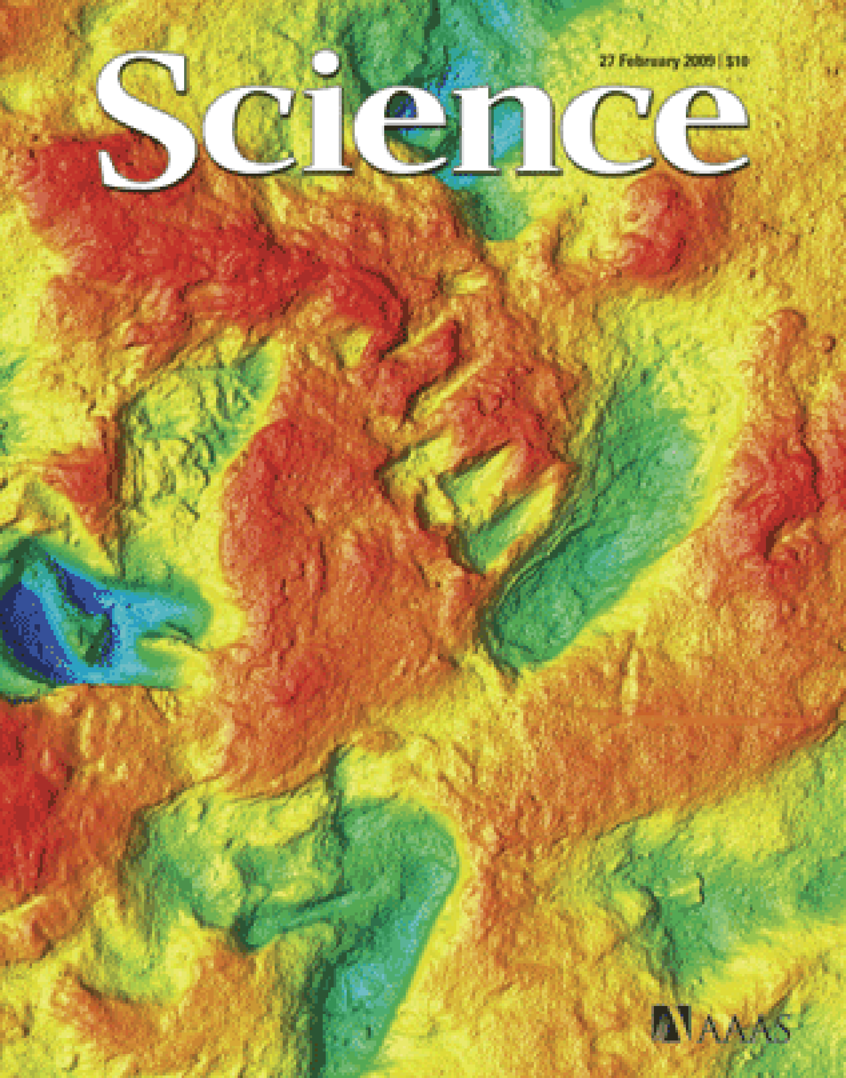
27 February 2009 | $10

# Science



AAAS

# CONTENTS

Volume 323 Issue 5918

## EDITORIAL

## NEWS OF THE WEEK

## NEWS FOCUS

## LETTERS

## BOOKS *ET AL.*

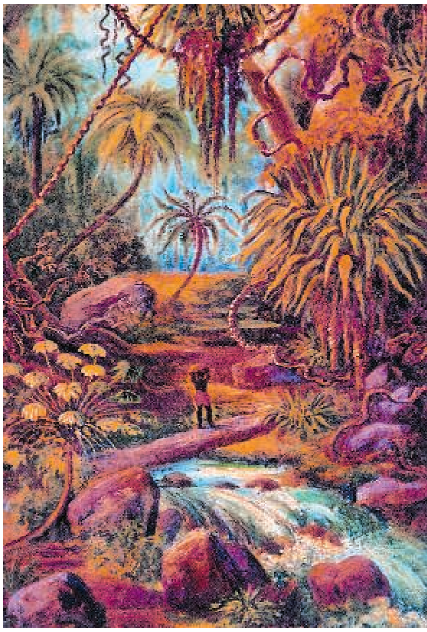
## EDUCATION FORUM

## PERSPECTIVES

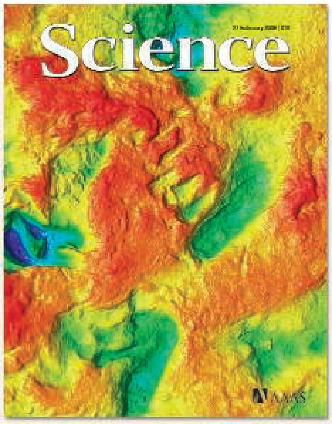
## BREVIA

**CONTENTS** continued **>>**

page 1162



page 1170



## COVER

Optical laser scan of early hominin footprints at Ileret, Kenya, color-rendered to illustrate depth; reds indicate areas of high elevation, blues lower elevation. The footprints are 1.5 million years old and were probably made by *Homo ergaster/erectus*. Two right footprints and a partial left are visible along with a range of animal prints. See page 1197.

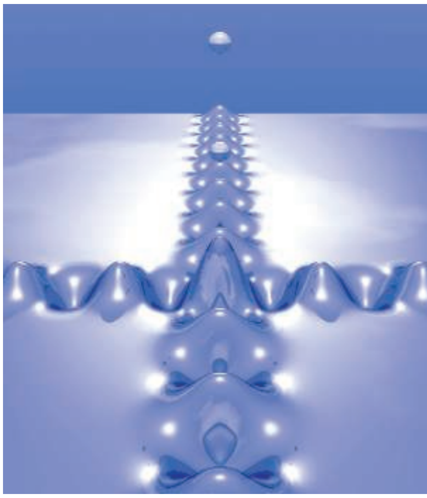
***Image processing in Rapidform$^{TM}$: Matthew Bennett/ Bournemouth University***

## DEPARTMENTS

## REPORTS

**CONTENTS** continued **>>**

page 1176



pages 1178 & 1190

page 1201

*Published by AAAS*

# SCIENCEONLINE

## SCIENCEXPRESS

www.sciencexpress.org

### Greatly Expanded Tropical Warm Pool and Weakened Hadley Circulation in the Early Pliocene

*C. M. Brierley* et al.

The warm tropics of the Early Pliocene, about 4 million years ago, extended much farther toward the poles than they do today.

10.1126/science.1167625

### RNA Pol II Accumulates at Promoters of Growth Genes During Developmental Arrest

*L. R. Baugh* et al.

Growth and development genes, poised for expression during developmental arrest in nematodes, respond rapidly to feeding.

10.1126/science.1169628

### A Transposon-Based Genetic Screen in Mice Identifies Genes Altered in Colorectal Cancer

*T. K. Starr* et al.

A functional screen in mice uncovers genes that are likely to drive the growth of gut-specific tumors.

10.1126/science.1163040

### Antibody Recognition of a Highly Conserved Influenza Virus Epitope

*D. C. Ekiert* et al.

A broadly neutralizing antibody binds the hemagglutinin stalk of pathogenic influenza viruses to block membrane fusion.

10.1126/science.1171491

## TECHNICALCOMMENTS

### Comment on "Multipartite Entanglement Among Single Spins in Diamond"

*B. W. Lovett and S. C. Benjamin*

full text at www.sciencemag.org/cgi/content/full/323/5918/1169c

### Response to Comment on "Multipartite Entanglement Among Single Spins in Diamond"

*P. Neumann* et al.

full text at www.sciencemag.org/cgi/content/full/323/5918/1169d

## SCIENCENOW

www.sciencenow.org

Highlights From Our Daily News Coverage

### Abuse Leaves Its Mark on the Brain

Parental abuse alters a stress gene in the brain.

### Permanent Protection Against the Flu?

A few antibodies may guard against a variety of influenza strains.

### X-ray Vision Reveals Intergalactic Medium

Missing matter has been spotted in a supersized wall of galaxies.

## SCIENCESIGNALING

www.sciencesignaling.org

The Signal Transduction Knowledge Environment

### RESEARCH ARTICLE: Dok-7 Activates the Muscle Receptor Kinase MuSK and Shapes Synapse Formation

*A. Inoue* et al.

The cytoplasmic protein Dok-7 is required for the full activation of the receptor tyrosine kinase MuSK.

### RESEARCH ARTICLE: IL-17 Receptor Signaling Inhibits C/EBPβ by Sequential Phosphorylation of the Regulatory 2 Domain

*F. Shen* et al.

The IL-17 receptor triggers dual, sequential phosphorylation of the transcription factor C/EBPβ, which represses the expression of target genes.

### REVIEW: Positive and Negative Modulation of Angiogenesis by VEGFR1 Ligands

*Y. Cao*

PlGF and VEGF-B exhibit dual effects on angiogenesis mediated by VEGF-A.

### PERSPECTIVE: Jasmonate—Preventing the Maize Tassel from Getting in Touch with His Feminine Side

*J. Browse*

Male identity in the maize tassel is controlled by jasmonate signaling.

### PERSPECTIVE: Chemoattractant Receptors and Leukocyte Recruitment—More Than Cell Migration

*E. M. Borroni* et al.

Atypical "chemoattractant receptors" that do not signal through G proteins contribute to leukocyte recruitment through indirect mechanisms.

### MEETING REPORT: Signal Transduction Molecules as Targets for Cancer Prevention

*A. M. Bode and Z. Dong*

Two conferences at The Hormel Institute highlight early intervention and chemoprevention in targeting cancer.

## SCIENCEPODCAST

www.sciencemag.org/multimedia/podcast

Free Weekly Show

Download the 27 February *Science* Podcast to hear about predicting election outcomes, drug-resistant influenza, your letters to *Science*, and more.

## ORIGINSBLOG

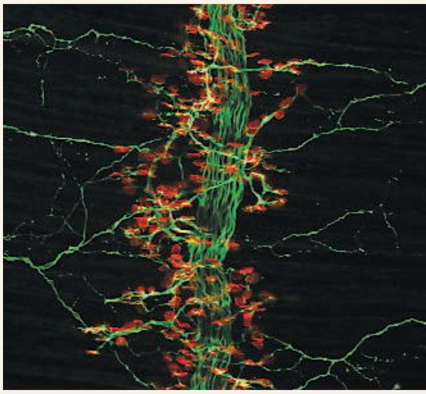
blogs.sciencemag.org/origins

A History of Beginnings

## SCIENCEINSIDER

blogs.sciencemag.org/scienceinsider

Science Policy News and Analysis



SCIENCE**SIGNALING**
Innervating diaphragm muscle.

CREDIT: (SCIENCE SIGNALING) AKANE INOUE AND YUJI YAMANASHI/INSTITUTE OF MEDICAL SCIENCE/UNIVERSITY OF TOKYO/JAPAN

## SCIENCECAREERS

www.sciencecareers.org/career_magazine

Free Career Resources for Scientists

### Special Feature: Two Young, African-American Women in Science

*J. Austin*

Our sample of African-American women reveals brilliance, scientific ambition, and progress against discrimination.

### When Ironies Make Perfect Sense

*A. Sasso*

Gina Wingood, a black Catholic girl raised in a white suburb, found love and her calling in San Francisco's ghettos.

### The Bigger Questions

*S. Gaidos*

Chemical engineer Kristala Jones Prather's career has taken her from academia to industry and back again.

### A Double Bind—Minority Women in Science in Europe

*E. Pain*

Minority women in European science confront an issue that remains taboo.

**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 © 2009 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): $146 ($74 allocated to subscription). Domestic institutional subscription (51 issues): $835; 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 $20.00 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.

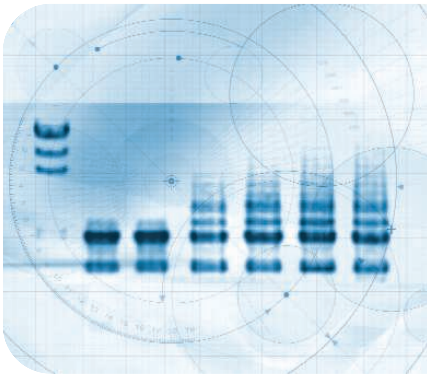
AAAS
ADVANCING SCIENCE, SERVING SOCIETY

Susan Hockfield is president of the Massachusetts Institute of Technology, Cambridge, MA. E-mail: Hockfield@mit.edu.

# The Next Innovation Revolution

U.S. FEDERAL INVESTMENTS IN BASIC RESEARCH TRANSFORMED LIFE AND COMMERCE IN THE 20th century. They sent us to the Moon and beyond, revolutionized communications, helped to feed the planet, reinvented work processes, and drove the remarkable economic growth of the post-1950s era in the United States. These advances and more grew out of the convergence between engineering and the early 20th-century discoveries in the physical sciences. The United States can anticipate comparable world-changing innovations in the 21st century if we adapt our education and research funding strategies to capitalize on new opportunities emerging at the convergence of the life sciences with the physical sciences and engineering.

This next convergence follows from the elucidation of the structure of DNA in the 1950s and from subsequent fundamental discoveries in molecular and cellular biology. These discoveries created a revolution in the life sciences and drove the development of recombinant DNA technology and the launch of the biotechnology industry. By the mid-1980s, the explosion of data from genomics and proteomics brought about a second revolution, further accelerating life science innovation.



These revolutions sowed the seeds of a third revolution that links the life sciences with engineering and the physical sciences in powerful new ways. Many of molecular biology's founders came from the physical sciences, bringing to biology new analytical strategies and technologies. With the evolution of data- and technology-based biology, biologists worked increasingly closely with mathematicians, engineers, and physical scientists. Yet too often the relationship focused simply on developing new tools, with engineers serving more as service providers rather than as true collaborators, separately developing analytical technologies or computational strategies for analyzing data. With the demands for analyzing increasingly large data sets in genomics, understanding the complex network of molecular interactions in cells, and increasing resolution and accuracy in measuring and manipulating molecular and cellular events, what began as a relationship of convenience has become a flourishing partnership that now is rapidly advancing life science discoveries toward practical uses for society.

At the Massachusetts Institute of Technology (MIT), a third of the almost 400 engineering faculty engage the life sciences in their research. In many fields, the research frontiers are populated by biologists, engineers, computational scientists, and chemists. MIT's new Koch Institute for Integrative Cancer Research includes biologists, chemists, and engineers working together to develop new strategies for cancer diagnosis, treatment, and prevention. Here and elsewhere, labs are, for instance, engineering nanoparticles to transport cancer-fighting agents selectively to tumor cells, thus minimizing injury to healthy tissue. Approaches like this could become a clinical reality within a decade. The evolving third revolution has, not surprisingly, produced a new academic discipline. The field is burgeoning: in 1996, U.S. universities awarded only 220 Ph.D.s in bioengineering and biomedical engineering; in 2006, that figure soared to 525.

Today, this convergence is spawning new discoveries and applications in areas from biomedicine to environmental science to energy technology. Accelerating these innovations will require many changes, from how we teach to how we fund research. Science and engineering students need a broad, more integrated education, so that they can work fluently across disciplines. We must also find more effective ways to evaluate and fund interdisciplinary research in the United States, including specifically promoting interdisciplinary work in the National Institutes of Health's (NIH's) programs for young investigators. Grant review committees will increasingly require multidisciplinary membership, and we will need to change the rules and practices that inhibit funding for projects that cut across federal funding agencies and NIH divisions. Finally, we need to change the way in which grants are awarded and administered: It's needlessly difficult to secure funding for research that involves multiple investigators, departments, or institutions. Above all, scientists must convince the public, the Congress, and the Obama administration that funding research that cuts across the life, physical, and engineering sciences is a vital investment in human health, environmental well-being, and economic prosperity.

**– Susan Hockfield**

10.1126/science.1170834

CREDITS: (TOP) MIT; (RIGHT) JUPITERIMAGES www.sciencemag.org

# Science

**1200 New York Avenue, NW**
**Washington, DC 20005**
Editorial: 202-326-6550, FAX 202-289-7562
News: 202-326-6581, FAX 202-371-9227
**Bateman House, 82-88 Hills Road**
**Cambridge, UK CB2 1LQ**
+44 (0) 1223 326500, FAX +44 (0) 1223 326501

**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** please call 202-326-6755 for any questions or information

**Reprints:** Author Inquiries 800-635-7181
Commercial Inquiries 803-359-4578

**Permissions** 202-326-7074, FAX 202-682-0816

**Member Benefits** AAAS/Barnes&Noble.com bookstore www.aaas.org/bn; AAAS Online Store www.apisource.com/aaas/ code MKB6; AAAS Travels: Betchart Expeditions 800-252-4910; Apple Store www.apple/eppstore/aaas; Bank of America MasterCard 1-800-833-6262 priority code FAA3YU; Cold Spring Harbor Laboratory Press Publications www.cshlpress.com/affiliates/aaas.htm; GEICO Auto Insurance www.geico.com/landingpage/go51.htm?logo=17624; Hertz 800-654-2200 CDP#343457; Office Depot https://bsd.officedepot.com/portalLogin.do; Seaburg & Smith Life Insurance 800-424-9883; Subaru VIP Program 202-326-6417; VIP Moving Services www.vipmayflower.com/domestic/index.html; Other Benefits: AAAS Member Services 202-326-6417 or www.aaasmember.org.

| | |
|---|---|
| science_editors@aaas.org | (for general editorial queries) |
| science_letters@aaas.org | (for queries about letters) |
| science_reviews@aaas.org | (for returning manuscript reviews) |
| science_bookrevs@aaas.org | (for book review queries) |

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

AAAS was founded in 1848 and incorporated in 1874. Its mission is to advance science, engineering, and innovation throughout the world for the benefit of all people. The goals of the association are to: enhance communication among scientists, engineers, and the public; promote and defend the integrity of science and its use; strengthen support for the science and technology enterprise; provide a voice for science on societal issues; promote the responsible use of science in public policy; strengthen and diversify the science and technology workforce; foster education in science and technology for everyone; increase public engagement with science and technology; and advance international cooperation in science.

**Information for Authors**
See pages 807 and 808 of the 6 February 2009 issue or access www.sciencemag.org/about/authors

EDITOR-IN-CHIEF **Bruce Alberts**
EXECUTIVE EDITOR **Monica M. Bradford**
DEPUTY EDITORS **R. Brooks Hanson, Barbara R. Jasny, Katrina L. Kelner**
NEWS EDITOR **Colin Norman**

**Editorial supervisory senior editor** Phillip D. Szuromi; **senior editor/perspectives** Lisa D. Chong; **senior editors** Gilbert J. Chin, Pamela J. Hines, Paula A. Kiberstis (Boston), Marc S. Lavine (Toronto), Beverly A. Purnell, L. Bryan Ray, Guy Riddihough, H. Jesse Smith, Valda Vinson; **associate editors** Kristen L. Mueller, Jake S. Yeston, Laura M. Zahn; **online editor** Stewart Wills; **associate online editors** Robert Frederick, Tara S. Marathe; **web content developer** Martyn Green; **book review editor** Sherman J. Suter; **associate letters editor** Jennifer Sills; **editorial manager** Cara Tate; **senior copy editors** Jeffrey E. Cook, Cynthia Howe, Harry Jach, Barbara P. Ordway, Trista Wagoner; **copy editors** Chris Filiatreau, Lauren Kmec; **editorial coordinators** Carolyn Kyle, Beverly Shields; **publications assistants** Ramatoulaye Diop, Joi S. Granger, Jeffrey Hearn, Lisa Johnson, Scott Miller, Jerry Richardson, Jennifer A. Seibert, Brian White, Anita Wynn; **editorial assistants** Carlos L. Durham, Emily Guise, Patricia M. Moore; **executive assistant** Sylvia S. Kihara; **administrative support** Maryrose Madrid

**News deputy news editors** Robert Coontz, Eliot Marshall, Jeffrey Mervis, Leslie Roberts; **contributing editors** Elizabeth Culotta, Polly Shulman; **news writers** Yudhijit Bhattacharjee, Adrian Cho, Jennifer Couzin, David Grimm, Constance Holden, Jocelyn Kaiser, Richard A. Kerr, Eli Kintisch, Andrew Lawler (New England), Greg Miller, Elizabeth Pennisi, Robert F. Service (Pacific NW), Erik Stokstad; **intern** Jackie D. Grom; **contributing correspondents** Dan Charles, Jon Cohen (San Diego, CA), Daniel Ferber, Ann Gibbons, Robert Koenig, Mitch Leslie, Charles C. Mann, Virginia Morell, Evelyn Strauss, Gary Taubes; **copy editors** Linda B. Felaco, Melvin Gatling, Melissa Raimondi; **administrative support** Scherraine Mack, Fannie Groom; **bureaus** New England: 207-549-7755, San Diego, CA: 760-942-3252, FAX 760-942-4979, Pacific Northwest: 503-963-1940

**Production director** James Landry; **senior manager** Wendy K. Shank; **assistant manager** Rebecca Doshi; **senior specialists** Steve Forrester, Chris Redwood; **specialist** Anthony Rosen; **Preflight director** David M. Tompkins; **manager** Marcus Spiegler

**Art director** Yael Kats; **associate art director** Laura Creveling; **illustrators** Chris Bickel, Katharine Sutliff; **senior art associates** Holly Bishop, Preston Huey, Nayomi Kevitiyagala; **art associate** Jessica Newfield; **photo editor** Leslie Blizard

*Science* International

**Europe** (science@science-int.co.uk) **editorial: international managing editor** Andrew M. Sugden; **senior editor/perspectives** Julia Fahrenkamp-Uppenbrink; **senior editors** Caroline Ash, Stella M. Hurtley, Ian S. Osborne, Peter Stern; **associate editor** Maria Cruz; **editorial support** Deborah Dennison, Rachel Roberts, Alice Whaley; **administrative support** John Cannell, Janet Clements; **news: europe news editor** John Travis; **deputy news editor** Daniel Clery; **contributing correspondents** Michael Balter (Paris), John Bohannon (Vienna), Martin Enserink (Amsterdam and Paris), Gretchen Vogel (Berlin); **intern** Sara Coelho

**Asia** Japan Office: Asca Corporation, Eiko Ishioka, Fusako Tamura, 1-8-13, Hirano-cho, Chuo-ku, Osaka-shi, Osaka, 541-0046 Japan; +81 (0) 6 6202 6272, FAX +81 (0) 6 6202 6271; asca@os.gulf.or.jp; **asia news editor** Richard Stone (Beijing: rstone@aaas.org); **contributing correspondents** Dennis Normile (Japan: +81 (0) 3 3391 0630, FAX +81 (0) 3 5936 3531; dnormile@gol.com); Hao Xin (China: +86 (0) 10 6307 4439 or 6307 3676, FAX +86 (0) 10 6307 4358; cindyhao@gmail.com); Pallava Bagla (South Asia: +91 (0) 11 2271 2896; pbagla@vsnl.com)

EXECUTIVE PUBLISHER **Alan I. Leshner**
PUBLISHER **Beth Rosner**

**Fulfillment Systems and Operations** (membership@aaas.org); **director** Waylon Butler; **senior systems analyst** Jonny Blaker; **customer service supervisor** Pat Butler; **specialists** Latoya Casteel, LaVonda Crawford, Vicki Linton, April Marshall; **data entry supervisor** Cynthia Johnson; **specialists** Eintou Bowden, Tarrika Hill, William Jones

**Business Operations and Administration director** Deborah Rivera-Wienhold; **assistant director, business operations** Randy Yi; **manager, business analysis** Michael LoBue; **manager, business operations** Jessica Tierney; **financial analysts** Priti Pamnani, Celeste Troxler; **rights and permissions: administrator** Emilie David; **associate** Elizabeth Sandler; **marketing director** Ian King; **marketing manager** Allison Pritchard; **marketing associates** Aimee Aponte, Alison Chandler, Mary Ellen Crowley, Julianne Wielga, Wendy Wise; **marketing executive** Jennifer Reeves; **marketing/member services executive** Linda Rusk; **director, site licensing** Tom Ryan; **director, corporate relations** Eileen Bernadette Moran; **publisher relations, eresources specialist** Kiki Forsythe; **senior publisher relations specialist** Catherine Holland; **publisher relations, east coast** Phillip Smith; **publisher relations, west coast** Philip Tsolakidis; **fulfillment supervisor** Iquo Edim; **fulfillment coordinator** Laura Clemens; **electronic media: manager** Lizabeth Harman; **project manager** Trista Snyder; **assistant manager** Lisa Stanford; **senior production specialists** Christopher Coleman, Walter Jones; **production specialists** Nichele Johnston, Kimberly Oster

**Advertising director, worldwide ad sales** Bill Moran

**Product** (science_advertising@aaas.org); **midwest/west coast/w. canada** Rick Bongiovanni: 330-405-7080, FAX 330-405-7081; **east coast/e. canada** Laurie Faraday: 508-747-9395, FAX 617-507-8189; **uk/europe/asia** Roger Goncalves: TEL/FAX +41 43 243 1358; **japan** Mashy Yoshikawa: +81 (0) 3 3235 5961, FAX +81 (0) 3 3235 5852; **senior traffic associate** Deiandra Simms

**Commercial Editor** Sean Sanders: 202-326-6430

**Project Director, Outreach** Brianna Blaser

**Classified** (advertise@sciencecareers.org); **inside sales manager: midwest/canada** Daryl Anderson: 202-326-6543; **inside sales representative** Karen Foote: 202-326-6740; **key account manager** Joribah Able; **northeast** Alexis Fleming: 202-326-6578; **southeast** Tina Burks: 202-326-6577; **west** Nicholas Hintibidze: 202-326-6533; **sales coordinators** Rohan Edmonson, Shirley Young; **international: sales manager** Tracy Holmes: +44 (0) 1223 326525, FAX +44 (0) 1223 326532; **sales** Susanne Kharraz, Dan Pennington, Alex Palmer; **sales assistant** Louise Moore; **japan** Mashy Yoshikawa: +81 (0) 3 3235 5961, FAX +81 (0) 3 3235 5852; **advertising production operations manager** Deborah Tompkins; **senior production specialist/graphic designer** Amy Hardcastle; **senior production specialist** Robert Buck; **senior traffic associate** Christine Hall; **publications assistant** Mary Lagnaoui

**AAAS Board of Directors retiring president, chair** James J. McCarthy; **president** Peter C. Agre; **president-elect** Alice Huang; **treasurer** David E. Shaw; **chief executive officer** Alan I. Leshner; **board** Alice Gast, Linda P. B. Katehi, Nancy Knowlton, Cherry A. Murray, Julia M. Phillips, Thomas D. Pollard, David S. Sabatini, Thomas A. Woolsey

AAAS
ADVANCING SCIENCE, SERVING SOCIETY

## SENIOR EDITORIAL BOARD

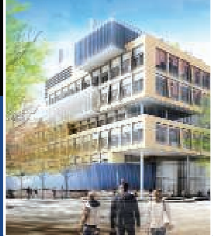
**John I. Brauman,** *Chair, Stanford Univ.*
**Richard Losick,** *Harvard Univ.*
**Robert May,** *Univ. of Oxford*
**Marcia McNutt,** *Monterey Bay Aquarium Research Inst.*
**Linda Partridge,** *Univ. College London*
**Vera C. Rubin,** *Carnegie Institution*
**Christopher R. Somerville,** *Univ. of California, Berkeley*

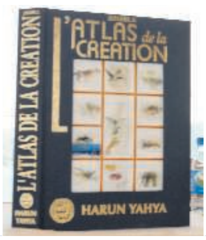
## BOARD OF REVIEWING EDITORS

**Joanna Aizenberg,** *Harvard Univ.*
**Sonia Altizer,** *Univ. of Georgia*
**David Altshuler,** *Broad Institute*
**Arturo Alvarez-Buylla,** *Univ. of California, San Francisco*
**Richard Amasino,** *Univ. of Wisconsin, Madison*
**Angelika Amon,** *MIT*
**Meinrat O. Andreae,** *Max Planck Inst., Mainz*
**Kristi S. Anseth,** *Univ. of Colorado*
**John A. Bargh,** *Yale Univ.*
**Cornelia I. Bargmann,** *Rockefeller Univ.*
**Ben Barres,** *Stanford Medical School*
**Marisa Bartolomei,** *Univ. of Penn. School of Med.*
**Facundo Batista,** *London Research Inst.*
**Ray H. Baughman,** *Univ. of Texas, Dallas*
**Stephen J. Benkovic,** *Penn State Univ.*
**Ton Bisseling,** *Wageningen Univ.*
**Mina Bissell,** *Lawrence Berkeley National Lab*
**Peer Bork,** *EMBL*
**Robert W. Boyd,** *Univ. of Rochester*
**Paul M. Brakefield,** *Leiden Univ.*
**Stephen Buratowski,** *Harvard Medical School*
**Joseph A. Burns**, *Cornell Univ.*
**William P. Butz,** *Population Reference Bureau*
**Mats Carlsson,** *Univ. of Oslo*
**Peter Carmeliet,** *Univ. of Leuven, VIB*
**Mildred Cho,** *Stanford Univ.*
**David Clapham,** *Children's Hospital, Boston*
**David Clary,** *Oxford University*
**J. M. Claverie,** *CNRS, Marseille*
**Jonathan D. Cohen,** *Princeton Univ.*
**Andrew Cossins,** *Univ. of Liverpool*
**Robert H. Crabtree,** *Yale Univ.*
**Wolfgang Cramer,** *Potsdam Inst. for Climate Impact Research*
**F. Fleming Crim,** *Univ. of Wisconsin*
**William Cumberland,** *Univ. of California, Los Angeles*
**Jeff L. Dangl,** *Univ. of North Carolina*
**Stanislas Dehaene,** *Collège de France*
**Edward DeLong,** *MIT*
**Emmanouil T. Dermitzakis,** *Wellcome Trust Sanger Inst.*
**Robert Desimone,** *MIT*
**Claude Desplan,** *New York Univ.*
**Dennis Discher,** *Univ. of Pennsylvania*
**Scott C. Doney,** *Woods Hole Oceanographic Inst.*
**W. Ford Doolittle,** *Dalhousie Univ.*
**Jennifer A. Doudna,** *Univ. of California, Berkeley*
**Julian Downward,** *Cancer Research UK*
**Denis Duboule,** *Univ. of Geneva/EPFL Lausanne*
**Christopher Dye,** *WHO*
**Gerhard Ertl,** *Fritz-Haber-Institut, Berlin*
**Mark Estelle,** *Indiana Univ.*
**Barry Everitt,** *Univ. of Cambridge*
**Paul G. Falkowski**, *Rutgers Univ.*
**Ernst Fehr,** *Univ. of Zurich*
**Tom Fenchel,** *Univ. of Copenhagen*
**Alain Fischer,** *INSERM*
**Scott E. Fraser,** *Cal Tech*
**Chris D. Frith,** *Univ. College London*
**Wulfram Gerstner,** *EPFL Lausanne*
**Charles Godfray,** *Univ. of Oxford*
**Diane Griffin,** *Johns Hopkins Bloomberg School of Public Health*
**Christian Haass,** *Ludwig Maximilians Univ.*
**Niels Hansen,** *Technical Univ. of Denmark*
**Dennis L. Hartmann,** *Univ. of Washington*
**Chris Hawkesworth,** *Univ. of Bristol*
**Martin Heimann,** *Max Planck Inst., Jena*
**James A. Hendler,** *Rensselaer Polytechnic Inst.*
**Ray Hilborn,** *Univ. of Washington*
**Kei Hirose,** *Tokyo Inst. of Technology*
**Ove Hoegh-Guldberg,** *Univ. of Queensland*
**Brigid L. M. Hogan,** *Duke Univ. Medical Center*
**Ronald R. Hoy,** *Cornell Univ.*
**Olli Ikkala,** *Helsinki Univ. of Technology*
**Meyer B. Jackson,** *Univ. of Wisconsin Med. School*
**Stephen Jackson,** *Univ. of Cambridge*
**Steven Jacobsen,** *Univ. of California, Los Angeles*
**Peter Jonas,** *Universität Freiburg*
**Barbara B. Kahn,** *Harvard Medical School*
**Daniel Kahne,** *Harvard Univ.*
**Gerard Karsenty,** *Columbia Univ. College of P&S*
**Bernhard Keimer,** *Max Planck Inst., Stuttgart*
**Elizabeth A. Kellog,** *Univ. of Missouri, St. Louis*
**Alan B. Krueger**, *Princeton Univ.*
**Lee Kump**, *Penn State Univ.*
**Mitchell A. Lazar**, *Univ. of Pennsylvania*
**Virginia Lee**, *Univ. of Pennsylvania*
**Olle Lindvall,** *Univ. Hospital, Lund*
**Marcia C. Linn,** *Univ. of California, Berkeley*
**John Lis,** *Cornell Univ.*
**Richard Losick,** *Harvard Univ.*
**Ke Lu,** *Chinese Acad. of Sciences*
**Andrew P. MacKenzie,** *Univ. of St Andrews*
**Raul Madariaga,** *École Normale Supérieure, Paris*
**Anne Magurran,** *Univ. of St Andrews*
**Charles Marshall,** *Harvard Univ.*
**Virginia Miller,** *Washington Univ.*
**Yasushi Miyashita,** *Univ. of Tokyo*
**Richard Morris,** *Univ. of Edinburgh*
**Edvard Moser,** *Norwegian Univ. of Science and Technology*
**Naoto Nagaosa**, *Univ. of Tokyo*
**James Nelson,** *Stanford Univ. School of Med.*
**Timothy W. Nilsen,** *Case Western Reserve Univ.*
**Roeland Nolte,** *Univ. of Nijmegen*
**Helga Nowotny,** *European Research Advisory Board*
**Eric N. Olson,** *Univ. of Texas, SW*
**Stuart H. Orkin,** *Dana-Farber Cancer Inst.*
**Erin O'Shea,** *Harvard Univ.*
**Elinor Ostrom,** *Indiana Univ.*
**Jonathan T. Overpeck,** *Univ. of Arizona*
**John Pendry**, *Imperial College*
**Simon Phillpot**, *Univ. of Florida*
**Philippe Poulin,** *CNRS*
**Mary Power,** *Univ. of California, Berkeley*
**Molly Przeworski,** *Univ. of Chicago*
**Colin Renfrew,** *Univ. of Cambridge*
**Trevor Robbins,** *Univ. of Cambridge*
**Barbara A. Romanowicz,** *Univ. of California, Berkeley*
**Edward M. Rubin,** *Lawrence Berkeley National Lab*
**Shimon Sakaguchi,** *Kyoto Univ.*
**Jürgen Sandkühler,** *Medical Univ. of Vienna*
**David W. Schindler,** *Univ. of Alberta*
**Georg Schulz,** *Albert-Ludwigs-Universität*
**Paul Schulze-Lefert,** *Max Planck Inst., Cologne*
**Christine Seidman,** *Harvard Medical School*
**Terrence J. Sejnowski,** *The Salk Institute*
**Richard J. Shavelson,** *Stanford Univ.*
**David Sibley**, *Washington Univ.*
**Joseph Silk**, *Univ. of Oxford*
**Montgomery Slatkin,** *Univ. of California, Berkeley*
**Davor Solter,** *Inst. of Medical Biology, Singapore*
**Joan Steitz,** *Yale Univ.*
**Elsbeth Stern,** *ETH Zürich*
**Jerome Strauss,** *Virginia Commonwealth Univ.*
**Jurg Tschopp,** *Univ. of Lausanne*
**Derek van der Kooy,** *Univ. of Toronto*
**Bert Vogelstein,** *Johns Hopkins Univ.*
**Ulrich H. von Andrian,** *Harvard Medical School*
**Bruce D. Walker,** *Harvard Medical School*
**Christopher A. Walsh,** *Harvard Medical School*
**Graham Warren,** *Yale Univ. School of Med.*
**Colin Watts,** *Univ. of Dundee*
**Detlef Weigel,** *Max Planck Inst., Tübingen*
**Jonathan Weissman,** *Univ. of California, San Francisco*
**Sue Wessler,** *Univ. of Georgia*
**Ellen D. Williams,** *Univ. of Maryland*
**Ian A. Wilson,** *The Scripps Res. Inst.*
**Jerry Workman,** *Stowers Inst. for Medical Research*
**Xiaoling Sunney Xie,** *Harvard Univ.*
**John R. Yates III,** *The Scripps Res. Inst.*
**Jan Zaanen,** *Leiden Univ.*
**Huda Zoghbi,** *Baylor College of Medicine*
**Maria Zuber,** *MIT*

## BOOK REVIEW BOARD

**John Aldrich,** *Duke Univ.*
**David Bloom,** *Harvard Univ.*
**Angela Creager,** *Princeton Univ.*
**Richard Shweder,** *Univ. of Chicago*
**Ed Wasserman,** *DuPont*
**Lewis Wolpert,** *Univ. College London*





SPACE SCIENCE

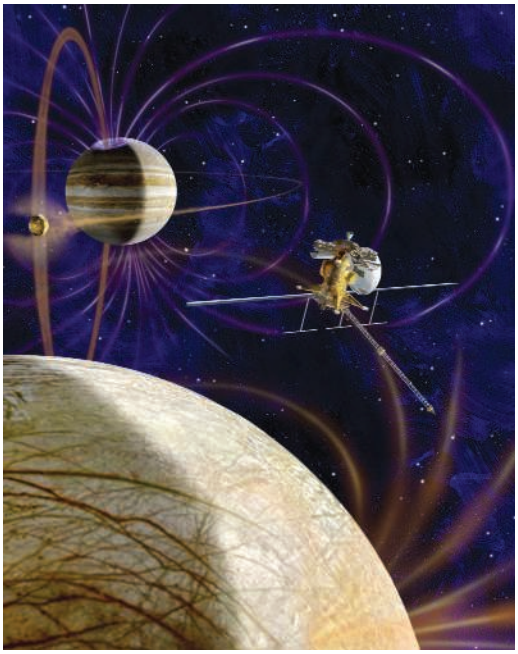
# NASA, ESA Choose King of Planets For Flagship Missions in 2020

In classical mythology, Jupiter abducted both Europa and Ganymede. But last week those jovian satellites turned the tables and seduced NASA and European Space Agency (ESA) officials into picking them, instead of Titan, an equally intriguing moon of Saturn, as the next major missions to explore the solar system beyond Mars. If all goes as planned, two separate spacecraft will head to the Jupiter system in 2020, with the NASA probe targeting Europa and the ESA probe heading for Ganymede.

The decision marks the end of a heated 2-year competition between the two systems (*Science*, 19 December 2008, p. 1780). The scientific payoff from the Galileo mission to Jupiter and the ongoing Cassini mission to Saturn made equally compelling cases for a follow-up visit, says Ronald Greeley, a planetary scientist at Arizona State University, Tempe, who co-chaired the team that defined the science for the Europa-Jupiter mission. But there's money for only one flagship planetary mission, say NASA officials, and the technical maturity of the Europa proposal clinched the deal. "The science was absolutely outstanding on both," says James Green, NASA's planetary sciences chief. "But a Europa mission has been studied for a long time."

Both the Europa and Ganymede missions still face hurdles. NASA must win approval from the Obama Administration and Congress for the estimated $3 billion cost of the Europa mission, while ESA will weigh its Ganymede mission—and its $1 billion price tag—against two astronomy projects in 2011 before it can start design work in earnest. Neither agency has room for additional scientific missions in its current budget, and the decision to launch separate probes allows either side to back out without torpedoing the entire mission. Not that they have any plans to do so, of course. ESA Director of Science David Southwood called the joint missions "a wonderful new exploration challenge and a landmark of 21st century planetary science."

Both Europa and Ganymede tantalize outer planetary scientists. The former has an icy crust concealing an ocean of water that could harbor the building blocks of life, or life itself. And Ganymede, the largest moon in the solar system, is the only satellite with an internally generated magnetic field. Its surface may also conceal an ocean far below.



**Home away?** NASA's Jupiter probe will investigate the habitability of Europa.

CREDIT: JPL/NASA

Neither probe would depend on the other for data, but officials are expecting synergistic results, notes Greeley. "Both probes could look at the magnetosphere of Jupiter at the same time from different positions," or image the volcanic activity on Io from different angles. The exact trajectory of the probes has yet to be determined, but each spacecraft would likely tour the Jupiter system before settling into orbit around its respective target.

Although initial plans for the Europa mission included a lander that would puncture the surface ice, that feature has been dropped after engineers deemed the task too difficult technically. But there will be challenges enough. One major obstacle is the intense radiation surrounding Europa. That hazard helped shoot down a Europa proposal a decade ago, but NASA's Green says there have been important advances since then in hardening sensitive instruments and spacecraft components against such radiation.

Even that technical hurdle, however, pales in comparison with what the backers of a Titan mission had proposed: a sophisticated lander and a French-built balloon to sail over the surface of the mysterious moon. Jonathan Lunine, a planetary researcher at the University of Arizona, Tucson, who worked on the Titan proposal, acknowledged its technical complexity but said "people were impressed by how far we took the study." One consolation is that Cassini, launched in 1997, could continue beaming back data from the Saturn system until 2016 or 2017.

As scientists move forward, the burning question is whether NASA and the Jet Propulsion Laboratory in Pasadena, California, which will lead the NASA portion of the effort, can keep costs under control. "We've all got [Mars Science Laboratory (MSL)] in the back of our minds," says Fran Bagenal, a planetary scientist at the University of Colorado, Boulder, who also chairs the board that advises NASA on outer planet science. MSL is well over budget and behind schedule, forcing NASA to make cuts in other planetary efforts. "We must not go down the same path," she warns.

Ed Weiler, NASA's science chief, emphasized that a mission to Titan is not off the table and that the agency has simply assigned Europa a higher priority. "The decision means a win-win situation for all parties involved," he added. Says Lunine: "Clearly, NASA wants to fly both."

Even if that's true, however, few of today's scientists are likely to be around to analyze data coming back from a Titan flagship. The sibling spacecraft won't reach Jupiter until 2026, and any mission to Saturn would presumably arrive even further into the future. "You have to think in terms of your students and your grad students," says Greeley. "This kind of work takes a lot of lead time."

**–ANDREW LAWLER**

*Published by AAAS*

Drug-resistant flu strain poses puzzle
1162

China's gender imbalance
1164

SCIENCE AND THE LAW

# Forensic Science Needs a Major Overhaul, Panel Says

In 1928, the National Academy of Sciences issued a report saying coroners' offices should be abolished in the United States as an "anachronistic institution"; it called for replacing coroners with medical examiners.

Now, 80 years later, more than half the states still have county coroners, an office for which no medical training is necessary. Last week, the National Academies' National Research Council (NRC) reiterated its message, this time in the context of a sweeping overhaul of all forensic sciences in the United States, which it says almost uniformly lack rigorous underpinnings. It is calling for top-down change, starting with a new National Institute for Forensic Science.

The 255-page report,* requested by Congress in 2005, also calls for removing crime labs from the direct control of law-enforcement agencies, mandatory accreditation for labs, certification and a code of ethics for forensic professionals, and a push for standardized procedures.

The country is "plagued by fractionated and inconsistent practices" in the way crimes and suspicious deaths are investigated, said study co-chair Harry T. Edwards, a former judge in the Washington, D.C., Court of Appeals, at a press conference last week in Washington, D.C. For example, Donald Murray of the National Association of Counties noted that there are only about 500 trained forensic pathologists in the nation's more than 3000 counties. In Minnesota, he said, "most autopsies are flown to two counties that happen to have forensic pathologists."

The committee's top priority is the establishment of an independent federal institute to put a scientific footing under myriad crime-fighting disciplines covering toxicology and analysis of fiber, hair, blood, coatings such as paint, shoeprints, tool marks, tooth marks, guns, fires, and fingerprints. The Justice Department is not the appropriate venue, says the report, and no other existing federal agency has "the capacity or appropriate mission" to do the job.

