using this well-characterized green fluorescence protein (GFP)–tagged, N-terminal construct of BicD2 (*20*), we isolated the DDB complex from pig brain (Fig. 1C and fig. S2A); this preparation did not contain detectable kinesin-1, Lis1, or Nudel (fig. S2B). The DDB complex eluted as a single peak by size-exclusion chromatography (fig. S2C), and EM revealed a single dynein dimer with its tail bound to a single dynactin, identifiable by its 37 nm Arp1 filament (Fig. 1C and fig. S3, A to C). The DDB complex appeared highly flexible, as evidenced by the variable orientations of dynactin's Arp1 filament (fig. S3B) and variable separation of the two motor domains (fig. S3D).

Next, we examined the motility of single DDB complexes by TIRF microscopy, using the GFP-tag on BicD2. In contrast to brain dynein, the DDB complexes moved robustly and processively along MTs (Fig. 1D and movie S3), although a small fraction (~15%) exhibited back-and-forth diffusive motion (fig. S2D). The DDB complexes also accumulated at MT minus-ends (Fig. 1E), indicating tenacious binding upon reaching the end of the MT track. Similar processive motility and minus-end accumulation was observed recently for dynein-driven transport of purified mRNA particles in vitro (*21*). By kymograph analysis, DDB processive movement appeared as extended diagonal lines (Fig. 1F). At 2 mM ATP, DDB moved at a mean velocity of 376 nm/s (Fig. 1G) and run length of 8.7 μm (Fig. 1H), considerably longer than yeast dynein (*8*). When measured at physiological temperature (37°C), the velocity was 892 nm/s (fig. S2E), very close to speeds of retrograde transport in vivo (*12*). Processive motility was completely abolished by the ATPase inhibitor vanadate (fig. S2F). The addition of BicD2 to purified brain dynein did not stimulate processive motility, indicating a requirement for dynactin (fig. S2G). To further confirm that motile DDB complexes contained single BicD2 dimers and not oligomers or aggregates, we added two fluorescently labeled BicD2 preparations (TMR- and 505-Star–labeled) to pig brain lysates followed by DDB complex purification. Two-color kymographs revealed that moving DDB complexes contained either TMR- or 505-Star-BicD2 but not a mixture (fig. S2H and movie S4), confirming that each DDB complex contains only a single BicD2 molecule. Thus, BicD2 links dynein and dynactin to form a stable complex that moves toward the MT minus-end with ultraprocessive run-lengths.

To examine each component of the DDB complex during motility, we purified a triple-colored DDB complex consisting of Alexa647-labeled, SNAPf-tagged BicD2 (expressed in *Escherichia coli*); GFP-tagged dynein intermediate chain (GFP-IC); and a tetramethylrhodamine (TMR)–labeled, Halo-tagged p62 dynactin subunit [tagged IC and p62 were coexpressed in human RPE-1 cells (*22*); fig. S4, A and B)]. This triple-color–labeled DDB moved processively, and many moving GFP-dyneins colocalized with a TMR-dynactin and an Alexa647-BicD2 molecule (Fig. 2A and movie S5); the mean velocity and run length of human DDB was 379 nm/s (Fig. 2B, left) and 8.84 μm (Fig. 2B, right), respectively. The fluorescence intensities of the motile GFP-labeled dynein molecules were similar to those of the well-characterized dimeric protein, GFP-kinesin K560 (fig. S4C). Additionally, the fluorescence intensities of isolated TMR-dynactin and Alexa647-BicD2 were similar to the intensities of these components within the motile DDB complex

Department of Cellular and Molecular Pharmacology and the Howard Hughes Medical Institute, University of California, San Francisco, CA 94158, USA.
*Corresponding author. E-mail: vale@ucsf.edu

(fig. S4, D and E). Thus, consistent with the EM data (Fig. 1B), a ternary complex, containing single copies of dynein, dynactin, and BicD2, is the active entity that moves processively along MTs.

Because dynactin has been suggested to activate dynein motility in the absence of other factors (*9*, *23*), we next examined the single-molecule behavior of dynein and dynactin in the absence of BicD2. GFP-dynein and TMR-dynactin could be separated from BicD2 by salt elution from a BicD2 affinity column (fig. S4F). Without ATP, human GFP-dynein molecules decorated MTs, but MT binding by human TMR-dynactin was not observed (Fig. 2C), as was also reported for yeast dynactin (*9*). The lack of MT binding by dynactin was surprising, because the recombinant N-terminal half of $p150^{Glued}$ (a subunit in the dynactin complex, Fig. 1B; herein termed

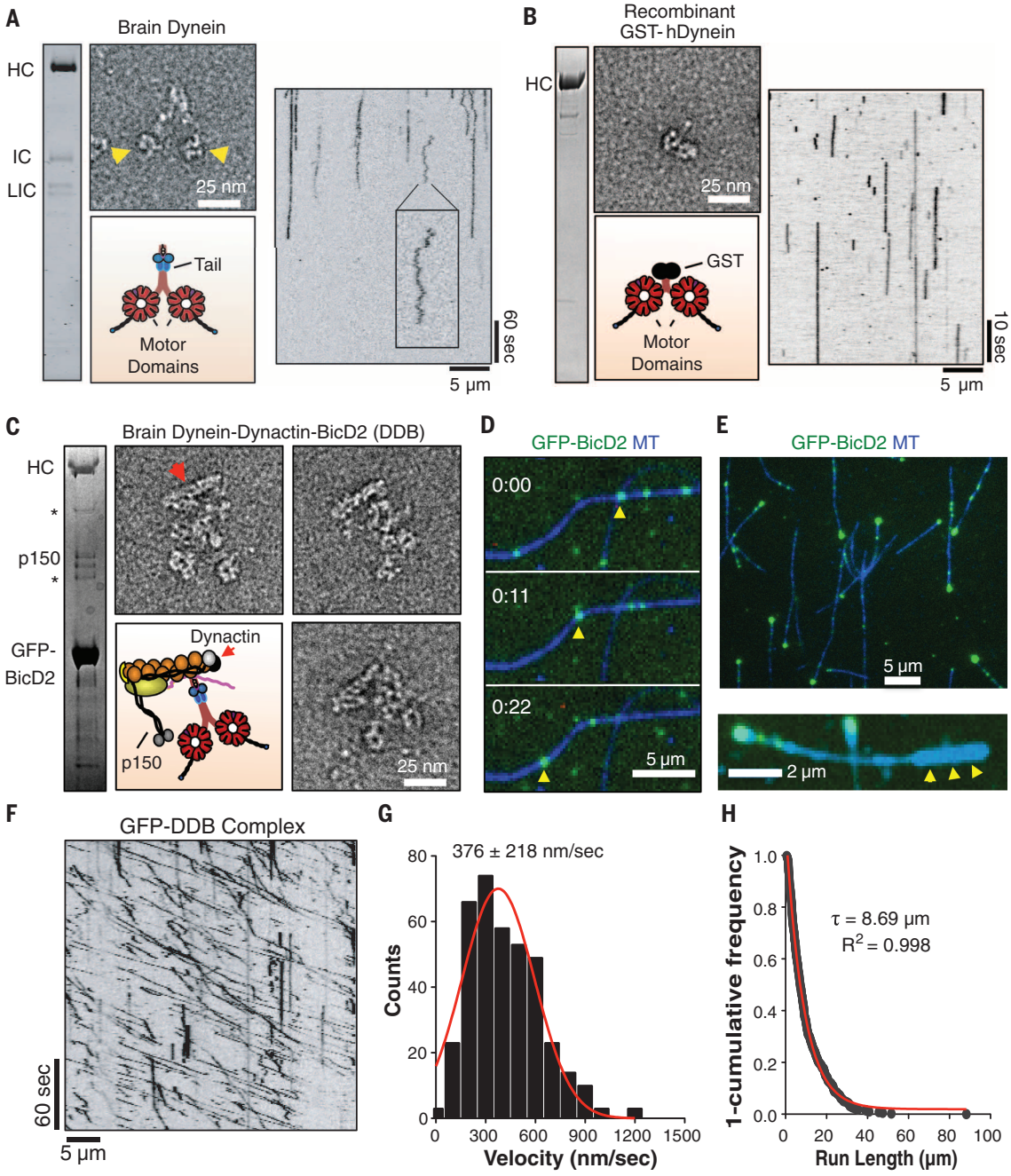


Fig. 1. Formation of a dynein-dynactin-BicD2 complex induces processive dynein motility. (**A**) SDS–polyacrylamide gel electrophoresis (SDS-PAGE) of purified rat brain dynein and negative-stain EM. Arrowheads denote motor domains. Kymograph analysis of Cy3-labeled rat brain dynein on MTs (100 µM ATP). Inset: back-and-forth motion, which is likely diffusion-driven. HC: heavy chain, LIC: light intermediate chain. (**B**) SDS-PAGE of recombinant human GST-dynein and negative-stain EM. Kymograph analysis of TMR-labeled GST-dynein on MTs revealing no movement (100 µM ATP). (**C**) SDS-PAGE and negative-stain EM of purified pig brain DDB complex. Asterisks mark nonspecific bands. Red arrowhead denotes dynactin's Arp1 filament. (**D**) Movie frames showing a single GFP-labeled DDB complex (yellow arrowhead) moving processively along a MT. (**E**) DDB complexes accumulate at one end of MTs. Below: polarity-marked MTs (arrowheads mark brightly labeled plus-end) reveal that accumulation occurs at the minus-end. (**F**) Kymograph analysis of pig brain DDB reveals diagonal lines reflecting long, unidirectional movements (100 µM ATP). (**G**) Velocity histogram (1 mM ATP) with a Gaussian fit (mean ± SD: 376 ± 218 nm/s, n = 379 molecules). (**H**) A "1-cumulative frequency distribution plot" run-lengths with fit to a one-phase exponential decay (red). Decay constant (τ, run length) and R^2 value of the fit are shown (n = 379 molecules, two independent preparations).