Reactions to the report have been largely positive. "I think this agency would be welcomed nearly universally," says Max Houck, director of the Forensic Science Initiative at West Virginia University, Morgantown. But how the Justice Department, which issued a terse statement saying it would evaluate the report, feels about it is not clear. Former *Science* editor-in-chief Donald Kennedy of Stanford University in Palo Alto, California, who co-chairs the NRC Committee on Science, Technology, and Law, says the Justice Department's research arm, the National Institute of Justice (NIJ), was leery of the study and declined to provide funding for it. NIJ officials did not return phone calls.

The report is hard on unnamed "members of the forensic science community [who] will not concede that there could be less than perfect accuracy either in given laboratories or in specific disciplines." It concludes: "With the exception of nuclear DNA analysis, no forensic method has been rigorously shown to have the capacity to consistently, and with a high degree of certainty, demonstrate a connection between evidence and a specific individual or source."

Fingerprints are Exhibit A. Committee member Sargur Srihari of the State University of New York, University at Buffalo, points out that DNA is available in only 5% to 10% of cases. That leaves fingerprints as the main way to nail a specific perpetrator, but recent cases have undermined the confidence of the courts. Most notable was that involving Brandon Mayfield, accused of being involved in the 2004 terrorist attacks on commuter trains in Madrid. The FBI ran digital images of prints sent from Spain through the Integrated Automated Fingerprint Identification System, which came up with several possible matches. Four experts both from within and outside the FBI narrowed the field to Mayfield. Subsequently, two FBI examiners changed their minds after going to Madrid to look at the original fingerprint images.

Srihari, a computer scientist, says he and others are working to boost the statistical reliability of fingerprint analysis with calculations of "random match probabilities." As with DNA analysis, he points out that matching a print to one in a database wouldn't be a problem if criminals left perfect prints. But authorities usually have to deal with partial prints, which, unlike DNA, have multiple variables—such as loops, whorls, ridges, and bifurcations—that are not clearly defined.

Authors of the NRC report emphasize that they do not intend to provide fodder for lawyers seeking to overturn convictions. But as lawyer Michael Saks of Arizona State University, Tempe, notes, defense attorneys could try to use the report to cast doubt on forensic evidence: "A capable lawyer will consider any and all information that might advance a client's case."

**–CONSTANCE HOLDEN**

**Loops and whorls.** Fingerprint analysis is still more art than science.

CREDITS: JEFFREY COOLIDGE/GETTY IMAGES; PHOTOS.COM

---

*Strengthening Forensic Science in the United States: A Path Forward.*

HIV/AIDS

# Tangled Patent Dispute Over 'Free' Drug-Resistance Database

A patent dispute among a small European biotechnology company, Stanford University, and one of its HIV/AIDS researchers has triggered a defamation lawsuit, raised issues of academic freedom, and led some observers to call it a battle out of *Don Quixote*.

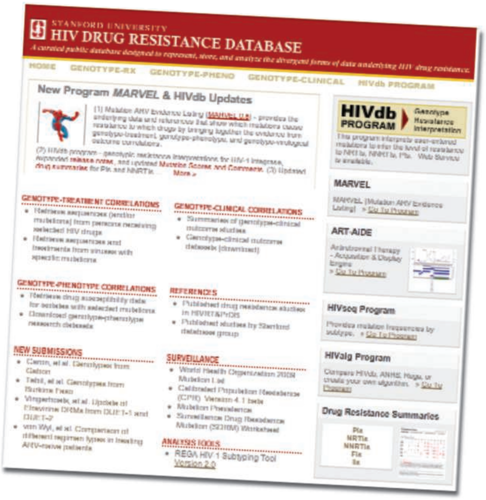
At the center of the dispute are U.S. patents owned by Advanced Biological Laboratories (ABL) S.A. in Luxembourg that involve computer methods to guide treatment of patients with HIV infection and other diseases. In early 2007, ABL notified Stanford that its popular Internet-accessible HIV Drug Resistance Database (HIVdb) possibly infringed two of the company's patents. The HIVdb helps clinicians, researchers, and drug developers around the world make sense of the complex array of mutations the virus has developed to dodge specific treatments. "It's a tremendous tool that takes full advantage of the power of the Web and marrying sequencing and clinical data," says HIV/AIDS clinician Daniel Kuritzkes of Brigham and Women's Hospital in Cambridge, Massachusetts. "I use it every time I'm in the clinic."

**Standing firm.** Stanford's Robert Shafer rejects ABL's assertion that it has patent rights over the popular database (*right*) he posts on the Web.

Robert Shafer, an infectious-disease specialist at Stanford who began building the database 10 years ago and receives support from the U.S. National Institutes of Health and unrestricted grants from industry, was outraged by ABL's claims. "It's so blatantly wrong," says Shafer. He claims the patents are overly broad and vague and notes that the European Patent Office "refused" a similar application in 2006.

Shafer says at first he saw eye to eye with Stanford's attorneys, who in October 2007 filed for "declaratory relief," essentially asking a judge to prevent litigation from moving forward. Stanford also told ABL that it might file reexamination requests with the U.S. Patent and Trademark Office (USPTO), which can correct or invalidate issued patents.

The company and the university continued to negotiate and in March 2008 reached a settlement—without Shafer's input. The "immunity agreement" states that ABL would not file patent-infringement claims against any party that used the database for noncommercial purposes. But only people affiliated with Stanford had the right to pursue commercial activities with the database; anyone else who used it "in activities for a fee or otherwise in exchange for monetary consideration" remained subject to ABL's patents. As part of the agreement, Stanford pledged to post a notice on the HIVdb that said as much.



STANFORD UNIVERSITY
HIV DRUG RESISTANCE DATABASE

HOME GENOTYPE-RX GENOTYPE-PHENO GENOTYPE-CLINICAL HIVdb PROGRAM

New Program *MARVEL* & HIVdb Updates

GENOTYPE-TREATMENT CORRELATIONS

GENOTYPE-CLINICAL CORRELATIONS

GENOTYPE-PHENOTYPE CORRELATIONS

REFERENCES

NEW SUBMISSIONS

SURVEILLANCE

ANALYSIS TOOLS

- REGA HIV 1 Subtyping Tool Version 2.0

Shafer says telling him what to post on his Web site violates academic freedom and he balked, noting that many companies use the HIVdb to develop drugs or to support their own commercial tests that help clinicians make treatment decisions. The agreement, says Shafer, "gives ABL a green light to go and sue other people. It makes Stanford complicit and makes me seem complicit."

Five months passed before the Web site posted a notice, which said, in part, "Stanford does not represent that use of this database would not infringe patent rights of other persons or entities." To ABL's dismay, this was not the agreed-upon language, and the notice made no mention of the company.

Shafer hired his own counsel and on 10 October 2008 filed reexamination requests with USPTO on the two patents. Twelve days later, he posted a notice on the Web site that mentioned the company's patent rights but noted that he had filed the reexamination petitions. Shafer also included a disclaimer that said he "considers these patents to be harmful to research on the use of computers in medicine and is concerned that the recent litigation resulting from these patents is harmful to the care of persons with HIV infection."

On 1 December, ABL sued both Stanford and Shafer for breach of contract and defamation. Chalom Sayada, a co-founder of ABL who has led the company's negotiations with Stanford, says he has "a very deep and sincere respect" for Shafer's scientific work and the HIVdb. And he stresses that ABL has attempted to solve these issues amicably.

Sayada, a geneticist who has served as CEO of several other biotechs, strongly objects to Shafer's allegation that the patents have harmed research and patients. "We are not aiming to prevent research," he says. "And we obviously are not looking to harm or prevent the care of anybody." He adds that clinicians, even though they charge their patients fees, are free to use the HIVdb without paying licensing fees to ABL. "We try to be realistic," he says.

Sayada contends that Stanford's and Shafer's actions have hurt ABL's business. "People are speaking very badly about my company," says Sayada. And he says USPTO's patent reexamination is the proper venue to resolve the debate. "Anything beyond this is mere politics, and I'm not sure who benefits."

Sayada thinks Stanford made a "big mistake" by not involving Shafer during the original agreement negotiations. Stanford counsel Patrick Dunkley sees this as a misstep, too. Although the university owns the database and "had the complete authority to enter into the agreement without consulting with ▶

CREDIT: COURTESY OF ROBERT W. SHAFER

Professor Shafer," says Dunkley, "in hindsight, it probably would have been more prudent to have involved him."

Shafer, who says he has spent more than $100,000 of his own money pursuing the patent reexaminations and defending himself against ABL's suit, contends that Stanford set a bad precedent by caving in to ABL's threats. "The university didn't back down at all," retorts Dunkley. "What the university did was protect the rights of the research community, and upon achieving its objective, there was no ongoing fight to have."

Shafer has peppered many colleagues with e-mails about the dispute and about restrictive patents in general. This has alienated some colleagues who see his impassioned battle against allegedly harmful patents as tilting at windmills. But he has also attracted some strong support. "As a faculty member who does work precisely in this area, I was surprised and concerned that the university took the position it took rather than trying to contest the patent," says Mark Musen, a Stanford researcher who specializes in biomedical informatics.

USPTO is expected to rule on the patent reexamination requests within the next few months. Stanford and Shafer have until the middle of March to respond to ABL's lawsuits. ABL's Sayada has attempted to engage Shafer in an informal mediation with leading HIV/AIDS researchers, but Shafer vows to fight to the end. "I'm not doing it out of spite," says Shafer. "It's a way of a vindication to say, 'Look, you guys were scared, but I fought this and was able to win this.' I want to show Stanford should have taken a stand."

**–JON COHEN**

UNIVERSITY RESEARCH

# Harvard Puts Science Campus on Slow Track

Harvard University hit the brakes last week as it was getting ready to build one of the country's largest new academic science centers—flagging another possible casualty of the economic downturn. The delay, described in an 18 February public letter by university president Drew Gilpin Faust, will affect plans for a massive research facility (50,000 square meters) in Allston, Massachusetts, across the river from Cambridge (*Science*, 11 July 2008, p. 190). Conceived as a science showcase and a way to bring together cutting-edge work in biology, medicine, and engineering, the project could be put on an indefinite "pause," Faust said. The university will build only the foundation in an existing hole in the ground; sometime later this year, it will decide what to do next.

**Distant vision.** Harvard's plans to build a massive science center on its campus in Allston, Massachusetts, could be delayed.

CREDIT: BEHNISCH ARCHITEKTEN

The setback shows that even huge endowments have their limits, says Harvard Provost Steven Hyman, a champion of the Allston campus. Harvard is planning for a 30% loss in its endowment for the fiscal year ending on 30 June (down from $36.9 billion last June). It's also expecting fewer and smaller gifts.

Similar but less focused cutbacks are affecting other academic institutions. Many private universities, anticipating endowment losses of 25% or more this year, are putting plans on hold. Stanford University in Palo Alto, California, for example, suspended $1.3 billion in capital construction early this year. Meanwhile, both Harvard and Stanford are raising tuition fees by 3.5% while increasing scholarship aid. State-funded schools are being hit hard by mandatory budget cuts. Last week, the University of California (UC) system was ordered to cut an extra $50 million from a budget that had already been cut by $65 million. Robert Price, UC Berkeley's interim vice chancellor for research, says this will almost certainly increase strains on the university's research infrastructure, from grants administration to the maintenance of animal facilities.

Because Harvard will not buy materials for the new research facility in Allston until it has finished a cost review and weighed options, the project almost certainly won't be completed by the target date of 2011. And delay can increase costs, too. One option being considered is to redesign the building for other uses, Faust and Hyman have said.

Some scientists who were planning to move into the Allston building—notably stem cell researcher Douglas Melton, co-chair with David Scadden of a new department of stem cell and regenerative biology—will be shifted to alternative digs. "Before we made this announcement, we made sure we could situate the [stem cell] department … in good, high-quality contiguous space," Hyman explains. Scouring available sites in Cambridge and Boston, Harvard planners came up with a solution that provides room to expand but doesn't include proximate labs for all the researchers it wanted to have in Allston. "We found approximately 75,000 square feet [7000 m$^2$] of space" for the department of stem cell and regenerative biology, Hyman explains, in an existing structure on Harvard's Cambridge campus, likely in the Sherman Fairchild Biochemistry Building. The site needs some renovation. Ironically, Hyman notes, because remodeling is quicker than construction, the researchers in this group will "coalesce in Cambridge a year earlier than they would have in Allston."

Other labs, however, will need to relocate out of the Fairchild building to make room for this large group. The emigrants have not yet been identified. Recognizing that these labs' work will be disrupted, Hyman says, the university plans to give them extra infrastructure support. Meanwhile, the new Wyss Institute for Biologically Inspired Engineering, sparked by a pledge last October of $125 million from philanthropist Hansjörg Wyss, will now be based in Harvard's medical area in Boston rather than in Allston as planned. Consolidation has been eased somewhat, Hyman says, by a grim side effect of the recession: The university has not been recruiting staff from outside.

**–ELIOT MARSHALL**

With reporting by Greg Miller.

**Grabbing an opportunity.** DOE is supporting a project to capture some of the carbon emissions from this North Dakota power plant.

ENERGY RESEARCH

# Stimulus Gives DOE Billions for Carbon-Capture Projects

Research on cutting $CO_2$ emissions from the world's biggest greenhouse polluters is getting an unprecedented boost from the U.S. government's $787 billion stimulus package. Much of the $3.4 billion designated for fossil fuel R&D—about five times what the Department of Energy (DOE) now spends annually on such research—will finance industrial-scale attempts to capture $CO_2$ from coal-burning power plants or oil refineries and lock it away deep underground.

Great hopes are riding on this technology. A major study from the Massachusetts Institute of Technology (MIT) in 2007 concluded that it was "the critical enabling technology" that would allow coal-rich nations such as the United States and China to reduce $CO_2$ emissions while still burning their most abundant fossil fuel. But it's also expensive: Several experts estimate that it would currently cost about $60 to capture and store each ton of $CO_2$. So it's unlikely that power companies will adopt it on a large scale until regulations—such as a carbon tax—make $CO_2$ financially worth capturing.

Technology already exists to capture $CO_2$ from the exhaust of power plants, transport it in pipelines, and inject it deep into Earth. M. Granger Morgan, head of the Engineering and Public Policy Department at Carnegie Mellon University (CMU) in Pittsburgh, Pennsylvania, says it's time to show that these processes can work together on a large scale. "We need to build an integrated system," he says.

The stimulus package may make that possible. It includes three large pots of money that DOE could use: $1 billion for cleaner fossil fuels, $800 million for a "clean coal" initiative to capture $CO_2$ from commercial power plants, and $1.5 billion to capture $CO_2$ or improve energy efficiency at industrial sites such as factories and refineries.

DOE officials are drafting plans to spend their windfall. "It sounds like a lot of money, but it doesn't go that far," said Secretary of Energy Steven Chu last week at a press conference organized by the energy company Platts. It's a recognition born of painful experience. Last year, the department canceled plans for a carbon-trapping power plant called FutureGen in Mattoon, Illinois, after the plant's price tag rose to $1.8 billion.

Even with the stimulus money, Chu hinted that FutureGen remains too expensive. "Right now, quite frankly, it is not a slam dunk which technology is the right one." Chu said he would prefer funding a "portfolio" of projects.

Howard Herzog, a specialist on carbon capture and sequestration at MIT, likes the idea of multiple demonstration projects because different kinds of power plants require completely different technologies. It's easier to capture $CO_2$ from power plants that convert coal into a gas before burning it, and Herzog thinks DOE should fund at least one project that includes coal gasification. But Morgan says demonstrating technology for capturing $CO_2$ from a conventional coal-burning plant is equally important: "There are an awful lot of existing plants around, some of which are relatively new, and you don't want to [discard] them." Some proponents of carbon capture are pushing for projects that will capture 90% of emissions, as FutureGen was designed to do. But Herzog says that's not necessary or even desirable, because such a high standard drives up the cost. "I'd rather see two demonstrations with 50% capture than one demonstration with 90% capture," he says. Capturing half the $CO_2$ would make coal power about as clean as electricity from natural gas.

DOE is also likely to use some stimulus money to expand its current research on $CO_2$ sequestration. It already funds seven regional carbon-sequestration partnerships that have been injecting $CO_2$ deep into Earth at various geologic sites and studying whether it stays there. Some researchers, such as John Kitchin, a chemical engineer at CMU, fault DOE for making such progress with sequestration while neglecting efforts to assess different capture technologies. "That's critically missing," says Kitchin. "We have to solve both of these problems at the same time—capture and injection."

Several of the regional partnerships are nevertheless starting ambitious collaborations with commercial power plants. The biggest so far would capture a million tons of $CO_2$ annually from the Antelope Valley Station, near Beulah, North Dakota. That amounts to about one-quarter of the $CO_2$ output from one of the plant's two 450-megawatt generating units. The project could break ground as early as this year and start operation in 2012.

The $CO_2$ will move through pipelines to Saskatchewan, Canada, where oil companies will pump it into depleting oil fields to push more oil to the surface. John Harju of the Energy and Environmental Research Center at the University of North Dakota, Grand Forks, says the entire project will cost $500 million. But DOE's share is only $67 million. Most of the project will be funded privately because of the extra revenue generated by using the $CO_2$ to extract more oil.

**–DAN CHARLES**

CREDIT: LARRY LEE PHOTOGRAPHY/CORBIS

Published by AAAS

EVOLUTION

# Creationist Beliefs Persist in Europe

**DORTMUND, GERMANY**—News coverage of the creationism-versus-evolution debate tends to focus on the United States, where surveys consistently show that less than half of Americans accept the theory of evolution. But in the past 5 years, political clashes over the issue have also occurred in countries all across Europe. In Italy, Silvio Berlusconi's government briefly tried to halt the teaching of evolution in schools in 2004. In 2006, a deputy Polish education minister called the theory of evolution "a lie." In 2007, the education minister of a major German state courted controversy by advocating that creationism and evolution be taught together in biology classes.

Indeed, creationism remains alive and well in Europe, according to researchers gathered here last week for a conference on the state of European science education and European teachers' attitudes toward and knowledge of evolution. European educators haven't yet conducted regionwide surveys, but snapshots from Germany, Turkey, and the United Kingdom presented at the meeting showed that creationist notions are more prevalent than researchers had expected, and that the understanding of evolution among teachers and teaching students—including biology teachers—is often problematic. "This isn't just an American problem," says Dittmar Graf of the Technical University of Dortmund, who organized the meeting.

**Fertile ground.** Turkish creationist Harun Yahya's antievolution tome went down well in parts of Europe.

CREDIT: PATRICK ROBERT/CORBIS

Even the birthplace of Charles Darwin is struggling with evolution, despite the myriad celebrations for the 150th anniversary of his *On the Origin of Species*. "Creationism is on the rise in the U.K.," says James Williams, a lecturer in science education at the University of Sussex. "Creationists have adopted the attitude that if you get to children young and early, you can indoctrinate them before they even start talking about evolution in schools." Williams cited a December 2008 Ipsos Mori poll of 923 primary and secondary schoolteachers in England and Wales: 37% of the respondents agreed that creationism should be taught in schools alongside evolution. Even among biology and science teachers, the number was 30%.

Within Europe, unity on the importance of teaching evolution has been hard to come by. In 2007, the Council of Europe—a body composed of 47 European countries that promotes the region's integration—had an unexpectedly close vote on a statement saying that it "firmly oppos[es] the teaching of creationism as a scientific discipline on an equal footing with the theory of evolution and in general the presentation of creationist ideas in any discipline other than religion."

The measure eventually passed but only after fierce opposition from some delegates and outside groups, says Luxembourg's Anne Brasseur, the rapporteur of the council's Committee for Culture, Science and Education. She says the council was lobbied intensively by the Vatican, which some believe has moved away from acceptance of evolution since the beginning of Benedict XVI's papacy.

In Western Europe, the reasons for creationism's tenacity lie as much in the classroom as in the churches, researchers suggested at the meeting. Graf reported a survey of 1228 German students planning to become teachers that evaluated their knowledge of and attitudes toward evolution. The results of the 108-question study—part survey, part quiz—revealed surprising gaps. Twenty percent of those studying to teach biology, for example, thought that evolution could be explained in part by Lamarckism, or the idea that traits acquired during a parent organism's life can be passed on to his or her offspring. And less than a third were able to answer basic questions about the role of reproductive fitness in evolution.

The survey also probed the students' belief in creationism, and Graf reported that the most likely predictor of creationist thinking wasn't religious belief but a lack of confidence in science, followed closely by a poor understanding of scientific principles. "What surprised me wasn't that religion correlated with antievolutionist thinking but that the correlation between a failure to understand science and not believing in evolution was very strong," Graf says.

Compared with Germany, where about 20% of the public holds creationist beliefs, creationism is even more deeply entrenched in the Muslim world, according to a study of teachers reported by Universite Lyon 1 researcher Pierre Clement. Well over half of biology teachers in Senegal, Lebanon, Morocco, Tunisia, and Algeria agreed with the statement: "It is certain that God created life." And Turkey has also seen a recent rise of an Islamic creationist movement (*Science*, 16 February 2007, p. 925).

In Turkey, a country pressing hard for membership into the European Union, schools are supposed to be secular institutions; teaching creationism is even officially banned at the university level. But many university students training to become elementary- and secondary-level teachers haven't gotten the message, apparently. Graf collaborated with Haluk Soran of Hacettepe University in Ankara, Turkey, to give those students the same survey given in Dortmund. More than 75% rejected the theory of evolution. "There's a minimal understanding of evolution in Turkey," Soran says. "The more religious people are, the more they forget about evolution."

What's the solution? Graf offered a prescription for Germany, one subsequently echoed by many of his fellow Europeans at the meeting. Science teachers in Germany have to compete with years of religious education, which is part of public school curriculum from first grade. Evolution, on the other hand, is generally first taught late in high school and not well integrated into the overall biology curriculum. "Many students who get evolutionary theory in science classes come as creationists," Graf says. "We need to be teaching evolution earlier and teaching the nature of science more intensively." **–ANDREW CURRY**

Andrew Curry is a freelance writer based in Berlin.

IMMUNOLOGY

# Flu Antibodies Stir New Hope for Treatment, Vaccine

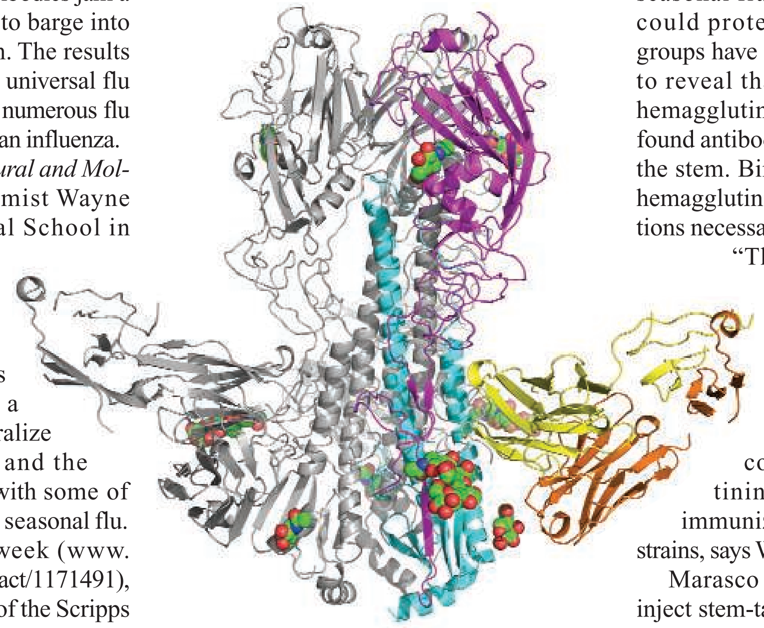
If you had the flu this year, that won't protect you from getting it again next year or the year after that. Thanks to its rapid evolution, the influenza virus usually manages to remain one step ahead of the immune system. But this week, two independent teams report that they've nabbed antibodies that disable multiple varieties of the flu virus. The antibodies jam a molecule that allows the virus to barge into cells, a crucial step in infection. The results might help researchers devise a universal flu vaccine or treatments that quash numerous flu strains, including the dreaded avian influenza.

On Sunday, in *Nature Structural and Molecular Biology*, immunochemist Wayne Marasco of Harvard Medical School in Boston and colleagues reported pinpointing 10 human antibodies that latch onto a section of the viral surface protein hemagglutinin that rarely changes between influenza strains. As a result, the antibodies can neutralize the viruses behind bird flu and the 1918–19 flu pandemic, along with some of the common strains that cause seasonal flu. And online in *Science* this week (www.sciencemag.org/cgi/content/abstract/1171491), structural biologist Ian Wilson of the Scripps Research Institute in San Diego, California, and colleagues described an antibody that apparently recognizes the same portion of the protein.

"They are both wonderful papers," says virologist Peter Palese of Mount Sinai School of Medicine in New York City. "They are reinforcing each other." Because the flu virus is so changeable, researchers thought a vaccine that protected against diverse strains was probably impossible, says viral immunologist James Crowe of Vanderbilt University Medical Center in Nashville, Tennessee. But these papers "raise the possibility of a vaccine that could induce broad cross-protective immunity." However, Crowe and other flu experts caution that producing such a vaccine will be tricky, and use of the antibodies themselves as a therapy may be prohibitively costly.

Still, the need for better flu treatments or preventive measures is dire. Seasonal flu kills more than 250,000 people every year worldwide. And many fear a pandemic like the one that swept the globe in 1918–19, when an estimated 50 million perished. Particularly worrying is the possibility that the avian influenza strain H5N1 could evolve to spread readily from person to person. Although H5N1 has infected only about 400 people, it's killed more than 60% of them.

The ability of the flu virus to change so fast explains why modern flu-fighting drugs such as Tamiflu are already losing power (*Science*, 26 January 2007, p. 448). And current flu vaccines, which are rejiggered every year to keep pace with viral adaptation, sometimes don't include the right mixture of strains to provide full protection.



**Catching the flu virus.** An antibody (yellow and orange bands) homes in on the constant region of the viral protein hemagglutinin.

Seeking more universal countermeasures, the two research teams set their sights on hemagglutinin, which sports a bulky head that fastens onto the surface of its target cell. Current flu vaccines mainly provoke antibodies that target the head. But that part of the protein changes rapidly, which often means those antibodies are useless against the next year's strain.

Now, using similar techniques, the two groups have revealed a section on the tail, or stem, of hemagglutinin that rarely changes among flu viruses. To find antibodies that latch onto the H5 version of hemagglutinin—the variant carried by the avian influenza virus—Marasco and colleagues trawled a gigantic library of more than 27 billion human antibodies that they had created. They identified 10 antibodies that recognized H5 and, in solution, could block it and eight of the 15 other hemagglutinin types, including the version used by the 1918–19 pandemic strain. The team tested three of these antibodies in mice dosed with lethal quantities of bird flu. Whether they received the antibodies before infection or up to 3 days afterward, most of the rodents survived, suggesting that the antibodies are therapeutic and preventive.

Some of Wilson's co-authors had last year reported isolating a single type of antibody from a person who'd been vaccinated against seasonal flu, and they too showed that it could protect mice from bird flu. Both groups have now used x-ray crystallography to reveal that instead of glomming onto hemagglutinin's variable head, these newfound antibodies nestle into three pockets on the stem. Binding in this position prevents hemagglutinin from performing the contortions necessary for the virus to enter a cell.

"These results could have a profound impact," says structural biologist Robert Liddington of the Burnham Institute in San Diego, who collaborated with Marasco. For example, vaccine makers might now target the constant section of hemagglutinin—and thus perhaps produce immunizations against a variety of flu strains, says Wilson.

Marasco envisions that physicians will inject stem-targeting antibodies directly into patients, producing short-term "passive immunity." Such an approach might protect people who often don't benefit from vaccines, such as the elderly and cancer patients. Stockpiles of antibodies might also allow health officials to quell a dangerous flu outbreak before it becomes a pandemic. Marasco's group is moving on to animal safety tests and hopes to begin clinical trials of antibody treatments by the 2011–12 flu season.

Developing a vaccine that elicits those antibodies could take longer. The constant section of hemagglutinin's stem typically doesn't provoke a strong antibody response, perhaps because it is fairly well hidden from immune surveillance. Vaccine makers, however, could find ways of presenting the constant region to boost an immune response.

The primary obstacle for therapeutic use of antibodies could be cost, if the large doses necessary to protect mice are also needed in people. But stockpiling such antibodies may be a wise investment against another flu pandemic. "If you know that 60% of the world is going to die, I don't think cost is an issue," says Wilson.

**–MITCH LESLIE**

CREDIT: D. C. EKIERT *ET AL.*, *SCIENCE*

SCIENTIFIC WORK FORCE

# NSF Restores Data on Minority Ph.D.s

It's not often a federal research agency does a U-turn after getting complaints from its constituents. But last week, the National Science Foundation (NSF) did—to the surprise and pleasure of groups working to increase the number of minorities in science and engineering.

Each year, NSF conducts a Survey of Earned Doctorates, asking newly minted Ph.D.s to provide a wealth of information on their educational history and career plans. The results can be broken down by field and by race, ethnicity, and gender (REG). In 2007, citing new federal privacy rules, NSF's statistical branch decided to suppress a considerable amount of information about underrepresented minorities (in particular, African-Americans, Hispanics, and Native Americans). Staffers feared that savvy data miners could make use of the small numbers reported in some subcategories—one Hispanic received a doctoral degree in astronomy in 2006, for example—to identify individuals.

Under its revised rule, any subcategory with fewer than six degree recipients went unreported. In practice, however, many more categories were also blanked out because NSF was concerned that the missing numbers could be calculated by a process of elimination. NSF also banned the use of zero, arguing that even a null set conveyed information—the absence of minorities in that category—that potentially compromised NSF's promise of anonymity to participants.

The news took some months to trickle down to researchers, institutions, and professional societies that use the data, including organizations running projects funded by NSF aimed at fostering broader participation in science and engineering. By last spring, however, they were bewildered and outraged. There were even rumors that sinister motives were at work. "Without evidence of underrepresentation, some people might wonder whether such programs are needed," notes Shirley McBay, president of the Quality Education for Minorities (QEM) Network.

Taken aback by the vociferous criticism, NSF asked QEM to hold a series of meetings that gave the community a chance to vent its anger and to suggest alternatives. (NSF had offered three options, which were universally panned.) Last week, McBay reported the results of those meetings to NSF's Committee on Equal Opportunities in Science and Engineering, an advisory body for issues affecting underrepresented minorities. Committee members also voiced their unhappiness with the changes and questioned why they were necessary.

That's when NSF announced it had had a change of heart. Lynda Carlson, head of NSF's Science Resources Statistics (SRS) division, followed McBay to the podium and shocked her audience by declaring that NSF was rescinding almost all of the new policy. From now on, only fields that award fewer than 25 total doctorates each year will be subject to any data suppression. That is likely to affect about 4% of the 280 subfields reported in the survey, estimates SRS's Mary Frase. (In those instances, subfields will be combined until the minimum is reached.) In all other cases, results will be reported by race, ethnicity, and gender, even if the result is zero in some categories.

"We listened," Carlson explained after the meeting. "We didn't realize the extent to which people are using the REG tables. We can't do everything the community wanted. But we've tried to meet as many of their needs as possible." McBay says she's pleased that SRS "heard the concerns expressed … and has reconsidered its approach."

**–JEFFREY MERVIS**

**A select group.** An NSF survey will reinstate data on the small number of minorities earning doctoral degrees.

CREDIT: CEASAR/HOWARD UNIVERSITY

*Science*Insider

## From the *Science* Policy Blog

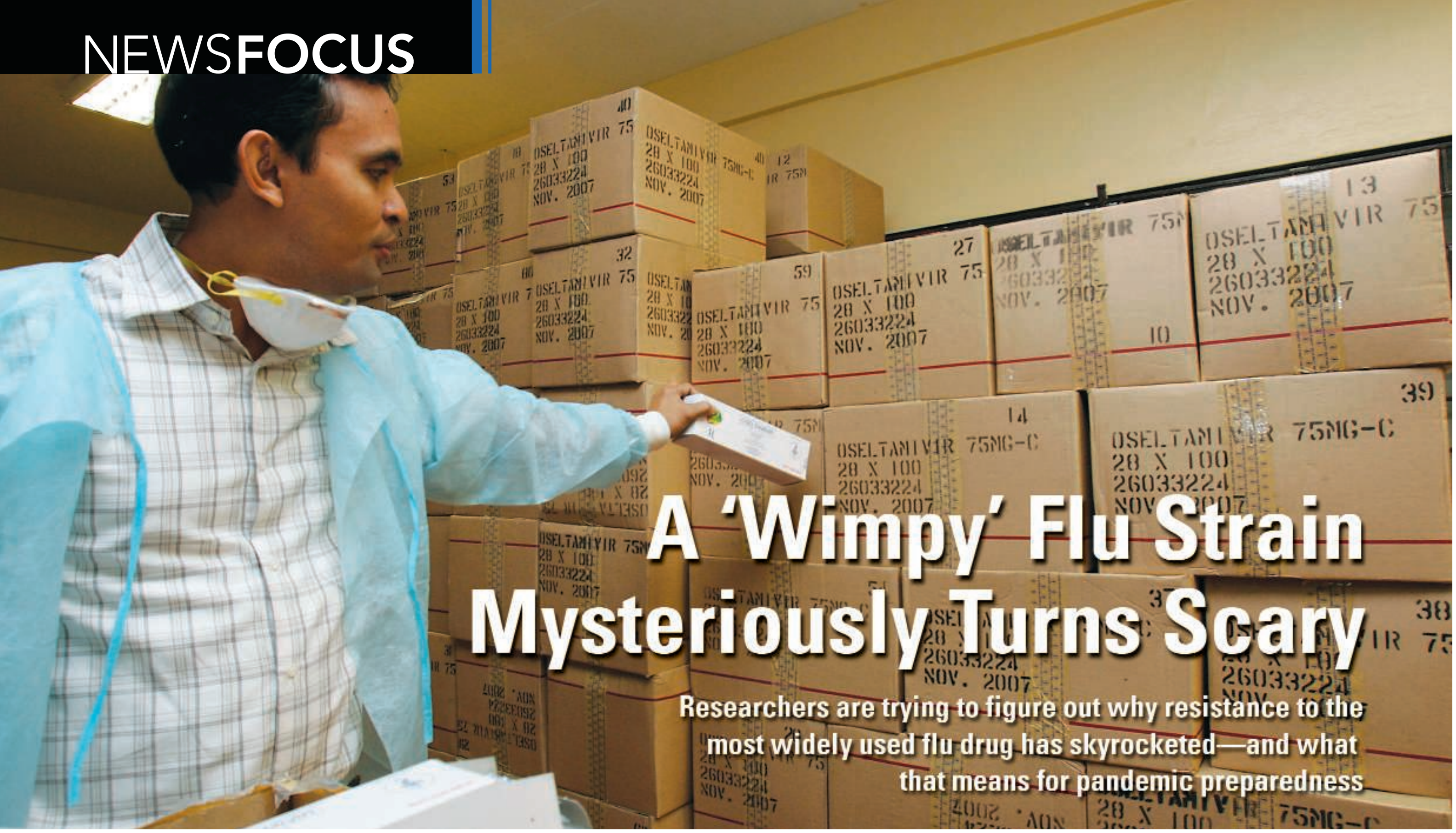
This past week has been a mixed bag for science and its proponents. There have been some big winners, some big losers, and a bunch of folks who don't know where they stand. Here's a roundup from *Science*'s policy blog, *Science*Insider:

As a divided Congress has shown, you can't pass a nearly trillion dollar stimulus bill and keep everyone happy. The same goes for the **2009 U.S. budget**, which is stingy when it comes to the National Institutes of Health (NIH), but kind to the National Science Foundation and positively munificent to the Department of Energy. The hefty stimulus itself has also proved divisive. Senator Arlen Specter (R–PA) is being hailed as a hero for his successful effort to keep $10 billion of the stimulus money for NIH. Meanwhile, other congressional supporters of science, such as Representative Vern Ehlers (R–MI), voted against the bill, claiming it was not good public policy.

In **non-U.S. news**, environmentalists are praising a proposed treaty to reduce mercury pollution. Among the 140 countries on board is India, though its science minister says the Asian nation has a lot of catching up to do when it comes to investing in science. Also catching heat is Japan, which the head of a green nongovernmental organization criticizes for not showing leadership on environmental issues.

And finally, a tale of **two viruses**. International health officials last week sought to reassure antsy staff of foreign embassies in Beijing that the recent spate of fatalities in China from the H5N1 strain of avian influenza is no cause for alarm. But they also noted that H5N1 remains as deadly and unpredictable as ever. And a different type of virus—this one of the computer variety—plagued, of all things, the International Meeting on Emerging Diseases and Surveillance. The good news: In contrast to many of the pathogens discussed at the conference, antiviral treatment is available for this one.

For the full postings and more, go to **blogs.sciencemag.org/scienceinsider**.



# A 'Wimpy' Flu Strain Mysteriously Turns Scary

Researchers are trying to figure out why resistance to the most widely used flu drug has skyrocketed—and what that means for pandemic preparedness

Everything is constantly changing in the world of influenza, but the latest development has taken even veteran researchers by surprise. A strain of flu resistant to the most widely used antiviral drug—oseltamivir, more commonly known as Tamiflu—has raced from country to country and has now become established around the globe. As winter took hold in Europe at the end of 2007, it began to rise to prominence, especially in Norway; it then jumped to the Southern Hemisphere, only to come back and conquer Asia and the United States this winter.

That wasn't supposed to happen. Based on lab studies and clinical experience, oseltamivir resistance had seemed a minor, manageable problem. Researchers are well-acquainted with the resistance-conferring mutation, called H274Y, that has now circled the planet; they have studied it for more than a decade. And almost everything they learned indicated that the mutation also made the virus less viable, or "wimpy," as one researcher puts it. The extraordinary spread of the drug-resistant strain has turned that wisdom upside down.

**Online**
**sciencemag.org**
Podcast interview with author Martin Enserink.

From a public health point of view, the spectacular rise of the strain is not as threatening as it sounds. In most countries, oseltamivir is used for only the sickest of patients or not at all; Japan and the United States are the drug's main users. Besides, the mutation affects only the influenza A virus H1N1, one of three human flu subtypes that circulate annually, and one that happens to be relatively mild. The other two, H3N2 and influenza B, are still susceptible. And other drugs, such as oseltamivir's lookalike zanamivir, still work against H1N1. That's why this year's flu season in the United States could even end up killing fewer people than last year's, when the more aggressive H3N2 was the most active subtype.

Scientifically, however, the spread of the strain with the H274Y mutation is "extremely interesting," says Keiji Fukuda, an influenza specialist at the World Health Organization (WHO) in Geneva, Switzerland. Researchers don't understand why a mutation that seemed to cripple the virus in lab studies now seems to have no effect at all on its viability in the real world. "A natural experiment comes along and gives you a very different answer," says John Skehel, an influenza scientist at the MRC National Institute for Medical Research in London.

The findings have also renewed the debate on how countries should prepare for an influenza pandemic. After the influenza subtype H5N1 started its romp through Asia, Africa, and Europe in 2004, infecting millions of birds and some 400 people so far, pandemic worries led many countries to buy massive amounts of oseltamivir. The specter that this particular mutation or another could take those precious supplies out of commission in a matter of months is troubling, says Angus Nicoll of the European Centre for Disease Prevention and Control (ECDC) in Stockholm. "This has changed the landscape," he says. "One now has to now imagine that you could have a pandemic strain that is drug-resistant."

## Unfortunate accident

Oseltamivir, which hit the market a decade ago, blocks the action of neuraminidase, an enzyme on the viral surface that cleaves a molecular structure called sialic acid sitting on the infected host cell membrane—the first step in the release of thousands of new virus particles from the cell. It's not the first flu drug to run into resistance problems. Over the past 4 decades, researchers have occasionally seen resistance against two older flu drugs, the adamantanes, pop up. Between 2002 and 2005, resistance to these drugs exploded in H3N2, a development some researchers have blamed on their excessive use in some countries (*Science*, 23 September 2005, p. 1976).

Flu scientists knew resistance could arise in oseltamivir as well. Studies had identified at least four mutations that seemed to lead to resistance, and H274Y, the one causing problems now, had been studied intensely because it occurred so frequently. Most scientists found the results quite reassuring.

In H274Y, one of neuraminidase's histidine amino acids is replaced by a tyrosine—the number 274 indicates the exact position of the amino acid in the protein. A study published in *Nature* last year showed that the replacement

CREDIT: BULLIT MARQUEZ/AP PHOTO

**Taking stock.** Resistance could render oseltamivir stockpiles, such as this one in the Philippines, useless during a flu pandemic.

tyrosine, slightly bulkier than a histidine, displaces a nearby amino acid a bit, thus disrupting the spot where oseltamivir normally binds. The same study also showed that the mutation didn't affect the binding of zanamivir (see figure), which explains why that drug has not lost its effectiveness, says one of the authors, Rupert Russell of the University of St. Andrews in the United Kingdom.

An extensive study published in 2002 tried to determine how big a threat the mutation might pose. The outcomes were hopeful: The oseltamivir-resistant virus didn't grow well in kidney cells or in mice; in ferrets, it caused less severe flu symptoms than a nonresistant, or "susceptible," strain. The drug-resistant strain was "severely compromised," the paper concluded. "Virus carrying a H274Y mutation," it went on to say, "is unlikely to be of clinical consequence in man."

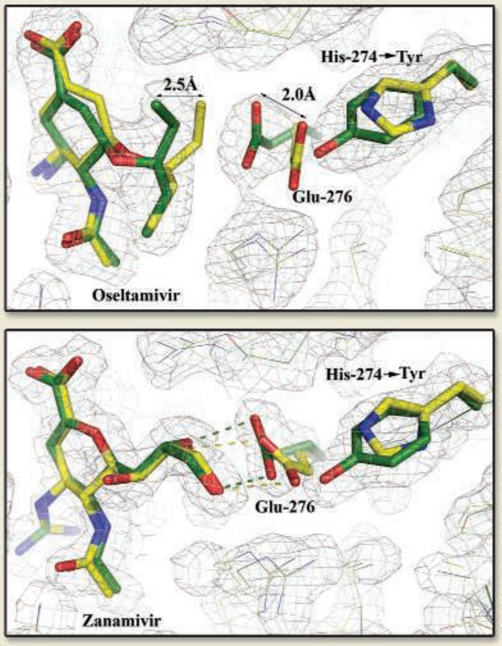
Other lab studies published later seemed to confirm that idea, as did clinical experience. In a 2006 paper, for instance, a global research network reported that during the first 3 years of the drug's use, resistant strains were found in only 0.33% of patients. Another study showed that resistance seemed most common among children in Japan, where as much as 5% of the population uses oseltamivir during a typical influenza season. But resistant viruses just didn't seem to spread.

Still, there were some warning signs, says Anne Moscona of Weill Cornell Medical College in New York City, who studies respiratory viruses and who wrote several opinion papers about flu for *The New England Journal of Medicine*. A 2004 study, for instance, found that H1N1 carrying the H274Y mutation was transmitted between ferrets—not as efficiently as a virus without the mutation, but still, it should have made flu researchers less complacent, says Moscona.

Why the mutant has suddenly become so adept at propagating itself is still unclear. The distribution of the resistant H1N1 virus shows that the sudden rise was not driven by oseltamivir use; indeed, "it's one of the ironies" that resistance in Japan and the United States has lagged, says Frederick Hayden of the University of Virginia, Charlottesville. More likely, another mutation, perhaps in the virus's other surface protein, hemagglutinin, compensated for the loss of fitness, says Peter Palese of the Mount Sinai School of Medicine in New York City. But if such compensatory mutations exist, researchers have yet to identify them.

Although far from wimpy, there's no evidence that the resistant strain causes more virulent disease than its wild-type counterpart, says Fukuda. In a paper in this month's issue of *Emerging Infectious Diseases*, a Norwegian research team reported that among 183 patients infected with the resistant strain last year, flu symptoms were no more severe and hospitalization rates no higher than among people infected with a susceptible strain.

Still, H1N1's resistance to oseltamivir does complicate treatment in countries that use the drug. Oseltamivir's advantage has been that it worked against all three seasonal flu subtypes. Now, doctors need to know first the subtype they're dealing with before they decide which drug to use. Yet the current flu drugs need to be given within 48 hours of the onset of symptoms to be effective, and often there's no time for lab tests that discern between subtypes.



**Out of shape.** The H274Y mutation causes structural changes to the viral protein neuraminidase, making it impossible for oseltamivir to bind that target (*top*). But the mutation does not affect a similar drug, zanamivir.

### What's in your country's stockpile?

As to the pandemic threat, it's not clear what the rise of oseltamivir resistance in H1N1 means. Over the past 5 years, most of the pandemic worries have focused on the chance that H5N1, a subtype that occurs widely in birds and sporadically infects people, could become transmissible among humans. Because no one has immunity to the virus, that could trigger a global outbreak that would be very hard to stop without effective drugs. So far, all H5N1 strains circulating in birds are oseltamivir-susceptible, but resistance caused by the H274Y mutation has occurred in a few human patients treated with the drug. And a study published in 2007 by Elena Govorkova and her colleagues at St. Jude Children's Research Hospital in Memphis, Tennessee, showed that at least some H5N1 viruses with the H274Y mutation remained just as fit as the wild-type virus.

Moreover, H5N1 is just one of many influenza subtypes that could trigger a future pandemic. There are nine known neuraminidase types, labeled N1 through N9, and oseltamivir is active against all of them. Very little research has been done on how easily resistance develops in most of these types, says Govorkova. The bottom line, says Hayden, is that influenza drug arsenals, like investment portfolios, should be diversified. Some countries have begun to do so; the United States and France, for instance, stockpile zanamivir in addition to oseltamivir, and the United Kingdom announced on 29 January that it would do the same. Further down the line, researchers hope that new antivirals or anti-influenza antibodies, such as those reported this week in both *Nature Structural and Molecular Biology* and online in *Science* (p. 1160), will help shore up defenses.

But the investment in new drugs has become a harder sell politically now that the pandemic threat has disappeared from the front pages of newspapers and the global recession has dried up available cash. Many countries are also facing the decision whether to replace their aging stockpiles of oseltamivir, which has an official shelf life of 5 years. Albert Osterhaus of Erasmus Medical Center in Rotterdam, the Netherlands, worries that the downturn may make pandemic preparation fall by the wayside, even though the risk has not decreased.

As to the oseltamivir-resistance mutation found in the H1N1 subtype, scientists say it's important to watch closely how it evolves. And perhaps it's not here to stay. New influenza strains come and go every year and usually spread from East Asia (*Science*, 18 April 2008, p. 340). The strains that win out are typically the most adept at beating human immune systems, and whether they are oseltamivir-resistant may not be the most important factor. In parts of China, an H1N1 strain is now circulating that does not have the H274Y mutation and is susceptible to oseltamivir, Hayden says. If that becomes dominant elsewhere in the world next year, oseltamivir resistance could simply subside again, he says—although nobody knows for how long. **–MARTIN ENSERINK**

CREDIT: P. J. COLLINS *ET AL*., *NATURE LETTERS*, 453 (26 JUNE 2008) © MCMILLAN PUBLISHERS LIMITED

**Generation XX.** Bao and Li received preferred land for settling in her hometown and cash for having a girl.

CREDIT: MARA HVISTENDAHL

DEMOGRAPHY

# Making Every Baby Girl Count

**China is ramping up a program that addresses the sociological factors behind an alarming trend: the country's rising sex ratio at birth**

**CHAOHU, CHINA—**When Bao Tiezhu and Li Qing married in 2002, they did everything backward. Instead of settling in his hometown, as most Chinese newlyweds would, Bao moved to Li's ancestral village and relinquished the right to inherit his family's land. When Li gave birth to two children in 5 years, they took her surname, not his. And—most importantly for the Chinese government workers who were watching the couple closely—the children are girls.

In discarding centuries of tradition, Bao and Li became the face of the new Chinese family: foot soldiers in a desperate program, called Care for Girls, to narrow China's gaping sex ratio at birth, among the highest in the world at 120 boys for every 100 girls. Care for Girls, which was developed by a team of social scientists following years of fieldwork in the Chinese countryside, has a variety of carrots and sticks at its disposal, including financial incentives, strict punishment for sex-selective abortion, and radical social restructuring.

In 1980, China, concerned about a population explosion, adopted a one-child policy—and enforced it through compulsory sterilizations and abortions. The policy clashed with the country's patrilineal tradition, in which sons carry on the ancestral line and care for elderly parents. Many families went to great lengths to ensure that their sole child was a boy. Confronted with a spike in the sex ratio and widespread resistance to coercive methods, Communist Party leaders relaxed the one-child policy in 1984. Most provinces subsequently allowed rural couples with one girl to try again, an exception sometimes termed the "1.5-child policy." But the change coincided with the introduction of ultrasound machines to rural China, which enabled couples to determine the sex of a fetus and abort females. The country's sex ratio continued to rise.

China is not the only country grappling with a skewed sex ratio. Across Asia, policymakers are scrambling to come up with ways to stem a tide of "missing girls." In India, the state of Haryana began paying parents for having girls in 1994, and last year the national government extended a similar plan to seven other states. But the powerful National Population and Family Planning Commission of China—the same body that enforces the one-child policy—may, ironically, be critical to bringing sex ratios back into biological harmony. "If [China] really focuses on it, they have the grassroots organizers who are capable of doing this," says Monica Das Gupta, a demographer at the World Bank who studies sex ratios. "They have a lot of information on the parents, and if they're going to punish people, they can."

Chaohu, in rural Anhui Province, is the pilot for Care for Girls. In 1999, local Family Planning Commission volunteers began augmenting their routine of birth monitoring and contraception distribution with the oversight of doctors operating ultrasound machines. The program also features microcredit grants to women and social security payments for parents of girls—Bao and Li received a small sum, $29, upon the birth of their first daughter—along with measures designed to change social customs.

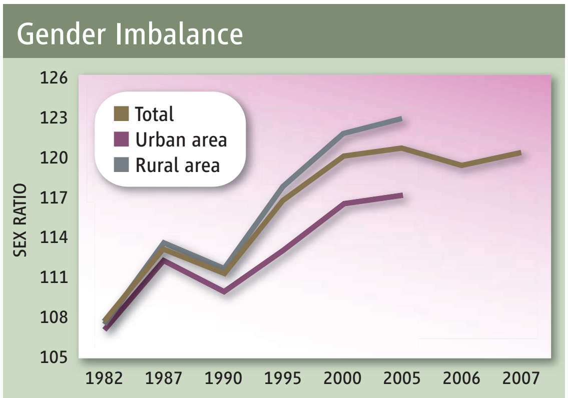
Bao and Li are one of four couples in their 600-person village to have espoused uxorilocal marriage, or living with the wife's family. Couples in some regions have opted for this lifestyle throughout Chinese history, but the practice is typically stigmatized. By rewarding daring couples with land and public praise, Care for Girls aims to remove the stigma. Bao says it worked: "People don't discriminate against you now."

## Unnatural selection

Down the street in Chaohu, a poster hanging in the entryway of Yang Chuanfeng's home

Published by AAAS

shows a smiling couple embracing a young girl next to the slogan "Let gender equality prevail." That is exactly the take-home message the family-planning volunteer and mother of two girls hopes to impart to women on house visits. "If it's a boy, have a boy," she says. "If it's a girl, have a girl. They're the same." To make her point, she hands out brochures in red bags emblazoned with "New Culture of Marriage." One pamphlet explains the concept of sex ratio at birth.



**Battle of the sexes.** The sex ratio at birth in China, now 120 boys for every 100 girls, has risen steadily.

Such materials are the result of Zhu Chuzhu's decades of research. Following China's 1982 census, Zhu, a tireless 76-year-old demographer at Xi'an Jiaotong University's Population Research Institute, noticed a spike in mortality among girls aged 1 to 4—the first sign something was wrong—and quietly began researching the causes.

Zhu didn't have funding to conduct extensive studies on child mortality until the mid-1990s, when the Ford Foundation backed fieldwork in rural Shaanxi Province. In a 2-year survey, she and Xi'an Jiaotong colleague Shuzhuo Li began investigating China's sex ratio at birth, which had jumped abruptly from a biologically stable level of 107 boys for every 100 girls in 1982 to 111 in 1990. (In the most recent census and in sample surveys, the sex ratio has continued to climb.) Son preference is ingrained in China, but Zhu and Li, a young star with postdoctoral experience at Harvard and Stanford universities, set out to see whether they could discern and combat the underlying causes.

In a study on infant mortality, the duo found that parents were more likely to avoid seeking medical help for sick girls who were second or third in birth order or only had sisters. The pattern mirrored other scholars' findings on sex ratio at birth. In countries with high sex ratios—including China, India, and, most recently, the Caucasus Mountains nations—the proportion of boys born rises significantly with birth order, presumably because mothers abort female fetuses.

The demographers realized that reversing the trend would require a major cultural shift. Undermining the patrilineal order, they suspected, might do the trick. With Marcus Feldman, an evolutionary biologist at Stanford, Zhu and Li surveyed two counties in China where a historically loose clan structure had led to a high percentage of men living with their wives' families. Both uxorilocal counties had a normal sex ratio at birth and low female child mortality. Moreover, matrilineality seemed to provide the same benefits as patrilineality: "We found that daughters provided economic and emotional support to their parents equal to that of sons," Li says.

Through fieldwork in rural China, Zhu and Li developed a framework that would make having girls economically and socially acceptable. At a national meeting in 1998, they approached the Family Planning Commission with their findings. The commission was moving away from coercive methods, repositioning itself as an arbiter of reproductive health, and Li and Zhu convinced Commissioner Zhang Weiqing to incorporate their findings into an upcoming gender-equality campaign. Still, the sex-ratio issue was sensitive because of its connection to the one-child policy—so much so that the demographers had trouble finding a pilot area for Care for Girls. In the end, Li says, they selected Chaohu simply because officials here agreed.

Some scholars question whether the Family Planning Commission, which continues to slap fines on parents who exceed the one or "1.5" child policy, is the best entity to tackle China's skewed sex ratio. Wang Feng, a demographer at the University of California, Irvine, says the issue would be better addressed by scrapping birth limits entirely. "The Family Planning Commission is … curing the disease by treating the symptoms rather than the root cause," he says. "They're evading the more difficult, and the more fundamental, part, which is the [one-child] policy." Other scientists point out that sex ratios have risen concurrently across Asia in countries that do not limit the number of children a couple can have. And proponents stress that the Family Planning Commission has an army of workers—a nationwide network of 300,000 volunteers—to help engineer rapid social change.

Li is among a group of demographers pushing for a further relaxation of the one-child policy. But he and Zhu argue that the policy exacerbates, but doesn't directly cause, China's high male-to-female sex ratio. "The essential problem is culture. And Chinese culture can be changed," says Zhu, who remembers hearing of infant girls being drowned in buckets of water while growing up in the 1930s and 1940s.

Deeper in Chaohu, in a fishing village at the end of a dirt road, 66-year-old Yang Peihua

**Demographic trendsetter.** China's rapid urbanization may help bring the skewed sex ratio closer to biological harmony, says Shuzhuo Li, because urban culture values sons and daughters more equally.

SOURCE: NATIONAL BUREAU OF STATISTICS OF CHINA; CREDIT: MARA HVISTENDAHL

**Anxious young men.** Rising numbers of restless men, like these jobless migrant workers in Nanjing, could augur hard times in China.

## CHINESE MEN: A RISING TIDE OF TROUBLEMAKERS?

Chinese officials are taking measures to encourage couples to value daughters (see main text), but they don't talk much about the flip side of the problem: men who can't find wives. The problem of guang gun, or "bare branches," as they are called, is "shrouded in secrecy," says Susan Greenhalgh, an anthropologist at the University of California (UC), Irvine.

With the current ratio of 120 boys born for every 100 girls, in the next couple of decades there will be at least 30 million Chinese men who cannot find wives, says Xi'an Jiaotong University demographer Shuzhuo Li. Men who lose out in the marriage sweepstakes, he says, tend to be at the bottom of the social ladder: rural, uneducated, marginally employed, and often transient.

A surfeit of frustrated, low-status males is bound to spell trouble for society, some experts argue. In a controversial 2004 book, *Bare Branches*, political scientists Valerie Hudson of Brigham Young University in Provo, Utah, and Andrea den Boer of the University of Kent in Canterbury, U.K., used historical examples from China and elsewhere to argue that China might soon be bedeviled by a rogue underclass of malcontented single males who could stir up political instability and even armed revolts.

Some experts dispute that forecast. Historically, up to 10% of Chinese males never marry because the sex ratio has long been skewed toward males, largely as a result of female infanticide, and rates of bachelorhood have been higher at times in some European countries as well as in parts of China, says demographer Wang Feng of UC Irvine. "So why only in China would this become such a big problem?" he says. "There is no convincing evidence that [young males] joined a peasant rebellion because they couldn't find a wife." Wang believes that the worst consequences of the sex ratio may be for individuals who suffer from loneliness and lack of social support.

Hudson counters that "we do see significantly rising crime" in China and rising social unrest (*Science*, 30 January, p. 574). But evidence that excess young males breed crime is only circumstantial, says Greenhalgh: "Nobody has done field research to see what's going on with these men."

That may soon change. Last August, Li and colleagues at Stanford University began administering questionnaires to 600 bare branches—defined as bachelors over the age of 28—as well as married men and women across China in which subjects can respond anonymously about their education, finances, physical and mental health, drinking, sex lives, and social connections. Village heads will be interviewed about the local crime situation. *Bare Branches* is a "very good book," says Li. The time has come, he says, to put its thesis to the test.

**–CONSTANCE HOLDEN**

---

represents the most intransigent part of that culture. Sitting on a bench in her mud-walled home, Yang recalls how she felt when her daughter-in-law gave birth to a second girl. "It was like I was dropped into a tub of ice cubes," she says. Under Care for Girls, Yang was required to attend a training session for elderly women, at which she says she learned, "Boy or girl, leave it to nature." But her attitude underscores the challenge facing China. "It's a very powerful cultural norm that they're addressing," says Feldman.

### Against the grain

When Bao told his family he would move to his wife's hometown, he recalls, "at first my parents were opposed. But I told them the environment in this village is more comfortable. And this kind of land is difficult to get otherwise." He gestures to the couple's courtyard home and its prime location on a busy road—a reward from the village government for their unconventional marriage.

As Bao tells his story, Shuzhuo Li looks on approvingly, satisfied that over a decade of work is paying off. The results of the trial here were promising. According to Family Planning Commission data and independent surveys, Chaohu's sex ratio dropped from 125 in 1999 to 114 in 2002. The next year, the Chinese government scaled up the program to 24 districts around the country. Results were similarly encouraging, with the sex ratio in those places declining from 134 in 2000 to 120 in 2005. In 2006, Care for Girls rolled out nationwide.

**Preaching equality.** Yang Chuanfeng emphasizes that girls are just as precious as boys.

Li and Zhu concede that the program's short-term gains may be the result of a crackdown on sex selection, not deeper cultural change. Further progress could be elusive: "The closer you get to a normal sex ratio, the more difficult it is," says Zhu. But unborn girls may get a boost from urbanization, which some scholars credit for a recent decline in South Korea's sex ratio. The flow of migrant workers to wealthy cities in eastern China raises women's earning power and introduces new ideas to the countryside. "Urban culture values sons and daughters more equally," Li says. As the sex ratio continues its upward march, China needs all the help it can get.

**–MARA HVISTENDAHL**

Mara Hvistendahl is a writer in Shanghai, China.

CREDITS: AP; MARA HVISTENDAHL

Published by AAAS

NUCLEAR PHYSICS

# Will Monju's Resurrection Give Breeders a Second Lease on Life?

**A long-shuttered fast breeder reactor could be the bridge to a new generation of such reactors, if it can overcome technical problems that have delayed its planned restart**

**TSURUGA, JAPAN—**A highlight of a tour of the Monju Fast Breeder Reactor is the opportunity to avert catastrophe. In a simulator room used to train operators, a lucky visitor throws a pair of emergency switches, setting off a cacophony of sirens and buzzers. Blinking lights and flashing numbers indicate that everything from coolant flow to the nuclear reaction in the core has come to a screeching halt.

This simulation is part of a campaign to convince Japanese officials and the wider public that the massive leak of sodium coolant that occurred in 1995—and the botched cover-up of the accident—will never happen again. For more than 13 years, Monju, a glistening white complex on a rugged promontory overlooking the Japan Sea, has sat idle, its future in doubt. Managers hoped to fire up the experimental nuclear plant this month, but a technological glitch has delayed the restart at least until later this year.


**Fast asleep.** After a 13-year pause, Japan intends to restart Monju later this year.

CREDIT: JAPAN ATOMIC ENERGY AGENCY

Proponents of the controversial fast-reactor technology—which burns plutonium, an ingredient in bombs—see resurrecting Monju as an urgent task. "Monju is a very important tool in the long-term strategy for nuclear energy," says Thierry Dujardin of the Organisation for Economic Co-Operation and Development's Nuclear Energy Agency in Paris. Fast reactors, he says, could reduce nuclear waste and cut greenhouse gas emissions. Proponents see their recycled, noncarbon fuel as a critical advantage. "It seems obvious we should have the technology ready when needed, and we cannot do that without tools like Monju," Dujardin says. With its 280-megawatt electric (MWe) capacity, Monju is a bridge between the last experimental fast reactor in the West—France's Phenix, which is about to be shut down—and a new generation of fast reactors planned or being built in China (a 20-MWe research reactor coming online in September), India (a 500-MWe prototype due for completion in 2012), and Russia (an 880-MWe demonstration reactor planned for 2012).

When Monju was conceived in the 1970s, scientists believed that fast neutron reactors were the next step in expanding nuclear power. Conventional reactors typically use water as a moderator to slow neutrons and increase the odds of energy-releasing collisions with uranium atoms in the fuel. Fast reactors do away with a moderator and use more-fissile plutonium, often combined with uranium. Configured as "breeders," fast reactors produce more plutonium than they consume. "Burners" transmute waste from conventional nuclear plants or from weapons into less radioactive materials.

The United States, the United Kingdom, Germany, France, Russia, and Japan built experimental fast reactors between the 1970s and 1990s. But they were plagued by technical problems, were extremely costly, and raised concerns about safeguarding plutonium. (The only road to Monju goes through a 900-meter-long tunnel guarded 24/7 by riot police.) The test reactors were shuttered one after another. The sole commercial fast reactor to date, France's 1200-MWe SuperPhenix, closed in 1985.

Japan never gave up on fast reactors. Monju achieved criticality in April 1994. Operations halted when a coolant loop sprang a leak in December 2005. More than 700 kilograms of molten sodium escaped through a broken temperature probe, releasing toxic fumes and damaging the plant. (No one was injured.) Engineers concluded that the flowing sodium induced oscillations in the probe, causing metal fatigue failure.

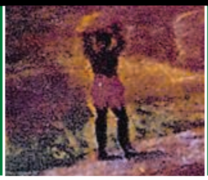
Plant managers created a PR fiasco when they misrepresented how much sodium escaped and the extent of the damage. "We lost the trust of the local people," admits Takehide Deshimaru, director of Monju development for the Japan Atomic Energy Agency (JAEA). "We learned the hard way that transparency is the most important thing." It took more than a decade to win approvals for repairs and modifications. But last month, days before the planned restart, engineers discovered rusted ventilation ducts in a plutonium-handling facility that will be used to replenish fuel. Repairs are under way.

Fast-reactor technology has not stood still during Monju's long dormancy. But JAEA officials are intent on squeezing value out of the aging complex by "verifying the plant's reliability through the experience of operation," says Deshimaru. Researchers from Japan, France, and the United States will use Monju to test new fuel combinations and transmutation. Work at Monju will also feed into the ongoing design of a 1500-MWe fast reactor. The Japanese government and utility companies are discussing who will pay for and build the new reactor after the design is completed in about 2015.

Monju's revival does not assure a bright future for breeders. Deshimaru says that because of the complexity, particularly of the cooling systems, a commercial plant based on Monju technology would cost 10 times that of a conventional light-water reactor of equivalent capacity. However, he says, recent designs with simpler cooling systems make fast reactors more cost-competitive.

Western countries are ambivalent. France intends to build a next-generation nuclear power plant by 2020. A sodium-cooled fast reactor is the leading candidate, but a final decision won't be made until 2012, says Jacques Bouchard, special adviser to France's Commissariat a L'Energie Atomique. In the United States, research on fast reactor concepts is continuing. "But in terms of moving quickly on a demonstration or prototype reactor, it's wait and see," says Mark Peters, a deputy associate lab director at Argonne National Laboratory in Illinois. In the meantime, U.S. and European scientists are lining up to visit Monju—not for simulations but to lay hands, at long last, on hard data.

**–DENNIS NORMILE**



# LETTERS

*edited by Jennifer Sills*

## Neuroscience and the Soul

SCIENCE AND RELIGION HAVE HAD A LONG RELATIONSHIP, BY TURNS COLLEGIAL AND ADVERSARIAL. In the 17th century Galileo ran afoul of the Church's geocentrism, and in the 19th century Darwin challenged the biblical account of creation. The breaches that open at such times often close again, as religions determine that the doctrine in question is not an essential part of faith. This is precisely what happened with geocentrism and, outside of certain American fundamentalist Christian sects, evolution.

A new challenge to the science-religion relationship is currently at hand. We hope that, with careful consideration by scientists and theologians, it will not become the latest front in what some have called the "culture war" between science and religion. The challenge comes from neuroscience and concerns our understanding of human nature.

Most religions endorse the idea of a soul (or spirit) that is distinct from the physical body. Yet as neuroscience advances, it increasingly seems that all aspects of a person can be explained by the functioning of a material system. This first became clear in the realms of motor control and perception (*1*, *2*). Yet, models of perceptual and motor capacities such as color vision and gait do not directly threaten the idea of the soul. You can still believe in what Gilbert Ryle called "the ghost in the machine" (*3*) and simply conclude that color vision and gait are features of the machine rather than the ghost.

However, as neuroscience begins to reveal the mechanisms underlying personality, love, morality, and spirituality, the idea of a ghost in the machine becomes strained. Brain imaging indicates that all of these traits have physical correlates in brain function. Furthermore, pharmacologic influences on these traits, as well as the effects of localized stimulation or damage, demonstrate that the brain processes in question are not mere correlates but are the physical bases of these central aspects of our personhood. If these aspects of the person are all features of the machine, why have a ghost at all?

By raising questions like this, it seems likely that neuroscience will pose a far more fundamental challenge than evolutionary biology to many religions. Predictably, then, some theologians and even neuroscientists are resisting the implications of modern cognitive and affective neuroscience. "Nonmaterialist neuroscience" has joined "intelligent design" as an alternative interpretation of scientific data (*4*). This work is counterproductive, however, in that it ignores what most scholars of the Hebrew and Christian scriptures now understand about biblical views of human nature. These views were physicalist, and body-soul dualism entered Christian thought around a century after Jesus' day (*5*, *6*).

To be sure, dualism is intuitively compelling. Yet science often requires us to reject otherwise plausible beliefs in the face of evidence to the contrary. A full understanding of why Earth orbits the Sun (as a consequence of the way the solar system was formed) took another century after Galileo's time to develop. It may take even longer to understand why certain material systems give rise to consciousness. In the meantime, just as Galileo's view of Earth in the heavens did not render our world any less precious or beautiful, neither does the physicalism of neuroscience detract from the value or meaning of human life.

**MARTHA J. FARAH[1]* AND NANCEY MURPHY[2]**

[1]Center for Cognitive Neuroscience, Department of Psychology, University of Pennsylvania, Philadelphia, PA 19104, USA. [2]School of Theology, Fuller Theological Seminary, Pasadena, CA 91182, USA.

*To whom correspondence should be addressed. E-mail: mfarah@psych.upenn.edu

### References

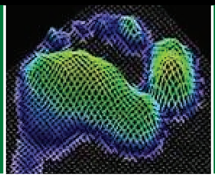
1. M. Jeannerod, *The Cognitive Neuroscience of Action* (Wiley-Blackwell, Hoboken, NJ, 1997).
2. M. J. Farah, *The Cognitive Neuroscience of Vision* (Wiley-Blackwell, Hoboken, NJ, 2000).
3. G. Ryle, *The Concept of Mind* (Univ. of Chicago Press, Chicago, 1949).
4. M. Beauregard, D. O'Leary, *The Spiritual Brain: A Neuroscientist's Case for the Existence of the Soul* (HarperCollins, New York, 2007).
5. N. Murphy, *Bodies and Souls, or Spirited Bodies?* (Cambridge Univ. Press, Cambridge, 2006).
6. J. B. Green, *Body, Soul, and Human Life* (Baker, Grand Rapids, MI, 2008).

## An Optimistic Read on Digital Libraries

IN THE NEWS STORY "NSF RETHINKS ITS DIGital library" (Special Section on Education & Technology, 2 January, p. 54), J. Mervis captures the uphill struggle digital libraries have faced as they try to sustain funding and gain visibility among educators and students. However, the picture the article paints of sustainability in general and the Digital Library for Earth System Education (DLESE) in particular is too pessimistic. The process by which the National Center for Atmospheric Research (NCAR) assumed responsibility for the DLESE collection is attracting a good deal of attention through articles, presentations, and grants addressing models of sustainability for digital efforts. The advantages of curated digital libraries are clear to many users. DLESE's collections of Earth science materials, now managed by the NCAR Library, have been selected by educators and scientists specifically because they are scientifically accurate, grade-level appropriate, and effective for teaching.

The University Corporation for Atmospheric Research, which operated the technical arm of DLESE and continues to run the NSDL Resource Center, is very proud of the impact of these digital library efforts. Although DLESE has lost its NSF funding, the open-source DLESE technical infrastructure underpins initiatives at NASA, NOAA, DOE, and scientific

CREDIT: PAUL COOKLIN/BRAND X/CORBIS



Fossilized behavior

1174

Extracting energy from the ocean

1176

data centers and digital libraries from Britain to China. It is notoriously difficult to demonstrate the impact of educational innovations, but a current project with Denver Public Schools leveraging DLESE collections holds promise for helping us evaluate the value of digital library collections in transforming teacher practice and learning outcomes. In many ways, NSF's investment in digital libraries has been transformative.

**JACK FELLOWS**

UCAR Office of Programs, University Corporation for Atmospheric Research, Boulder, CO 80307, USA. E-mail: jfellows@ucar.edu

## Developing Psychomotor Skills the Wii Way

IN HIS PERSPECTIVE "IMMERSIVE INTERFACES for engagement and learning" (Special Section on Education & Technology, 2 January, p. 66), C. Dede only briefly touched upon the particularly important area of psychomotor skill development. Hand-eye coordination and fine three-dimensional control of limbs in space are areas that are usually neglected by the educational system, despite the fact that they are crucial to a huge number of professions. There is evidence to suggest that surgical simulations can improve speed and accuracy (*1*). Moreover, the opportunity for motor skill development by immersive interfaces is not limited to specialist training—the Nintendo Wii console is one example of a commercially available games system that integrates hand and, in some cases, whole body movements in a range of simulated environments. Games systems provide opportunities to improve motor skills both for laypeople and for professional trainees in places that cannot afford full-scale simulators.

**CIARAN SCOTT HILL**

Institute of Neurology, University College London, Queen Square, London WC1N 3BG, UK. E-mail: ciaran.hill@ucl.ac.uk

**Reference**

1. J. C. Rosser *et al.*, *Arch. Surg.* **142**, 181 (2007).

## Explaining the Reasoning-Fact Gap

THE EDUCATION FORUM BY L. BAO *ET AL.* ("Learning and scientific reasoning," 30 January, p. 586) raises the question, why doesn't knowledge of scientific facts (in physics specifically) seem to correlate with skill in reasoning? I suspect that one factor is that American curricula generally emphasize a broad range of technical and nontechnical subjects, only one year of which is devoted to physics. Reasoning skills learned from these other subjects likely compensate for the dearth of physics, relative to that in Chinese schools. A second factor is probably the way in which science is taught. Science education in the United States is not focused on national college entrance exams (neither the SAT nor the ACT test factual science knowledge). Requirements and exams vary from state to state, but guidelines from the National Academy of Sciences emphasize reasoning over facts (*1*). Overall, educators have greater leeway to focus on reasoning and other less testable skills.

Teaching American students more scientific knowledge is a laudable goal, but we should be careful about what gets sacrificed in the process. It is worth considering how much value most people will derive from advanced specific knowledge (Gauss's Law, for example) compared to broader background in a variety of subjects. Science education in the United States is woefully inadequate in many respects (*2*), and the number of failing schools is embarrassing. We must correct these failures but also remember what we've done right. The broad education characteristic of both K–12 and higher education in the United States has done an excellent job fostering creativity and innovation, measured by publications, patents, and growth in the science and technology work force (*2*, *3*). We should work to improve scientific education by building on that success.

**SCOTT M. CARLSON**

Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. E-mail: scottmc@mit.edu

**References**

1. National Research Council, *National Science Education Standards* (National Academy Press, Washington, DC, 1996).
2. National Science Board, *Science and Engineering Indicators 2008* (National Science Foundation, Arlington, VA, 2008), chs. 3, 5, 7.
3. World Intellectual Property Organization, *World Patent Report—A Statistical Review* (2008); www.wipo.int/freepublications/en/patents/931/wipo_pub_931_2008.pdf.

## TECHNICAL COMMENT ABSTRACTS

### COMMENT ON "Multipartite Entanglement Among Single Spins in Diamond"

**Brendon W. Lovett and Simon C. Benjamin**

Neumann *et al.* (Reports, 6 June 2008, p. 1326) recently reported the preparation of multiparticle entanglement among single spins in diamond. However, two of the system's nuclear eigenstates are incorrectly described as product states when they are inherently entangled. Consequently, three of the six states reported, namely the odd-parity Bell states and the *W* state, were not actually produced.

Full text at www.sciencemag.org/cgi/content/full/323/5918/1169c

### RESPONSE TO COMMENT ON "Multipartite Entanglement Among Single Spins in Diamond"

**P. Neumann, N. Mizuochi, F. Rempp, P. Hemmer, H. Watanabe, S. Yamasaki, V. Jacques, T. Gaebel, F. Jelezko, J. Wrachtrup**

Our study reported entanglement among single spins in diamond. Lovett and Benjamin argue that three of six described entangled states were not achieved. Here, we explain our choice of entangled states and discuss their importance for quantum information processing. We also show that the eigenstates discussed by Lovett and Benjamin, although formally entangled and routinely generated in our experiments, cannot be used to detect nonlocal correlations.

Full text at www.sciencemag.org/cgi/content/full/323/5918/1169d

## Letters to the Editor

Letters (~300 words) discuss material published in *Science* in the previous 3 months or issues of general interest. They can be submitted through the Web (www.submit2science.org) or by regular mail (1200 New York Ave., NW, Washington, DC 20005, USA). Letters are not acknowledged upon receipt, nor are authors generally consulted before publication. Whether published in full or in part, letters are subject to editing for clarity and space.

# Comment on "Multipartite Entanglement Among Single Spins in Diamond"

**Brendon W. Lovett[1]* and Simon C. Benjamin[1,2]**

Neumann *et al*. (Reports, 6 June 2008, p. 1326) recently reported the preparation of multiparticle entanglement among single spins in diamond. However, two of the system's nuclear eigenstates are incorrectly described as product states when they are inherently entangled. Consequently, three of the six states reported, namely the odd-parity Bell states and the *W* state, were not actually produced.

Neumann *et al*. (*1*) reported the creation of quantum entanglement among the spins in the vicinity of a single nitrogen-vacancy (NV) defect in diamond. The work demonstrates the complex interplay between different degrees of freedom in this solid state system. However, we believe that certain important eigenstates of the NV system have been wrongly classified as products of single-spin states. We argue that these eigenstates are inherently entangled and, therefore, although the authors have demonstrated manipulation of these states, they did not introduce entanglement.

Neumann *et al*. made an unambiguous claim as to the two-spin states they generated. They stated that "All four maximally entangled states, namely the Bell states $\Phi^{\pm} = {}^1\!/_{\sqrt{2}}(|00\rangle \pm |11\rangle)$ and $\Psi^{\pm} = {}^1\!/_{\sqrt{2}}(|01\rangle \pm |10\rangle)$ are generated, where '0' and '1' denote the two possible nuclear spin orientations $(m_I = -{}^1\!/_2 \mapsto |0\rangle,\ m_I = +{}^1\!/_2 \mapsto |1\rangle,\ |N_1N_2\rangle)$."

The spin product states $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$ thus defined are indeed the appropriate choice for a study aimed at entanglement generation; entanglement is a physical resource that exists between distinct entities (*2*), and the relevant entities here are the nuclear (and electron) spins. The quoted definition directs the reader to figure 1B in (*1*), where these product states are assigned to four eigenstates of the system. As we show here, however, two of these assignments are incorrect; consequently, in creating a superposition of the system's eigenstates, the authors do not actually create the $\Psi^{\pm}$ states. In actuality, those very Bell states are system eigenstates before any manipulations, and the applied manipulations evidently disentangle the spins. The authors also discuss tripartite-entangled quantum states, in which the third party is the electron spin associated with the NV defect. They assert that a *GHZ* state and a *W* state are generated. However, the latter claim, because it involves the same odd-parity nuclear spin states, is also incorrect.

We first classify the eigenstates of the electron-nuclear spin system. The supporting online material for (*1*) presents the following Hamiltonian for the single electron spin ($S = 1$) coupled equally to two surrounding $^{13}$C spins ($I = \tfrac{1}{2}$):

$$H = g_e\beta_e B_z\hat{S}_z + \hat{S}_z D\hat{S}_z + \sum_i(\hat{S}\cdot A\cdot\hat{I}_i - g_n\beta_n B_z\hat{I}_{zi}) \qquad (1)$$

The first term describes the Zeeman splitting of the electron spin and the second term is the zero field splitting, which lifts the degeneracy of the $S_z = \pm1$ states and the $S_z = 0$ state. The third term is the hyperfine coupling between each nuclear spin and the electron, which is described by a tensor $A$. The tensor is assumed uniaxial with principal components lying along the three Cartesian axes $x,y,z$ and $A_\perp = A_{xx} = A_{yy}$, $A_{zz} = A_\parallel$. The fourth term is the Zeeman splitting for the nuclear spins.

The Hamiltonian commutes with the operator $J_z \equiv S_z + I_{z1} + I_{z2}$ (our definition); hence this quantity is conserved. Therefore the Hamiltonian splits into separate uncoupled subspaces characterized by distinct values of $J_z$. Referring to figure 1B in (*1*), we would like to discuss the middle two states of the four states within the $m_s = -1$ manifold ($m_s$ in that figure corresponds to the $S_z$ used in the supplementary information). These two states have $J_z = -1 + \tfrac{1}{2} - \tfrac{1}{2} = -1$.

The Hamiltonian for the $J_z = -1$ subspace spans a three dimensional Hilbert space. Using the notation $|S_z, I_{z1}, I_{z2}\rangle$, a suitable basis is $\{|-1, \tfrac{1}{2}, -\tfrac{1}{2}\rangle, |-1, -\tfrac{1}{2}, \tfrac{1}{2}\rangle, |0, -\tfrac{1}{2}, -\tfrac{1}{2}\rangle\}$. In a matrix form in this basis the Hamiltonian becomes

$$H_{J_z=-1} = \begin{pmatrix} E_0 & 0 & K \\ 0 & E_0 & K \\ K & K & E_1 \end{pmatrix} \qquad (2)$$

where $K = A_\perp\sqrt{2}$, $E_0 = D - g_e\beta_e B_z$, and $E_1 = g_n\beta_n B_z$. However, using the typical values quoted in the supporting online material, we find $E_0 - E_1 = 2.64$ GHz, and $K = 174$ MHz. Thus, $K \ll (E_0 - E_1)$, and we can apply degenerate perturbation theory.

Because in the limit $K \ll (E_0 - E_1)$, the third state $|0, -\tfrac{1}{2}, -\tfrac{1}{2}\rangle$ is an approximate eigenstate itself, we can derive an effective approximate Hamiltonian for just the $\{|-1, \tfrac{1}{2}, -\tfrac{1}{2}\rangle, |-1, -\tfrac{1}{2}, \tfrac{1}{2}\rangle\}$ subspace. It is

$$H_{eff} = \begin{pmatrix} E_0 - \dfrac{K^2}{E_0 - E_1} & \dfrac{K^2}{E_0 - E_1} \\ \dfrac{K^2}{E_0 - E_1} & E_0 - \dfrac{K^2}{E_0 - E_1} \end{pmatrix} \qquad (3)$$

The diagonal components are equal, and so we find that the eigenstates are

$$|\psi_+\rangle = \frac{1}{\sqrt{2}}\left(|-1,\tfrac{1}{2},-\tfrac{1}{2}\rangle + |-1,-\tfrac{1}{2},\tfrac{1}{2}\rangle\right),$$
$$|\psi_-\rangle = \frac{1}{\sqrt{2}}\left(|-1,\tfrac{1}{2},-\tfrac{1}{2}\rangle - |-1,-\tfrac{1}{2},\tfrac{1}{2}\rangle\right) \qquad (4)$$

with eigenenergies $E_0$ and $E_0 - \frac{2K^2}{E_0 - E_1}$ respectively. Using the values stated above, this gives a splitting of 23 MHz. This is very close to the splitting observed in figure 1C. In the reduced notation for the nuclear spins used in (*1*), defined for the $m_s = -1$ subspace, these eigenstates become

$$|\psi_+\rangle = \frac{1}{\sqrt{2}}(|10\rangle + |01\rangle),$$
$$|\psi_-\rangle = \frac{1}{\sqrt{2}}(|10\rangle - |01\rangle) \qquad (5)$$

Importantly, these are not the separable states shown in the scheme of figure 1B in (*1*); rather, they are entangled Bell states. We thus conclude that the eigenstates of the system in this subspace are already maximally entangled. The result of making an equal superposition of these two states $\frac{1}{\sqrt{2}}(|\psi_+\rangle \pm |\psi_-\rangle)$, for which Neumann *et al.* present evidence in figure 2, B and C (*1*), is nothing but a construction of the separable states $|10\rangle$ and $|01\rangle$. The definition of the nuclear spin eigenstates also impacts on the definition of the *W* state discussed as the conclusion of (*1*) and in figure 4.

### References

1. P. Neumann *et al*., *Science* **320**, 1326 (2008).
2. L. Amico, R. Fazio, A. Osterloh, V. Vedral, *Rev. Mod. Phys.* **80**, 517 (2008).

22 August 2008; accepted 2 February 2009
10.1126/science.1168458

---

[1] Department of Materials, Oxford University, Oxford OX1 3PH, UK. [2] Centre for Quantum Technologies, National University of Singapore, 117543 Singapore.

*To whom correspondence should be addressed. E-mail: brendon.lovett@materials.ox.ac.uk

# Response to Comment on "Multipartite Entanglement Among Single Spins in Diamond"

**P. Neumann,[1]*† N. Mizuochi,[2]* F. Rempp,[1] P. Hemmer,[3] H. Watanabe,[4] S. Yamasaki,[5] V. Jacques,[1] T. Gaebel,[1] F. Jelezko,[1] J. Wrachtrup[1]**

Our study reported entanglement among single spins in diamond. Lovett and Benjamin argue that three of six described entangled states were not achieved. Here, we explain our choice of entangled states and discuss their importance for quantum information processing. We also show that the eigenstates discussed by Lovett and Benjamin, although formally entangled and routinely generated in our experiments, cannot be used to detect nonlocal correlations.

We recently reported on multipartite entanglement among single spins in diamond (*1*). In their comment, Lovett and Benjamin (*2*) claim that three of the six demonstrated entangled states are in fact not entangled, namely the odd-parity Bell states $\Psi^+$ and $\Psi^-$ and the *W* state. They support this contention by approximating the eigenbasis of our system and deriving that two of the nuclear spin eigenstates are already entangled [$|01\rangle$ and $|10\rangle$ in figure 1B (*1*)] and that we disentangled these states. Although it would have been easier for us to prepare the eigenstates $|01\rangle$ and $|10\rangle$, which Lovett and Benjamin call the entangled ones, we created the states $\Psi^+ = \frac{1}{\sqrt{2}}(|01\rangle+|10\rangle)$ and $\Psi^- = \frac{1}{\sqrt{2}}(|01\rangle-|10\rangle)$, which we believe are the correct choice to show entanglement in this system, as justified below.

We welcome the opportunity to address this conceptual problem because it led to some initial confusion in the field of quantum computing. Although this issue is now largely resolved, those not familiar with the experimental details of how quantum logic is performed in spin systems may not be aware of it. The resolution of this problem is also closely related to concepts such as decoherence-free subspaces and logical qubits in spin systems, which are now very important in the field of quantum computing.

Lovett and Benjamin (*2*) say that entanglement is a physical resource that exists between physical entities, which they assign to the nuclear spins. In their product spin basis, they claim that some eigenstates look like Bell states, namely $|01\rangle = \frac{1}{\sqrt{2}}(|-\frac{1}{2},\frac{1}{2}\rangle + |\frac{1}{2},-\frac{1}{2}\rangle)$ and $|10\rangle = \frac{1}{\sqrt{2}}(|-\frac{1}{2},\frac{1}{2}\rangle - |\frac{1}{2},-\frac{1}{2}\rangle)$. Although this state can be formally written like a Bell state, there is doubt that it can be called such. This is because, at the heart of entanglement, there is the correlation between two physically uncoupled systems (*3*)—a requirement that is not fulfilled here. Rather, the eigenstate is a mixing of the product spin states. Similarly, it also does not make sense to call the spin states of ortho and para hydrogen entangled. If, however, the spins are later physically separated, for example by breaking the molecular bonds and allowing the hydrogen atoms to fly apart (or the $^{13}C$ nuclei in the case of diamond), so that their strong coupling is removed and product states become eigenstates, then there will be entanglement. In the same way, Berkley *et al*. prepared an eigenstate of two strongly coupled Josephson-junction qubits that was formally entangled (*4*). This arose in a commentary by Wójcik *et al*. (*3*), which noted that such entanglement cannot be used to observe quantum correlations. However, Wójcik *et al.* further pointed out that there would be entanglement if the strong coupling were removed and the two qubits could be measured independently, as shown by Pashkin *et al*. (*5*). Thus, although it is theoretically possible to construct operators revealing the correlations among two strongly coupled spins as Bell-like, experimentally the strong coupling usually is removed, that is, the Hamiltonian is changed so that product states are eigenstates, and only then potential correlations can be monitored.

A particularly useful study of qubits is given in a coupled quantum dot spin system in (*6*). Unlike the NV diamond system, the Hamiltonian in the quantum dot system can be changed far more extensively by applying bias voltages. In a later study, Petta *et al*. also considered a case where it was possible to go from a situation where spin Bell states were eigenstates to a case where spin product states were eigenstates and back again (*7*), as in the case of Pashkin (*5*). This paper (*1*) is a particularly good illustration of the importance of the Hamiltonian in determining the correct choice for the unentangled spin qubit states, that is, the eigenstates.

In this sense, it is useless for us to show that we can prepare the eigenstates $|01\rangle$ and $|10\rangle$ to prove alleged entanglement because we cannot decouple the single spins/qubits from each other to achieve entanglement of the single nuclear spins, as is possible for quantum dots or superconducting qubits. In our basis, we drove four selective radio frequency (rf) transitions between all four eigenstates $|00\rangle\leftrightarrow|01\rangle$, $|00\rangle\leftrightarrow|10\rangle$, $|10\rangle\leftrightarrow|11\rangle$, and $|01\rangle\leftrightarrow|11\rangle$ [see figure 1 B (*1*)]. In this way we encoded two qubits that we can manipulate. We created entanglement among the states $|00\rangle$, $|01\rangle$, $|10\rangle$, and $|11\rangle$ of the NV center's $m_S = -1$ level shown in figure 1B in (*1*). This entanglement has been proven, for example, by doing Ramsey fringe experiments (*8*). This is the preparation of Bell states, a subsequent free precession period, and a final readout of their phases. In figure 2B(iv) in (*1*), Ramsey fringes following a phase dependence characteristic for multiple quantum transitions are clearly visible, thus demonstrating entanglement among the eigenstates $|01\rangle$ and $|10\rangle$. In addition to the Ramsey fringe experiments, state tomography has been performed on all produced states that revealed the correlations that are characteristic for entangled states; the same holds for the *W* states prepared. As such, we have a system at hand that is useful for quantum information application. Qubit information can be encoded and read out, as well as the correlations for which Bell states are famous. This is not the case for the eigenstates $|01\rangle$ and $|10\rangle$.