"p150") avidly binds to MTs (*24*, *25*). However, upon addition of Alexa647-BicD2 in the absence of ATP, the intensity of GFP-dynein and TMR-dynactin fluorescence on MTs increased by ~2- and 50-fold, respectively (Fig. 2C). With ATP, GFP-tagged dynein bound statically or diffused back-and-forth on MTs (Fig. 2D), and TMR-dynactin did not bind to the MTs (Fig. 2D). Re-addition of Alexa647-BicD2 strongly stimulated processive motility of TMR-dynactin and GFP-labeled dynein (Fig. 2D and movie S6) without evidence of aggregation (fig. S4G). Thus, dynein and dynactin have very low affinity for one another in the absence of BicD2, in agreement with prior biochemical studies (*20*), and BicD2 is required for processive motion. Furthermore, MT binding by the p150 subunit appears to be masked in the isolated dynactin complex, perhaps reflecting an autoinhibited state.

The above experiments raised the question of whether MT binding by p150 plays a role in DDB motility. Previous studies showed that removal of tubulin's C-terminal tails with the protease subtilisin (Δ-CTT MTs) markedly reduced MT binding of a purified p150 construct (*25*), which we confirmed (fig. S5, A and B). In the absence of ATP, we found that GFP-human dynein bound almost equally well to both control and Δ-CTT MTs, whereas human DDB complex exhibited greatly reduced binding to Δ-CTT MTs, thus behaving more like p150 than dynein (Fig. 3, A and B, and fig. S5, A and B). The difference was even more pronounced in the presence of ATP. GFP-human dynein bound to both types of MTs (but did not move), whereas DDB moved robustly on untreated MTs, but rarely moved on Δ-CTT MTs that were placed on the same slide (Fig. 3C and movie S7). Similar findings were made for DDB from bovine brain (fig. S5, A to C). In contrast, *Saccharomyces cerevisiae* dynein, which is not regulated by BicD2, moves almost as well on Δ-CTT as on untreated MTs (*26*). p150 contains a well-defined MT binding site (a CAP-Gly domain flanked by a basic rich region) at its N terminus (*24*). To test the role of this domain, we overexpressed Halo-tagged versions of p150 or p135, a naturally occurring splice form lacking the CAP-Gly domain, in RPE cells (*27*), and then isolated DDB complexes and fluorescently labeled them with Halo-TMR. MT-bound p135-containing DDB complexes displayed one-third as many processive movements versus p150-containing complexes (Fig. 3D). Of the TMR-p135 complexes that moved processively, their velocity (498 ± 226 nm/s) and run-length (8.9 μm) were similar to those of moving TMR-p150 complexes (417 ± 147 nm/s and 12.19 μm; fig. S5, D and E). The significant fraction (~15%) of DDB-p135 complexes that exhibited ultraprocessivity suggests either that the CAP-Gly domain is not absolutely required for motility or that the residual motion observed with TMR-p135 could be due to heterodimerization with endogenous p150. Further work is required to understand the complex interplay between dynactin's CAP-Gly domain and dynein activation (*9*, *28*, *29*).

In addition to BicD2, several other coiled-coil proteins have been implicated in linking dynein to cargoes, including Rab11-FIP3 on Rab11-positive recycling endosomes (*4*), Spindly (hSpindly) on kinetochores (*3*), and Hook proteins on early endosomes (*30*, *31*). We asked if these cargo adapter proteins might similarly initiate processive motion by increasing dynein's affinity for dynactin.