Another point made by Lovett and Benjamin (*2*) is that we made the a priori assumption that the eigenstates can be assigned as follows ($m_I = -\frac{1}{2} \mapsto |0\rangle$, $m_I = +\frac{1}{2} \mapsto |1\rangle$). This is the standard introduction to spin states, because it is conceptually simple to understand; these are the eigenstates in the absence of spin coupling. However, it often happens that the Hamiltonian of the total spin system has eigenstates that are not product states of the individual spins involved. When the Hamiltonian of the actual experimental spin system is stated, this physical picture must be modified [as was done in the supporting online material (SOM) (*1*)], but the notation is still correct as long as we are always working with eigenstates. We regret that we did not make this point more clearly in the original text.

If Lovett and Benjamin's derivation of eigenstates ($|01\rangle = \frac{1}{\sqrt{2}}(|-\frac{1}{2},\frac{1}{2}\rangle+|\frac{1}{2},-\frac{1}{2}\rangle)$ and $|10\rangle = \frac{1}{\sqrt{2}}(|-\frac{1}{2},\frac{1}{2}\rangle-|\frac{1}{2},-\frac{1}{2}\rangle)$) (i.e., singlet and triplet states) is correct, why were we able to drive the nuclear spin transitions shown in figure 1B in (*1*) at all? The resolution of this question is that their derivation is only an approximation, similar to the one we described in the SOM for (*1*). In their comment (*2*), the eigenstates were derived from a Hamiltonian that was commuting with an operator $J_z = S_z + I_{z1} + I_{z2}$ (their definition). They assumed from the SOM that the hyperfine tensors for the interaction of the electron with each of the

---

[1] 3. Physikalisches Institut, Universität Stuttgart, Pfaffenwaldring 57, D-70550 Stuttgart, Germany. [2] Graduate School of Library, Information, and Media Studies, University of Tsukuba, 1-2 Kasuga, Tsukuba-City, Ibaraki 305-8550, Japan. [3] Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX 77843, USA. [4] Diamond Research Center, National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba Central 2, Tsukuba 305-8568, Japan. [5] Nanotechnology Research Institute, AIST, Tsukuba Central 2, Tsukuba 305-8568, Japan.

*These authors contributed equally to this work.

†To whom correspondence should be addressed. E-mail: p.neumann@physik.uni-stuttgart.de

two nuclei was collinear with the symmetry axis of the NV center, which is in fact not the case (*9*), as referenced in the SOM but not clearly mentioned. The symmetry axis of the hyperfine tensor is collinear with the vacancy-$^{13}$C axis that is tilted against the NV symmetry axis by the tetrahedral angle 109.5° [figure 1A (*1*)]. Thus, their operator does not commute with the Hamiltonian, and the singlet and triplet descriptions for these states are not correct. The wrong axis treatment also causes the incorrect derivation of 23 MHz for the energy splitting. Nevertheless, as we have stated in our SOM, the eigenstates $|01\rangle$ and $|10\rangle$ indeed have some singlet and triplet character.

In conclusion, for the case of Bell states, we chose a set of basis states among which we encoded two qubits. There, we are able to drive four selective radio frequency transitions, allowing us to manipulate nuclear spin states, that is, to perform quantum gates. In this way, we have been able to create and read out those particular correlations for which Bell states are famous and what is demanded for quantum computation. That's why we think this is the right basis and that we have created the right Bell states. The same holds for the *W* state.

## References and Notes

1. P. Neumann *et al.*, *Science* **320**, 1326 (2008).
2. B. W. Lovett, S. C. Benjamin, *Science* **323**, 1169 (2009); www.sciencemag.org/cgi/content/full/323/5918/1169c.
3. A. Wójcik *et al.*, *Science* **301**, 1183 (2003).
4. A. J. Berkley *et al.*, *Science* **300**, 1548 (2003).
5. Y. A. Pashkin *et al.*, *Nature* **421**, 823 (2003).
6. J. R. Petta *et al.*, *Science* **309**, 2180 (2005).
7. J. R. Petta *et al.*, *Physica E* **35**, 251 (2006).
8. M. Mehring, J. Mende, W. Scherer, *Phys. Rev. Lett.* **90**, 153001 (2003).
9. J. H. N. Loubser, J. A. van Wyk, *Diamond Res.* **11**, 4 (1977).

10. We acknowledge A. Wójcik for comments on this manuscript.

10 October 2008; accepted 3 February 2009
10.1126/science.1168459

HISTORY OF SCIENCE

# Making German Evolution: Translation and Tragedy

**Lynn K. Nyhart**

In this year of Darwin anniversaries (the 200th year of his birth and the 150th anniversary of *On the Origin of Species*), *The Tragic Sense of Life* and *H. G. Bronn, Ernst Haeckel, and the Origins of German Darwinism* remind us that the history of evolutionary thought in the 19th century extended well beyond Darwin himself. Darwin did not launch his theory onto an unprepared public and scientific community, nor was the evolutionism that developed after 1859 a mere extension of his views—it was not even one thing. How, then, should we think about the history of evolution in the 19th century? What sorts of accounts best help us understand the reception of Darwin's theory, its relations to earlier ideas about nature, the directions that evolutionary investigation subsequently took, and the relations of all of these to the broader social, cultural, and religious concerns scientists shared with their contemporaries?

These questions become especially pointed when one considers German Darwinism, and especially Germany's best-known follower of Darwin, Ernst Haeckel. Most often remembered by biologists as the author of the biogenetic law ("ontogeny recapitulates phylogeny"), Haeckel has also been accused of promoting European fascism via his monistic philosophy and of presenting a eugenic, biologically determinist vision of humanity that led to Hitler's "final solution." Can one scientist be responsible for so much? Most historians would say no, arguing that it takes a community, rather than an individual, to make a movement; that single-cause explanations are insufficient to account for something as broad as fascism; and that an individual cannot be held responsible for the ways in which others (such as Hitler) took up his ideas and molded them to new agendas after his death. But that still leaves open the questions of how to write responsibly about what Haeckel actually believed and how we should situate him in the history of evolutionary thought.

The historians under consideration here have chosen two radically different strategies to understanding Haeckel's place within German evolutionism, and both have produced important books. Robert J. Richards, the director of the University of Chicago's Fishbein Center for the History of Science and Medicine and a much-published author on Darwin and German Romantic biology, has written a biography of Haeckel. Sander Gliboff, a professor in Indiana University's Department of History and Philosophy of Science, places Haeckel at the end of a study that examines the larger process through which Darwin's words were translated, and his ideas modified, in the context of German biology. Both illuminate the twists and turns that evolutionary thought took in Germany, but they do so in dramatically different ways.

**The Tragic Sense of Life**
Ernst Haeckel and the Struggle over Evolutionary Thought
***by Robert J. Richards***
University of Chicago Press, Chicago, 2008. 579 pp. $39, £27. ISBN 9780226712147.

**H. G. Bronn, Ernst Haeckel, and the Origins of German Darwinism**
A Study in Translation and Transformation
***by Sander Gliboff***
MIT Press, Cambridge, MA, 2008. 271 pp. $35, £22.95. ISBN 9780262072939.

Richards's book, though over twice as long as Gliboff's, is the more entertaining read of the two. In his characteristically rich and rolling prose, Richards weaves a compelling story of a life marked by tragedy and of an intense, larger-than-life figure whose passions drove his scientific research and philosophy. In Richards's rendering, the scientific Haeckel cannot be understood separately from the man's personality and private circumstances. His love of nature was surpassed only by his love for his first wife, Anna Sethe, who died in abdominal agony on his 30th birthday. Over the next year, he wrote his way through the despair that enveloped him, producing his foundational work, *Generelle Morphologie* (*1*). Although he remarried, the union was not happy, and passionate love would elude him until his sixties, when he had a secret affair that ended tragically with the death of his lover. Science remained his salvation and refuge.

His professional life was also filled with drama, much of which centered on his philosophy of evolutionary monism—a science-centered faith that became one of the most successful alternatives to the Judeo-Christian religion among those searching for a secular spirituality. Haeckel could not turn down a fight: He battled the physician-statesman Rudolf Virchow over the role of evolution in the schools (Haeckel argued that it should replace religious education), sparred with religiously conservative scientists and with former students who challenged his ideas as they gained intellectual independence, and debated the pro-evolution (but anti-Haeckel) Jesuit priest and entomologist Erich Wasmann—the list could go on and on. These were not isolated episodes but rather moments in a lifelong campaign to advance his philosophy, which was accompanied by a bitter hostility to organized religion.

Richards does not neglect Haeckel's science proper, treating us to fascinating and original discussions of his pathbreaking systematic and phylogenetic work on radiolaria and other marine organisms, the importance of linguistic analysis to his phylogenetic trees of the races of humans, and his remarkable experimental work with siphonophores. These constitute important contributions to our understanding of the technical development of evolutionary biology.

The big picture here, however, is an argument about the power of personality—at least one personality—to shape the course of science. In Richards's presentation, German evolutionism was profoundly shaped by both Haeckel's charisma and his combativeness. Perhaps the late-19th-century opposition of evolutionary science to Christianity would not have been so fiery, he suggests, had Haeckel not continually fanned its flames. And although Richards absolves Haeckel of personal responsibility for fascism and Nazism, in part by situating him firmly in his time and place, he does show how the scientist's ardent temperament led him to the occasional intemperate statement that could be taken up by extreme thinkers. One cannot leave this book without a deep appreciation for Haeckel as a tragic figure and for the force of personality in shaping the direction science may take.

Gliboff's account is of a completely different order. His is not a story of personalities or private lives (although he mentions salient details), but of German academics seeking to live up to the highest (if changing) ideals of Wissenschaft and of the ways in which Darwin's theory was translated into this environment. He thus situates Haeckel at the end of a revised intellectual history of 19th-century German evolutionism. Central to his account is the idea of translation, which he uses both synchronically, especially in treat-

---

The reviewer is at the Department of the History of Science, University of Wisconsin, Madison, WI 53706–1393, USA. E-mail: lknyhart@wisc.edu

ing the translation of *Origin of Species* into German, and (more intriguingly) diachronically, as scientists reworked older words such as "perfection" and "type" to lend them new meanings. Gliboff's own clear, crisp prose is key to the success of this analysis, as he deftly leads his reader through dense philosophical and terminological thickets with nary a thorn scratch. This is some of the best close reading I have seen. It also represents a profound challenge to our standard picture of 19th-century German biology.

The old story, crudely put, is that Haeckel's version of evolution was a Darwinism in name only, best understood as an update on early-19th-century idealistic morphologists such as Carl F. Kielmeyer and J. F. Meckel that retained their teleology, their typological emphasis on form, and their linear recapitulationism. This story, emphasizing the long persistence of a German transcendental approach to nature, has been deeply entrenched in the history of biology.

Gliboff challenges this history right from the beginning. The ascription of simple linear recapitulationism to the views of Romantic embryologists, he notes, owes much to a caricature developed by Karl Ernst von Baer in a polemical context, then adopted uncritically by influential historians such as E. S. Russell and Stephen Jay Gould. Gliboff's fresh reading of the original sources interprets Kielmeyer and Meckel as far less rigidly typological in their orientation and much more attentive to nature's variability than has been seen before. Both for these early-19th-century naturalists and for their intellectual heirs, Gliboff argues, the critical issue was to understand nature's manifold variety while seeking out underlying strict natural laws to account for it.

This provides a new starting point for analyzing Darwin's first translator, the prominent paleontologist H. G. Bronn—a figure little attended to in the standard story but the lynchpin of Gliboff's. Intriguingly and plausibly, Gliboff argues that Bronn's use of terms like "vervollkommnet" (perfect) as translations for Darwin's "improved" or "favored" were not about dragging Darwin backward into a German teleological view of nature (as has been claimed by those who have paid attention to Bronn at all). Instead, Gliboff asserts, Bronn's translations involved an attempt to recast existing German terms in a newer, more up-to-date mode that encompassed selection yet tamed Darwin's emphasis on unpredictability to meet the more rigorous requirements of a German academic scientist's understanding of a "law" of organic nature. Simultaneously, Bronn sought to translate Darwin's ideas about selection into a language without an exact equivalent for the term, and for an academic audience lacking the gentlemanly traditions of breeding pigeons and dogs so central to Darwin's exposition. The selection metaphor was further fraught with an anthropomorphism foreign to Germans, who were not brought up on British natural-theological assumptions about a personified God who had created a perfectly adapted nature. Bronn's translation, though it altered key ideas to make Darwin comprehensible to a German academic audience, was not a conservative throwback. It represented the dynamic engagement of a leading paleontologist who had also long been working on many of the questions Darwin claimed as his own—a critical yet generous equal, who saw himself as moving science forward through the modifications he made to Darwin's flawed theory. Bronn's death in 1862 afforded him little chance to steer the conversation further.


**A painter, too.** Haeckel's oil landscape of highlands in Java, from *Wanderbilder* (1905).

CREDIT: ERNST HAECKEL/FROM *WANDERBILDER* (W. KOEHLER, GERA-UNTERMHAUS, 1905)

And so, finally, we come to Haeckel. Gliboff's key insight here is that Haeckel originally read Bronn's translation of Darwin, not Darwin in the original. Gliboff shows Haeckel as both echoing and responding to Bronn's concerns, rather than either reflecting directly on Darwin's original writing or reaching directly back to the Romantic embryologists. (Although Gliboff acknowledges the centrality of monism to Haeckel's thought, he focuses on the working evolutionary theorist, not the popular ideologue.) Like Bronn himself, Haeckel made further amendments both terminological and intellectual, and Gliboff rereads Haeckel's research program as one not dominated by a typological and linear-recapitulationist mindset but rather as continuing to wrestle with the need to account for variability and unpredictable change in terms of mechanistic laws of nature—among which Haeckel included, at the top of his list, natural selection. Haeckel's Darwinism thus shows continuity with early-19th-century concerns, mediated through Bronn. But those concerns were always more flexible than has been acknowledged, and their articulation changed over time. Of course Haeckel's Darwinism was not Darwin's own, but it was not an aberration or a distortion of some true theory, any more than any other post-Darwinian additions or adjustments were. It was science moving on.

Gliboff's overall picture of scientific advance, in contrast to Richards's emphasis on charisma and passion, is one of scientists building and innovating incrementally, working with what their predecessors have handed them and sculpting it into something new yet understandable to those around them. His sensitive reading allows us to see post-1859 German evolutionists as rational actors rather than irrationally stuck in some early-19th-century moment with unmodern commitments. By challenging the very foundations of the standard narrative of German morphology, this careful, compelling account does at least as much as Richards's to undermine the association of 19th-century German Darwinism with a dangerously exceptional view of nature. But the two books offer very different reads. Is scientific progress a matter of personal anguish and triumph, or of intellectual chugging along? Our concept of it should be capacious enough to include both.

### References and Notes

1. E. Haeckel, *Generelle Morphologie der Organismen* (Georg Reimer, Berlin, 1866).
2. The reviewer previously served as a press reader for both books at the manuscript stage.

10.1126/science.1169621

ECOLOGY

# A National Ecological Network for Research and Education

A proposal links continental-scale ecological data with education projects.

**Margaret Lowman,[1]* Charlene D'Avanzo,[2] Carol Brewer[3]**

In 1804, Lewis and Clark described nearly 300 new plants and animals from Missouri to Oregon, setting a baseline for observational ecology in the United States (*1*). Forty years later, Thoreau made some of the earliest field measurements at Walden Pond (*2*). Now, field measurements by scientists, students, and citizens are being taken to a new level as ecologists build the National Ecological Observatory Network (NEON), a continental-scale research platform, to study the nation's ecological challenges (*3*). Ecologists will use a distributed network of sensors linked by advanced cyberinfrastructure to predict responses of the biosphere to changes in land use, invasive species, and climate over the next 30 to 50 years. Dubbed the "ecologists' Hubble Telescope," this large-scale sentinel for environmental change is funded by the National Science Foundation's Major Research Equipment and Facilities Construction (MREFC) program. During a time of national economic downturns, NEON will foster jobs and student research opportunities at its sites and will contribute data sets to encourage "best practices" to solve environmental challenges on a continental scale.

**The future: A K–12 scenario.** In the year 2020, sixth graders in Miami, Florida, grab their personal digital assistants (PDAs, handheld computers) and log into an ecology lesson. The students download a database on insect pests of subtropical forests, which they can explore as an ecosystem visually and quantitatively. They then engage in simulation "games" involving insect outbreaks that threaten the southeastern forest canopies. Students use case studies and manipulate real-time data (including forest cover, foliage production, and caterpillar population dynamics) to create future scenarios for urban and rural forests. Their PDAs enable them to exchange ecosystem and biodiversity tours that they have created with classes 2000 miles away. Science education researchers simultaneously measure the effectiveness of this combination of cyberinfrastructure, gaming, and hands-on ecosystem research.

CREDIT: M. LOWMAN

Citizen science programs have greatly increased in recent years, in part owing to the availability of the Internet for data input and access (*4*). In the Jason Expedition (www.jason.org), fifth to eighth grade students have collected ecological data to share with other schools. Citizen Science Central began in 2006 as a "clearinghouse for ideas, news, and resources in support of citizen science" and developed a toolkit to facilitate that process (www.birds.cornell.edu/citscitoolkit). The National Institute of Invasive Species Science (NIISS, www.niiss.org) is a consortium of organizations that relies on citizen-science monitoring and reporting of invasive species. Project Budburst (www.budburst.org), a partner with the National Phenology Network (www.usanpn.org), engages the public in observations of seasonal events, such as first leaf, first flower, and first ripe fruit for a diverse array of plant species found in the United States.

NEON differs from other national science programs in its goals to launch ecological research and education platforms simultaneously in a way that improves ecological forecasting. Planning for NEON has taken a decade with hundreds of ecologists collaborating (www.neoninc.org/). Projected to become fully operational in 2016, the NEON national network includes 20 domain sites of matching stationary and mobile field sensors, laboratories, archival data, synthesis and analysis centers, and education nodes, all linked through high-capacity cyberinfrastructure (*3*, *5*–*7*). Each candidate site, located in a wild land area, represents the vegetation, soils, landforms, climate, and ecosystems performance of its given domain (Mid-Atlantic, Southern Plains, Pacific Northwest, and so forth). Candidate sites are located across the continental United States, including two in Alaska and one each in Hawaii and Puerto Rico. Work has already begun on a prototypic site north of Boulder, Colorado. These sites are expected to be operational for ~30 years;

---

[1] Departments of Biology and Environmental Studies, New College of Florida, Sarasota, FL 34243, USA; NEON senior design team, NEON Education Tiger team. [2] School of Natural Science, Hampshire College, Amherst, MA 01002, USA; Chair of NEON Education Tiger team. [3] Division of Biological Sciences, University of Montana, Missoula, MT 59812, USA; NEON senior design team, NEON Board of Directors.

*Author for correspondence. E-mail: canopymeg@gmail. com

they will allow scientists to monitor ecosystems and forecast ecological changes with regard to six of the National Academies' grand challenges: biodiversity, climate change, ecohydrology, biogeochemistry, infectious disease, invasive species, and land use (*8*).

### Research and Education for Public Policy

NEON aims to integrate ecological research with education (*3*, *9*) (see figure, page 1176). Over the short term, educators and cyberinfrastructure engineers will collaborate to design user-appropriate interfaces, data processing, search services, and other technologies. The plan is that ecologists, the public, policy-makers, students, teachers and other user groups will be able to benefit from and contribute to NEON data streams, products, and learning experiences.

What NEON education and outreach might look like over the next 30 years will ultimately be shaped by a staff that will include an education chief, NEON domain education coordinators, and advisory committees. As a preliminary guide, a NEON Education and Outreach Strategic Plan was released in 2007 (*9*). One of the goals described in that plan is to train a new generation of scientists not only to utilize NEON's cyberinfrastructure and technologies, but also to communicate better with the public, as well as with each other. NEON education aspires to serve as a test bed for the efficacy of science education linked directly to the science.

A priority in the education program is to include underrepresented groups, such as ethnic minorities, women, and first-generation college students, and to ensure both excellence and applicability of the questions addressed to the general public. NEON plans to foster a generation of researchers trained to work collaboratively with new technologies on a continental scale that integrates multiple factors over time (*10–12*). In NEON user workshops, graduate students and veteran researchers alike will learn novel approaches to mathematical modeling, simulation, and digital mapping of complex ecological phenomena.

For the public, NEON will bring observations about regional and national environmental changes into schools, town halls, and households, facilitating better understanding of the links between human health and healthy ecosystems. NEON education will develop links between publicly available data sets and user groups through activities such as citizen-scientist investigations, ecological field trips for students from kindergarten through high school (K–12), visualizations of near-real-time data from across the continent, and conversations between decision-makers and scientists.

### Future Questions

At continental, long-term scales, NEON provides an opportunity for ecology education research. Educators will have access to data bases to study how NEON data and materials influence regional decision-makers, K–12 amd college students, and ecology researchers (*13–15*). Examples of questions include: How do we best train graduate students to become science educators as well as researchers? What is the relation between people's knowledge of NEON forecasts and their decisions or actions related to the environment? Can the integration of child-friendly simulations, gaming, and personal technology serve to jump-start student proficiency in science? What is the appropriate balance between technological simulations and hands-on, nature-based activities?

Because NEON science and education are both enormous undertakings, the importance of their integrative capacity could become lost over time. Critical questions that will define the success of NEON education include: Will the NEON cyberinfrastructure network serve multiple audiences including educators, citizens, diverse populations, and leaders? How can the transformational element of parallel NEON education and science platforms enhance science education in the United States?

Certainly NEON will need to collaborate with other organizations. For example, in building on the scenario envisaged in the figure below, NIISS is developing a Global Organism Detection and Monitoring system (GODM) that allows users to browse, upload, download, and analyze data on invasive species of all taxa. Field data are collected by members of the public trained to follow research protocols. Together NEON and NIISS could build a national program of geographically distributed "human sensors." Ecologists could then configure the data these citizen-scientists gather for use by local policy-makers. How to make such collaborations successful is another question that will need to be addressed.

NEON represents an ambitious project, much of which, especially the education component, is still on the drawing boards. Will the NEON education vision be achieved, or will it become a pile of dusty papers on a bureaucrat's desk? Only time will tell, as NEON continues to unfold. However, given the urgent environmental challenges we face as a nation, raising ecological literacy in the United States is crucial.

**The future: Informed policy-making.** Sarasota County's regional government in Florida has a problem. Pythons (first documented in 2001 in the Everglades after pets were carelessly released into the wild) are rapidly moving northward. Using the NEON database on invasive species and specifically targeting subtropical watersheds, county officials execute a risk analysis and estimate costs for controlling this new invasive species over the next two decades.

CREDIT: M. LOWMAN

### References and Notes

1. D. Botkin, *Our Natural History: The Lessons of Lewis and Clark* (Oxford Univ. Press, New York, 1995).
2. H. D. Thoreau, *Walden Pond* (Ticknor & Fields, Boston, MA, 1865).
3. NEON Integrated Science and Education Plan, Boulder, CO; www.neoninc.org.
4. For additional information, supporting materials are available at *Science* Online.
5. S. Senkowsky, *Bioscience* **55**, 106 (2005).
6. J. A. Teeri, P. H. Raven, *Science* **298**, 1893 (2002).
7. J. Mervis, J.Kaiser, *Science* **300**, 1869, (2003).
8. National Research Council, *Grand Challenges in Environmental Sciences* (National Academies Press, Washington, DC, 2001).
9. C. D'Avanzo *et al.*, *NEON Education and Outreach Strategic Plan* (NEON, Boulder, CO, 2008); www.neoninc.org.
10. National Research Council, *National Science Education Standards* (National Academies Press, Washington, DC, 1996).
11. AAAS, *Benchmarks for Science Literacy* (Oxford Univ. Press, New York, 1993).
12. National Research Council, *Inquiry and the National Science Education Standards* (National Academies Press, Washington, DC, 2000).
13. C. D'Avanzo, *Front. Ecol. Environ.* **1**, 533 (2003).
14. C. Brewer, *Conserv. Biol.* **20**, 689 (2006)
15. M. Keller, D. Schimel, W. Hargrove, F. Hoffman, *Front. Ecol. Environ.* **6**, 282 (2008).
16. Funding from TREE Foundation (www.treefoundation.org) to M.L.

### Supplemental Online Material

www.sciencemag.org/cgi/content/full/323/5918/1176/DC1

10.1126/science.1166945

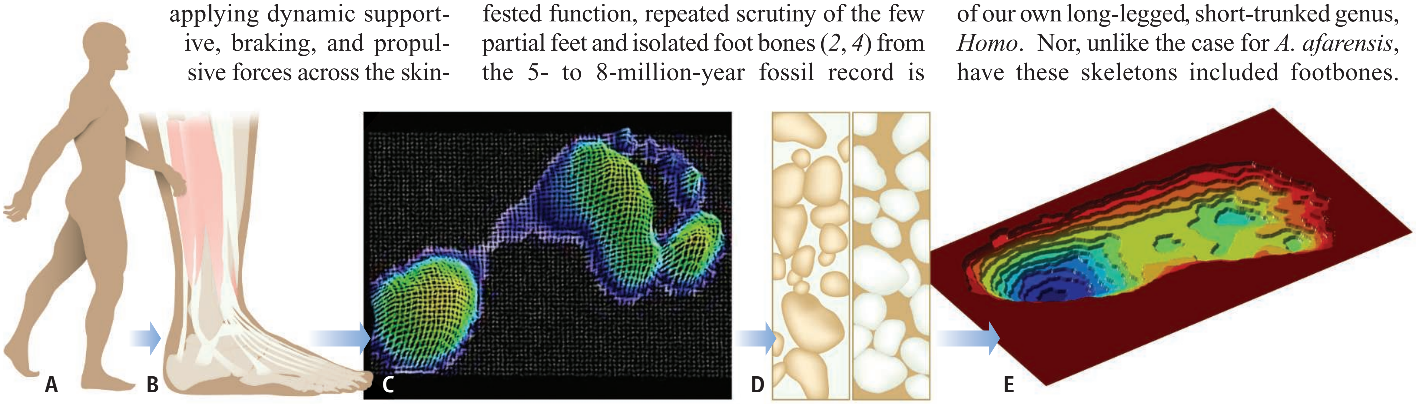
ANTHROPOLOGY

# Stepping Out

What can fossil footprints reveal about the evolution of the human foot?

**Robin Huw Crompton* and Todd C. Pataky**

The human foot is our most distinctive adaptation. Twenty-six bones, moving against each other over complex—sometimes multibone—joint surfaces are held in a three-dimensional latticework of ligaments and tendons, cushioned below by fibrous and fatty pads and thick skin (*1*). Together, these features ultimately keep us upright and drive our walking and running by applying dynamic supportive, braking, and propulsive forces across the skin-ground interface (*1*, *2*). How and when did our foot, this most distinctive human adaptation, evolve? On page 1197 of this issue, Bennett *et al.* (*3*) shed light on these questions by reporting the exciting discovery, at Ileret, east of Lake Turkana in Kenya, of a ~1.5 million-year-old fossil footprint trail, the oldest attributable to our own genus, *Homo*.



**How to read the footprints.** Bipedal walking (**A**) transmits forces via the anatomical structures of the feet (**B**) to generate pressure (**C**) against the substrate (**D**). The interaction between foot pressure and substrate material properties produces a footprint (**E**). If we understand the relation between foot pressure, foot anatomy, substrate properties, and footprint form, we should be able to reconstruct the gait of early human ancestors from their fossil footprints.

Visual comparison with the feet of other living apes identifies obvious differences. Humans have shorter toes, and the human big toe lies with the others, rather than angling away from them. Other apes' feet lack a medial longitudinal arch, and the midfoot flexes under load to apply propulsive forces to the ground (*2*, *4*). At least at rest, the feet of other apes maintain an inturned posture. Finally, unlike other apes, human feet change compliance during gait, acting as a shock absorber as the heel contacts the ground but as a stiff lever during push-off (*5*). This capability allows propulsive forces to be applied by the forefoot in humans, whereas they are applied by the midfoot in other apes, giving humans greater propulsive leverage about the ankle (*1*, *2*, *4*).

It is easy to conclude that the permanently compliant, inturned, grasping ape foot is better adapted to movement on tree branches, whereas the human foot is better adapted to bipedal movement on level ground. However, until we understand the relation between internal foot mechanics and externally manifested function, repeated scrutiny of the few partial feet and isolated foot bones (*2*, *4*) from the 5- to 8-million-year fossil record is unlikely to significantly improve our understanding of how the human foot became so different. Fortunately, because fossil footprints were formed by the interaction between dynamic external foot pressure fields and the material properties of the substrate (*6*–*8*), they are as close as we can get to a record of a given gait—in short, "fossilized behavior" (see the figure)—giving us an additional source of functional information.

The three-dimensional morphology of footprints could, for example, show whether a medial longitudinal arch was present or absent. Although archlike features might be created by large propulsive forces exerted during forefoot push-off, they might also suggest the presence of springlike ligaments, which contribute, on recoil, to power running in humans (*9*). The presence of an impression made by the ball of the big toe would suggest a foot held, like our own, in a relatively outturned posture compared with that of other apes, and the orientation of a big toe impression would show the extent to which the makers had abandoned use of the big toe as a grasping, thumblike organ.

Footprint trails from early stages of human evolution have long been known to exist at Laetoli, Tanzania, but the human ancestor present at Laetoli when those footprints were made was the diminutive, short-legged, long-trunked biped *Australopithecus afarensis*. Until now, there have been no footprint trails contemporaneous with the appearance of the first skeletons of members of our own long-legged, short-trunked genus, *Homo*. Nor, unlike the case for *A. afarensis*, have these skeletons included footbones. This is a tantalizing lack, because the new body proportions are often thought to be adaptations for long-distance striding and running (*10*).

The discovery by Bennett *et al.* is thus particularly important. Using modern laser scanners and software for surface reconstruction, they have analyzed the morphology of the Ileret footprints with methods that allow direct, quantitative comparison with footprints made by the Laetoli hominin and by modern humans.

With few exceptions, previous studies of the Laetoli footprints have been qualitative, with the result that the footprints have been variously interpreted (*2*) as fully modern-human in aspect (*11*) or as made by an inverted foot with a divergent, grasping big toe (*12*). Quantitative, statistical comparison of these complex three-dimensional shapes is far from straightforward. There are two main problems, common to all analyses of shape (*13*): First, the prints have arbitrary orientation and size, so one cannot, for example, directly compare perimeter tracings. Second, many alternative features can be isolated from the shapes, and identifying the features

CREDIT: ADAPTED BY P. HUEY/SCIENCE

---

School of Biomedical Sciences, University of Liverpool, Liverpool L69 3GE, UK. E-mail: rhcromp@liverpool.ac.uk

most suitable for distinguishing footprint samples is challenging.

To overcome these problems, Bennett *et al.* use coordinate transformations to remove size and orientation differences from the footprint shapes and then conduct discriminant analysis to identify the combination of features that best differentiates their footprint samples. They conclude that, although variables including forefoot-midfoot width ratio and divergence of the big toe are sufficient to statistically distinguish the Ileret footprints from the Laetoli prints, the Ileret footprints are broadly indistinguishable from those of modern humans.

Were the Ileret footprint makers' feet the first to function just like ours? Do the Laetoli prints represent more "apelike" foot function? Do all regions of a footprint record local maxima in foot pressure, or do some record how pressure changes over time, as braking forces change to propulsive ones? None of these questions can be answered at present. It is not even clear whether a nondivergent big toe is important for the extended push-off typical of human walking or just a by-product of other anatomical changes (*1*, *2*, *4*, *8*). The effect of substrate recoil or of later abrasion on the reliability of footprint measurements must also be established (*6*–*8*). But the findings of Bennett *et al.* herald an exciting time for studies of the evolution of human gait.

### References

1. L. Klenerman, B. Wood, *The Human Foot: A Companion to Clinical Studies* (Springer, Heidelberg, 2006).
2. R. H. Crompton *et al.*, *J. Anat.* **212,** 501 (2008).
3. M. R. Bennett *et al.*, *Science* **323**, 1197 (2009).
4. W. E. H. Harcourt-Smith, L. C. Aiello, *J. Anat.* **204**, 403 (2004).
5. F. Bojsen-Moller, *J. Anat.* **129**, 165 (1979).
6. K. Bates *et al.*, *Palaeontology* **51,** 999 (2008).
7. J. R. L. Allen, *Phil. Trans. R. Soc. London B* **352**, 481 (1997).
8. D. J. Meldrum, in *From Biped to Strider*, D. J. Meldrum, C. E. Hilton, Eds. (Kluwer, New York, 2004), pp. 63–83.
9. R. F. Ker *et al.*, *Nature* **325**, 147 (1987).
10. D. M. Bramble, D. E. Lieberman, *Nature* **432**, 345 (2004).
11. L. M. Robbins, in *Laetoli: A Pliocene Site in Northern Tanzania*, M. D. Leakey, J. M. Harris, Eds. (Oxford Univ. Press, Oxford, 1987), pp. 496–502.
12. Y. Deloison, in *Origine(s) de la Bipédie chez les Hominidés*, B. Senut, Y. Coppens, Eds. (CNRS, Paris, 1991), pp. 177–186.
13. C. G. Small, *The Statistical Theory of Shape* (Springer, New York, 1996).

10.1126/science.1170916

---

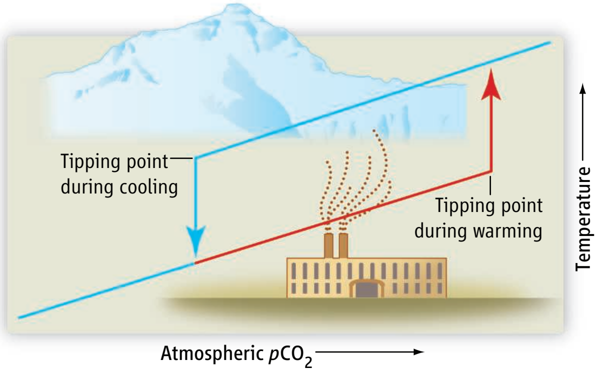
ATMOSPHERIC SCIENCE

# Tipping Pointedly Colder

Data from multiple ocean basins elucidate an ancient climate transition from greenhouse to icehouse.

**Lee R. Kump**

Analysis of the geologic record shows that Earth's temperature has skyrocketed suddenly, and plunged precipitously, many times in its history. These abrupt climate transitions presumably occurred because tipping points in the climate system were exceeded (*1*). What are these tipping points, and are we pushing the climate system to another with fossil-fuel burning? On page 1187 of this issue, Liu *et al.* (*2*) bring new tools to bear on these questions. Using data from multiple ocean basins, they study the relation between temperature and ice-sheet initiation during an ancient tipping point: a transition from greenhouse to icehouse.



**How to get into and out of the icehouse.** As the atmospheric partial pressure of $CO_2$ ($pCO_2$) declines during a cooling phase of an ice-free greenhouse interval (red path), perhaps driven by mountain building and enhanced continental weathering, the climate system approaches a tipping point beyond which it shifts abruptly from greenhouse to icehouse. The escape from the icehouse (blue path) requires $pCO_2$ to rise above the level that triggered the ice age because of hysteresis in ice-sheet dynamics. Eventually the melting occurs, and when it does, it does so abruptly.

For much of Earth history, the climate has been considerably warmer than it is today. But 33.7 million years ago, at the Eocene-Oligocene boundary, the world became trapped in the glacial state that continues to this day. Within just 200,000 years, Antarctica went from being a rather hospitable place to a polar continent buried under kilometers of ice. The transition was abrupt, but also overshot the new equilibrium with a superglaciation—dubbed the Oi-1 climate event—that lasted a few hundred thousand years.

In an earlier study, Lear *et al.* (*3*) found evidence from deep-water sites that there was little if any cooling during the transition. This finding challenged common sense: Surely the planet needs to cool if ice sheets are to grow. Recent work on better preserved shallow-water sites by Lear and colleagues (*4*) and others (*5*), together with the results presented by Liu *et al.* from a geographically diverse suite of locations, shows that cooling did indeed accompany the growth of ice sheets on Antarctica.

The most commonly used indicator of ancient climates is the oxygen isotopic composition of seawater, recorded by proxy in marine calcium carbonate sediments. Unfortunately, this proxy is ambiguous. Both cooling and ice-sheet growth drive the oxygen isotopic composition of marine carbonates in the same direction (increasing the standardized ratio of $^{18}$O to $^{16}$O, or $\delta^{18}$O). What part of this dramatic shift was a result of cooling, and what part a result of ice-sheet growth?

Two unknowns and one known demand another known for solution. Liu *et al*. use two independent temperature proxies—referred to as TEX$_{86}$ and U$^{K'}_{37}$—to determine what part of the oxygen isotope signal was a result of cooling; the rest must have been the result of ice-sheet growth. By extracting alkenones from marine sediments spanning the Eocene-Oligocene boundary and performing the TEX$_{86}$ and U$^{K'}_{37}$ analyses, the authors are able to develop a proxy record of temperature change during the transition. They then use these temperature estimates to subtract the temperature component from the $\delta^{18}$O record, leaving the part that reflects the growth of Antarctic ice. They find that the Antarctic ice sheet swelled between 40 and 120% of its modern size during the transition. Lack of substantial cooling in the earlier analysis by Lear *et al.* (*5*) led to much larger ice-volume estimates, seemingly requiring ice-sheet growth in the Northern Hemisphere, a suggestion not well supported by geologic data.

---

Department of Geosciences and Earth and Environmental Systems Institute, Pennsylvania State University, University Park, PA 16802, USA. E-mail: lkump@psu.edu

CREDIT: P. HUEY/*SCIENCE*

Liu *et al.* find that mid-to-high–latitude sites (from 40° to 70° in both hemispheres) cooled by 5° to 6°C across the event into the Oi-1 superglacial and then warmed slightly, resulting in an overall cooling of ~3°C. Tropical data are less consistent: Most sites show little evidence for cooling. In the case of those that do indicate substantial cooling (more than 10°C), the authors conclude that poorly understood ecological effects on the $TEX_{86}$ proxy (*6*) make them untenable. The authors do not address the possibility that some higher-latitude data might be subject to similar non–temperature-related effects.

Simultaneous cooling at high northern and southern latitudes points strongly to a greenhouse gas reduction. It is puzzling, though, that the tropics show little cooling. Climate model simulations, including those conducted by the present authors, cool both tropics and higher latitudes when atmospheric carbon dioxide concentrations are reduced from greenhouse climate conditions.

Using a climate model, Liu *et al.* can only reproduce the cooling indicated by the temperature proxies when they impose a reduction in atmospheric carbon dioxide concentrations from 8 to 2 times the preindustrial level across the transition. This reduction is much larger than that previously interpreted from atmospheric $CO_2$ proxies, from 4 to 2 times the preindustrial level (*7*). If the atmospheric $CO_2$ proxies are correct, then the models are missing something that amplifies the climate sensitivity to changes in atmospheric $CO_2$ (*8*).

Given the strong evidence that cooling accompanied the transition into the glacial world (*2*–*4*), the playing out of the onset of Antarctic glaciation follows the script of a tipping-point climate transition. The event was foreshadowed in the Late Eocene, 34.1 million years ago, with ice-sheet growth and sea-level fall, and another event at 33.8 million years ago with both cooling and sea-level fall (*4*). In coupled climate and ice-sheet models (*9*), Antarctic ice sheets begin to grow when atmospheric carbon dioxide levels fall, from more than four times preindustrial carbon dioxide concentrations in the Late Eocene to 2.5 to 3.0 times those concentrations in the early Oligocene. The transition into the glacial state proceeded through an overshoot and adjustment, the Oi-1 event, driven by a rapid drop in atmospheric carbon dioxide concentrations (*7*); a slower forcing may have led to a smoother transition, without overshoot (*10*).

A substantial sea-level fall accompanied the growth of the ice sheets, and shallow-water tropical calcium carbonate deposition was severely reduced as the platforms on which they grew were exposed. Carbonate deposition was displaced offshore, as evidenced by an accumulation of calcium carbonate in regions of the deep sea that were carbonate-free in the Late Eocene. This is the expected consequence of an abrupt sea-level fall (*11*). Yet, even larger drops in sea level during the Pleistocene (between 1.8 million and 10,000 years ago) had very little effect on global deep-water carbonate accumulation. This poses a problem for our understanding of carbonate deposition today (*12*), which is thought to be roughly equally distributed between shallow- and deep-water settings. The lack of response in deep-ocean sedimentation to sea-level fall indicates that shallow-water carbonate deposition during the Pleistocene and the current Holocene (since 10,000 years ago) is a small fraction of global carbonate deposition. Given that knowledge of the carbonate budget is critical to our projections of the effects of fossil-fuel burning on ocean chemistry, it may be time to revisit these estimates.

In thinking about the future, we must recognize that threshold behavior in one direction—like the Eocene-Oligocene boundary studied by Liu *et al.*—is normally accompanied by threshold behavior in the reverse direction, although the barriers to transition can differ in size (see the figure). For example, modern climate/ice-sheet models exhibit considerable hysteresis, requiring atmospheric carbon dioxide concentrations to rise well above the original initiation level to melt the Antarctic ice sheet (*13*). Future tipping behavior into a warm Eocene-like climate state may thus be delayed, but if and when it does occur, the transition will likely be abrupt.

## References and Notes

1. R. B. Alley *et al.*, *Science* **299**, 2005 (2003).
2. Z. Liu *et al.*, *Science* **323**, 1187 (2009).
3. C. H. Lear, H. Elderfield, P. A. Wilson, *Science* **287**, 269 (2000).
4. M. E. Katz *et al.*, *Nat. Geosci.* **1**, 329 (2008).
5. C. H. Lear *et al.*, *Geology* **36**, 251 (2008).
6. C. Turich *et al.*, *Geochim. Cosmochim. Acta* **71**, 3272 (2007).
7. M. Pagani, Z. Liu, M. Huber, *Geochim. Cosmochim. Acta* **72**, A716 (2008).
8. L. R. Kump, D. Pollard, *Science* **320**, 195 (2008).
9. R. M. Deconto, D. Pollard, *Nature* **421**, 245 (2003).
10. J. C. Zachos, L. R. Kump, *Global Planet. Change* **47**, 51 (2005).
11. H. K. Coxall *et al.*, *Nature* **433**, 53 (2005).
12. J. D. Milliman, *Global Biogeochem. Cycles* **7**, 927 (1993).
13. D. Pollard, R. M. DeConto, *Global Planet. Change* **45**, 9 (2005).

14. Supported by NSF grants EAR-0652020 and EAR-0628486.

10.1126/science.1170613

---

ENGINEERING

# Harvesting Ocean Wave Energy

**Jeff Scruggs[1] and Paul Jacob[2]**

Offshore structures that generate electrical power from ocean waves have been deployed but engineering challenges remain.