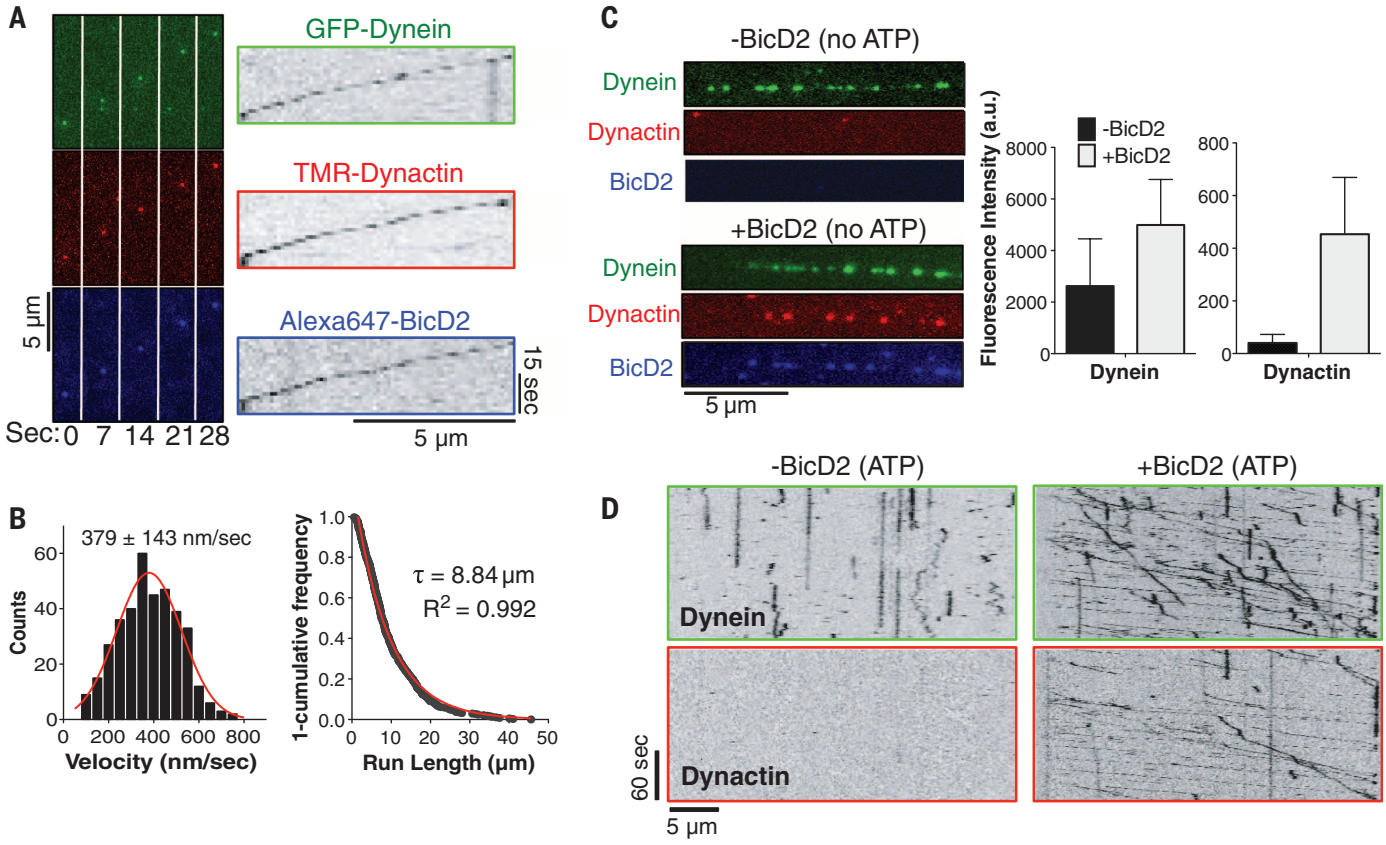


Fig. 2. Three-color single-molecule analysis of the DDB complex and requirement of BicD2 for motility. (**A**) Successive frames from movie S5 showing a processive DDB complex purified from human RPE-1 cells with all three components fluorescently labeled (see text and fig S4A). Right: corresponding kymographs. (**B**) Histogram of human DDB velocities with a Gaussian fit (mean ± SD; *n* = 374 molecules). Right: a "1-cumulative frequency distribution plot" of human DDB run-lengths fit to a one-phase exponential decay (red). Decay constant (τ) and R^2 value of the fit are shown (*n* = 374 molecules). (**C**) Binding of fluorescently labeled human dynein and dynactin without ATP to MTs (not visible) in the absence and presence of BicD2. Note that not all DDBs are triple-labeled. Right: quantification of fluorescence along MTs (mean ± SD, *n* = 16 and 11 MTs before and after BicD2 addition, respectively). (**D**) Kymograph analysis in the presence of 100 μM ATP shows that BicD2 addition is required for human dynein motility and dynactin binding to the MTs. Note that not all dynactins in the DDB particles (+BicD2) are labeled, owing to incomplete Halo tag labeling and/or photobleaching.

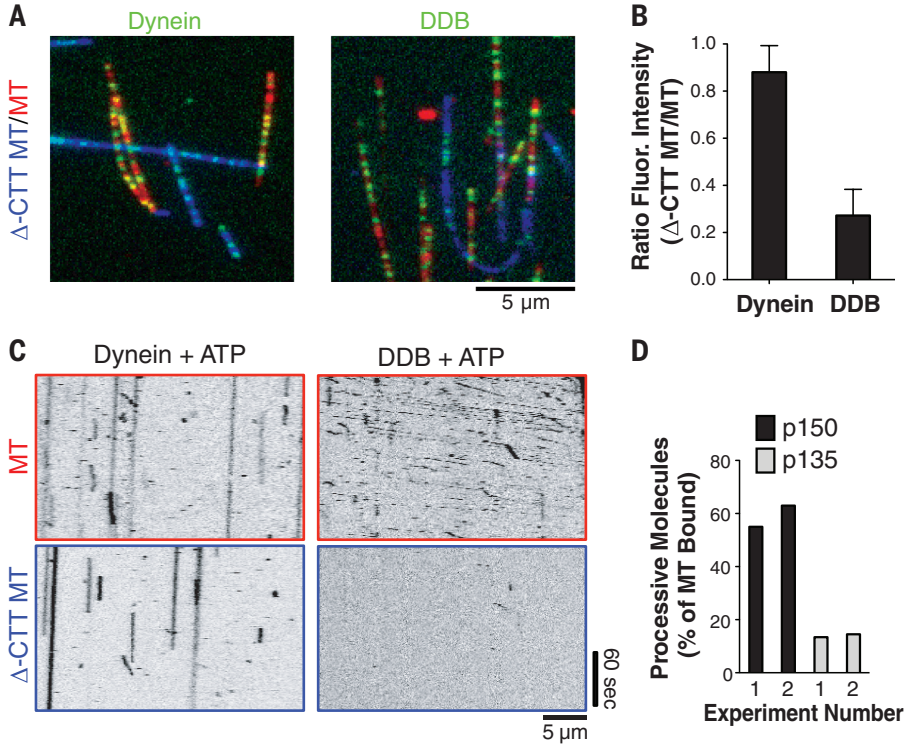


Fig. 3. MT binding and processivity of DDB requires the C-terminal tails of tubulin. (**A**) MT-binding behavior of GFP-tagged dynein (no BicD2), or GFP-DDB from human RPE cells, on normal MTs (red) and Δ-CTT MTs (blue) in the absence of ATP. (**B**) Quantification of the fluorescence intensity ratios (GFP-Dyn: 24 MTs, 22 Δ-CTT MTs; DDB: 59 MTs, 38 Δ-CTT MTs, mean ± SD). (**C**) Kymograph analysis of MT binding and motility with ATP (2 mM) on normal MTs (upper row) and on Δ-CTT MTs (lower row). (**D**) Quantification of processively moving (>2 μm) human DDB complexes with incorporated TMR-labeled p150-Halo or p135-Halo subunits (*22*) (percent of total MT bound; others were statically bound or diffusing). Results from two independent experiments are shown.

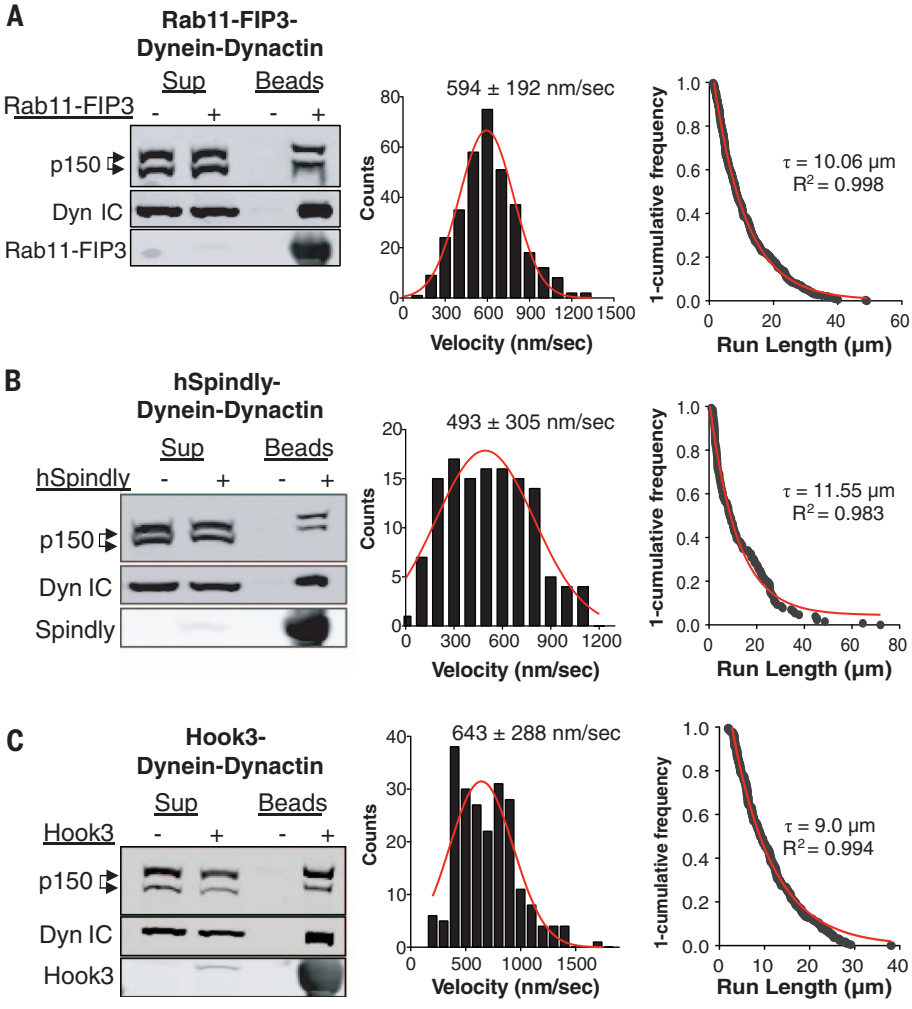


Fig. 4. Rab11-FIP3, hSpindly, and Hook3 activate brain dynein motility by linking together dynein and dynactin. (**A**) Recombinant full-length Rab11-FIP3, (**B**) full-length human Spindly, and (**C**) human N-terminal Hook3 (amino acids 1 to 552) were prepared with an N-terminal SNAPf tag for TMR labeling (*22*). Left panels: All three adapters attached to beads specifically bound to dynein and dynactin from pig brain lysate (immunoblots show dynein intermediate chain and dynactin p150). Middle panels: Velocity histograms of the indicated brain dynein-dynactin-adapter complexes [Gaussian fits (red) and mean ± SD velocities are shown]. Right panels: "1-cumulative frequency distribution plot" of run-lengths for the indicated dynein-dynactin-adapter complexes. One-phase exponential decay (red), decay constants (τ) to give the run-length, and R^2 values of fit are shown [*n* values for both middle and right panels: Rab11-FIP3 (333); hSpindly (129); Hook3 (219)].