First proposed more than 30 years ago (*1*), systems to harvest utility-scale electrical power from ocean waves have recently been gaining momentum as a viable technology. The potential for this resource is promising, especially on west-facing coasts in either hemisphere with latitudes between 40° and 60°. In the United Kingdom, for example, the Carbon Trust recently estimated the extent of the economically viable offshore resource at 55 TWh per year, about 14% of current national demand (*2*). Across Europe, the technically achievable resource has been estimated to be at least 280 TWh per year (*3*). In 2003, the U.S. Electric Power Research Institute (EPRI) estimated the viable resource in the United States at 255 TWh per year (6% of demand), comparable to the energy currently generated in the United States by conventional hydropower (*4*). EPRI also found that wave energy could become economically competitive with other renewables in the Pacific Northwest and Hawaii (*5*).

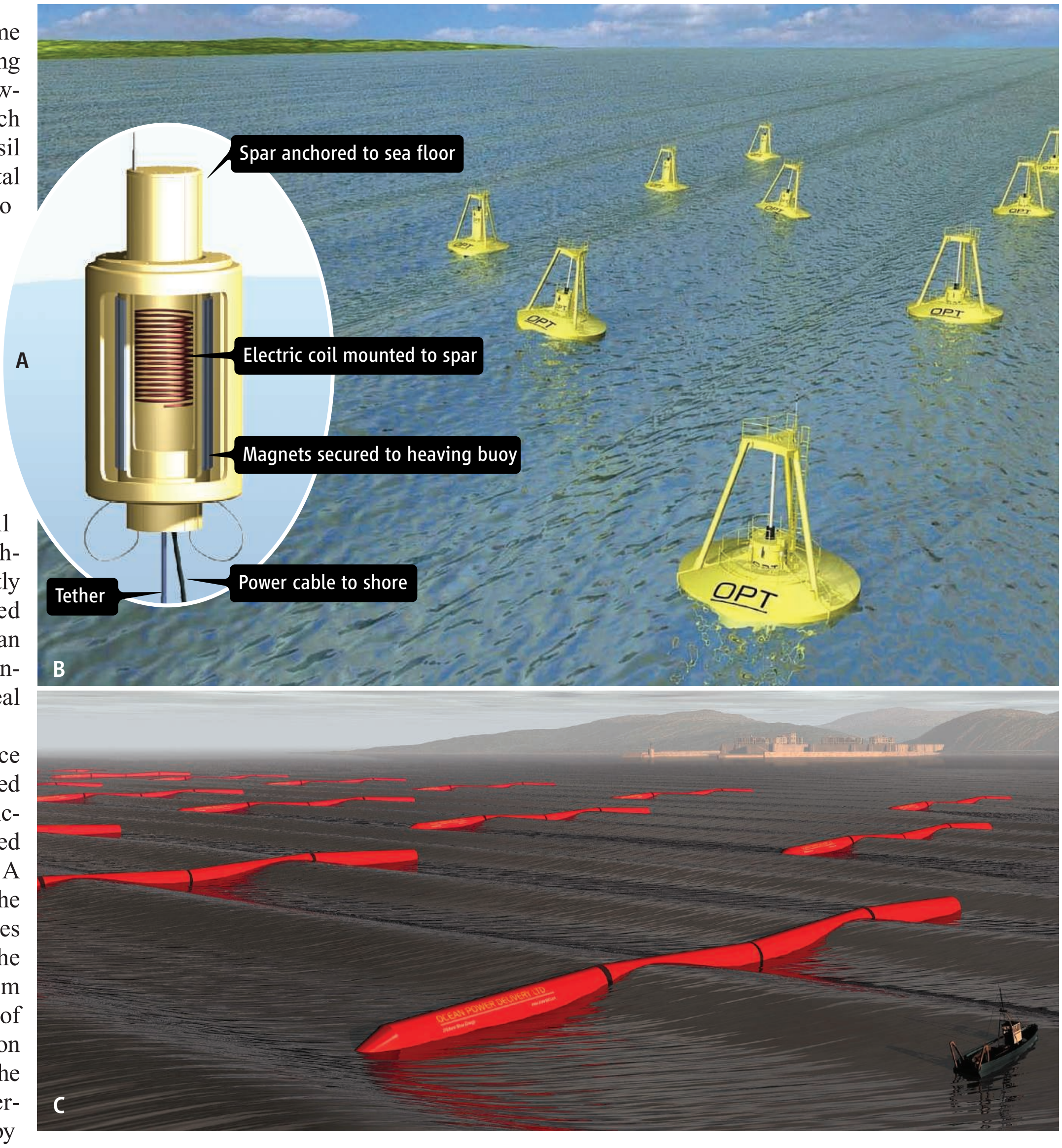
The amount of available energy in an ocean wave is proportional to both its oscillatory period and the square of its amplitude. Thus, optimal locations for power generation are often at depths between 40 and 100 m, where periods and amplitudes are both high. In more shallow depths, the energy tends to dissipate because of interactions with the sea bed. In deeper waters, it is often impractical and uneconomical to deploy a wave power plant and connect it to a shoreline power station.

Typically, wave energy converters (WECs) are nearly invisible from shore. Wave climates

---

[1] Department of Civil and Environmental Engineering, Duke University, Durham, NC 27708, USA. E-mail: jeff.scruggs@duke.edu [2] MMI Engineering, Houston, TX 77077, USA. E-mail: pjacob@mmiengineering.com

are also more predictable than some other renewable sources, including wind. Nonetheless, like all renewables, wave energy has a much smaller energy density than fossil fuels and requires higher capital investments (*6*). There are also considerable engineering challenges associated with deploying a WEC in the ocean environment that must be economically overcome in order for the technology to be reliable.



**Do the wave energy.** (**A**) Schematic of a direct-drive point absorber. (**B**) An array of Ocean Power Technologies point absorber WECs. (**C**) An array of Pelamis WECs.

CREDIT: ILLUSTRATIONS COURTESY OF OREGON STATE UNIVERSITY (A), OCEAN POWER TECHNOLOGIES INC. (B), AND PELAMIS WAVE POWER LTD. (C)

A wide variety of WECs have been proposed to extract power from the ocean (*7*). These designs differ in their structural configuration, as well as in the energy conversion technology they use. It is only recently that any have been demonstrated at full scale in depths of more than 40 m, and there is still no consensus on which type of device is ideal for which circumstances.

One of the most basic device types is the point absorber, depicted in panel A of the figure; this particular point absorber was developed at Oregon State University (*8*). A buoyant float, when excited by the heave motion of the waves, moves relative to a spar anchored to the ocean floor. Typically, the system resonates at the dominant period of the waves, and this resonant motion generates electricity. Panel B of the figure depicts an array of PowerBuoy point absorbers developed by New Jersey–based Ocean Power Technologies (*9*). The company is currently deploying such arrays at a number of sites, including 1.4- and 2.0-MW commercial projects off the coasts of Spain and Oregon, respectively.

Another type of WEC, called an attenuator, uses a slender structure that is comparable in length to an ocean wave and oriented along the waves' propagation direction. One such device, developed by Edinburgh-based Pelamis Wave Power, has reached commercial availability. It consists of four 30-m-long articulated cylinders connected by universal joints (see the figure, panel C). Each joint houses hydraulic systems that generate energy from wave-induced bending motions. In 2008, Pelamis deployed three attenuators with a combined rating of 2.25 MW off the coast of Portugal (*10*).

A third class of WEC, generally called a terminator, is a wide structure oriented along the axis of the wave crest that, like a breakwater, absorbs energy from an incoming wave. One example is the floating "duck" converter (so named for its cam-shaped cross section), proposed in (*1*) and subsequently developed at the University of Edinburgh. Other types of terminators include oscillating water columns and overtopping devices, which process wave power in an entirely different manner from the others discussed here.

Power conversion in a WEC can be accomplished by a direct-drive generator (i.e., panel A of the figure), which uses permanent magnets to generate energy directly from the primary mechanical motion. However, many WECs make use of intermediary hydraulic systems to transfer wave energy to pressure accumulators, which in turn drive rotary generators.

Wave energy is concentrated at low frequencies and at low, alternating velocities, which makes efficient conversion and transmission to a grid difficult and limits the options for efficient power takeoff technology. Historically, hydraulic systems have had the advantage: Direct-drive generators require permanent magnets and a large amount of iron, so they are considerably heavier. However, this weight disparity is being reduced as a result of advances in both efficient power electronics and magnetic materials. Direct-drive generators have been used in some full-scale technologies, such as the AWS point absorber (*7*) developed in the Netherlands by Teamwork Technologies BV, in part because their mechanical simplicity makes them more reliable.

Generally, power takeoff systems require a control system in order to maximize energy transfer. Since the early 1980s, a well-understood theory has existed for the optimal absorption of energy by WEC arrays of various geometries, for waves oscillating at a single frequency (*11*, *12*). Fundamental limits exist for a WEC's capture width (the lateral width of an incoming wave crest with incident power equal to that being absorbed). Optimal control theory has shown that the capture width can be many times the converter's physical width (i.e., the converter can attract energy propagating on either side of it, in addition to incident energy), and that multiple converters can exploit constructive interference to enhance power absorption.

WECs are typically tuned to extract power from only one frequency. Real waves, however, exhibit random oscillation with available power over a range of frequencies. Mathematically, the optimal manner in which to harvest power from random waves requires knowledge of their future behavior. The use of a statistical estimator to predict future incident waves, and its incorporation into a WEC's power generation control system, therefore has the potential to greatly enhance its capture width in random waves.

The power takeoff and control system form the core of a WEC. They must be supported and protected by a structure that is either founded on the seabed or held on station by a mooring system. The marine environment is aggressive—seawater is corrosive, marine organisms cause fouling, and extreme storms impart very large loads to the WECs. Structural and mooring system designs need to address survivability and reliability, without compromising conversion efficiency.

The wave energy industry faces a number of engineering and economic challenges. More efficient designs often have greater initial costs, so development will necessarily be incremental and will take both time and investment. Beyond power plant design, current needs include investigations of the environmental impacts of WECs, as well as policy and permitting issues. Despite these challenges, wave energy is a large and viable source of renewable power, and its development deserves serious attention.

**References**

1. S. H. Salter, *Nature* **249**, 720 (1974).
2. J. Callaghan, R. Boud, *Future Marine Energy*, Rep. CTC601, 2006 (www.carbontrust.co.uk/Publications/publicationdetail.htm?productid=CTC601&metaNoCache=1).
3 M. T. Pontes, *J. Offshore Mech. Arctic Eng.* **120**, 226 (1998).
4. G. Hagerman, R. Bedard, *E2I EPRI Specification: Guidelines for Preliminary Estimation of Power Production by Offshore Wave Energy Conversion Devices*, 2003 (oceanenergy.epri.com/attachments/wave/reports/001_WEC_Power_Production.pdf).
5. R. Bedard *et al.*, *Final Summary Report: Project Definition Study: Offshore Wave Power Feasibility Demonstration Project*, 2005 (oceanenergy.epri.com/attachments/wave/reports/009_Final_Report_RB_Rev_2_092205.pdf).
6 J. Baxter, *Energy for Life*, 122nd George Stevenson Presidential Address to the Institution of Mechanical Engineers (http://presidentschoice.imeche.org.uk/president122.htm).
7. J. Cruz, Ed., *Ocean Wave Energy: Current Status and Future Perspectives* (Springer, Berlin, 2008).
8. A. von Jouanne, *Mech. Eng. Mag.* **128**, (12), 24 (2006).
9. Ocean Power Technologies Inc. (www.oceanpowertechnologies.com).
10. Pelamis Wave Power Ltd. (www.pelamiswave.com).
11. D. V. Evans, *Annu. Rev. Fluid Mech.* **13**, 157 (1981).
12. J. Falnes, *Ocean Waves and Oscillating Systems, Linear Interaction Including Wave Energy Extraction* (Cambridge Univ. Press, Cambridge, 2002).

10.1126/science.1168245

---

APPLIED PHYSICS

# Looking Below the Surface

Scanning tunneling microscopy can now be used to determine electronic properties of bulk materials.

**Andreas J. Heinrich**

Every once in a while, experiments are reported that apply existing tools to apparently well-understood scientific concepts and come up with tantalizing, novel results. On page 1190 of this issue, such a case is beautifully demonstrated by Weismann *et al.* (*1*). They use scanning tunneling microscopy (STM), a standard surface science technique, to visualize electron flow in the bulk of a piece of copper.

Primarily based on its atomic resolution imaging capability, the STM has had phenomenal success in the field of surface science. How can a truly surface-sensitive technique be used to measure a bulk property? The key trick applied by Weismann *et al.* is to exploit the wave nature of the electrons in copper and study their interference patterns on the surface caused by scattering centers in the bulk of the material. Their technique opens the door to a real-space investigation of electron propagation in materials and to the scattering of electrons at defects well below the surface.

A finite-size piece of copper contains a huge number of atoms and an even larger number of electrons. It might, therefore, seem impossible to understand the behavior of such a complex system. Yet, solid-state physics enables most measurable properties to be characterized by the properties of the Fermi surface—the dividing line between the states occupied by an electron and the empty states. The Fermi surface is most often presented in reciprocal space, where one can characterize a spatially extended state such as an electron wave with a single parameter, its wave number.

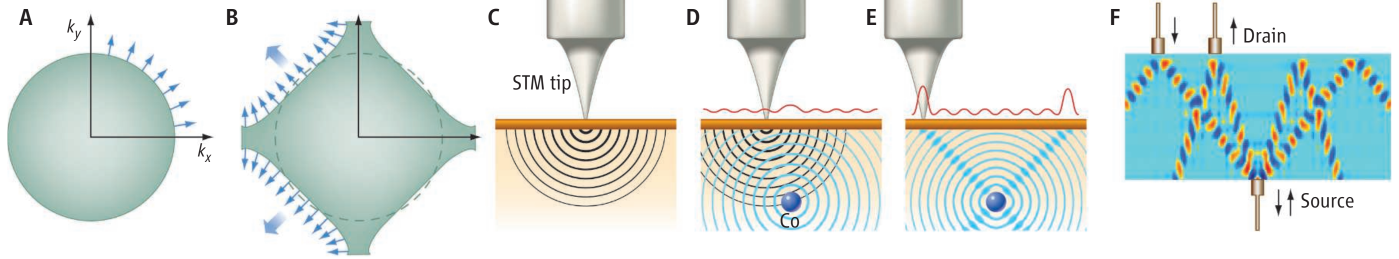
The simplest model system for such a solid is the nearly-free-electron gas (*2*), which is spatially homogeneous and, as a result, its Fermi surface is a perfect sphere (see the figure, panel A). Prominent examples of nearly-free-electron gas metals are gold, silver, and copper. The primary deviation from the spherical shape in these materials is an exclusion of allowed states along certain crystal directions (see the figure, panel B). It is this small deviation from the spherical that allows Weismann *et al.* to measure the shape of the bulk Fermi surface.

Most frequently, the shape of a Fermi surface is measured with quantum oscillations and angle-resolved photoemission spectroscopy and calculated with density functional theory. Because only electrons in a small band around the Fermi surface are responsible for electronic transport, the shape of the Fermi surface therefore plays a dominant role in the electrical conductance of solids. It is not surprising, then, that recent years have seen a large body of work measuring the Fermi surface of high-temperature superconductors (*3*) and other advanced materials.

The main reason why the STM can be used to measure these bulk properties is conceptually rather simple: The STM consists of two components, the sample (comprising the surface and the bulk material under the surface) and the tip. The atomic resolution of the STM at the surface stems from the extreme sensitivity of the tunneling current to changes in the overlap of the electron wave function at the tip and the wave functions at the surface. Once an electron has tunneled and arrives in the sample as a hot electron (a state outside of the Fermi surface), it propagates as a wave inside the material and eventually scatters or arrives in the back contact to complete the electrical circuit (see the figure, panel C).

---

Almaden Research Center, IBM Research Division, San Jose, CA 95120, USA. E-mail: heinrich@almaden.ibm.com



**Scattered electrons.** (**A**) A nearly-free-electron gas has a spherical Fermi surface. The blue arrows indicate the direction of electron propagation at the Fermi surface. (**B**) In the cartoon model of the Fermi surface of Cu, certain directions become preferred due to the nonspherical shape of the Fermi surface. The thick arrows indicate directions of electron focusing. (**C**) In a typical STM experiment on a metal, an electron tunnels into the surface and becomes a bulk electron wave whose amplitude decays with distance. (**D**) When a scatterer is present under the surface, the electron wave can be reflected. For a spherical Fermi surface, this results in a weak interference pattern at the surface. (**E**) When the Fermi surface is not spherical, electron focusing is observed along certain directions, which can give rise to a pronounced interference pattern observable at the surface. (**F**) Theoretical prediction of separate interference patterns for different spin channels in a magnetic material [from Weismann *et al*. (*1*)].

This wave behavior of the electrons in the bulk of the sample is not visible in most STM images and is thus typically neglected in the analysis of STM experiments.

The situation changes dramatically when a point defect is incorporated under the surface. Such a defect can scatter the electron waves emanating from the tunneling tip. The reflected wave can interfere with the incoming wave, giving rise to a standing-wave pattern that can be seen at the surface. For the case of a spherical Fermi surface, the amplitude of the scattered electron wave decays rapidly, and only a very weak interference pattern can be expected on the surface (see the figure, panel D). Weismann *et al.* see a dramatic increase of this interference pattern at the surface for Co atoms buried several layers underneath (see the figure, panel E), and argue that this can be understood from the shape of the Fermi surface: Along certain spatial directions, the amplitude of the scattered wave decays very slowly (arrows in panel B; see supplementary movies S1 and S2). In essence, the electrons are scattered along beams of electron waves, a phenomenon referred to as electron focusing. When these beams intersect the surface of the material, a strong and characteristically shaped interference pattern is observed. This interference pattern reflects information about the propagation of electrons through the bulk of the material—and hence on the shape of the Fermi surface—and the strength and type of scattering potential below the surface. Weismann *et al.* show that these interference patterns can be accurately calculated by incorporating a very large number of atoms in the sample.

The observation of electron interference patterns on surfaces with STM goes back to the beautiful standing-wave patterns of electrons confined to the inside of a quantum corral on copper (*4*). More recently, the wave nature of electrons in two-dimensional electron gases at surfaces has been used to perform electron holography (*5*) and to study the electron propagation in high-temperature superconductors (*6*). In the latter case, one can deduce a plethora of spatially resolved information on the electron behavior in such partially disordered systems with complex electron-electron interactions.

Weismann *et al.* also use their calculational approach to highlight a wide range of exciting future experiments. They discuss the fact that electrons of different spin character in magnetic materials generally have differing Fermi surfaces. This should enable the observation of separate interference patterns for injecting minority spin versus majority spin electrons (see the figure, panel F). The technique may also be used to study buried interfaces with high spatial resolution. The system used in the present study is a prototypical Kondo system—a single magnetic impurity in a sea of electrons—and one should be able to obtain deeper insights into electron scattering above and below its characteristic magnetic transition temperature. Interpreted correctly, one can therefore judge a book by its cover.

**References**

1. A. Weismann *et al.*, *Science* **323**, 1190 (2009).
2. N. W. Ashcroft, N. D. Mermin, *Solid State Physics* (Harcourt, New York, 1976).
3. Z.-X. Shen *et al.*, *Phys. Rev. Lett.* **70**, 1553 (1993).
4. M. F. Crommie, C. P. Lutz, D. M. Eigler, *Science* **262**, 218 (1993).
5. C. R. Moon, L. S. Mattos, B. K. Foster, G. Zeltzer, H. C. Manoharan, *Nat. Nanotechnol.* 10.1038/nnano.2008.415 (2009).
6. Y. Kohsaka *et al.*, *Science* **315**, 1380 (2007).

**Supporting Online Material**

www.sciencemag.org/cgi/content/full/323/5718/1178/DC1
Movies S1 and S2

10.1126/science.1170744

CREDIT: ADAPTED BY P. HUEY/SCIENCE

---

PSYCHOLOGY

# From Oral to Moral

**Paul Rozin,[1] Jonathan Haidt,[2] Katrina Fincher[1]**

Is moral disgust an elaboration of a food rejection system?

The term "disgusting" is applied to bad tastes, cockroaches, incest, and proposing an unfair division of money in an ultimatum game. Is the emotional response the same in all four cases? On page 1222 of this issue, Chapman *et al*. (*1*) show that there is activation of a muscle central to the facial expression of disgust in response to unfair treatment (divisions of money), and argue that it "elicits the same disgust as disease vectors and bad tastes." What does that mean, and how would you demonstrate it?

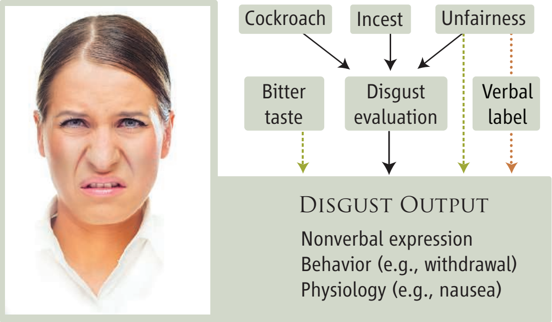
One possible model to consider is a temporal analysis of disgust comprising three layers. At the top are the elicitors of disgust. To one degree or another, these trigger a set of mental activities that can be considered a "disgust evaluation system" (see the figure) that appraises the elicitor, generates a sense of offensiveness and revulsion, and leads to thoughts of "contamination." Psychological contamination refers to the feeling or belief that when something offensive touches something else, the offensiveness is transferred to the contacted object (thus, when a

---

[1]Department of Psychology, University of Pennsylvania, Philadelphia, PA 19104, USA. [2]Department of Psychology, University of Virginia, Charlottesville, VA 22904, USA. E-mail: rozin@psych.upenn.edu

sterilized cockroach is dipped into a glass of juice, the juice becomes offensive). This neural system in turn triggers a disgust output program, including emotional expressions, behaviors, and physiological responses such as nausea.

According to the principle of preadaptation, a system that evolves for one purpose is later used for another purpose. From this viewpoint, disgust originates in the mammalian bitter taste rejection system, which directly activates a disgust output system. This primal route (e.g., bitter and some other tastes) evokes only the output program, without a disgust evaluation phase. During human evolution, the disgust output system was harnessed to a disgust evaluation system that responded not to simple sensory inputs (such as bitter tastes) but to more cognitively elaborated appraisals (e.g., a cockroach). Initially, the evaluation system was a food rejection system that rejected potential foods on the basis of their nature or perceived origin. This was the first "true disgust," because it engaged this evaluation system. Later, through some combination of biological and cultural evolution, the eliciting category was enlarged to include reminders of our animal nature, as wel as some people or social groups (*2*). This process had adaptive value, because by making things or thoughts disgusting a culture could communicate their negativity and cause withdrawal from them.



**Domains of disgust.** The schematic represents routes by which eliciting situations may trigger the disgust output program. Those that run through the disgust evaluation system—which includes appraisal of the elicitor, feelings, and contamination ideation—trigger the full disgust emotion. Solid lines represent routes through which an elicitor can activate the disgust evaluation-output program. Dashed lines (green) represent direct elicitation of the disgust output program. The dotted line (brown) represents a metaphoric, indirect route.

It has also been proposed that the disgust evaluation system was further extended (*2*, *3*) to a class of moral offenses involving violations of purity and sacredness, described by anthropologist Richard Shweder as "the ethics of divinity" in a taxonomy of three widely found clusters of moral meanings (*4*). However, recent evidence indicates that disgust may also be elicited by violations of fairness and justice (Shweder's ethic of autonomy). Autonomy violations are typically associated with anger (*3*).

According to a possible three-layer scheme of disgust analysis, there are three pathways through which an elicitor could activate the disgust output program (see the figure). The core route elicits a set of disgust evaluations (appraisals, feelings, and contamination cognitions), which in turn lead to the disgust output. What about routes involving moral violations of incest and unfairness? It may be that incest and other corporeal (divinity) violations activate the disgust evaluation system, just as do elicitors of core disgust. If unfairness and other moral violations that have no corporeal element trigger the disgust evaluation system, then they represent the furthest expansion of the "oral to moral" evolution of disgust. But does the evidence of Chapman *et al.* indicate such an expansion?

Alternative views by a number of scholars propose that the link between morality and disgust is largely a metaphor (*5*), construed as such because it bypasses the disgust evaluation system. But the link is not "just" a metaphor. Unfairness and other moral violations may directly affect the disgust output system, after processing by some other evaluation system, or these violations might simply activate the verbal label "disgust," which would then activate the disgust output system. The outcome of either route would include the facial expression of disgust. The Chapman *et al.* observations are consistent with both these alternative routes as well as the one that uses the disgust evaluation system. But only if evidence is found for a route from unfairness to the disgust evaluation system can it be concluded that disgust at unfairness is "the same" as disgust that is elicited through the core route (such as in response to cockroaches).

There is evidence that violations of the ethics of divinity (especially violations of food and sex taboos) engage the full disgust evaluation output. People feel disgust for divinity violations (*3*). There is also a link between incest and oral inhibition (such as nausea, gagging, and loss of appetite) (*6*). And contaminating cognitions accompany divinity violations (*7*).

A few studies suggest that fairness violations might indeed activate the disgust evaluation program, at least to some degree. Cleansing actions (related to purity and divinity) influence moral judgments about autonomy as well as divinity violations (*8*). Priming disgust, through exposure to disgust-eliciting material, makes subsequent moral judgments of both divinity and autonomy violations more severe (*9*). Divinity and fairness violations activate parts of the brain (particularly the anterior insula) that are also activated by core disgust (*10*), but the anterior insula is not uniquely associated with disgust (and vice versa).

If Chapman *et al.* are correct in that unfair divisions of money activate "the same" disgust as is activated by cockroaches, then it is almost surely mixed with anger, the prototypical emotion for autonomy violations (*3*). Anger is also an emotion that often activates the raised upper lip (*11*), the main disgust marker that Chapman *et al.* relied on. Until studies examine the effects of a variety of elicitors on a variety of dependent measures (e.g., contamination, appraisals, and feelings), it is unclear whether it's "the same" disgust, or just some common elements in the output system. Moreover, there are probably important variations in the evaluative and output systems for different types of disgust (*12*).

Even if the evolutionary and development history of disgust is indeed "oral to moral," it does not follow that a modern person's experience of moral disgust has to have an oral aura. But it appears that there is quite a bit of oral in moral experience, almost certainly for divinity violations and perhaps even for autonomy violations.

### References and Notes

1. H. A. Chapman, D. A. Kim, J. M. Susskind, A. K. Anderson, *Science* **323**, 1222 (2009).
2. P. Rozin, J. Haidt, C. R. McCauley, in *Handbook of Emotions*, M. Lewis, J. Haviland, Eds. (Guilford, New York, 1993), pp. 575–594.
3. P. Rozin, L. Lowery, S. Imada, J. Haidt, *J. Pers. Soc. Psychol.* **76**, 574 (1999).
4. R. A. Shweder, N. C. Much, M. Mahapatra, L. Park, in *Morality and Health*, A. Brandt, P. Rozin, Eds. (Routledge, New York, 1997), pp. 119–169.
5. E. B. Royzman, J. Sabini, *J. Theory Soc. Behav.* **31**, 29 (2001).
6. E. B. Royzman, R. F. Leeman, J. Sabini, *Motiv. Emotion* **32**, 100 (2008).
7. P. Rozin, L. Millman, C. Nemeroff, *J. Pers. Soc. Psychol.* **50**, 703 (1986).
8. C.-B. Zhong, K. Liljenquist, *Science* **313**, 1451 (2006).
9. S. Schnall, J. Haidt, G. L. Clore, A. H. Jordan, *Pers. Soc. Psychol. Bull.* **34**, 1096 (2008).
10. A. G. Sanfey, J. K. Rilling, J. A. Aronson, L. E. Nystrom, J. D. Cohen, *Science* **300**, 1755 (2003).
11. P. Rozin, L. Lowery, R. Ebert, *J. Pers. Soc. Psychol.* **66**, 870 (1994).
12. J. Simpson, S. Carter, S. H. Anthony, P. G. Overton, *Motiv. Emotion* **30**, 31 (2006).
13. We thank E. Royzman for valuable comments on the manuscript and the ideas in it.

10.1126/science.1170492

PHOTO CREDIT: SUPERSTOCK

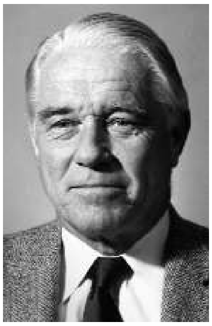
RETROSPECTIVE

# Frederic M. Richards (1925–2009)

A pioneer of biochemistry and structural biology overturned conceived wisdom about protein structure.

**Thomas A. Steitz**

Frederic Middlebrook Richards, a world figure in the field of structural biology, died at his home in Guilford, Connecticut, on 11 January at the age of 83. Sterling Professor of Molecular Biophysics and Biochemistry emeritus at Yale, he made seminal contributions to our understanding of protein structure and function, served in an innovative leadership capacity in many organizations, and was a leader in the development of structural biology at Yale.



In his paradigm-shifting experiment, published in 1958, Fred Richards showed that two fragments of the enzyme ribonuclease A (RNase A) that are separately inactive could be reconstituted to form an active enzyme. During postdoctoral research at the Carlsberg laboratory in 1952, he had demonstrated that the addition of the protease, subtilisin, to RNase A resulted in a cleaved product, RNase S, that remained active. After starting his own lab at Yale, Fred was able to separate the cleaved enzyme into a 20-residue S peptide and a 102-residue S protein, which lacked enzymatic activity. However, when the S peptide was added back to a solution of the S protein, the activity was recovered.

This experiment overturned then current views of protein structure, providing the first evidence that the amino acid sequence of a protein determines the three-dimensional structure it adopts. It anticipated the subsequent discovery that the unfolded whole RNase A protein could be refolded in a test tube to form an active enzyme, a finding that resulted in the award of a Nobel Prize to Christian Anfinsen in 1972—a prize that Richards could arguably have shared. In the 1960s, Richards teamed up with a faculty colleague, Harold Wyckoff, to determine the atomic structure of RNase S, which in 1967 tied with three other proteins as the third protein structure to be established after myoglobin and lysozyme.

Fred Richards was born in New York and obtained his B.S. degree at the Massachusetts Institute of Technology in 1948 and his Ph.D. at Harvard in 1952, working with Barbara Low in the medical school department led by Edsall and Cohen. Thereafter, he pursued postdoctoral research at the Carlsberg laboratory in Copenhagen, Denmark, and at Cambridge University in England. In 1955, he joined the faculty of the Biochemistry Department in the Medical School of Yale University, and in 1963, at the age of 38, he moved to the Science Hill campus at the request of President Kingman Brewster to chair the Biophysics Department, which at the time had a radiological focus. His mandate was to redirect the department to the new field of molecular biology. After a sabbatical at Oxford in 1967–1968, and again at the request of Brewster, he took on the task of merging the Departments of Biochemistry and of Biophysics to form the Department of Molecular Biophysics and Biochemistry.

Fred was an inspiring and charismatic leader, always with a great smile. In his two terms as chair of Biophysics (1963 to 1967) and then Molecular Biophysics and Biochemistry (1969 to 1973), he initiated the development of one of the major centers in the world for the study of biophysics and structural biology, hiring 10 faculty, 7 of whom are now members of the National Academy of Sciences. In doing so he brought together a group of structural biologists that pushed the frontiers of this emerging field.

Fred Richards's most seminal discoveries were often obtained with approaches that were uniquely different from the general direction of the field, and were often executed by himself. While on sabbatical in Oxford in 1968, he constructed a large device employing a half-silvered mirror that greatly simplified the task of fitting atomic models of proteins into experimentally obtained electron density maps. Perhaps because of its unusual construction, it was initially known as "Fred's folly"; but when the rest of the structural biology community found how useful it was, it was respectfully renamed "The Richards optical comparator."

Richards also developed computational methods for understanding the nature, stability, and function of proteins from the knowledge of their atomic structures. His method of calculating their solvent-exposed surface areas has been used ever since to analyze the interactions between proteins and between proteins and ligands. Another of his methods used Voronoi polyhedra to compute volumes occupied by all atoms of a protein and to evaluate the tightness of its packing density. Additional insights into how proteins might interact with ligands large and small were provided by methods developed in his lab to calculate the electrostatic charge distribution on a protein surface, which were used initially to predict how two proteins—flavodoxin and cytochrome c—interact.

Richards was a member of the National Academy of Sciences, the American Academy of Arts and Sciences, and the American Philosophical Society. He took on many tasks that served the community at large, including the scientific directorship of the Jane Coffin Childs Memorial Fund for Medical Research (1976–1991), which awards postdoctoral fellowships, and the presidency of the Biophysical Society and of the now named American Society for Biochemistry and Molecular Biology. He was awarded the Connecticut Medal of Science in 1995.

Throughout his life, Fred had an intense love of sailing, a love shared by Chris Anfinsen (as well as their mutual interest in RNase A). On one occasion, they sailed together from Washington, DC, to Bermuda, and found they had with them only a U.S. road map as a chart. Fred also sailed from New England to England in 1967 on his way to Oxford to spend his sabbatical year in David Phillips's lab. He religiously took at least a month's holiday every August, usually sailing his boat off the coast of Canada with his wife, Sally, and friends (competing with the icebergs). Fred felt that in order to do outstanding science, it was necessary to get away and clear the brain. Fred's passion for making and fixing things using an enormous workshop in his garage enabled him to build a boat he named "Sally's Baarge," reflecting his wife's Maine roots. Fred also had a great concern for the conservation of the environment, both offshore and onshore, and contributed land, time, and effort to the Guilford Land Trust. He is survived by his wife, three children, and four grandchildren.

10.1126/science.1171157

CREDIT: M. MARSLAND/YALE UNIVERSITY

---

Molecular Biophysics and Biochemistry, Howard Hughes Medical Institute, Yale University, New Haven, CT 06520, USA. E-mail: thomas.steitz@yale.edu

# AAAS NEWS & NOTES

EDITED BY EDWARD W. LEMPINEN

SCIENCE POLICY

## Climate Change Already Dramatic in U.S., Experts Warn Congress

**Extreme weather.** Soaring greenhouse gas levels could spur more destructive weather like the 2008 U.S. Midwest floods.

Droughts, floods, fires, and crop failures—the extreme climate events of today—will become commonplace in the next 100 years if levels of greenhouse gas emissions continue to soar, a panel of prominent climate experts warned at two Capitol Hill briefings.

With U.S. President Barack Obama and Congress moving to shape new climate policy, the speakers on the panel assembled by AAAS and four other science groups urged swift, substantive action. Hoping to counteract any remaining doubts, they stressed that warming is already having a dramatic impact in the United States and around the world.

Reserves of mountain snow in the Western United States are declining, and most of the snowpack could be gone by the end of the century. Drought in the American Southwest, already decades old, could soon be exacerbated as rainfall plummets to levels last seen during the years of the Dust Bowl, the experts said.

"The impacts of climate change are not only unavoidable, they are already occurring, posing significant threats to critical resources like the availability and quality of water in the United States," said Peter Gleick, president and cofounder of the Pacific Institute for Studies in Development, Environment and Security.

At the briefings for Senate and House staffers, Gleick and other top climate scholars outlined the scientific perspective on climate change and offered some concrete suggestions for revising current law and environmental oversight to help humans mitigate and adapt to the effects of climate change.

"The beginning of a new Congress and administration provided an opportunity to have these leading scientists inform the policy debates that are to come, with an update on the current state of climate science and options," said Kasey White, a senior program associate at AAAS's Center for Science, Technology and Congress, which helped organize the briefings.

Although mitigation strategies are still being pursued, adaptation may be the watc word of the near future as the effects of climate change become more apparent. Eleven of the highest global mean temperatures on record have occurred within the past 12 years, said Susan Solomon, a former co-chair of the Intergovernmental Panel on Climate Change. Recent studies by Solomon and colleagues suggest that if atmospheric carbon dioxide levels reach concentrations of 450 parts per million by midcentury—a likely scenario, according to the researchers—rising sea levels and severe fluctuations in rainfall are inevitable.

"Every year of emissions means a commitment to climate change for more than 30 generations," said Solomon, now a senior scientist at the U.S. National Oceanographic and Atmospheric Administration.

Ted Parsons, a University of Michigan law and environment professor, warned the staffers that the United States is "currently going in the wrong direction" in limiting climate change emissions. To bring emissions down, he noted, nations must transform the global energy structure through renewable fuels, carbon capture technologies, and more efficient energy consumption.

U.S. lawmakers could speed this process by supporting a "blended" plan that mixes a carbon cap-and-trade system with some form of emissions tax, Parsons said. He suggested that "markets should choose" which renewable energy sources and climate-friendly technologies become part of an emissions control plan, but that governments could help by raising the performance efficiency standards of cars and appliances.

The rest of the world will be watching closely as the Obama administration and the new Congress decide on a course for climate change policy, said former U.S. Senator Timothy Wirth, president of the United Nations Foundation and the Better World Fund. The Kyoto Protocol, a global emissions reduction treaty never ratified by the U.S. Senate, expires in 2012. Wirth said U.N. climate talks in Copenhagen, Denmark, in December 2009 will be a key event in negotiating a new emissions treaty.

The 9 January briefings were moderated by AAAS CEO Alan I. Leshner and cosponsored by the Ecological Society of America, Geological Society of America, American Meteorological Society, and the Pew Center on Global Climate Change. The talks were held in conjunction with the House Science and Technology Committee, the Senate Energy and Natural Resources Committee, and the Senate Commerce, Science and Transportation Committee.

*—Benjamin Somers and Becky Ham*

EDUCATION

## AAAS Poll Seeks Ideas for White House Lectures

U.S. President Barack Obama has proposed a White House science lecture series, and users of a AAAS education resources Web site have some ideas for subjects that should be on the calendar: alternative energy, personalized medicine, and nanotechnology.

The AAAS Science NetLinks site created an informal poll in response to Obama's plan, and it has generated a strong response, said project manager Suzanne Thurston.

"We were really excited that Obama talked about science and specifically science education and public outreach, and we wanted to make sure that our users were aware of it," said Thurston. The program will forward the poll's results to the White House, she said.

Using resources at the site, students can research and write letters to the president in support of their poll vote. Other resources include a "speed-dating" game in which students seek funders for specific research projects, and a science café to debate the science policy priorities of the new administration.

Since 1997, Science NetLinks (www.sciencenetlinks.com) has been a collaborating partner of Thinkfinity, a free collection of educational resources and teacher training supported by the Verizon Foundation. The NetLinks project contributes materials tied to specific standards of learning such as *Benchmarks for Science Literacy,* developed by AAAS's Project 2061. *—Becky Ham*

CREDIT: TONY CAMPBELL/ISTOCK.COM

# Predicting Elections: Child's Play!

John Antonakis* and Olaf Dalgas

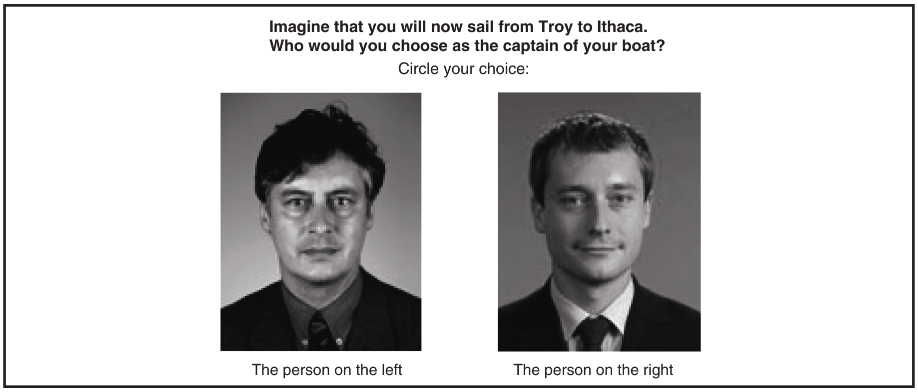
In *The Republic*, Plato states, "Imagine ... a ship in which there is a captain who is taller and stronger than any of the crew, but he is a little deaf and has a similar infirmity in sight, and his knowledge of navigation is not much better" (*1*). Plato argues that the crew (i.e., voters) cannot select a competent captain (i.e., ruler) because the crew is beguiled, in part, by appearances. Plato uses this allegory to suggest that voters lack the rational faculties and knowledge to elect competent rulers.

Ideally, democracies should elect politicians on their competence. Intellectual (or learning) ability predicts effective performance in complex domains (*2*) and in the U.S. presidency (*3*). Presidents, though, are not elected on ability (*4*). Among other factors, voters are biased by facial appearances; naïve raters can predict elections after simply rating the competence (i.e., intelligence, leadership, and competence per se) of political candidates from their photographs (*5*). On a general level, individuals automatically infer characteristics of social targets based on facial appearances (*6*). Voters acting rationally, however, should change these initial classifications as they receive information about the target individual's values, performance, political affiliation and the like. Unfortunately, voters are anchored in first impressions and do not appropriately correct initial inferences; additional information on the candidates does not change choices by much (*5*). Perhaps voters are acting knowledgeably if, after experience, they have learned that facial appearance correlates with competence or performance. However, intelligence of adults cannot be predicted from facial appearance (*7*), and there is great variation in the competence of politicians (*3*).

Why do naïve ratings and actual votes correlate? Are voters using the same rudimentary decision heuristics that children use? Facial stereotypes and other classification schemes are well developed even in infancy (*8*), probably stemming from an innate template and rapid early learning (*9*). We hypothesized that voters might still be using the same cues that children do to categorize individuals on competency, which explains why voters may largely ignore additional information on candidates. We tested our claim by examining whether naïve voters predict actual voter preferences in the same way that children do.

We recruited adults and children in Switzerland to rate pairs of faces (the winner and runner-up) from the run-off stages of the 2002 French parliamentary election (*10*). In experiment 1 ($N$ = 684 adults), results of a logistic regression showed that the probability of predicting an election result correctly on the basis of ratings of competence was 0.72. Ratings of competence also predicted margin of victory (standardized beta = 0.32, $P < 0.001$).

Using the same materials in experiment 2, 841 individuals—of whom 681 were children aged 5 to 13 years (mean age = 10.31, SD = 1.81)—participated in a game involving a computer-simulated trip from Troy to Ithaca. Thereafter, participants chose from two faces the captain of their boat (Fig. 1A). For the children, results from a logistic regression showed that probability of predicting an election result correctly on the basis of choice of captain was 0.71. The results did not differ when including the other participants ($N$ = 160, mean age = 30.49, SD = 16.32); prediction accuracy did not depend on age (fig. S1).



**Imagine that you will now sail from Troy to Ithaca.**
**Who would you choose as the captain of your boat?**
Circle your choice:

**Fig. 1.** An example of a pair of faces we used from the Meurthe-et-Moselle electoral district (number 1). Jean-Jacques Denis (left) lost to Laurent Hénart (right). Of the participants who rated this pair, 77% of children and 67% of adults in experiment 1 chose Hénart. Over the 57 pairs of faces and across all raters, the adults in experiment 1 chose correctly 60% of the time, likelihood ratio $\chi^2(1) = 28.86$, $\phi = 0.20$, $P < 0.001$; for experiment 2, both children and adults chose correctly 64% of the time, likelihood ratio $\chi^2(1) = 68.10$, $\phi = 0.28$, $P < 0.001$. These effects become stronger when controlling for covariates and fixed effects for pairs of faces (*10*).

Next, we compared the adults in experiment 1 to the children. We used the mean (i.e., at the pair-level) predicted probabilities for each pair of faces for both children and adults in a random-effects regression model. The variable indicating adults was unrelated to the predicted probabilities; again, child-adult response patterns were indistinguishable. Furthermore, children ratings strongly predicted the adult ratings (standardized regression beta = 0.61, $P < 0.001$). Face effects appear to be age-invariant, suggesting that adults and children use similar cues in judging competence from facial appearance.

Evidently, young children, who are less experienced than are adults in observing performance in complex domains, playing an innocuous game can predict election results retrospectively (*11*). These findings suggest that voters are not appropriately weighting performance-based information on political candidates when undertaking one of democracy's most important civic duties.

### References and Notes

1. Plato, *The Republic*, B. Jowett, Transl. (Collier, New York, 1901).
2. F. L. Schmidt, J. E. Hunter, *Psychol. Bull.* **124**, 262 (1998).
3. D. K. Simonton, in *Advances in Psychological Research*, S. P. Shohov, Ed. (Nova Science, Hauppauge, NY, 2002), vol. 14, pp. 143–153. Note, in Simonton's analysis, the standardized partial beta for relation between intelligence and presidential greatness is 0.29 (the zero-order correlation was 0.55); removing an endogenous predictor (i.e., the number of years in office) increases the beta of intelligence to 0.41.
4. If politicians were selected on ability, the correlation between ability and performance would be zero or very weak (given the range restriction in ability).
5. A. Todorov, A. N. Mandisodza, A. Goren, C. C. Hall, *Science* **308**, 1623 (2005).
6. R. Hassin, Y. Trope, *J. Pers. Soc. Psychol.* **78**, 837 (2000).
7. L. A. Zebrowitz, J. A. Hall, N. A. Murphy, G. Rhodes, *Pers. Soc. Psychol. Bull.* **28**, 238 (2002).
8. J. L. Ramsey, J. H. Langlois, R. A. Hoss, A. J. Rubenstein, A. M. Griffin, *Dev. Sci.* **7**, 201 (2004).
9. A. Slater, P. C. Quinn, *Infant Child Dev.* **10**, 21 (2001).
10. Materials, methods, and extended results are available as supporting material on *Science* Online.
11. In experiment 2, we also presented children with two pairs of faces (one from the recent Democratic party primary and the other from the U.S. election). The children correctly predicted the Obama-Clinton [likelihood-ratio $\chi^2(1) = 3.94$, $\phi = 0.54$, $P < 0.05$] and the Obama-McCain [likelihood-ratio $\chi^2(1) = 8.45$, $\phi = 0.81$, $P < 0.01$] election results.
12. We thank our university rectorate for support and J. P. Bonardi, M. Brulhart, F. Butera, S. Faulk, U. Hoffrage, P. Jacquart, F. Krings, R. Lalive, G. Palazzo, B. Sattler, C. Zehnder, and the OB group for helpful comments.

**Supporting Online Material**

www.sciencemag.org/cgi/content/full/323/5918/1183/DC1
Materials and Methods
SOM Text
Fig. S1

27 October 2008; accepted 16 December 2008
10.1126/science.1167748

---

Faculty of Business and Economics, University of Lausanne, 1015 Lausanne, Switzerland.

*To whom correspondence should be addressed. E-mail: john.antonakis@unil.ch

# Inducing a Magnetic Monopole with Topological Surface States

Xiao-Liang Qi,[1] Rundong Li,[1] Jiadong Zang,[2] Shou-Cheng Zhang[1]*

**Existence of the magnetic monopole is compatible with the fundamental laws of nature; however, this elusive particle has yet to be detected experimentally. We show theoretically that an electric charge near a topological surface state induces an image magnetic monopole charge due to the topological magneto-electric effect. The magnetic field generated by the image magnetic monopole may be experimentally measured, and the inverse square law of the field dependence can be determined quantitatively. We propose that this effect can be used to experimentally realize a gas of quantum particles carrying fractional statistics, consisting of the bound states of the electric charge and the image magnetic monopole charge.**

The electromagnetic response of a conventional insulator is described by a dielectric constant ε and a magnetic permeability μ. An electric field induces an electric polarization, whereas a magnetic field induces a magnetic polarization. As both the electric field **E**($x$) and the magnetic induction **B**($x$) are well defined inside an insulator, the linear response of a conventional insulator can be fully described by the effective action $S_0 = \frac{1}{8\pi}\int d^3xdt\left(\varepsilon\mathbf{E}^2 - \frac{1}{\mu}\mathbf{B}^2\right)$, where $d^3xdt$ is the volume element of space and time. However, in general, another possible term is allowed in the effective action, which is quadratic in the electromagnetic field, contains the same number of derivatives of the electromagnetic potential, and is rotationally invariant; this term is given by $S_\theta = \left(\frac{\theta}{2\pi}\right)\left(\frac{\alpha}{2\pi}\right)\int d^3xdt\mathbf{E}\cdot\mathbf{B}$. Here, $\alpha = \frac{e^2}{\hbar c}$ (where $\hbar$ is Planck's constant $h$ divided by 2π and $c$ is the speed of light) is the fine-structure constant, and θ can be viewed as a phenomenological parameter in the sense of the effective Landau-Ginzburg theory. This term describes the magneto-electric effect (*1*), where an electric field can induce a magnetic polarization, and a magnetic field can induce an electric polarization.

Unlike conventional terms in the Landau-Ginzburg effective actions, the integrand in $S_\theta$ is a total derivative term, when **E**($x$) and **B**($x$) are expressed in terms of the electromagnetic vector potential (where $\partial^\mu$ denotes the partial derivative; μ, ν, ρ, and τ denote the spacetime coordinates; $F^{\mu\nu}$ is the electromagnetic field tensor; and $A^\mu$ is the electromagnetic potential)

$$\begin{aligned} \mathrm{S}_\theta &= \frac{\theta}{2\pi}\frac{\alpha}{16\pi}\int d^3xdt\varepsilon_{\mu\nu\rho\tau}F^{\mu\nu}F^{\rho\tau} \\ &= \frac{\theta}{2\pi}\frac{\alpha}{4\pi}\int d^3xdt\partial^\mu(\varepsilon_{\mu\nu\rho\sigma}A^\nu\partial^\rho A^\tau) \end{aligned} \tag{1}$$

Furthermore, when a periodic boundary condition is imposed in both the spatial and temporal directions, the integral of such a total derivative term is always quantized to be an integer; i.e., $\frac{S_\theta}{h} = \theta n$ (where $n$ is an integer). Therefore, the partition function and all physically measurable quantities are invariant when the θ parameter is shifted by 2π times an integer (*2*). Under time-reversal symmetry, $e^{i\theta n}$ is transformed into $e^{-i\theta n}$ (here, $i^2 = -1$). Therefore, all time-reversal invariant insulators fall into two general classes, described by either θ = 0 or θ = π (*3*). These two time-reversal invariant classes are disconnected, and they can only be connected continuously by time-reversal breaking perturbations. This classification of time-reversal invariant insulators in terms of the two possible values of the θ parameter is generally valid for insulators with arbitrary interactions (*3*). The effective action contains the complete description of the electromagnetic response of topological insulators. Topological insulators have an energy gap in the bulk, but gapless surface states protected by the time-reversal symmetry. We have shown (*3*) that such a general definition of a topological insulator reduces to the $Z_2$ topological insulators described in (*4*–*6*) for non-interacting band insulators; this finding is a three-dimensional (3D) generalization of the quantum spin Hall insulator in two dimensions (*7*–*10*). For generic band insulators, the parameter θ has a microscopic expression of the momentum space Chern-Simons form (*3*, *11*). Recently, experimental evidence of the topologically nontrivial surface states has been observed in $\mathrm{Bi}_{1-x}\mathrm{Sb}_x$ alloy (*12*), which supports the theoretical prediction that $\mathrm{Bi}_{1-x}\mathrm{Sb}_x$ is a $Z_2$ topological insulator (*4*).

With periodic temporal and spatial boundary conditions, the partition function is periodic in θ under the 2π shift, and the system is invariant under the time-reversal symmetry at θ = 0 and θ = π. However, with open boundary conditions, the partition function is no longer periodic in θ, and time-reversal symmetry is generally broken (but only on the boundary), even when θ = (2$n$ + 1)π. Our work in (*3*) gives the following physical interpretation: Time-reversal invariant topological insulators have a bulk energy gap but have gapless excitations with an odd number of Dirac cones on the surface. When the surface is coated with a thin magnetic film, time-reversal symmetry is broken, and an energy gap also opens up at the surface. In this case, the low-energy theory is completely determined by the surface term in Eq. 1. As the surface term is a Chern-Simons term, it describes the quantum Hall effect on the surface. From the general Chern-Simons-Landau-Ginzburg theory of the quantum Hall effect (*13*), we know that the coefficient θ = (2$n$+1)π gives a quantized Hall conductance of $\sigma_{xy} = \left(n + \frac{1}{2}\right)\frac{e^2}{h}$. This quantized Hall effect on the surface is the physical origin behind the topological magneto-electric (TME) effect. Under an applied electric field, a quantized Hall current is induced on the surface, which in turn generates a magnetic polarization and vice versa.

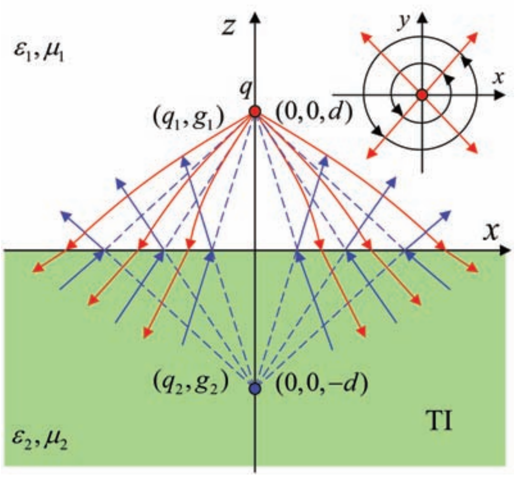


**Fig. 1.** Illustration of the image charge and monopole of a point-like electric charge. The lower-half space is occupied by a topological insulator (TI) with dielectric constant $\varepsilon_2$ and magnetic permeability $\mu_2$. The upper-half space is occupied by a topologically trivial insulator (or vacuum) with dielectric constant $\varepsilon_1$ and magnetic permeability $\mu_1$. A point electric charge $q$ is located at (0, 0, $d$). When seen from the lower-half space, the image electric charge $q_1$ and magnetic monopole $g_1$ are at (0, 0, $d$); when seen from the upper-half space, the image electric charge $q_2$ and magnetic monopole $g_2$ are at (0, 0, −$d$). The red solid lines represent the electric field lines, and blue solid lines represent magnetic field lines. (Inset) Top-down view showing the in-plane component of the electric field at the surface (red arrows) and the circulating surface current (black circles).

---

[1] Department of Physics, Stanford University, Stanford, CA 94305–4045, USA. [2] Department of Physics, Fudan University, Shanghai, 200433, China.

*To whom correspondence should be addressed. E-mail: sczhang@stanford.edu

We propose a manifestation of the TME effect. When a charged particle is brought close to the surface of a topological insulator, a magnetic monopole charge is induced as a mirror image of the electric charge. The full set of electromagnetic field equations can be obtained from the functional variation of the action $S_0 + S_\theta$ (*14*), and they can be presented as conventional Maxwell's equations but with the modified constituent equations describing the TME effect (*3*)

$$\begin{aligned} \mathbf{D} &= \mathbf{E} + 4\pi\mathbf{P} - 2\alpha P_3\mathbf{B} \\ \mathbf{H} &= \mathbf{B} - 4\pi\mathbf{M} + 2\alpha P_3\mathbf{E} \end{aligned} \qquad (2)$$

where $P_3(x) = \theta(x)/2\pi$ is the magneto-electric polarization (*3*), **D** is the electric displacement, **P** is the electric polarization, **H** is the magnetic field, and **M** is the magnetization. It takes the value of $P_3 = 0$ in vacuum or conventional insulators and $P_3 = \pm 1/2$ in topological insulators, with the sign determined by the direction of the surface magnetization.

Now consider the geometry as shown in Fig. 1. The lower-half space ($z < 0$) is occupied by a topological insulator with a dielectric constant $\varepsilon_2$ and a magnetic permeability $\mu_2$, whereas the upper-half space ($z > 0$) is occupied by a conventional insulator with a dielectric constant $\varepsilon_1$ and a magnetic permeability $\mu_1$. A point electric charge $q$ is located at $(0, 0, d)$ with $d > 0$. The Maxwell equations, along with the modified constituent equations and the standard boundary conditions, constitute a complete boundary value problem. To solve this problem, the method of images (*15*) can be used. We assume that, in the lower-half space, the electric field is given by an effective point charge $q/\varepsilon_1$ and an image charge $q_1$ at $(0, 0, d)$, whereas the magnetic field is given by an image magnetic monopole $g_1$ at $(0, 0, d)$. In the upper-half space, the electric field is given by $q/\varepsilon_1$ at $(0, 0, d)$ and an image charge $q_2$ at $(0, 0, -d)$; the magnetic field is given by an image magnetic monopole $g_2$ at $(0, 0, -d)$. The above ansatz satisfies the Maxwell equations on each side of the boundary. At the boundary $z = 0$, the solution is then matched according to the standard boundary condition, giving

$$\begin{aligned} q_1 &= q_2 \\ &= \frac{1}{\varepsilon_1}\frac{(\varepsilon_1 - \varepsilon_2)(1/\mu_1 + 1/\mu_2) - 4\alpha^2P_3^2}{(\varepsilon_1 + \varepsilon_2)(1/\mu_1 + 1/\mu_2) + 4\alpha^2P_3^2}q \\ g_1 &= -g_2 \\ &= -\frac{4\alpha P_3}{(\varepsilon_1 + \varepsilon_2)(1/\mu_1 + 1/\mu_2) + 4\alpha^2P_3^2}q \end{aligned} \qquad (3)$$

We will first take $\varepsilon_1 = \varepsilon_2 = \mu_1 = \mu_2 = 1$ below and then recover the $\varepsilon_{1,2}$, $\mu_{1,2}$ when discussing the experimental proposals later. The solution shows that, for an electric charge near the surface of a topological insulator, both an image magnetic monopole and an image electric charge will be induced, as compared with conventional electromagnetic media where only an electric image charge will be induced. It is notable that the magnitudes of the image magnetic monopole and image electric charge satisfy the relation $q_{1,2} = \pm(\alpha P_3)g_{1,2}$. This is just the relation $q = (\theta/2\pi)g$ for the electric and magnetic charges of a dyon inside the θ vacuum (*16*), with $\theta/2\pi = \pm P_3$ here.

The physical origin of the image magnetic monopole is understood by rewriting part of the Maxwell equations as

$$\nabla \times \mathbf{B} = 2\alpha P_3\delta(z)\hat{\mathbf{n}} \times \mathbf{E} \qquad (4)$$

with $P_3 = \pm 1/2$ the value for the topological insulator [in the above equation, $\nabla$ is the derivative vector, $\hat{\mathbf{n}}$ is the normal vector of the surface, and $\delta(z)$ is the Dirac δ function]. The right-hand side of the above equation corresponds to a surface current density $\mathbf{j} = \sigma_{xy}(\hat{\mathbf{n}} \times \mathbf{E})$, which is induced by the in-plane component of the electric field and is perpendicular to this component. This current is nothing but the quantized Hall current mentioned earlier. For the problem under consideration, the surface current density is calculated as

$$\mathbf{j} = P_3\left(\frac{e^2}{h}\right)\left(\frac{q}{1 + \alpha^2P_3^2}\right)\frac{r}{(r^2 + d^2)^{\frac{3}{2}}}\hat{\mathbf{e}}_\varphi \qquad (5)$$

which is circulating around the origin (inset of Fig. 1) (here, $r$ is the radial distance and $\hat{\mathbf{e}}_\varphi$ is the tangental unit vector). Physically, this surface current is the source that induces the magnetic field. On each side of the surface, the magnetic field induced by the surface current can be viewed as the field induced by an image magnetic monopole on the opposite side.

According to the above calculation, the image magnetic monopole field indeed has the correct magnetic field dependence expected from a monopole, and it can be controlled completely through the position of the electric charge. As we started with the Maxwell's equation, which includes $\nabla \cdot \mathbf{B} = 0$, the magnetic flux integrated over a closed surface must vanish. We can check that this is the case by considering a closed surface—for example, a sphere with radius $a$—that encloses a topological insulator. The detailed calculation is presented in the supporting online material (*17*). Inside the closed surface, there is not only a image magnetic monopole charge, but also a line of magnetic charge density whose integral exactly cancels the point image magnetic monopole. However, when the separation between the electric charge and the surface ($d$) is much smaller than the spherical radius ($a$), the magnetic field is completely dominated by the image magnetic monopole, and the contribution due to the line of magnetic charge density is vanishingly small. Therefore, we propose here to experimentally observe the magnetic monopole in the same sense that we can experimentally observe other fractionalization, or de-confinement, phenomena in condensed matter physics. In any closed electronic system, the total charge must be quantized to be an integer. However, one can separate fractionally charged elementary

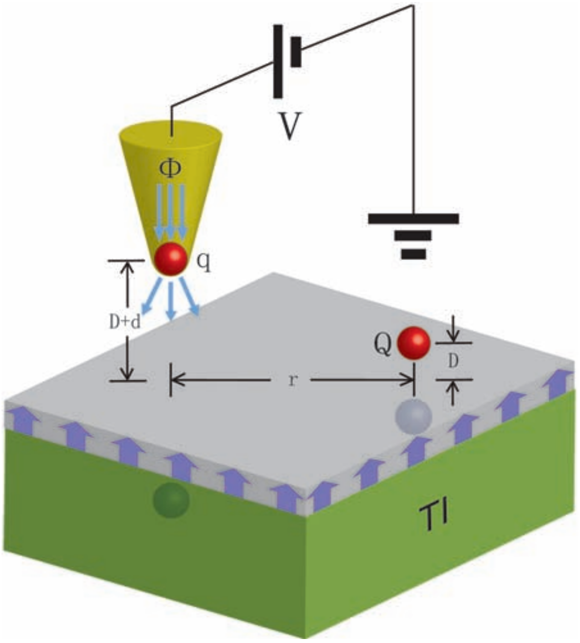


**Fig. 2.** Illustration of the experimental setup to measure the image monopole. A magnetic layer is deposited on the surface of the topological insulator, as indicated by the layer with blue arrows. (The same layer is drawn in Figs. 3 and 4.) A scanning MFM tip carries a magnetic flux ϕ and a charge *q*. A charged impurity is confined on the surface with charge *Q* and distance *D* out of the surface. By scanning over the voltage *V* and the distance *r* to the impurity, the effect of the image monopole magnetic field can be measured (see text).

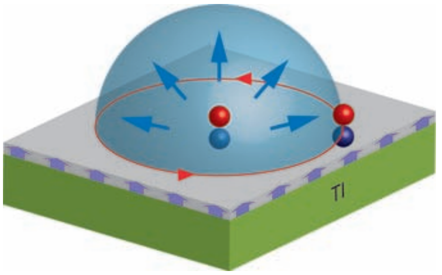


**Fig. 3.** Illustration of the fractional statistics induced by image monopole effect. Each electron forms a dyon with its image monopole. When two electrons are exchanged, an AB phase factor is obtained (which is determined by half of the image monopole flux) and leads to statistical transmutation.

excitations arbitrarily far from each other, so that fractional charge can have well-defined meaning locally. A similar situation occurs in spin-charge separation. Whereas the total charge and the total spin of a closed system must be linked to each other, spin and charge can occur as separated local excitations. In our case, as long as $d$ is much smaller than the radius of curvature of a topological surface $a$, the local magnetic field is completely determined by a single image magnetic monopole.

Such an image monopole can be observed experimentally by a magnetic force microscope (MFM). Consider the surface of the topological insulator with a localized charged impurity (Fig. 2). A scanning MFM tip can be applied to detect the magnetic field distribution of the image monopole. However, the charge of the impurity also generates an electric force to the tip. The contribution of the image monopole can be distinguished from other trivial contributions by scanning both the tip position $r$ and the tip voltage $V$. For a given position $r$, $f_{\min}(r)$ defines the minimal force applied to the tip when scanning the voltage, and $D$ is the distance of the charged impurity to the surface. In the limit of $r >> D$, the conventional charge interaction leads to a $1/r^6$ dependence of $f_{\min}(r)$. The image monopole magnetic field leads to more dominant contribution

$$f_{\min}(r) \simeq \frac{4\alpha P_3}{(1+\varepsilon_2/\varepsilon_1)(1/\mu_1+1/\mu_2)-4\alpha^2 P_3^2}\frac{Q\phi}{r^3} \tag{6}$$

in which $Q$ is the impurity charge and $\phi$ is the net flux of the magnetic tip. For the estimated parameters $\varepsilon_2 \cong 100$ for the $Bi_{1-x}Sb_x$ alloy (*18*), $\varepsilon_1 = 1$, $\mu_1 \cong \mu_2 \cong 1$, $\phi \cong 2.5hc/e$, and a typical distance $r = 50$ nm, the force is $f_{\min}(r) \cong 0.3$ pN/μm, which is observable in the present experiments. In a real experimental system, the surface magnetic layer can have defects such as domains and steps that induce an inhomogeneous fringe magnetic field acting on the MFM tip. However, all surface roughness effects can only induce magnetic field of dipolar or higher order. Consequently, the contribution of surface roughness decays faster than $1/r^3$ so that the monopole contribution proportional to $1/r^3$ still dominates the long-range behavior. Further detail on the calculations is presented in the supporting online material (*17*).

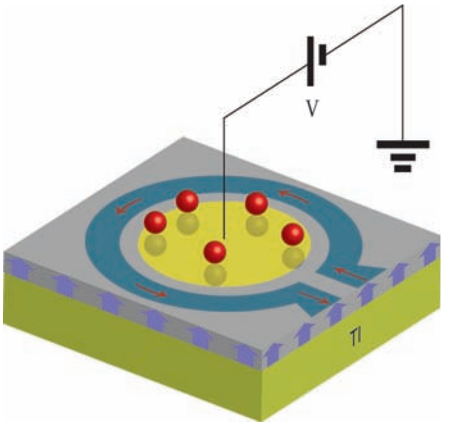


**Fig. 4.** Illustration of the experimental proposal measuring the fractional statistics of the dyons. When a gate voltage is applied to the central metallic island, the number of electrons in the central island can be tuned, which in turn changes the net flux threaded in the superconducting ring and leads to a supercurrent.

Because the position of the image monopole is determined by that of the charge above the surface, the pure monopole is not an elementary excitation of the system, and this is different from the monopole proposed in some spin ice models (*19*). However, if we consider the electrons moving on the topological surface as part of the system, the bound state of an electron and its image monopole does become a dynamical object with nontrivial properties. Such a bound state of charge and monopole is known in high-energy physics as a "dyon" (*16*)": a composite particle with both electric and magnetic charges. Besides contributing a monopole-monopole Coulomb interaction, the image monopole also induces a statistical interaction between two dyons. When two dyons are exchanged, each of them will obtain an Aharonov-Bohm (AB) phase due to the magnetic field of the other one. The net AB phase obtained by the two-particle system during an exchange process is independent of the path of the particles on the 2D plane, which thus can be interpreted as a statistical angle of the dyon (Fig. 3). Therefore, this setup provides a condensed matter realization of the anyon concept (*20*, *21*) in the absence of any external magnetic field. Anyons have unusual quantum statistics; interchanging a pair of them gives rise to a complex phase factor different from zero (as in the case of bosons) or π (as in the case of fermions). This effect also provides a mechanism of statistical transmutation in (3+1) dimensions (*22*, *23*). If each dyon has charge $q$ and monopole flux $g$, then the statistical angle is given by $\varphi = \left(\frac{gq}{2\hbar c}\right)$. For example, the binding of an electron with charge $e$ and a monopole with flux $hc/e$ leads to $\varphi = \pi$, in which case the dyon is a boson. However, there is a difference between the dyon studied here and that in high-energy physics. In high-energy physics, the charged particle and monopole are point-like particles moving in three dimensions. Consequently, the monopole flux must be quantized in unit of $hc/e$, and correspondingly, the statistical angle $\varphi$ can only be 0 or π modular 2π. This is consistent with the fact that there are no anyonic statistics in (3+1) dimensions. Our dyon can only be defined if an electron is moving close to the surface of a topological insulator. Consequently, the dyon is always confined to a 2D surface. In this case, $\nabla \cdot \mathbf{B} = 0$ is satisfied everywhere and the flux of the image monopole does not have to be quantized. Correspondingly, the dyon can have anyonic statistics. [A possibly related anyon model has been studied in the fractional quantum Hall context (*24*).] According to Eq. 3, the statistical angle of an electron-induced dyon is

$$\varphi = \frac{2\alpha^2 P_3}{(\varepsilon_1+\varepsilon_2)(1/\mu_1+1/\mu_2)+4\alpha^2 P_3^2} \tag{7}$$

which is of order $\alpha^2$ and can generally be irrational (*25*). It should be noticed that the image charge beneath the electron does not contribute additional AB phase because its motion corresponds to a displacement current. For $P_3 = 1/2$ and $\varepsilon_1, \mu_1, \mu_2 \sim 1$, $\varepsilon_2 \sim 100$, we obtain $\varphi \cong \alpha^2/200 \cong 2.6 \times 10^{-7}$ rad. Though the statistical angle is quite small, it is physically observable. With the geometry of Fig. 4, a quasi-1D superconducting ring and a metallic island surrounded by the ring are deposited on top of the topological insulator surface (which has already been gapped by a magnetic layer). By tuning the gate voltage of the central island, the number of electrons $N$ on the island can be tuned. Because of the statistical angle $\varphi$ given by Eq. 7, each electron in the central island induces a flux of $\varphi$ seen by the electrons in the ring. Consequently, the net flux through the ring is $N\varphi$ in unit of the flux quanta, which generates a supercurrent in the superconducting ring as in a standard superconducting quantum interference device. For a typical electron density of $n = 10^{11}/\text{cm}^2$ and island size $R = 1$ μm, the net magnetic flux is $n\varphi\pi R^2 \hbar c/e \simeq 2.6 \times 10^{-4} hc/2e$, and the corresponding magnetic field is $B = n\varphi\hbar c/e \simeq 1.7 \times 10^{-3}$G, which is observable. It is possible to have some other topological insulator material with a smaller dielectric constant $\varepsilon_2$, in which the statistical angle of the dyon can be larger.

We have shown theoretically that the topological surface states of a 3D topological insulator can act as a mirror that images an electron as a magnetic monopole. Such a transmutation between electric field and magnetic field is a direct manifestation of the TME effect (*3*). Because of this effect, a 2D electron gas in the neighborhood of the surface will become a dyon gas with fractional statistics. We have presented realistic experimental setups to observe the field of the image magnetic monopole and the statistical angle of the dyon.

### References and Notes

1. L. D. Landau, E. M. Lifshitz, *Electrodynamics of Continuous Media* (Pergamon, Oxford, 1984).
2. E. Witten, *Sel. Math. New Ser.* **1**, 383 (1995).
3. X.-L. Qi, T. L. Hughes, S.-C. Zhang, *Phys. Rev. B* **78**, 195424 (2008).
4. L. Fu, C. L. Kane, *Phys. Rev. B* **74**, 195312 (2006).
5. J. E. Moore, L. Balents, *Phys. Rev. B* **75**, 121306 (2007).
6. R. Roy, preprint available at http://arxiv.org/abs/cond-mat/0604211.
7. C. L. Kane, E. J. Mele, *Phys. Rev. Lett.* **95**, 226801 (2005).
8. B. A. Bernevig, S.-C. Zhang, *Phys. Rev. Lett.* **96**, 106802 (2006).
9. B. A. Bernevig, T. L. Hughes, S.-C. Zhang, *Science* **314**, 1757 (2006).
10. M. König *et al.*, *Science* **318**, 766 (2007); published online 19 September 2007 (10.1126/science.1148047).
11. A. M. Essin, J. E. Moore, D. Vanderbilt, preprint availble at http://arxiv.org/abs/0810.2998.

12. D. Hsieh *et al.*, *Nature* **452**, 970 (2008).
13. S.-C. Zhang, T. H. Hansson, S. Kivelson, *Phys. Rev. Lett.* **62**, 82 (1989).
14. F. Wilczek, *Phys. Rev. Lett.* **58**, 1799 (1987).
15. J. D. Jackson, *Classical Electrodynamics* (Wiley, New York, 1999).
16. E. Witten, *Phys. Lett. B* **86**, 283 (1979).
17. See supporting material on *Science* Online.
18. O. Madelung, U. Rössler, M. Schulz, in *Landolt-Börnstein, Substance/Property Index, Semiconductors, Non-Tetrahedrally Bonded Elements and Binary Compounds I, Subvolume III/17E-17F-41C* (Springer, Berlin, 1998), pp. 1–11.
19. C. Castelnovo, R. Moessner, S. L. Sondhi, *Nature* **451**, 42 (2008).
20. J. Leinaas, J. Myrheim, *Nuovo Cim. B* **37**, 1 (1977).
21. F. Wilczek, *Phys. Rev. Lett.* **49**, 957 (1982).
22. R. Jackiw, C. Rebbi, *Phys. Rev. Lett.* **36**, 1116 (1976).
23. P. Hasenfratz, G. ′t Hooft, *Phys. Rev. Lett.* **36**, 1119 (1976).
24. E. Fradkin, S. Kivelson, *Nucl. Phys. B* **474**, 543 (1996).
25. C. Chamon *et al.*, *Phys. Rev. Lett.* **100**, 110405 (2008).

26. We thank O. M. Auslaender, T. L. Hughes, S. Kivelson, and L. Luan for insightful discussions. This work is supported by the U.S. Department of Energy, Office of Basic Energy Sciences under contract DE-AC03-76SF00515.

**Supporting Online Material**
www.sciencemag.org/cgi/content/full/1167747/DC1
SOM Text
Figs. S1 to S5
References

27 October 2008; accepted 16 January 2009
Published online 29 January 2009;
10.1126/science.1167747
Include this information when citing this paper.

---

# Global Cooling During the Eocene-Oligocene Climate Transition

**Zhonghui Liu,[1,2]* Mark Pagani,[1]* David Zinniker,[1] Robert DeConto,[3] Matthew Huber,[4] Henk Brinkhuis,[5] Sunita R. Shah,[6] R. Mark Leckie,[3] Ann Pearson[6]**

**About 34 million years ago, Earth's climate shifted from a relatively ice-free world to one with glacial conditions on Antarctica characterized by substantial ice sheets. How Earth's temperature changed during this climate transition remains poorly understood, and evidence for Northern Hemisphere polar ice is controversial. Here, we report proxy records of sea surface temperatures from multiple ocean localities and show that the high-latitude temperature decrease was substantial and heterogeneous. High-latitude (45 degrees to 70 degrees in both hemispheres) temperatures before the climate transition were ~20°C and cooled an average of ~5°C. Our results, combined with ocean and ice-sheet model simulations and benthic oxygen isotope records, indicate that Northern Hemisphere glaciation was not required to accommodate the magnitude of continental ice growth during this time.**

The abrupt shift to glacial conditions near the Eocene-Oligocene (E-O) boundary ~33.7 million years ago (Ma) is characterized by a ~+1.5 per mil (‰) change in oxygen isotopic ($\delta^{18}O$) values of benthic foraminifera (*1*–*3*) in ~300,000 years, which is indicative of continental ice accumulation and high-latitude cooling, and an ~1-km deepening of the global calcite compensation depth (CCD) (*2*). Proposed causes for this fundamental change in Earth's climate state include changes in ocean circulation due to the opening of Southern Ocean gateways (*4*), a decrease in atmospheric $CO_2$ (*5*–*8*), and a minimum in solar insolation (*2*).

How Earth's temperature changed during ice expansion is poorly defined, largely because benthic $\delta^{18}O$ records do not distinguish between ice volume and temperature. Deep-sea temperature records based on foraminiferal Mg/Ca ratios show little change during ice expansion (*9*–*11*). As a result, benthic $\delta^{18}O$ records imply E-O ice volumes that must be accommodated by Northern Hemisphere glaciation (*2*, *9*, *12*). This conclusion is nearly untenable given scant physical evidence for Northern Hemisphere ice sheets before the latest Miocene (*7*, *12*–*15*). Deep-water foraminiferal Mg/Ca ratios could be affected by factors other than temperature (*9*, *11*), including a deepening of the CCD (*2*) and changes in deep-water carbonate ion concentration that occurred during the E-O climate transition. Indeed, shallow-water Mg/Ca–based temperatures, from exceptionally well-preserved foraminifera deposited above the CCD, indicate ~2.5°C of cooling in the tropics (*14*, *15*) and cast further suspicion on deep-water Mg/Ca–based temperatures across this major CCD deepening event.

Here, we report E-O sea surface temperature (SST) changes, which were determined with alkenone unsaturation index ($U^{K'}_{37}$) and tetraether index ($TEX_{86}$) (*16*, *17*), from 11 globally dispersed ocean localities. These localities include Ocean Drilling Program/Deep Sea Drilling Project (ODP/DSDP) sites 277, 336, 511, 913, and 1090, with paleolatitudes between ~45° and 70° in both hemispheres (*18*), and sites 628, 803, 925, 929, 998, and 1218 in the tropics (Fig. 1 and table S1) (*19*). Chronologies for these sites were previously established or refined and/or determined in this study (table S2 and fig. S1). $TEX_{86}$ indices were converted to SST by use of a modified temperature calibration based on all published ocean surface sediment data (fig. S2) (*20*). Nonetheless, older calibrations would yield qualitatively similar results over the temperature ranges observed.

Both $U^{K'}_{37}$ and $TEX_{86}$ SSTs show substantial high-latitude cooling between ~34 and 33 Ma (Fig.

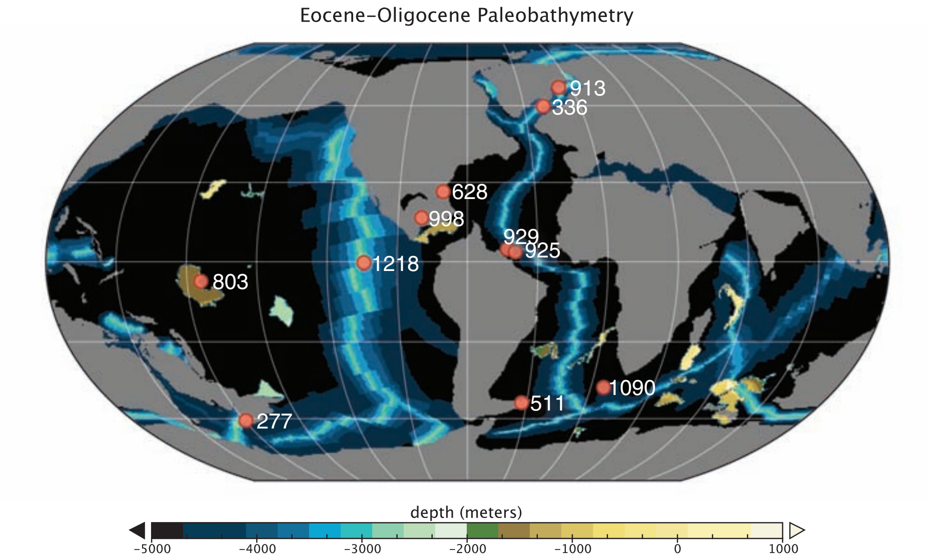


**Fig. 1.** Paleolocations of DSDP and ODP sites used in this study, superimposed on a map of reconstructed E-O bathymetry and geography (*18*).

---

[1] Department of Geology and Geophysics, Yale University, New Haven, CT 06520, USA. [2] Department of Earth Sciences, University of Hong Kong, Hong Kong, People's Republic of China. [3] Department of Geosciences, University of Massachusetts-Amherst, Amherst, MA 01003, USA. [4] Department of Earth and Atmospheric Sciences, Purdue University, West Lafayette, IN 47907, USA. [5] Laboratory of Palaeobotany and Palynology, Utrecht University, Budapestlaan 4, 3584 CD Utrecht, Netherlands. [6] Department of Earth and Planetary Sciences, Harvard University, Cambridge, MA 02138, USA.

*To whom correspondence should be addressed. E-mail: zhliu@hku.hk (Z.L.); mark.pagani@yale.edu (M.P.)

2). At sites 511 and 336, for which E-O transition records are relatively complete, SST cooled >5°C from the late Eocene to the early and mid-Oligocene, whereas maximum cooling at sites 277, 913, and 1090 occurred ~33.5 Ma. High-latitude cooling averaged ~4.8°C (using a direct average of all high-latitude data ranging from ~2° to ~8°C at individual sites) from the late Eocene (~34 to 37 Ma) to the earliest Oligocene (~33 to 34 Ma) (table S3 and fig. S4), or ~5.4°C at 33.4 Ma when determined by a 10-point running average of the combined high-latitude records (Fig. 2B). Our estimate of SST cooling is most likely a minimum value, given that the full magnitude of SST change is probably not expressed because of low resolution, core gaps, and sediment hiatuses close to the E-O boundary (fig. S4).

In general, the timing and pattern of high-latitude SST reconstructions from the late Eocene to early and mid-Oligocene correlate well with benthic $\delta^{18}$O changes (Fig. 2) (*21*, *22*). SST reached the lowest values near the same time as the maximum $\delta^{18}$O excursion at ~33.5 Ma, particularly at site 511 (Fig. 2C). Given the resolution of our entire data set, high-latitude SST changes appear to be approximately synchronous with benthic $\delta^{18}$O changes during the climate transition.

Our records also indicate that late Eocene high-latitude SST was substantially warmer than previous estimates. Before the E-O climate transition (~34 to 37 Ma), high-latitude SST was ~20°C [±2.7°C for $U^{K'}_{37}$ and ±3.7°C for $TEX_{86}$ (*19*)] (Fig. 2A), which is ~10°C warmer than temperatures derived from benthic and planktonic $\delta^{18}$O records from deep-sea cores (*21*, *23*). Some localities, such as site 277, exhibit extraordinarily warm SST, with temperatures reaching ~27°C. Differences between our records and benthic/planktonic $\delta^{18}$O records could be explained in various ways: diagenesis, seasonality, and/or locations of deep-water production. Diagenesis could potentially alter both $U^{K'}_{37}$ (*24*) and $TEX_{86}$ (*25*) indices, resulting in temperature estimates warmer than original SSTs. However, there is no consistently strong evidence that $U^{K'}_{37}$ (*24*) or $TEX_{86}$ (*26*) values are radically altered by diagenesis. In addition, diagenetic pathways differ for these two distinct proxies, yet SST estimates yield remarkably similar results (Fig. 2A). Two relatively complete records from sites 511 and 336 indicate that the coldest early Oligocene $U^{K'}_{37}$ SST was only ~3° to 6°C warmer when compared with the same locations today (table S1). Although these warmer-than-modern temperatures are anticipated for the early Oligocene, they also constrain warm-temperature biases introduced by diagenetic effects for Oligocene and late Eocene to smaller than 3° to 6°C. Finally, our late Eocene proxy records from the high latitudes are broadly consistent with other documented marine and terrestrial biotic changes (*27*, *28*) and terrestrial temperature records (*29*, *30*).

Discrepancies between foraminiferal $\delta^{18}$O– and organic-based temperature estimates could also reflect differences in the locations of our sites relative to deep-water sources as well as the seasonality of foraminifera versus organic production. Paired benthic and planktonic $\delta^{18}$O records during this time show systematic offsets to at least 65°S, and SST gradients of 7° to 10°C between mid- to high-latitude sites such as 511 and 1090, and polar site 690 (*23*), indicating that deep-water formation was probably focused across latitudes higher than those represented by our site localities. Further, benthic $\delta^{18}$O values reflect deep-water production during winter months, whereas $U^{K'}_{37}$ and $TEX_{86}$ values capture mean annual SSTs that were probably biased toward late spring/early autumn temperatures near the poles. Our model results also

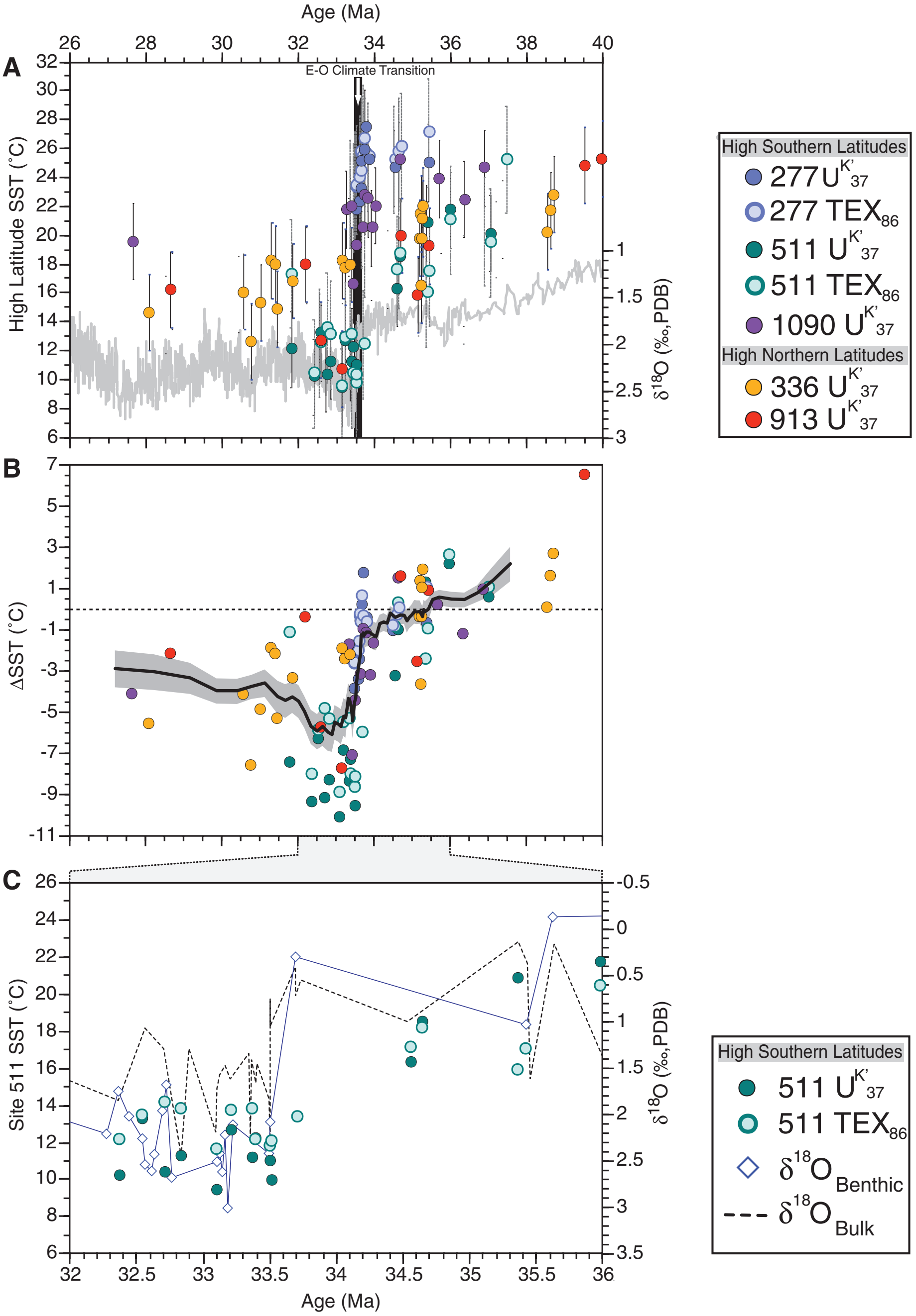


**Fig. 2.** High-latitude SST records during the late E-O. (**A**) SST reconstructions from five high-latitude sites. The gray line represents a composite benthic $\delta^{18}$O record (*21*). SST calibration equations for $U^{K'}_{37}$ and $TEX_{86}$ are $SST = (U^{K'}_{37} - 0.039)/0.034$ (*16*) and $SST = 50.47 - 16.33/TEX_{86}$, respectively (*19*). SST uncertainty is based on 1σ SE of their respective calibration regressions (fig. S3). (**B**) Changes in SST represented as temperature deviations from mean values at each site before the E-O transition (~34 to 37 Ma). The solid black line represents a 10-point running average of SST changes from all high-latitude sites and yields ~5.4°C of cooling at 33.4 Ma. The shaded region brackets the SE of the 10-point running average. The dotted line marks zero SST change relative to the pre–E-O (~34 to 37 Ma) SST average. (**C**) A detailed comparison of $U^{K'}_{37}$ and $TEX_{86}$ SST values with $\delta^{18}$O records of benthic foraminifera (*22*) and bulk carbonate during the E-O transition at site 511.

show that it is possible to sustain ~20°C annual mean temperatures in mid- to high-latitude regions, with near-freezing winter temperatures in the polar coastal waters that reflect deep-water formation regions (fig. S5). Recent evidence suggests that Mg/Ca and $\delta^{18}$O SST estimates from well-preserved early and late Eocene planktonic foraminifera are broadly consistent with $TEX_{86}$ estimates but ~10°C warmer than other $\delta^{18}$O estimates from the tropics (*14*, *31*) and high latitudes (*32*), suggesting that the primary planktonic $\delta^{18}$O values (*23*) were altered by early diagenesis (*31*). Although our late Eocene SST estimates from site 277 and probably site 1090 (Fig. 2A) appear particularly warm when compared with other high-latitude sites investigated here, they are in accord with early and mid-Eocene SST estimates from New Zealand, based on the $TEX_{86}$ proxy, Mg/Ca, and $\delta^{18}$O values from extraordinarily well-preserved planktonic foraminifera (*32*). Thus, we suggest that our high-latitude estimates, because of biases introduced by diagenesis and seasonal production, are probably at the warm end of but remain close to the E-O mean annual temperatures. Physical interpretation of exceptional high-latitude warmth during the late Eocene remains an important challenge to climate scientists.