Recombinant SNAPf-tagged Rab11-FIP3, human Spindly, and Hook3 (fig. S6, A and B) all efficiently coprecipitated dynein and dynactin from pig brain lysates (Fig. 4, A to C). The dynein-dynactin-adapter complexes, labeled with TMR on the SNAPf tag and which did not contain contaminating BicD2 (fig. S6C), all displayed fast (~500 nm/s) and ultraprocessive (9- to 11-μm run-length) motility (Fig. 4, A to C), although the number of moving Spindly-dynactin-dynein complexes was somewhat lower than for Rab11-FIP3, Hook3, and BicD2 (movie S8). Triple-color single-molecule assays using material from human RPE cells confirmed that Alexa647-labeled Rab11-FIP3, Spindly, and Hook3 colocalized with moving GFP-dynein and TMR-dynactin (fig. S6, D to F; movie S9). Thus, four adapter proteins that link dynein to cargoes can induce the formation of highly processive dynein-dynactin-adapter complexes.

Our results suggest a general model for regulating mammalian cytoplasmic dynein motility (fig. S7). In the cytoplasm, without attachment to a cargo, dynein adopts an inactive conformation that cannot undergo processive motion. Activation of motility requires the simultaneous binding of dynein to an adapter protein that defines a particular cargo for transport (e.g., BicD2, Rab11-FIP3, Spindly, Hook3) and dynactin, a universal activator involved in the transport of many dynein cargos. *S. cerevisiae* dynein is constitutively active on its own and dynactin only modestly increases its processivity (*9*), suggesting that the yeast motor may have lost the switch-like regulation of motility displayed by mammalian dynein. Processive motility appears to involve tethering of the DDB complex to the tubulin C-terminal tails, which might be facilitated by the flexibility of dynactin within the DDB complex (Fig. 1C), although allosteric activation mechanisms might also be involved (*9*, *28*, *29*). Our results also suggest that dynactin may use a yet undiscovered autoregulatory mechanism, because its MT binding appears to be masked until it becomes incorporated into a dynein-cargo complex. Thus, mammalian dynein and dynactin both become activated for transport in a process that is coupled to cargo selection.

REFERENCES AND NOTES

1. V. J. Allan, *Biochem. Soc. Trans.* **39**, 1169–1178 (2011).
2. I. Grigoriev *et al*., *Dev. Cell* **13**, 305–314 (2007).
3. E. R. Griffis, N. Stuurman, R. D. Vale, *J. Cell Biol.* **177**, 1005–1015 (2007).
4. C. P. Horgan, S. R. Hanscom, R. S. Jolly, C. E. Futter, M. W. McCaffrey, *J. Cell Sci.* **123**, 181–191 (2010).
5. R. B. Vallee, R. J. McKenney, K. M. Ori-McKenney, *Nat. Cell Biol.* **14**, 224–230 (2012).
6. R. J. McKenney, M. Vershinin, A. Kunwar, R. B. Vallee, S. P. Gross, *Cell* **141**, 304–314 (2010).
7. J. Huang, A. J. Roberts, A. E. Leschziner, S. L. Reck-Peterson, *Cell* **150**, 975–986 (2012).
8. S. L. Reck-Peterson *et al*., *Cell* **126**, 335–348 (2006).
9. J. R. Kardon, S. L. Reck-Peterson, R. D. Vale, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 5669–5674 (2009).
10. B. M. Paschal, R. B. Vallee, *Nature* **330**, 181–183 (1987).
11. Z. Wang, S. Khan, M. P. Sheetz, *Biophys. J.* **69**, 2011–2023 (1995).
12. K. M. Ori-McKenney, J. Xu, S. P. Gross, R. B. Vallee, *Nat. Cell Biol.* **12**, 1228–1234 (2010).
13. D. L. Coy, W. O. Hancock, M. Wagenbach, J. Howard, *Nat. Cell Biol.* **1**, 288–292 (1999).
14. J. L. Ross, K. Wallace, H. Shuman, Y. E. Goldman, E. L. Holzbaur, *Nat. Cell Biol.* **8**, 562–570 (2006).
15. M. E. Tanenbaum, R. D. Vale, R. J. McKenney, *eLife* **2**, e00943 (2013).
16. M. Miura, A. Matsubara, T. Kobayashi, M. Edamatsu, Y. Y. Toyoshima, *FEBS Lett.* **584**, 2351–2355 (2010).
17. M. Trokter, N. Mücke, T. Surrey, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 20895–20900 (2012).
18. C. C. Hoogenraad *et al*., *EMBO J.* **20**, 4041–4054 (2001).
19. N. Stuurman *et al*., *Eur. J. Cell Biol.* **78**, 278–287 (1999).
20. D. Splinter *et al*., *Mol. Biol. Cell* **23**, 4226–4241 (2012).
21. H. C. Soundararajan, S. L. Bullock, *eLife* **3**, e01596 (2014).
22. Materials and methods are available as supplementary materials on *Science* Online.
23. S. J. King, T. A. Schroer, *Nat. Cell Biol.* **2**, 20–24 (2000).
24. T. L. Culver-Hanlon, S. A. Lex, A. D. Stephens, N. J. Quintyne, S. J. King, *Nat. Cell Biol.* **8**, 264–270 (2006).
25. J. E. Lazarus, A. J. Moughamian, M. K. Tokito, E. L. Holzbaur, *PLOS Biol.* **11**, e1001611 (2013).
26. W. B. Redwine *et al*., *Science* **337**, 1532–1536 (2012).
27. M. K. Tokito, D. S. Howland, V. M. Lee, E. L. Holzbaur, *Mol. Biol. Cell* **7**, 1167–1180 (1996).
28. H. Kim *et al*., *J. Cell Biol.* **176**, 641–651 (2007).
29. A. J. Moughamian, E. L. Holzbaur, *Neuron* **74**, 331–343 (2012).
30. E. Bielska *et al*., *J. Cell Biol.* **204**, 989–1007 (2014).
31. J. Zhang, R. Qiu, H. N. Arst Jr., M. A. Peñalva, X. Xiang, *J. Cell Biol.* **204**, 1009–1026 (2014).

ACKNOWLEDGMENTS

We thank K. Ori-McKenney for helpful discussions and analysis. This work was supported by NIH F32 postdoctoral grant F32GM096484 (to R.J.M.) and R01GM097312 (to R.D.V.). G.B. is the Merck Fellow of the Damon Runyon Cancer Research Foundation (DRG-2136-12).

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/345/6194/337/suppl/DC1
Materials and Methods
Figs. S1 to S7
Movies S1 to S9
References (*32*–*37*)

1 April 2014; accepted 9 June 2014
Published online 19 June 2014;
10.1126/science.1254198

Transfection Reagent

The FectoPRO is an optimized transfection reagent which has been designed for high yields in medium- and large-scale production of proteins and antibodies in mammalian cells. FectoPRO delivers very high protein yields in CHO and HEK cell line systems. FectoPRO will also produce up to three times more protein using half the DNA required by other leading commercial transfection reagents. In addition, scientists working with transient gene expression systems using FectoPRO will be able to benefit from reproducibility experiment-to-experiment and between reagent batches. The FectoPRO transfection reagent offers scientists working in pharmaceuticals and biotechnology specific advantages over other commercial transfection reagent offerings for bio-production. The reagent is well characterized in terms of chemical structure and is supplied ready-to-use. FectoPRO is also manufactured and formulated using a highly specific and controlled production process in Polyplus' dedicated manufacturing facility.

Polyplus-transfection
For info: +33-(0)-390-406180
www.polyplus-transfection.com

Oligonucleotide Delivery Agent

A novel vitamin modifier has shown in initial tests to have the potential of improving cell delivery of oligonucleotides. The new reagent, 5'-Niacin-CE Phosphoramidite, is easily incorporated during solid phase oligo synthesis. The niacin-based modifier offers several advantages over the use of conventional lipophilic delivery agents, including reduced risk of in vivo toxicity, and removes the necessity of cleaving the delivery reagent once in the cell. These benefits make vitamins an attractive method for the delivery of therapeutic oligonucleotides. While lipophilic modifier reagents have been shown to enhance cell penetration, vitamin-mediated cell delivery offers a distinct advantage due to the fact that vitamins are required, but not produced by cells. As such, it is believed that interaction with a specific binding protein is required before the vitamin-oligo conjugate is internalized. Not only does this enhance delivery and overcome the risk of toxicity, but it also offers some exciting potential for cell targeting.

Link Technologies
For info: +44-(0)-1698-849911
www.linktech.co.uk

RNA FFPE Kit

The Maxwell 16 LEV RNA FFPE Kit for RNA isolation from mammalian formalin-fixed, paraffin-embedded (FFPE) tissue samples. The new kit is designed to rapidly extract pure, high yield RNA from thin sections of FFPE tissue samples, with consistency and time saving provided by the automated Maxwell 16 instrument. The purified RNA is suitable for downstream applications including reverse transcription polymerase chain reaction (RT-PCR), RT-qPCR, endpoint RT-PCR, cDNA library synthesis, and gene expression analysis. Specifically designed for optimal purification of RNA from 5 µm sections of FFPE samples, the new kit maximizes RNA yield and eliminates the need to concentrate the extract prior to amplification, minimizing sample waste and the need for reruns. The kit is very simple to use, with the automated walkaway system able to extract from 16 samples in a single run. Lysates are placed directly into the cartridges, and RNA is ready for amplification in approximately 60 minutes (after proteinase K digest).

Promega
For info: 608-274-4330
www.promega.com/maxwellffperna

Target-Enrichment Kit

The SureSelect^{QXT} Reagent Kit is a revolutionary next generation sequencing target enrichment solution that produces sample-to-sequencing-ready-libraries in just seven hours with only 50 ng of input gDNA. Designed to meet the exacting needs of clinical researchers for a fast, easy, same-day sample-to-sequencing-workflow, SureSelect^{QXT} kits are three times faster than existing transposase-based methods and require 30% less hands-on time. They are also optimized for use with samples of limited availability and provide superior coverage of genomic targets for confident variant calling. Combined with the fastest sequencers on the market, these kits make sample-to-data in 24–36 hours a reality. SureSelect^{QXT} kits' unprecedented speed is achieved by coupling transposase-based library preparation with major advancements in hybridization chemistry, significantly decreasing hybridization time from 16 hours to 90 minutes while maintaining high performance. Its proven hybridization technology enables the sensitivity and specificity required for the complete and accurate variant analysis of exomes or custom genomic regions.