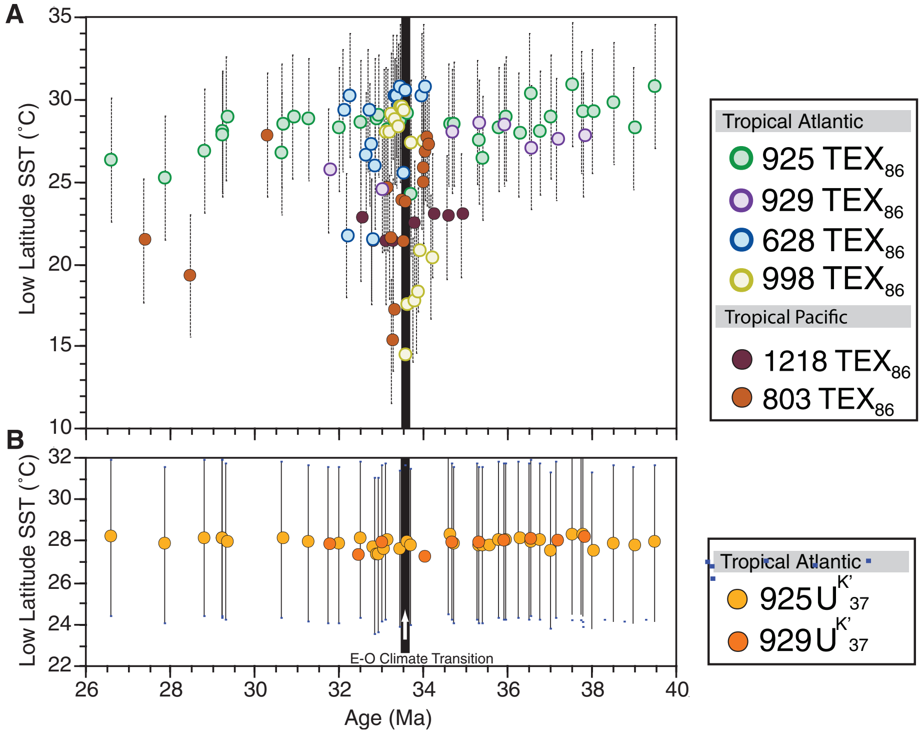


**Fig. 3.** Tropical SST records during the late E-O. (**A**) $TEX_{86}$ SST records from six tropical sites. (**B**) $U^{K'}_{37}$-SST records from sites 925 and 929. The $TEX_{86}$ proxy probably overestimates the magnitude of tropical SST changes because of tropical oceanographic perturbations associated with high-latitude cooling.

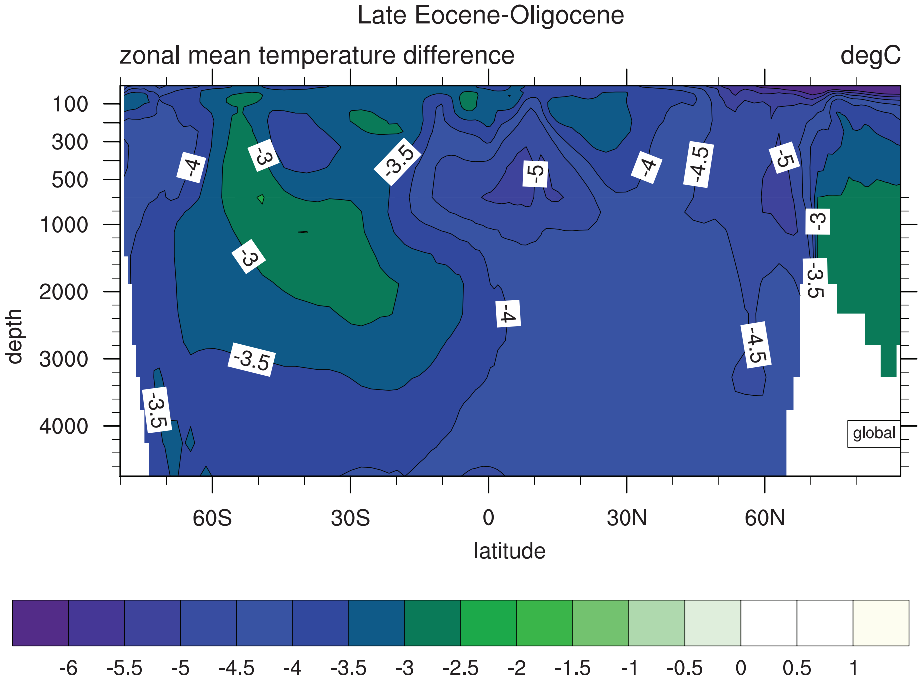


**Fig. 4.** Fully coupled E-O model results. The zonally and annually averaged temperature difference between the high and low partial pressure of $CO_2$ cases described in (*19*), representing late Eocene and early Oligocene conditions, is shown. Cooling is maximized in the equatorial upper thermocline below the mixed-layer, high-latitude oceans in the upper 2 km and in the abyss, predominately north of the equator. The vertical scale (in meters) is stretched.

Compared with high-latitude E-O SST, tropical $TEX_{86}$-SST records are more difficult to interpret. Tropical SST predominantly range between 27° to 31°C before ~34 Ma and after ~32 Ma. A brief interval (between ~33 to 34 Ma) of apparent colder temperatures is expressed in most of our tropical $TEX_{86}$-SST records, with inferred temperature changes varying from 0° to 15°C (Fig. 3A). However, such large tropical SST variations require physically improbable atmospheric and oceanic circulations when considered within the context of spatial temperature gradients that drive wind fields (*8*). Further, large tropical SST changes are unsupported by other tropical Mg/Ca-based SST reconstructions from well-preserved foraminifera (*14*, *15*). Indeed, $U^{K'}_{37}$ values from sites 925 and 929 in the tropical Atlantic indicate temperatures >27°C and do not support the large temperature variability suggested by some of our $TEX_{86}$ records (Fig. 3B). Given these considerations, we suggest that large tropical $TEX_{86}$ perturbations do not solely record SST variations. Instead, extreme $TEX_{86}$ values probably reflect changes in tropical water-column properties, associated with high-latitude surface cooling identified here, and variations in the population and production depths of Crenarchaeota (*20*, *33*). If we consider only $TEX_{86}$ records with temperature variations constrained by coeval $U^{K'}_{37}$ records (sites 925 and 929), tropical cooling does not appear to exceed ~3°C (table S3), which is consistent with other tropical temperature records (*14*, *15*).

Our results affect interpretations of ice volume and the potential for Northern Hemisphere glaciation during the E-O climate transition. Ice-volume calculations, based on benthic $\delta^{18}$O records, require an estimate of deep-ocean cooling. However, the full change in SST from sparsely sampled localities is not necessarily translated to abyssal depths. To evaluate how surface temperature change was translated into the deep ocean, we ran coupled atmosphere-ocean simulations with Eocene boundary conditions (*8*) and $CO_2$ levels representative of pre- and posttransition atmospheric mixing ratios (*5*, *6*) (*19*). Model simulations reproduce ~5°C of high-latitude cooling consistent with our data, a 3.8°C volumetric-mean ocean cooling, and global-mean surface cooling of 4.4°C (Fig. 4). The model provides clear evidence of spatially heterogenous benthic cooling, averaging ~4°C (that is, somewhat less than the surface temperature change). However, because (i) these model simulations do not account for temperature feedbacks related to growing ice sheets and (ii) our estimate of ~5°C of high-latitude cooling might represent a minimum value, model results for deep-water temperature change also potentially slightly underestimate the magnitude of cooling. Therefore, the combination of our model and proxy results provides a range between 3° and 5°C of benthic cooling during the E-O climate transition, with our best estimate converging on 4°C.

Assuming that the average $\delta^{18}$O value of Oligocene Antarctic ice sheets was ~20 ‰ more positive relative to today (an assumption that maximizes ice-volume estimates) (*34*), the growth of continental ice implied by a +1.5 ‰ shift in benthic $\delta^{18}$O and a benthic cooling of 3° to 5°C is between ~$10 \times 10^6$ and $30 \times 10^6$ km$^3$. This estimate is equivalent to ~50 to 140% of the volume of a simulated early Oligocene East Antarctic ice sheet (*5*, *34*) or ~40 to 120% of modern Antarctic ice volume and 30 to 100% of Antarctic ice volume during the Last Glacial Maximum. Although our results do not contradict evidence of small localized glaciers in the northern high latitudes (*13*), estimated ice volumes could be easily accommodated on Antarctica alone (*34*) and do not require Northern Hemisphere glaciation (*12*) to explain the magnitude of the benthic $\delta^{18}$O shift. Finally, between ~33 and 34 Ma, our data support a pattern of systematic cooling in both the northern and southern high latitudes (Fig. 2A) that was approximately in phase with benthic $\delta^{18}$O changes (Fig. 2C)—a temporal pattern also observed in other tropical records (*14*, *15*). These results are consistent with a scenario of global cooling forced by a reduction in greenhouse gas concentration rather than the more regionalized effects of ocean gateways (*5*–*8*).

## References and Notes

1. K. G. Miller, R. G. Fairbanks, G. S. Mountain, *Paleoceanography* **2**, 1 (1987).
2. H. K. Coxall, P. A. Wilson, H. Palike, C. H. Lear, J. Backman, *Nature* **433**, 53 (2005).
3. N. J. Shackleton, J. P. Kennett, *Initial Rep. Deep Sea Drill. Proj.* **29**, 743 (1975).
4. J. P. Kennett, *J. Geophys. Res.* **82**, 3843 (1977).
5. R. M. DeConto, D. Pollard, *Nature* **421**, 245 (2003).
6. M. Pagani, J. C. Zachos, K. H. Freeman, B. Tipple, S. Bohaty, *Science* **309**, 600 (2005).
7. R. DeConto, D. Pollard, D. Harwood, *Paleoceanography* **22**, Pa3214 (2007).
8. M. Huber, D. Nof, *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **231**, 9 (2006).
9. C. H. Lear, Y. Rosenthal, H. K. Coxall, P. A. Wilson, *Paleoceanography* **19**, Pa4015 (2004).
10. C. H. Lear, H. Elderfield, P. A. Wilson, *Science* **287**, 269 (2000).
11. K. Billups, D. P. Schrag, *Earth Planet. Sci. Lett.* **209**, 181 (2003).
12. A. K. Tripati *et al.*, *Earth Planet. Sci. Lett.* **265**, 112 (2008).
13. J. S. Eldrett, I. C. Harding, P. A. Wilson, E. Butler, A. P. Roberts, *Nature* **446**, 176 (2007).
14. C. H. Lear, T. R. Bailey, P. N. Pearson, H. K. Coxall, Y. Rosenthal, *Geology* **36**, 251 (2008).
15. M. E. Katz *et al.*, *Nat. Geosci.* **1**, 329 (2008).
16. F. G. Prahl, L. A. Muehlhausen, D. L. Zahnle, *Geochim. Cosmochim. Acta* **52**, 2303 (1988).
17. S. Schouten, E. C. Hopmans, E. Schefuss, J. S. Sinninghe Damsté, *Earth Planet. Sci. Lett.* **204**, 265 (2002).
18. R. D. Müller, M. Sdrolias, C. Gaina, B. Steinberger, C. Heine, *Science* **319**, 1357 (2008).
19. Materials and methods are available as supporting material on *Science* Online.
20. J.-H. Kim, S. Schouten, E. C. Hopmans, B. Donner, J. S. Sinninghe Damsté, *Geochim. Cosmochim. Acta* **72**, 1154 (2008).
21. J. Zachos, M. Pagani, L. Sloan, E. Thomas, K. Billups, *Science* **292**, 686 (2001).
22. J. P. Muza, D. F. Williams, S. W. Wise, *Initial Rep. Deep Sea Drill. Proj.* **71**, 409 (1983).
23. J. C. Zachos, L. D. Stott, K. C. Lohmann, *Paleoceanography* **9**, 353 (1994).
24. J. O. Grimalt *et al.*, *Geochem. Geophys. Geosys.* **1**, 2000GC000053 (2000).
25. S. R. Shah, G. Mollenhauer, N. Ohkouchi, T. I. Eglinton, A. Pearson, *Geochim. Cosmochim. Acta* **72**, 4577 (2008).
26. S. Schouten, E. C. Hopmans, J. S. Sinninghe Damsté, *Org. Geochem.* **35**, 567 (2004).
27. H. Brinkhuis, H. Visscher, *Soc. Econ. Paleontol. Mineral. Spec. Publ.* **54**, 295 (1995).
28. D. R. Prothero, T. H. Heaton, *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **127**, 257 (1996).
29. S. Schouten *et al.*, *Geology* **36**, 147 (2008).
30. A. Zanazzi, M. J. Kohn, B. J. MacFadden, D. O. Terry, *Nature* **445**, 639 (2007).
31. P. N. Pearson *et al.*, *Geology* **35**, 211 (2007).
32. C. J. Hollis *et al.*, *Geology* **37**, 99 (2009).
33. C. Turich *et al.*, *Geochim. Cosmochim. Acta* **71**, 3272 (2007).
34. R. M. DeConto *et al.*, *Nature* **455**, 652 (2008).

35. The authors thank S. Schouten and E. Hopmans for their assistance in establishing an interlaboratory calibration, J. Eldrett for advice on the site 913 age model, M. Woodruff of the University of Massachussetts Stable Isotope Facility for analyses of site 511 material, and Integrated Ocean Drilling Program for providing samples. The reviews of P. Wilson and one anonymous referee greatly improved the manuscript. This work was supported by a postdoctoral fellowship provided by Yale University and a Major Research Instrumentation grant from NSF. Computing was performed on Rosen Center for Advanced Computing resources within Information Technology at Purdue. M.H.'s contribution is partially supported by the New Zealand's Global Change Through Time Programme at GNS Science.

**Supporting Online Material**

www.sciencemag.org/cgi/content/full/323/5918/1187/DC1
Materials and Methods
Figs. S1 to S5
Tables S1 to S3
References

24 September 2008; accepted 23 January 2009
10.1126/science.1166368

---

# Seeing the Fermi Surface in Real Space by Nanoscale Electron Focusing

**Alexander Weismann,[1,2] Martin Wenderoth,[1]* Samir Lounis,[3] Peter Zahn,[4] Norbert Quaas,[1] Rainer G. Ulbrich,[1] Peter H. Dederichs,[3] Stefan Blügel[3]**

**The Fermi surface that characterizes the electronic band structure of crystalline solids can be difficult to image experimentally in a way that reveals local variations. We show that Fermi surfaces can be imaged in real space with a low-temperature scanning tunneling microscope when subsurface point scatterers are present: in this case, cobalt impurities under a copper surface. Even the very simple Fermi surface of copper causes strongly anisotropic propagation characteristics of bulk electrons that are confined in beamlike paths on the nanoscale. The induced charge density oscillations on the nearby surface can be used for mapping buried defects and interfaces and some of their properties.**

The coherent propagation of electrons in solids is central for a variety of phenomena that are at the core of modern physics. Scanning tunneling microscopy (STM) has been used to manipulate atoms and create structures that allow standing electron wave patterns to be visualized (*1*). C. R. Moon *et al.* (*2*) extended this line of investigations to the retrieval of quantum-phase information in nanostructures with the scanning tunneling microscope. Another facet of electron propagation has been revealed by measurements (*3*) of exchange interaction between adatoms and wires mediated through the Ruderman-Kittel-Kasuya-Yosida (RKKY) mechanism (*4*) on a platinum surface. All of these effects depend on a fundamental property of the electron sea: It rearranges itself to minimize the disturbance caused by foreign atoms. These Friedel oscillations (*5*) may cause technologically important effects such as the formation of diluted magnetic semiconductors, spin-glasses, or the interlayer exchange coupling between magnetic layers (*6*) exploited in read heads of magnetic hard discs.

---

[1] IV Physikalisches Institut, Universität Göttingen, 37077 Göttingen, Germany. [2] Courant Research Center Göttingen, 37077 Göttingen, Germany. [3] Institut für Festkörperforschung and Institute for Advanced Simulation, Forschungszentrum Jülich, 52425 Jülich, Germany. [4] Fachbereich Physik, Martin-Luther-Universität Halle-Wittenberg, 06099 Halle, Germany.

*To whom correspondence should be addressed. E-mail: wenderoth@ph4.physik.uni-goettingen.de

In contrast to surface impurities, research on subsurface defects has been less intense (*7–10*) because of the inherent experimental and theoretical difficulties involved in the investigation process. We show that a single buried impurity can be used to visualize an unexpected directional propagation of electrons, even in a rather simple metal such as copper. Combining an STM experiment and density functional calculations, we demonstrate the existence of unusually strong anisotropies in the resulting charge densities. This peculiar behavior is caused by the shape of the Fermi surface, which focuses electrons in very narrow directions. These findings are general and should be observable in other crystalline solids.

The presence of the Fermi surface in metals is the ultimate signature that electrons of a system can be treated as quasi-particles within the Fermi liquid theory. The determination of the Fermi surface of a material is crucial to understanding its thermal, electrical, magnetic, and optical properties. Experimentally, the Fermi surface is usually measured with the De Haas–van Alphen effect, after the pioneering work of D. Shoenberg (*11*), or with angle-resolved photoemission spectroscopy. The importance of the Fermi surface on macroscopic electron transport through single crystals has been demonstrated by Y. V. Sharvin *et al.* (*12–14*). Their explanation uses the electrons group velocities $\boldsymbol{v}_{\mathrm{Gr}} = \hbar^{-1}\nabla_{\mathbf{k}}E(\boldsymbol{k})$ that are normal to the Fermi surface of the particular system (here, $\hbar$ is Planck's constant $h$ divided by $2\pi$, $\nabla_{\mathbf{k}}$ is the gradient with respect to the wave vector $\mathbf{k}$, and $E(\boldsymbol{k})$ is the band structure). An accumulation of these vectors in a certain direction results in enhanced electron flux, which is also called electron focusing. Our investigations suggest that an unforeseen tool, the scanning tunneling microscope, can be used to provide a real space visualization of Fermi surfaces.

Furthermore, we propose different possible applications of this electron focusing effect. We prepared isolated subsurface Cobalt impurities in single crystals of Cu(100) and Cu(111) and investigated their influence on the local density of states (LDOS) with the use of low-temperature STM. The impurities were embedded in different depths below the surface. We added 0.1% Co to the topmost 15 monolayers by simultaneous deposition of a host metal and the impurity compound from two electron-beam evaporators.

A 9-by-9–nm STM-topography (Fig. 1A) illustrates the influence of different Co atom depths (6, 7, 9, and 10 monolayers) below the Cu(111) surface on the electron's spectral density at the surface. Apart from a long-wavelength standing-wave pattern caused by surface state electrons scattered at a monoatomic step-edge in the upper left corner, four ringlike structures can be observed. These correspond to short-wavelength oscillations with an amplitude of ~2 pm, varying in diameter but being constant in their radial envelope width (~1.5 oscillations). As the diameter expands with increasing depth of the impurity below the surface, the LDOS within a certain angle does not seem to be perturbed by the impurity.

The experiments performed on the (100) surface elucidate the propagation of electrons in Cu from another perspective (Fig. 1B). Here we observe the LDOS being influenced only in a squarelike region above the impurity and within four beamlike regions extending diagonally outside the four vertices. It is also observed that the wavelength of the standing wave pattern is very close to the lattice constant. This similarity occurs because the corresponding Fermi wave vectors are near the boundary of the Brillouin zone.

From these results, we can demonstrate that the surface LDOS—especially at greater distances from the impurity atoms—is influenced only in narrowly confined directions. These regions are not perfectly circular-symmetric, but they reflect the symmetry of the underlying band structure [i.e., threefold symmetry for (111) and fourfold symmetry for (100) surfaces].

All of these observed effects are related to the propagation of electrons and can be theoretically described with one fundamental concept: the quantum-mechanical propagator [i.e., the one-electron retarded Green function $G(\boldsymbol{x},\boldsymbol{x}',\varepsilon)$ of the given system]. It describes how electrons of energy $\varepsilon$ propagate from a point source at $\boldsymbol{x}'$ to other positions $\boldsymbol{x}$ in the crystal. The ripple extending from a stone impact in still water and the Huygens wavelet in optics are comparable examples of propagators in other areas of physics. Huygens principle states that the behavior of extended systems (the diffraction pattern of a grating, for example) can be expressed as a superposition of multiple Huygens-waves—that is, by the integration over all possible propagation paths. The single electron propagator in vacuum is a spherical wave, in full analogy to the Huygens wavelet

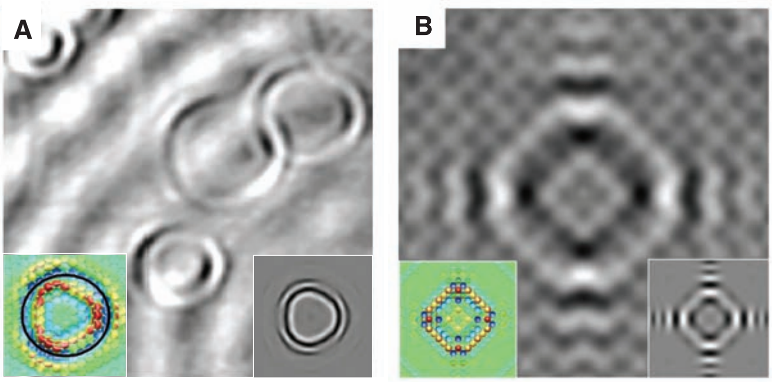


**Fig. 1.** STM topographies of (**A**) four Co-Atoms below the Cu(111) surface (9 by 9 nm, –80 mV, 1 nA) and (**B**) one Co Atom below the Cu(100) surface (3.5 by 3.5 nm, 10 mV, 2 nA). The right insets show (4 by 4 nm) calculated LDOS using Eq. 2, whereas the left insets refer to DFT calculations.

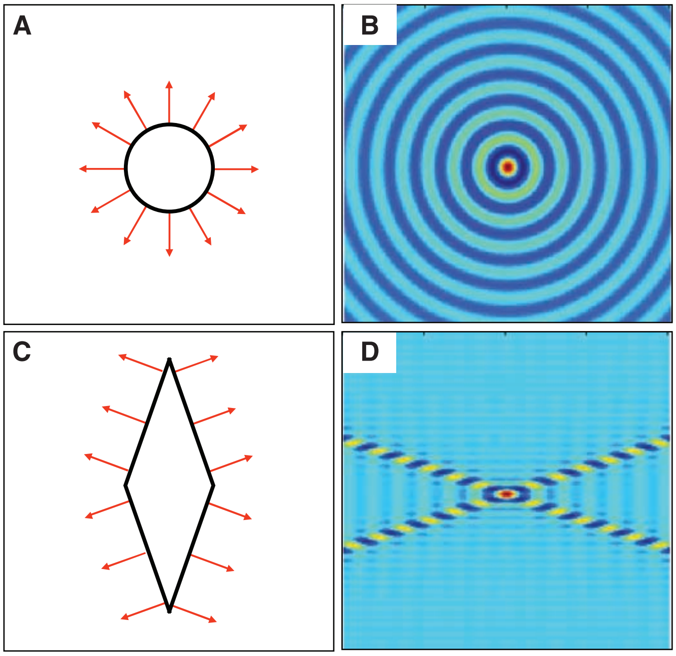


**Fig. 2.** Illustration of the relation between Fermi surface [black contours in (**A**) and (**C**)], group-velocities [red arrows in (A) and (C)], and the corresponding Huygens-wave [(**B**) and (**D**)] of a given system. The source (position $\boldsymbol{x}'$) was chosen in the center of (B) and (D).

in isotropic optical media. If an electron propagates within a crystal potential, the propagation becomes anisotropic. With the use of spectral representation, the single-electron propagator of a given system is connected to its electronic structure [the wave functions $\Psi_{\boldsymbol{k}}(\boldsymbol{x})$ and the band structure $E(\boldsymbol{k})$] via:

$$G_0(\boldsymbol{x},\boldsymbol{x}',\varepsilon) = \lim_{\eta\to 0}\int_{\boldsymbol{k}} d^3k\, \frac{\Psi_{\boldsymbol{k}}(\boldsymbol{x})\Psi^*_{\boldsymbol{k}}(\boldsymbol{x}')}{\varepsilon + i\eta - E(\boldsymbol{k})} \qquad (1)$$

In the above equation, $\eta$ is an infinitesimal positive quantity and $i^2 = -1$. The imaginary part of Eq. 1 can be transformed into a superposition of all wave functions having energy $\varepsilon$. Thus, the propagator is directly associated with the k-space geometry of the corresponding constant energy-surface. In the case of free electrons, the Fermi surface is a sphere (Fig. 2A) with radius of the Fermi wave vector $k_F$, and the corresponding Green function is a spherical wave decaying with $|\boldsymbol{x} - \boldsymbol{x}'|^{-1}$ in amplitude (Fig. 2B). Interesting effects arise if the Fermi surface deviates from a spherical shape, which is the case for the vast majority of crystals. In the most extreme case of a Fermi surface being composed of flat areas (Fig. 2C), the wave functions interfere constructively in beamlike regions perpendicular to these facets and destructively elsewhere in space (Fig. 2D). In strong contrast to the isotropic case, the amplitude within the beams does not decay with increasing distance, as is the case for the Green function in one dimension. Thus, whenever electrons are emitted from pointlike-sources or scattered at pointlike defects, they do not propagate in a spherical wave as in free space, but instead are, in general, focused in preferential directions and are detectable at much larger distances from the source than in the isotropic case.

For copper, the Fermi surface is rather spherical but shows areas having strongly reduced curvature (Fig. 3A). The propagator for such a metal can be calculated quickly by considering plane waves as wave functions and using $E(\boldsymbol{k})$ obtained by a simple linear combination of atomic orbitals (LCAO) technique with band structure parameters taken from (*15*). The calculated Green function (Fig. 3B) demonstrates the premises of a substantial electron focusing in copper. In contrast to the spherical wave of a free electron, a preferred propagation can be observed along eight slightly distorted hollow cones around the [111] directions. The explicit connection with the shape of the Fermi surface is settled when evaluating the LDOS change induced by a subsurface impurity

$$\Delta LDOS(\boldsymbol{x},\varepsilon) = -\frac{1}{\pi}\mathrm{Im}\iint d\boldsymbol{x}_{\mathrm{i}}\, d\boldsymbol{x}_{\mathrm{j}}\, G_0(\boldsymbol{x},\boldsymbol{x}_{\mathrm{i}},\varepsilon)\, t(\boldsymbol{x}_{\mathrm{i}},\boldsymbol{x}_{\mathrm{j}},\varepsilon)\, G_0(\boldsymbol{x}_{\mathrm{j}},\boldsymbol{x},\varepsilon) \qquad (2)$$

with $\boldsymbol{x}$, $\boldsymbol{x}_{\mathrm{i}}$, and $\boldsymbol{x}_{\mathrm{j}}$ being arbitrary positions in the system, whereas Im stands for the imaginary part. For our problem, $\boldsymbol{x}$ corresponds to a vacuum site probed by STM. Here, $t$ is the t-matrix describing the scattering of electrons by the impurity caused by the potential change $\Delta V = V_{\mathrm{new}} - V_{\mathrm{host}}$.

At large distances $R$ between the impurity and the vacuum site, we can apply the stationary phase approximation and end up with a result $\Delta LDOS(\boldsymbol{x},\ \varepsilon) \propto \frac{1}{|c|R^2}$, similar to that given by the well-known theory of interlayer exchange coupling (*16*). Here, $c$ is the Gaussian curvature (*17*) of the isoenergy surface at energy $\varepsilon$. Thus, the observed charge density oscillations (CDO) depend strongly on the shape of the Fermi surface: A small value of the curvature means that the Fermi surface has a flat region leading to big values of the LDOS and to strong focusing of intensity in this space region.

This statement is reinforced after comparing the results of Eq. 2 using the simple LCAO model (shown in the lower right insets in Fig. 1) with the density functional theory (DFT) calculations (depicted in the left insets of Fig. 1) [spin-polarized full-potential Korringa-Kohn-Rostoker Green function method (*18*)] that are in total accordance with our experimental measurements. The lower left inset in Fig. 1A shows the case of a substitutional Co impurity sitting at the 6th layer below the surface. The induced CDO has been calculated around the Fermi energy $E_F$ in the vacuum region at ~6.1 Å above the center of the surface layer after removing the ideal surface background. The same period of oscillations was also noticed in the experiment. Figure 3C shows the DFT Fermi surface of Cu, on which the

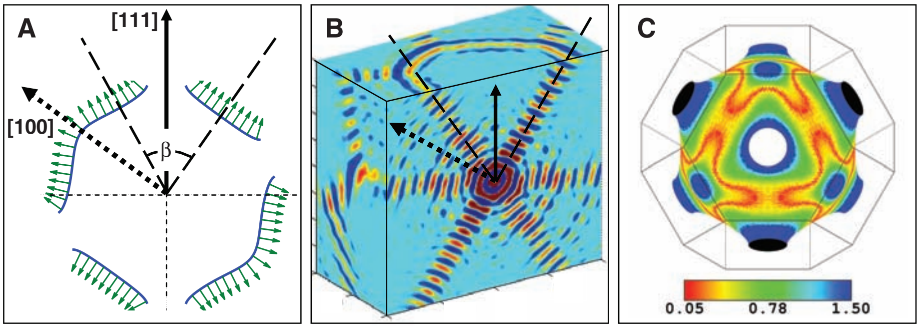
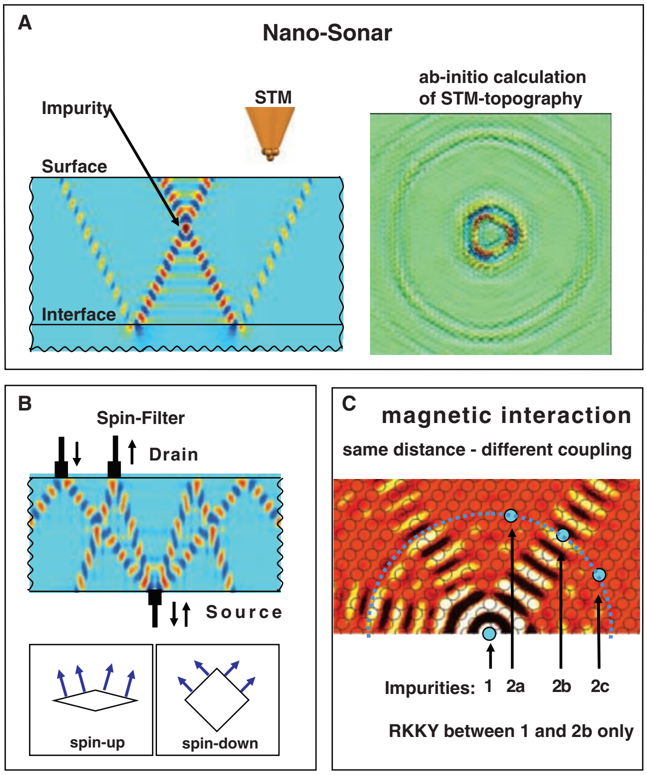


**Fig. 3.** (**A**) Cross section of the Cu Fermi surface showing areas of reduced curvature and bandgaps in [111] directions. (**B**) Corresponding propagator [$-\mathrm{Im}G_0(\boldsymbol{x},\boldsymbol{x}',\varepsilon_F)$] with strong electron focusing onto hollow-conelike beams around [111]. (**C**) The Gaussian curvature of the Cu Fermi surface is represented with color. The drawing plane is oriented perpendicular to the [111] direction. Small curvatures represented in red lead to high amplitudes of the LDOS oscillations.

**Fig. 4.** Applications based on focusing properties of the single electron propagator. (**A**) Nano-sonar: A buried interface below the impurity produces a second concentric LDOS modulation. An STM investigation of the surface LDOS would gain access to the depths and the reflectivity of both impurity and interface. (**B**) Spin filter: In a ferromagnetic material, a spin-unpolarized current emitted from a point-contact ($S_{\uparrow\downarrow}$) splits up because of the different Fermi surfaces and thus different propagators of majority and minority spins. (**C**) Conduction electron–related indirect exchange interaction between two magnetic atoms in copper: strong and long-range interaction only between impurities 1 and 2b (i.e., in the directions of the focused beams).



Gaussian curvature is represented with color. The low curvatures depicted in red correspond to less curved regions of the Fermi surface and are in accordance with the high-intensity points calculated for the charge variation in the vacuum. Diagonally opposite of the high intensities, we get only small oscillations resulting from the stronger curved regions shown in Fig. 3C. This result explains the triangular shape with a three fold symmetry instead of the six fold symmetry of the face-centered cubic-(111) surface.

Knowing the cone angle β of the focusing beams (Fig. 3A; β~60°) and the depth $Z_{\mathrm{imp}}$ of the impurity allows us to evaluate the diameter $D$ of the LDOS pattern and vice versa. This result allows us to experimentally determine the position of impurities below any noble metal surface. As a proof, a circle having the theoretical diameter $D \sim 21$ Å [obtained from $D = 2Z_{\mathrm{imp}}\tan(\beta/2)$] is shown in the left inset of Fig. 1A. In addition, we performed the identical procedure to determine the depths of the impurities observed in the STM measurements.

The strong directionality of electron propagation is not unique to copper and has many consequences for other materials. Once the shape of the propagator is known, either by STM investigation, calculations based on the band-structure, or quick estimates from a knowledge of the Fermi surface, many physical effects can be easily predicted. We have demonstrated and explained the very anisotropic CDO caused by subsurface point defects. Of course, similar behavior will be found for both spin channels if the impurity is magnetic, in the arrangement of the Kondo-screening cloud if the system is below the Kondo-temperature, and for the indirect exchange (RKKY)-interaction between two magnetic atoms (Fig. 4C).

By scrutinizing the real-space properties of the single-electron propagator and settling the explicit relation between the Fermi surface shape and the anisotropic charge oscillations, we demonstrate a new application of STM: visualization of Fermi surfaces in real space. The LDOS pattern as function of energy is determined by (i) the dispersion of host-metal band structure, (ii) the energy-dependent coherence length of the quasiparticles, and (iii) the energy-dependent scattering phase shift of the impurity. With the use of spectroscopic STM techniques, our approach can provide access to constant-energy surfaces apart from the Fermi surface. For energies close to $E_F$, effects due to decoherence can be neglected, and even for $|\varepsilon - E_F| = 1$ eV, the mean free path is much larger than the distances between impurity and surface (*19*, *20*). Thus, for magnetic impurities, for instance, the information in the focusing pattern is dominated by the scattering behavior of the impurity, and the focusing effect can be applied to investigate the electronic properties of subsurface impurities spectroscopically.

The beamlike propagation paths could, for example, be used to construct a "nano-sonar" (Fig. 4A). Such a device could determine not only the depth and reflectivity of buried interfaces but also the position and magnetism of the interface. Here, in addition to the LDOS-oscillations of Fig. 1 caused by processes of direct propagation between impurity and surface, larger concentric rings should appear that are produced by electrons propagating from the impurity toward the interface and being reflected toward the surface. In fact, we extracted CDO and spin-dependent CDO for an interface of Co with Cu(111) using DFT calculations. These results allowed us to determine the position as well as the magnetism of the interface (supporting online material).

In ferromagnets, the Fermi surfaces and, consequently, the electron propagators are obviously different for the two spin channels. This should allow the design of effective spin filters (Fig. 4B). A mixture of both spin species enters the ferromagnet at a pointlike source contact ($S\uparrow\downarrow$) from which spin up– and spin down–electrons propagate in different directions and are collected at different drain contacts $D_\uparrow$ and $D_\downarrow$. According to recent theoretical calculations (*21*), body-centered cubic europium may be a good candidate for such an application.

### References and Notes

1. M. F. Crommie, C. P. Lutz, D. M. Eigler, *Science* **262**, 218 (1993).
2. C. R. Moon *et al.*, *Science* **319**, 782 (2008).
3. F. Meier, L. Zhou, J. Wiebe, R. Wiesendanger, *Science* **320**, 82 (2008).
4. M. A. Ruderman, C. Kittel, *Phys. Rev.* **96**, 99 (1954).
5. J. Friedel, *Nuovo Cim.* **7** (suppl. 2), 287 (1958).
6. P. Grünberg, R. Schreiber, Y. Pang, M. B. Brodsky, H. Sowers, *Phys. Rev. Lett.* **57**, 2442 (1986).
7. M. Schmid, W. Hebenstreit, P. Varga, S. Crampin, *Phys. Rev. Lett.* **76**, 2298 (1996).
8. P. T. Sprunger, L. Petersen, E. W. Plummer, E. Lægsgaard, F. Besenbacher, *Science* **275**, 1764 (1997).
9. N. Quaas, M. Wenderoth, A. Weismann, R. Ulbrich, K. Schönhammer, *Phys. Rev. B* **69**, 201103 (2004).
10. M. Rose, A. Borg, T. Mitsui, D. F. Ogletree, M. Salmeron, *J. Chem. Phys.* **115**, 10927 (2001).
11. D. Shoenberg, *Nature* **183**, 171 (1959).
12. Y. V. Sharvin, L. M. Fisher, *JETP Lett.* **1**, 152 (1965).
13. V. S. Tsoi, *JETP Lett.* **19**, 70 (1974).
14. J. Heil, M. Primke, K. U. Würz, P. Wyder, *Phys. Rev. Lett.* **74**, 146 (1995).
15. D. A. Papaconstantopoulos, *Handbook of the Band Structure of Elemental Solids* (Plenum, New York, 1986).
16. P. Bruno, C. Chappert, *Phys. Rev. Lett.* **67**, 1602 (1991).
17. The denominator is also directly proportional to the effective mass tensor $\sqrt{\frac{d^2E}{dK_x^2}\cdot\frac{d^2E}{dK_y^2}}^{-1}$ (here, $K_x$ and $K_y$ are orthogonal wave vector coordinates that are located in a tangent plane on the Fermi surface). If one or both principal curvatures are zero (i.e., if the surface is cylindrical or flat), this quantity diverges and the expression $\Delta LDOS(\mathbf{x},\varepsilon) \propto \frac{1}{|c|R^2}$ breaks down. The LDOS in certain distances $R$ then follows a different decay law $\sim R^{-1}$ and $\sim R^0$, respectively.
18. N. Papanikolaou, R. Zeller, P. H. Dederichs, *J. Phys. Condens. Matter* **14**, 2799 (2002).
19. E. Knoesel, A. Hotzel, M. Wolf, *Phys. Rev. B* **57**, 12812 (1998).
20. I. Campillo, A. Rubio, J. M. Pitarke, A. Goldmann, P. M. Echenique, *Phys. Rev. Lett.* **85**, 3241 (2000).
21. J. Kunes, R. Laskowski, *Phys. Rev. B* **70**, 174415 (2004).
22. This work was supported by the European Science Foundation EUROCORES Programme Self-Organized NanoStructures under contract N. ERAS-CT-2003-980409, the Deutsche Forschungsgemeinschaft Priority Programme SPP1153, and the Deutsche Forschungsgemeinschaft Collaborative Research Centre SFB602.

**Supporting Online Material**

www.sciencemag.org/cgi/content/full/323/5918/1190/DC1
Materials and Methods
SOM Text
Fig. S1
References

19 November 2008; accepted 27 January 2009
10.1126/science.1168738

---

# Conductance of a Single Conjugated Polymer as a Continuous Function of Its Length

**Leif Lafferentz,[1] Francisco Ample,[2] Hao Yu,[3] Stefan Hecht,[3] Christian Joachim,[2] Leonhard Grill[1]***

The development of electronic devices at the single-molecule scale requires detailed understanding of charge transport through individual molecular wires. To characterize the electrical conductance, it is necessary to vary the length of a single molecular wire, contacted to two electrodes, in a controlled way. Such studies usually determine the conductance of a certain molecular species with one specific length. We measure the conductance and mechanical characteristics of a single polyfluorene wire by pulling it up from a Au(111) surface with the tip of a scanning tunneling microscope, thus continuously changing its length up to more than 20 nanometers. The conductance curves show not only an exponential decay but also characteristic oscillations as one molecular unit after another is detached from the surface during stretching.

A key challenge in molecular electronics (*1*) is the detailed understanding of charge transport through a single molecular wire (*2*–*4*). One of the first conductance measurements of a metal–single molecule–metal junction used a scanning tunneling microscope (STM); the STM tip made electrical contact to a single $C_{60}$ molecule adsorbed on a metal surface (*5*). This result was soon followed by break junction experiments (*6*–*8*), in which an ultrathin metallic wire junction is mechanically broken, providing a two-electrode junction where a statistical number of molecules can be located and where conductance histograms are collected to reveal the average conductance of presumably single molecules in this molecular junction (*9*, *10*). These two techniques demonstrated that electrons can easily tunnel through a molecule with a low energy gap between the highest occupied molecular orbital (HOMO) and lowest unoccupied molecular orbital (LUMO). This approach can be used to provide statistical information about the conductance of single or very few molecules. However, they do not allow the conductance of a single and long molecular wire to be determined as the function of the distance between the two contacts on the same molecule. We present an experimental procedure to measure the conductance of a single and the same molecular wire with a well-defined and defect-free chemical structure as a function of the distance between the two contact points on the wire while precisely imaging the molecule's conformation before and after the measurement with submolecular resolution.

Recently, it was reported that a single molecule can be picked up from a solid surface with scanning probe techniques by approaching the scanning tip toward the surface and subsequently lifting up a single or several molecules when retracting the tip. This method was applied to molecules (*11*–*17*) by STM and to oligomers (*18*) or DNA strands (*19*) by atomic force microscopy (AFM), although the latter one is not capable of conductance measurements. With such a pulling manipulation, one end of a molecule is bound to the tip of an STM while maintaining surface adsorption at the other end. The pulling was mostly practiced at random on many molecules at the same time, with no characterization of molecular conformations in the junction before and after (*11*–*16*). A few studies compared the statistically obtained conductances of an oligomer series of different lengths up to 7 nm (*13*–*15*). However, in the controlled pulling experiment on a single perylene-tetracarboxylic-dianhydride (PTCDA) molecule, the length of the molecule in the junction was fixed to only 1 nm (*17*).