Agilent Technologies
For info: 800-227-9770
www.agilent.com/genomics/qxt

Tumor Panel

Fully validated on formalin-fixed, paraffin-embedded samples, the new 60-gene next generation sequencing hybridization-based SureSeq Solid Tumor Panel offers researchers accurate and reliable solid tumor profiling for both known and novel variants. The content of the panel has been defined by recognized cancer experts, covering key genes for a range of cancer types. All exons of these genes are fully covered, including mutation hotspots, enabling both detection and discovery of known and novel variants, respectively. The hybridization-based SureSeq Solid Tumour Panel minimizes polymerase chain reaction bias and duplications commonly associated with alternative enrichment methodologies, enabling greater run-to-run consistency. This is particularly important in situations where there is limited sample or where the ability to detect minor allele frequencies is required. Such sample types require a highly uniform and sensitive enrichment and OGT's expert bait design ensures this by providing efficient and improved uniformity of coverage of the targeted regions.

Oxford Gene Technology
For info: +44-(0)-1865-856826
www.ogt.com

Electronically submit your new product description or product literature information! Go to www.sciencemag.org/products/newproducts.dtl 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.

Downloaded from www.sciencemag.org on August 4, 2014

By Fanuel Muindi

Tell the negative committee to shut up

When I started out in Stanford University's biology doctoral program, I didn't feel ready. My feeling that I was poorly prepared was corroborated by a committee that told me often how underprepared and unqualified I was. I attempted to argue my case, but the committee held to its position: I was unworthy. Fast forward 2 years: Just before I walked into my qualifying exam, the committee convinced me that I was going to fail. I succeeded, but that didn't change its low opinion of me. Then, at the end of my third year, I was selected to receive a Diversifying Academia, Recruiting Excellence (DARE) fellowship, which targets students in their final 2 years of graduate school who are interested in pursuing academic careers. The committee now argued that I didn't belong in the company of the other fellows, who, it insisted, were way more accomplished than I was.

Later, when I interviewed for postdoc positions, committee members had nothing positive to say. "You are not ready," they said. "You are not good enough." When the time came to defend my Ph.D. thesis, they were at it again: "You don't deserve this," they said, after I had succeeded. "Your work isn't that good. You won't get it published anywhere decent." It was a continuous barrage of criticism aimed at undermining my self-confidence.

"The committee is no longer in charge–I am."

Fortunately, this was not my doctoral committee or any other committee that existed at Stanford University, or any other place in the real world. My Ph.D. committee at Stanford was constructive and supportive. The one I'm referring to—call it the negative committee—existed, and continues to exist, only inside my own brain.

For years, the members of the negative committee have tried to make me feel like an impostor, and for years they succeeded. They didn't believe in the person I was becoming: an accomplished academic who is critical, visionary, confident, and authentic. They have made it difficult, as Valerie Young, an expert on impostor syndrome, puts it, "to internalize or feel deserving of any success." The negative committee drowns out other voices, making it difficult to hear anything but its own incessant harping—difficult, but not impossible.

In time, I gained the confidence needed to stand up to the negative committee. Intense self-reflection through mindful meditation helped me learn more about what the committee was after—my self-confidence—and how to fight it. The committee never quite succeeded because even as it was working to undermine my confidence, I was working to build it up. My accomplishments in graduate school, and the DARE fellowship, helped build my confidence. I also had help from a lot of people: formal and informal mentors, friends, lab mates, and family. I draw on this confidence today when I tell the negative committee to "sit down and shut up," as writer Ann Bradford has suggested.

The negative committee isn't all bad. As with all things, it depends on how you look at it. Yes, it has tried to undermine me, but I have learned from it. I have learned to turn its negatives into positives, to view its criticisms as opportunities for growth. The negative committee helps me stay grounded and be realistic about my true abilities. It keeps me on my toes.

You, too, probably have a negative committee that you need to learn how to manage. It takes time, and it isn't easy, but it isn't that hard, either. Use your mentors, friends, and family—your positive committee, which has the advantage of existing in the real world—to help rein in your negative committee. I think the Dalai Lama said it best: "The way to overcome negative thoughts and destructive emotions is to develop opposing, positive emotions that are stronger and more powerful."

Now that I'm a postdoc at the Massachusetts Institute of Technology in Cambridge, my negative committee continues to criticize my efforts. But things are different now. The committee is no longer in charge—I am. I now know how to not take its criticisms seriously. I know how and when to shut it up. And when I can't shut it up, I just prove it wrong. ■

Fanuel Muindi is a postdoctoral fellow at the Massachusetts Institute of Technology in Cambridge. For more on careers and life in science, visit http://www.sciencecareers.org.

ILLUSTRATION: MARC ROSENTHAL

Downloaded from www.sciencemag.org on August 4, 2014

Published by AAAS