Our measurements require long $\pi$-conjugated oligomers adsorbed on a surface that cannot be deposited intact by conventional sublimation techniques under clean conditions because of their large molecular weight (*20*). To access long one-dimensional molecular chains, polyfluorene composed of conjugated fluorene repeat units was targeted by our recently developed in situ polymerization (*21*). For this purpose, we used dibromoterfluorene (DBTF) monomers, consisting of three fluorene units, carrying lateral methyl groups and a Br atom at each end (Fig. 1A). These terminal groups are dissociated from the terfluorene (TF) molecular core in the first activation step of our on-surface synthesis (*21*). At a surface temperature of 10 K (*22*), single DBTF molecules on a Au(111) surface appeared in constant-current STM as three intense lobes corresponding to the dimethyl groups (Fig. 1B). DBTF molecules adopt a zig-zag shape on the

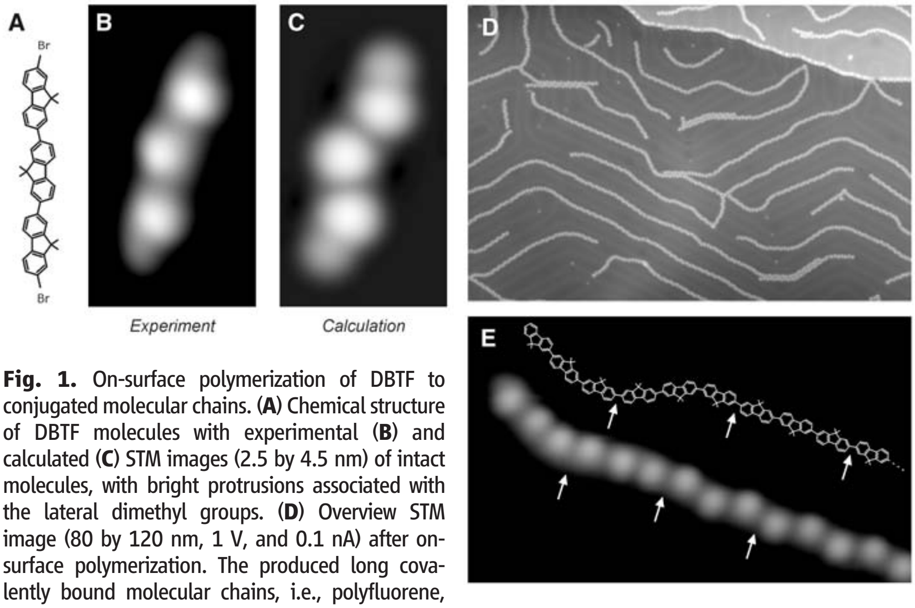


**Fig. 1.** On-surface polymerization of DBTF to conjugated molecular chains. (**A**) Chemical structure of DBTF molecules with experimental (**B**) and calculated (**C**) STM images (2.5 by 4.5 nm) of intact molecules, with bright protrusions associated with the lateral dimethyl groups. (**D**) Overview STM image (80 by 120 nm, 1 V, and 0.1 nA) after on-surface polymerization. The produced long covalently bound molecular chains, i.e., polyfluorene, follow the herringbone reconstruction of the substrate. (**E**) STM image (5.9 by 3.6 nm) of a single polyfluorene chain end with its chemical structure superimposed (using a different scaling). The arrows indicate three identical (in the STM image and the chemical structure), newly formed covalent bonds between individual building blocks.

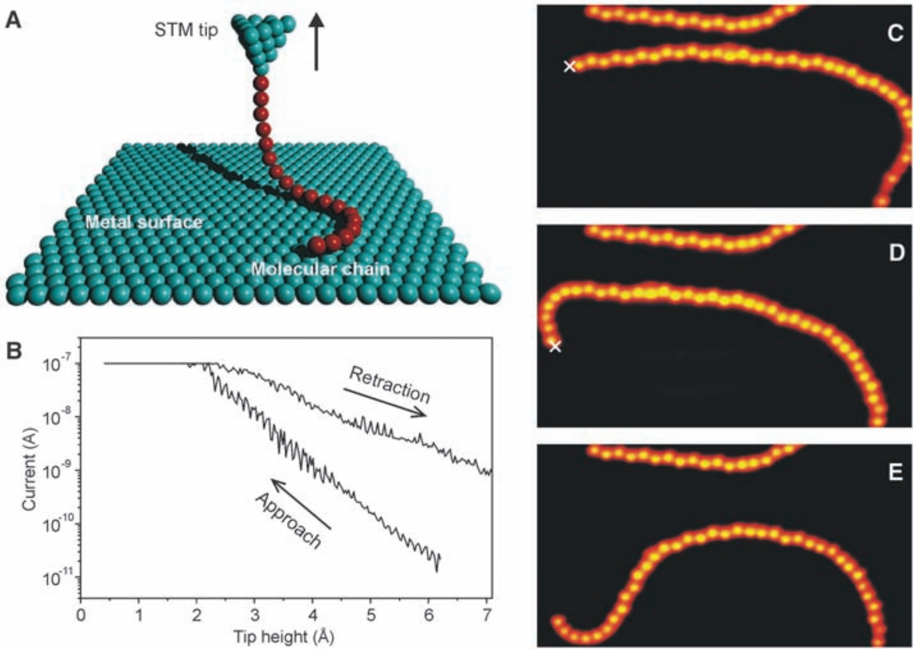


**Fig. 2.** Lifting a single molecular chain with the STM tip. (**A**) Scheme of the chain pulling procedure: After contacting a molecular chain to the STM tip, it can be lifted from the surface in a ropelike manner upon retraction because of its flexibility and weak interaction with the substrate. (**B**) Tunneling current as a function of the tip height during a vertical manipulation (approach and retraction with different slopes are marked by arrows). The maximum current is limited in this experiment to 100 nA (by the current preamplifier). (**C** to **E**) STM images (25.4 by 13.7 nm) of the same surface area during a vertical manipulation series (the cross indicates the position of tip approach and retraction). The manipulated chain, extending beyond the lower image border, changes its shape during the pulling processes. The fixed chain at the upper image border serves as reference.

---

[1] Physics Department, Freie Universität Berlin, 14195 Berlin, Germany. [2] Nanosciences Group, Centre d'Elaboration des Materiaux et d'Etudes Structurales (CEMES)–CNRS, 31055 Toulouse, France. [3] Department of Chemistry, Humboldt-Universität zu Berlin, 12489 Berlin, Germany.

*To whom correspondence should be addressed. E-mail: leonhard.grill@physik.fu-berlin.de

Au(111) surface, corresponding to the energetically favored alternate conformation of the methyl groups with respect to the molecular board (*22*). This result is in very good agreement with STM image calculations (Fig. 1C) that allow us to extract the exact position of each dimethyl group along a DBTF unit. Because of the low evaporator temperature during DBTF deposition, the Br atoms should still be attached to the TF core; a comparison between the observed and calculated images indicates that this is the case.

After heating the Au(111) surface up to 520 K for 5 min, the DBTF's Br atoms dissociate and covalent bonds are formed between different activated TF monomers, which diffuse randomly on the surface (Fig. 1D). The high mobility of the monomers, being a prerequisite for polymerization, is favored by the presence of lateral methyl groups on each monomer. Hence, this on-surface synthesis allows the formation of long, well-defined chains on the Au(111) surface; oligomer lengths greater than 100 nm were observed.

These polyfluorenes, carrying only methyl side groups, cannot be prepared by conventional polymerization processes in solution because of the absence of sufficiently solubilizing side chains. The homogeneous appearance of each molecular chain demonstrates the extreme regularity of their chemical composition, meaning that the newly formed covalent bonds are equivalent to the existing bonds, connecting the three fluorene units within each DBTF monomer, and that defect-free polymers were synthesized on the surface (Fig. 1E). The resulting polymers, commonly following the herringbone substrate reconstruction (Fig. 1D), are mobile enough on the Au(111) surface to be manipulated laterally with an STM tip for chains as long as 25 nm (*22*). Furthermore, the manipulation proves their high flexibility, enabling different curvatures of the chain without breaking the chemical bonds between the different TF monomers.

After selecting an isolated oligomer chain on the surface on the basis of a first STM image, this chain was pulled upward by the STM tip apex (Fig. 2A). For this purpose, the STM tip was first positioned at one end of the chain where, because of the Br dissociation during the on-surface synthesis, a chemical radical is presumably located. Subsequently, the tip apex was gently brought into close proximity of the selected chain to establish the electronic contact and then progressively retracted to lift it upward. The precise tip location on the molecular chain end is of great importance in this process, as evidenced by characteristic differences of the pulling success rate (*22*). The variation of the tunneling current *I* as a function of vertical distance *z* during this procedure was recorded, giving detailed insight into the vertical manipulation (Fig. 2B). After a characteristic *I*(*z*) exponential increase during the tip approach to the surface (*23*), a different *I*(*z*) dependence was recorded during retraction, when the molecular chain has been bonded to the tip apex. The current was much greater at the same value of *z*, when the tip is retracted very far from the surface, as compared to the case where no molecular chain was attached. Apparently, the tip-molecule interaction, which is less defined than the contact on the surface (that is imaged before the pulling experiment), is stronger than the chain segment's interaction with the substrate, keeping the chain attached to the tip during the pulling process. Although it is possible to lift entire chains from the surface, the tip-chain connection sometimes ruptures during the retraction (but not at constant tip-surface distance), which manifests as an abrupt drop of the tunneling current.

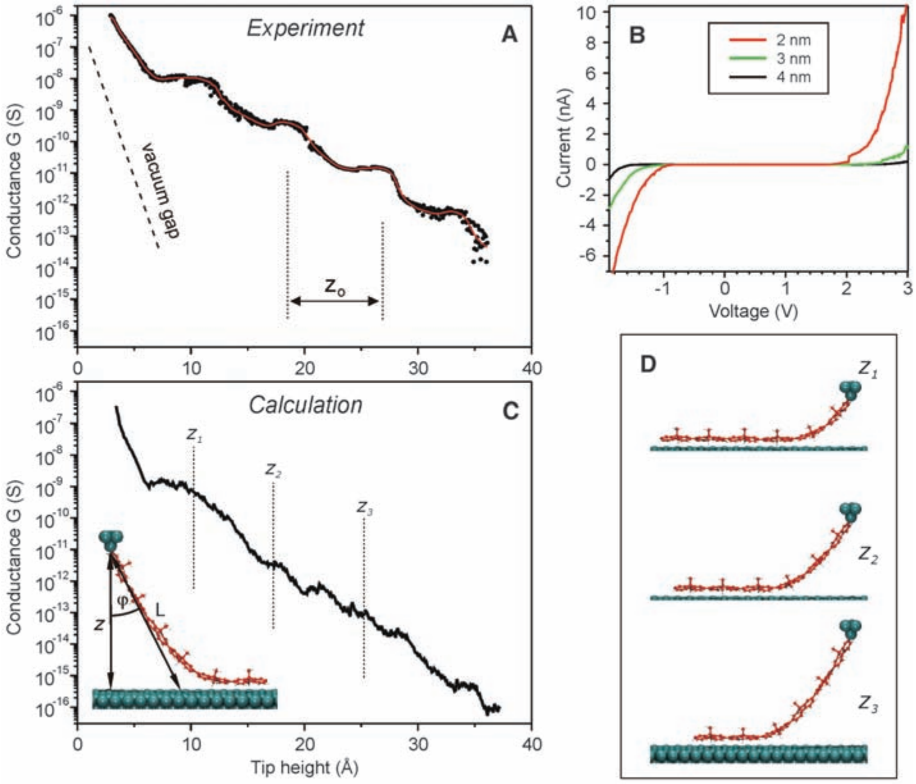


**Fig. 3.** Conductance as a function of the length of the molecular wire. Experimental (**A**) and calculated (**C**) *G*(*z*) curves (equally scaled), both exhibiting characteristic oscillations with a period of $z_o$ (the decay of a vacuum gap is plotted for comparison). The experimental curve is composed of two data sets from measurements below and above about 20 Å, respectively, using different setups and thus ranges for current detection (each about four orders of magnitude). (**B**) *I-V* curves (of single wires and thus not averaged) at three tip-surface distances (2, 3, and 4 nm). (**D**) Schematic views of characteristic conformations during the pulling process, just before the detachment of another molecular unit ($z_1$ = 10.2 Å, $z_2$ = 17.2 Å, and $z_3$ = 25.2 Å). The inset in (C) shows a sketch with the characteristic parameters *z*, *L*, and φ.

Complementary to the *I*(*z*) curve, STM images recorded before and after a vertical manipulation can aid in the visualization of the changes in the chain conformation and position on the surface. The high mechanical stability of the STM junction (lateral and vertical drift of less than 7 pm/min) renders such a vertical STM manipulation extremely reliable for conductance measurements of long molecular wires. STM images before and after two vertical manipulations are shown in Fig. 2C and Fig. 2, D and E, respectively. In each experiment, the STM tip picked up a single chain and dropped it down (after an arbitrary retraction distance), as monitored by the *I*(*z*) curve. The polymer in Fig. 2D is not imaged at the same position as before the pulling (Fig. 2C) but has moved to the left, following the pulled chain end in a ropelike manner. The consequence of the pulling-release sequence is that the initial linear conformation of the chain first adopts a left- and then a right-handed curvature with a large lateral displacement of the contacted chain end (Fig. 2D). Such a modification of the molecular conformation is hardly possible to obtain by lateral manipulation when the full chain remains adsorbed on the surface. Furthermore, we have not observed chain scission during any of the several hundreds of vertical manipulation procedures.

To measure the charge transport through a constructed surface–single molecule–tip junction as a function of the distance, the junction conductance variation *G*(*z*) was recorded at a 100 mV bias voltage while pulling a given molecular chain (Fig. 3A). At low temperature, *G*(*z*) reflects the ability of the chain to electronically couple the two electrodes of the STM junction through its more or less delocalized electronic structure as a function of the distance between the two electronic contact points on the chain (the surface and the tip apex) (*24*). Fitting *G*(*z*) by the exponential relationship $G = G_o \times e^{-\beta z}$ (25), a damping factor $\beta$ of $0.38 \pm 0.09$ Å$^{-1}$ is obtained experimentally. This value, a measure for the conductance of the molecular chain, is in very good agreement with the calculated one of 0.36 Å$^{-1}$ [calculations were done with the Extended Hückel Molecular Orbital Theory–Elastic Scattering Quantum Chemistry (EHMO-ESQC) method (*22*)]. The average value of $G_o$, which is determined by the contacts to the electrodes, is 300 nS (with a variation between about 10 and 500 nS). The $z$ distance is shorter than the real length of the molecular chain between the two contact points because of the ropelike bending of the chain. Correcting $\beta$ by taking into account the relation between $z$ and the effective molecular wire length $L$, between the tip apex and the part of the molecular wire remaining adsorbed flat on the surface (Fig. 3C, inset), leads to $z = L.\cos\varphi$, that is, $\beta = 0.3$ Å$^{-1}$. The decay factor $\beta$ is much less than the 2.4 Å$^{-1}$ inverse decay length through a vacuum gap (indicated by a dashed line in Fig. 3A) (*2*), confirming that the electrons are tunneling through a single molecular chain. Other examples of molecular wires with smaller $\beta$ values have been reported. However, they also display smaller HOMO-LUMO gaps (*26*) than the polyfluorene molecular wire investigated here [a gap between 3 and 4 eV has been determined for similar poly(dioctylfluorene) molecules (*27*)].

With our pulling technique, the current through the wire can also be measured as a function of the bias voltage at any fixed wire length, that is, tip-sample distance, by stopping the pulling process. The characteristic asymmetric $I(V)$ curves of such single molecular wires, taken, for instance, at $z$ of 2, 3, and 4 nm in Fig. 3B, confirm the presence of two electronic resonance thresholds. We attribute these to the first reduced and oxidized states of the molecular wire, corresponding in a first approximation to its HOMO and LUMO levels, with the HOMO closer to the Fermi level than the LUMO, which leads to asymmetric *I-V* characteristics.

The exponential conductance decay of a molecular chain with increasing length should follow a linear curve on a logarithmic scale (*4*). However, the one presented in Fig. 3A is not straight but shows characteristic oscillations with maxima at retraction distances of 10, 19, 29, and 41 Å (with an error of about ±3 Å), that is, a typical oscillation period $z_o$ of about 10 Å. Using the EHMO-ESQC calculation method (*22*), $G(z)$ was calculated as presented in Fig. 3C, taking into account the mechanical deformation of the molecular wire, its sliding over the surface during the pulling process, and the full junction valence electronic structure, including the interaction with the tip apex and the physisorption of the molecular chain on the Au(111) surface at the other end, which defines the electronic contact of the molecular wire with the surface.

At small $z$ values, a steep $G(z)$ decay is observed both in the experimental and the calculated curve, due to the tunneling through vacuum at low tip heights. Aside from its general exponential trend at larger $z$ values, the calculated $G(z)$ (Fig. 3C) shows regular oscillations with a $z_o$ between 6 and 8 Å, reproducing the experimental ones. Given the known optimized molecular chain conformation of a calculated pulling, we can attribute each oscillation maximum to a given monomer unit being adsorbed flat on the surface and about to be pulled away from the surface after a short sliding motion (see schemes in Fig. 3D). Hence, superimposed on the exponential decay trend, the contact conductance is oscillating because of the chainlike behavior that involves one monomer unit after the other being removed from contact with the surface by the pulling process. The maxima in the $G(z)$ curve reflect the conformations slightly before the detachment of another chain unit (shown for three different heights in Fig. 3D) that in turn leads to an increased chain length between tip and surface and thus a drop of $G(z)$. Interestingly, the electronic contact between one monomer of the chain and the surface is ensured by its $\pi$ system and not by its methyl leg, which would cause a $G(z)$ oscillation period of the order of the 2.88 Å Au(111) surface lattice constant (*22*). Instead, it must be of the order of the monomer unit length (8 to 9 Å), but not exactly equal to it because of the absent registry between the spatial extension of the monomer $\pi$ system weakly overlapping with the surface and the Au(111) surface corrugation. The difference between the 10 Å experimental and our 7 Å calculated periods originates from the difficulty in theoretically reproducing in detail the full molecular mechanics of the total junction that arises from the large number of atoms involved. For example, upon pulling, the molecule is sliding but also snaking over the surface, a mechanical motion we have found difficult to reproduce in the calculation and which easily shifts the conductance oscillation by a few angstrom in the $z$ direction.

Very often, the chain end to tip apex bond breaks before a complete vertical pulling, but in some cases, it remains attached to the tip until the entire chain is lifted up vertically. We measured the conductance of such extremely long molecular wires up to 20 nm in length. Figure 4A shows such a case, in which a single chain is lifted up at the left end (the position of the tip approach is marked by a cross). The $I(z)$ curve during lifting (Fig. 4B) exhibits an abrupt decrease at the end of the pulling process and shows that the tip to surface electronic interaction through the molecular wire is lost exactly at a $z = 20$ nm distance. Hence, the chain is in a fully elongated conformation before the break, while being contacted vertically between the tip apex and the surface (Fig. 4C). To record the $I(z)$ curve for large tip–molecular wire–surface junctions in Fig. 4B, it was necessary to increase the junction bias voltage continuously from –100 mV up to –4 eV during the pulling, compensating for the tunneling current exponential decay at low bias voltage. This result confirms that $\beta$ is energy dependent, with a lower value in the energy gap

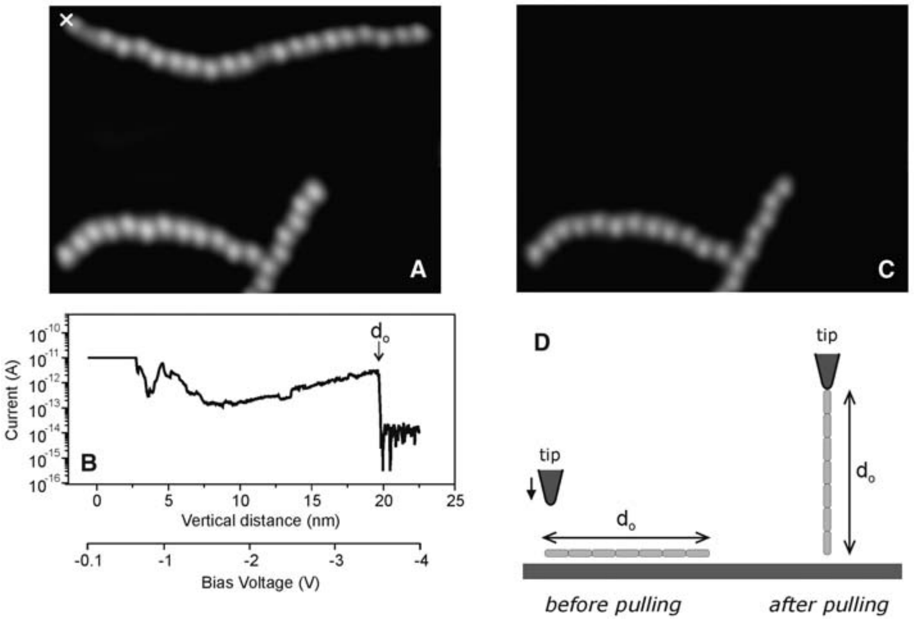


**Fig. 4.** Complete detachment of a long single chain. STM images (20 by 15 nm) before (**A**) and after (**C**) a vertical manipulation process in which the entire chain is lifted from the surface. (**B**) Current curve as a function of the vertical tip distance from the surface (the corresponding bias voltage scale is shown below). The abrupt decrease at a tip height of about 20 nm reflects the disruption of the tip-surface connection through the chain and hence its complete removal at a vertical distance equal to its length $d_o$. (**D**) Schematic views of the tip-chain configurations before and after this pulling process. At the end, the chain is attached vertically to the tip, which is thus elongated by $d_o$, and therefore removed from the surface [see (C)].

when approaching the HOMO or LUMO resonance of the molecular wire (*28*). The increased conductance at higher bias voltages then compensates for the molecular wire length increase in Fig. 4B (a factor of 40 is not sufficient for the achieved current increase). Thus, such a setup allows the determination of the small conductance ($8.6 \times 10^{-13}$ S) of a single and the same molecular wire with 20 nm length (the conductance at small bias voltages cannot be measured over such a large distance, due to the extremely low current—below the detection limit—passing through the polymer in this case). In this regard, it would be interesting to prepare and study conjugated polymers with smaller HOMO-LUMO gaps. Such molecular wires should exhibit higher conductances and allow charge transport to be determined over even larger distances.

### References and Notes

1. C. Joachim, J. K. Gimzewski, A. Aviram, *Nature* **408**, 541 (2000).
2. A. Nitzan, *Annu. Rev. Phys. Chem.* **52**, 681 (2001).
3. A. Nitzan, M. A. Ratner, *Science* **300**, 1384 (2003).
4. N. J. Tao, *Nat. Nanotechnol.* **1**, 173 (2006).
5. C. Joachim, J. K. Gimzewski, R. R. Schlittler, C. Chavy, *Phys. Rev. Lett.* **74**, 2102 (1995).
6. M. A. Reed, C. Zhou, C. J. Muller, T. P. Burgin, J. M. Tour, *Science* **278**, 252 (1997).
7. R. M. H. Smit *et al.*, *Nature* **419**, 906 (2002).
8. J. Reichert *et al.*, *Phys. Rev. Lett.* **88**, 176804 (2002).
9. S. Wu *et al.*, *Nat. Nanotechnol.* **3**, 569 (2008).
10. M. Kiguchi *et al.*, *Phys. Rev. Lett.* **101**, 046801 (2008).
11. B. Xu, N. J. Tao, *Science* **301**, 1221 (2003).
12. W. Haiss *et al.*, *J. Am. Chem. Soc.* **125**, 15294 (2003).
13. X. Xiao, B. Xu, N. Tao, *J. Am. Chem. Soc.* **126**, 5370 (2004).
14. J. He *et al.*, *J. Am. Chem. Soc.* **127**, 1384 (2005).
15. S. H. Choi, B. Kim, C. D. Frisbie, *Science* **320**, 1482 (2008).
16. W. Haiss *et al.*, *Nat. Mater.* **5**, 995 (2006).
17. R. Temirov, A. Lassise, F. B. Anders, F. S. Tautz, *Nanotechnology* **19**, 065401 (2008).
18. F. Kühner, M. Erdmann, H. E. Gaub, *Phys. Rev. Lett.* **97**, 218301 (2006).
19. S. K. Kufer, E. M. Puchner, H. Gumpp, T. Liedl, H. E. Gaub, *Science* **319**, 594 (2008).
20. T. Zambelli *et al.*, *Int. J. Nanoscience* **3**, 331 (2004).
21. L. Grill *et al.*, *Nat. Nanotechnol.* **2**, 687 (2007).
22. Materials and methods are available as supporting material on *Science* Online.
23. L. Limot, J. Kröger, R. Berndt, A. Garcia-Lekue, W. A. Hofer, *Phys. Rev. Lett.* **94**, 126102 (2005).
24. C. Joachim, M. A. Ratner, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 8801 (2005).
25. M. Magoga, C. Joachim, *Phys. Rev. B* **56**, 4722 (1997).
26. M. Magoga, C. Joachim, *Phys. Rev. B* **57**, 1820 (1998).
27. D. Neher, *Macromol. Rapid Commun.* **22**, 1365 (2001).
28. C. Joachim, M. Magoga, *Chem. Phys.* **281**, 347 (2002).
29. Financial support from the European Integrated Project PICO INSIDE and the Deutsche Forschungsgemeinschaft (DFG) through SFB 658 and contract GR 2697/1-2 is gratefully acknowledged.

**Supporting Online Material**
www.sciencemag.org/cgi/content/full/323/5918/1193/DC1
Materials and Methods
Figs. S1 to S6
References

10 November 2008; accepted 16 January 2009
10.1126/science.1168255

---

# Early Hominin Foot Morphology Based on 1.5-Million-Year-Old Footprints from Ileret, Kenya

**Matthew R. Bennett,[1]* John W.K. Harris,[2] Brian G. Richmond,[3,4] David R. Braun,[5] Emma Mbua,[6] Purity Kiura,[6] Daniel Olago,[7] Mzalendo Kibunjia,[6] Christine Omuombo,[7] Anna K. Behrensmeyer,[8] David Huddart,[9] Silvia Gonzalez[9]**

**Hominin footprints offer evidence about gait and foot shape, but their scarcity, combined with an inadequate hominin fossil record, hampers research on the evolution of the human gait. Here, we report hominin footprints in two sedimentary layers dated at 1.51 to 1.53 million years ago (Ma) at Ileret, Kenya, providing the oldest evidence of an essentially modern human–like foot anatomy, with a relatively adducted hallux, medial longitudinal arch, and medial weight transfer before push-off. The size of the Ileret footprints is consistent with stature and body mass estimates for *Homo ergaster/erectus*, and these prints are also morphologically distinct from the 3.75-million-year-old footprints at Laetoli, Tanzania. The Ileret prints show that by 1.5 Ma, hominins had evolved an essentially modern human foot function and style of bipedal locomotion.**

Bipedalism is a key human adaptation that appears in the fossil record by 6 million years ago (Ma) (*1*). Considerable debate continues over when and in what context a modern human–like form of bipedalism evolved, because of a fragmentary record and disagreements over the functional interpretations of existing fossils and footprints (*2*–*7*). Modern human footprints reflect the specialized anatomy and function of the human foot, which is characterized by a fully adducted hallux, a large and robust calcaneus and tarsal region, a pronounced medial longitudinal arch, and short toes (*2*). Footprints reflect the pressure distribution as the foot makes contact with the substrate, but also the sediment's geomechanical properties (*8*). During normal walking, the weight-bearing foot undergoes a highly stereotypical movement and pressure distribution pattern in which the heel contacts the ground first, making a relatively deep impression on the substrate. This is followed by contact with the lateral side of the foot and metatarsal heads, after which weight transfers to the ball of the foot with peak pressure under the medial metatarsal heads, and finally ending with toe-off pressure under the hallux (*9*, *10*). As a consequence, the deepest part of a footprint often occurs beneath the first and second metatarsal heads, that along with a deep hallucal impression corresponds to the peak pressures at toe-off (*10*). The extent to which any pressure, or footprint impression, occurs medially varies with the anatomy of the midfoot, including the height of the longitudinal arch and other factors (*11*), and the extent to which lateral toes leave impressions depends on factors such as foot orientation relative to the direction of travel, transverse versus oblique push-off axes, and substrate properties. This contrasts with the less stereotypical pattern of footfall observed in African apes during quadrupedal and bipedal locomotion. Here the heel and lateral mid-foot make contact with the ground first, followed by contact with the lateral toes that are often curled and with a hallux that is often widely abducted. Lift-off in the African apes is variable, but it usually involves relatively low pressure during final contact by both the lateral toes and widely abducted hallux, in stark contrast to modern human foot function (*11*).

Here, we report hominin footprints from the Okote Member of the Koobi Fora Formation (*12*), second in age only to the mid-Pliocene (3.7 Ma) Laetoli prints (*13*), located close to Ileret, Kenya (Fig. 1; site FwJj14E; latitude 4°18′44″N, longitude 36°16′16″E). The footprints are found in association with animal prints on two stratigraphically separated levels and were digitized with an optical laser scanner (Fig. 1) (*14*). The upper surface contains three hominin footprint trails comprising two trails of two prints and one of seven prints, as well as a number of isolated prints (Figs. 2 and 3 and figs. S3 and S6 to S11). The lower surface, approximately 5 m below, preserves one trail of two prints and a single isolated hominin print (Fig. 3). The footprints occur within a 9-m-thick sequence of fine-grained, normally

---

[1] School of Conservation Sciences, Bournemouth University, Poole, BH12 5BB, UK. [2] Department of Anthropology, Rutgers University, 131 George Street, New Brunswick, NJ 08901, USA. [3] Center for the Advanced Study of Hominid Paleobiology, Department of Anthropology, George Washington University, Washington, DC 20052, USA. [4] Human Origins Program, National Museum of Natural History, Smithsonian Institution, Washington, DC 20013–7012, USA. [5] Department of Archaeology, University of Cape Town, Private Bag, Rondebosch 7701, South Africa. [6] National Museums of Kenya, Post Office Box 40658-00100, Nairobi, Kenya. [7] Department of Geology, University of Nairobi, Post Office Box 30197, Nairobi, Kenya. [8] Department of Paleobiology, MRC 121, National Museum of Natural History, Smithsonian Institution, Washington, DC, 20013–7012, USA. [9] School of Biological and Earth Sciences, Liverpool John Moores University, Liverpool, L3 3AF, UK

*To whom correspondence should be addressed. E-mail: mbennett@bmth.ac.uk

graded, silt and sand units deposited as overbank flood deposits with evidence of paleosol development. Interbedded within this succession are three fluvially reworked volcanic ashes; the upper ash (Northern Ileret Tuff) forms a prominent landscape bench that correlates with other nearby sites where traces of hominin activity have been recovered (*15*) and is unconformably overlain by the Galana Boi Formation of Holocene age (*12*). The ash layers are correlated geochemically to dated tuffs within the Turkana Basin, thereby providing an age of 1.51 to 1.52 Ma for the upper tuff and 1.53 Ma for the lower tuff (Fig. 1) (*14*, *16*).

The prints from both the upper and lower levels at FwJj14E have a well-defined, deeply depressed and adducted hallux; visible lateral toe impressions that vary in depth of indentation; a well-defined ball beneath the first and second metatarsal heads; and a visible instep reflecting a medial longitudinal arch. The angle of hallux abduction, relative to the long axis of the foot, is typically 14° compared to, and statistically distinct from (table S4), 8° for the modern reference prints and 27° for the Laetoli prints (Fig. 4A). The morphology of the Ileret prints suggests that the feet of these hominins had functional medial longitudinal arches. In prints FUT1-2 and FLI1, for example, the medial side of the mid-print is slightly raised, indicating a lack of substrate impression in this area (Figs. 2 and 3). A comparison of the instep width relative to the width in the metatarsal head region shows that the upper prints at FwJj14E fall within the modern human range and are distinct from the relatively wider insteps characterizing the Laetoli prints (Fig. 4C). The FwJj14E lower prints show more variability; the two prints that differ from modern human prints appear to have undergone taphonomic mediolateral compression (fig. S13). Most of the Ileret prints are similar in length to the longest modern human prints (fig. S19), although the isolated footprint on the lower level (FLI1; Fig. 3) is significantly smaller, despite a similar gross anatomy, and may represent a subadult.

The contours of the upper- and lower-level FwJj14E prints suggest a modern human–like toe-off mechanism, in contrast to the more ambiguous evidence from the Laetoli footprints

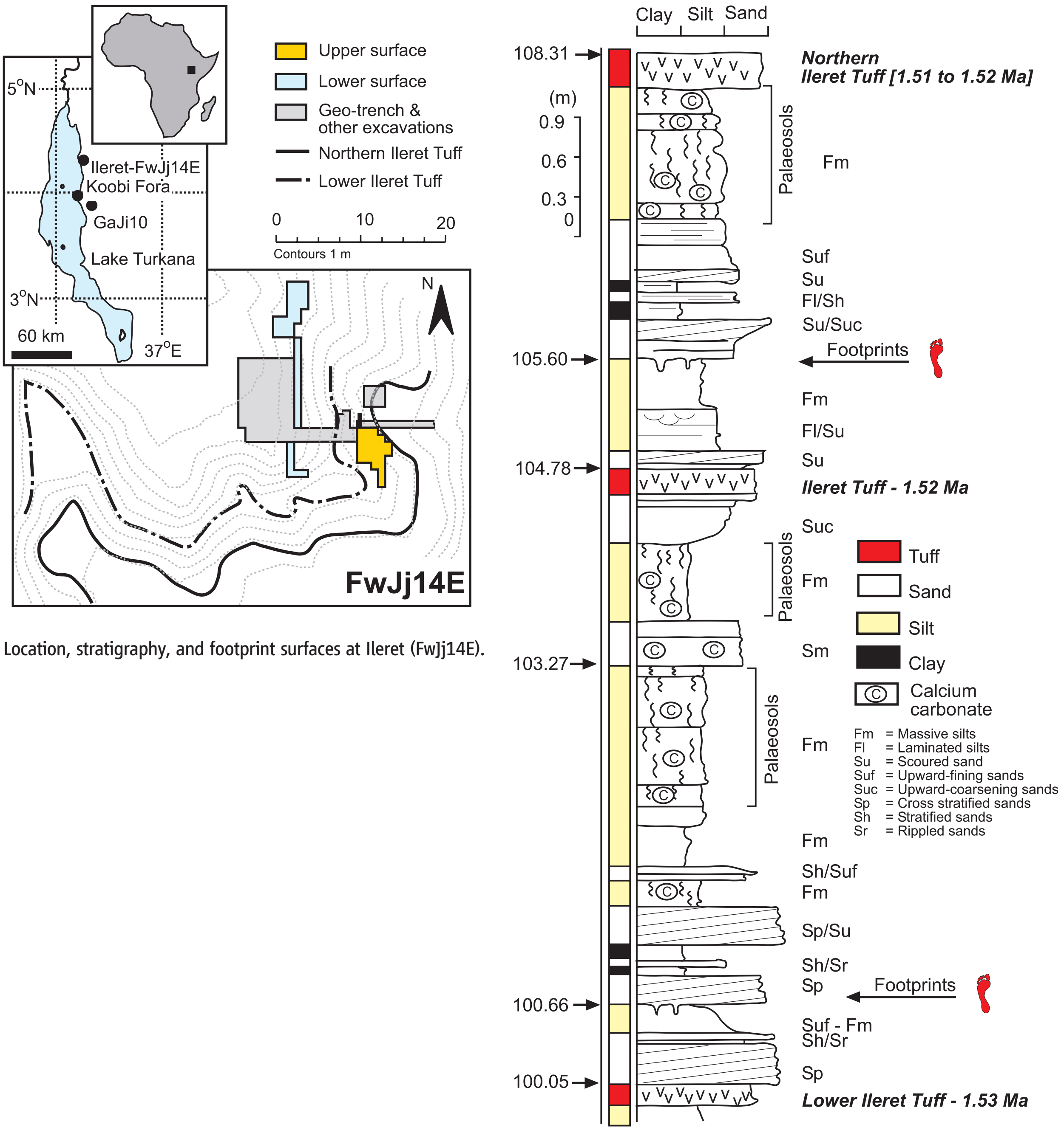


**Fig. 1.** Location, stratigraphy, and footprint surfaces at Ileret (FwJj14E).

(Fig. 4B). Both the upper and lower Ileret prints show the greatest depth in the metatarsal head region to be typically medially located, in contrast to its lateral position in many of the Laetoli prints (fig. S19). In the majority of the modern human prints, the normal shift in pressure from lateral to medial in late stance phase, with peak medial pressures, is registered in the greater medial depth. However, the variability shows that fully modern feet during walking can also produce prints with greater lateral depth (fig. S19) (*3*, *6*), demonstrating variability in the foot function, something that is further complicated by the geomechanical properties of the substrate. Therefore, a laterally concentrated depth cannot, in itself, rule out (for instance, in the Laetoli prints) a foot structure capable of medial weight transfer in late stance phase. However, the modern human–like relative instep width (Fig. 4C) and medially concentrated ball impression in the Ileret prints provide compelling positive evidence of a medial longitudinal arch and the medial pressure shift and push-off from the ball of the foot beneath the medial metatarsal heads.

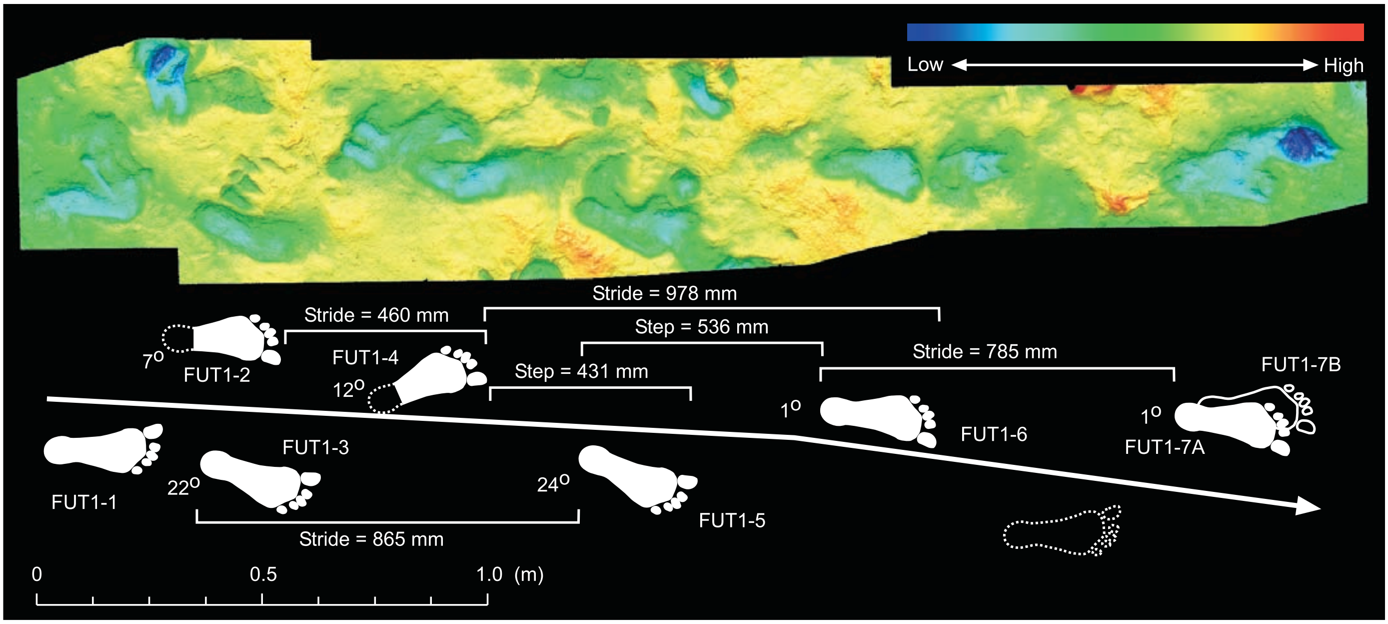


**Fig. 2.** Tessellated swath of optical laser scans of the main footprint trail on the upper footprint surface at FwJj14E. Color is rendered with 5-mm isopleths.

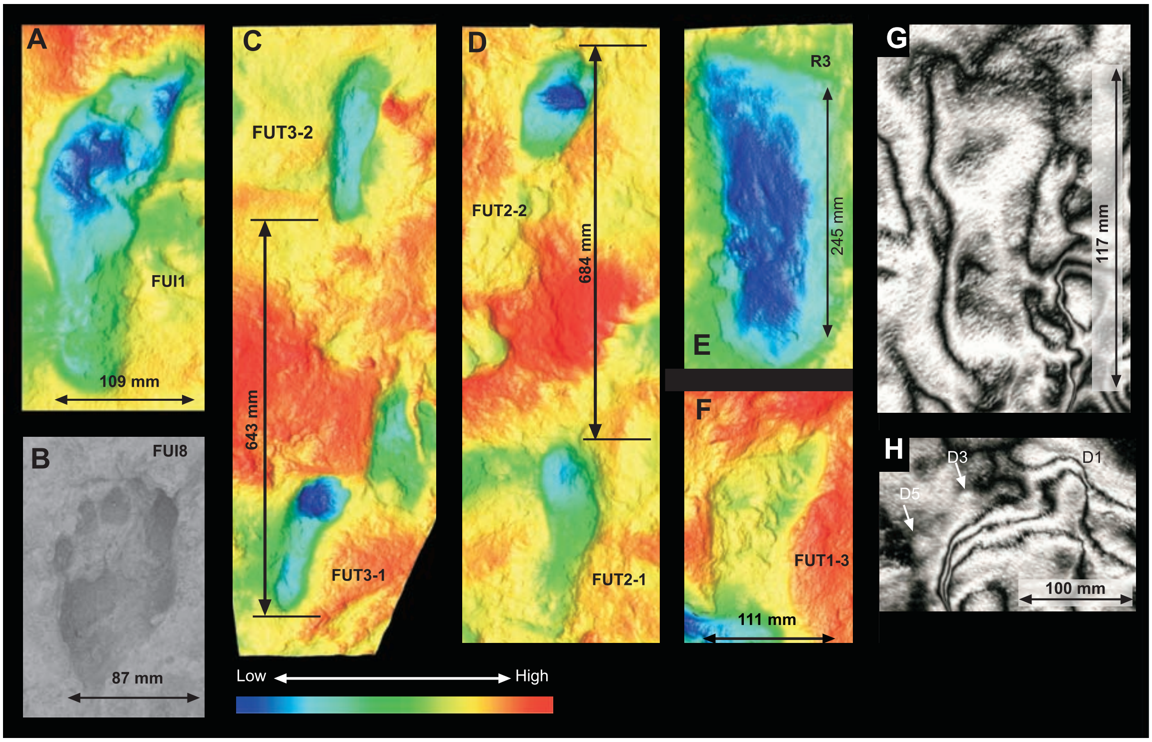


**Fig. 3.** Optical laser scan images color-rendered with 5-mm isopleths for footprints at both FwJj14E and GaJi10. (**A**) Isolated left foot (FUI1) on the upper footprint surface at FwJj14E. (**B**) Photograph of FUI8 on the upper footprint surface at FwJj14E, showing good definition of the toe pads; the second toe is partially obscured by the third toe. (**C**) Second trail on the upper footprint surface at FwJj14E, showing two left feet. (**D**) Third trail on the upper footprint surface at FwJj14E, showing a right and a left foot. (**E**) Print R3 from GaJi10 (*22*), re-excavated and scanned as part of this investigation. (**F**) Partial print (FUT1-2) on the upper footprint surface at FwJj14E; the heel area has been removed by a later bovid print. (**G**) Print FLI1 on the lower footprint surface at FwJj14E, rendered with 5-mm alternating black and white isopleths. (**H**) Inverted image of the toe area of print FUT1-1 with alternating 5-mm black and white isopleths. Note the locations of the pads of the small toes and the presence of a well-defined ball beneath the hallucial metatarsophalangeal joint. The first, third, and fifth toes are marked D1, D3, and D5, respectively.

These are hallmarks of modern human walking that can be related to a long stride with extended lower limbs, key to the energetic efficiency of human walking (*17*).

Inferences about stride are possible from the short trails at FwJj14E. The longest trail (FUT1) consists of seven prints, initially with three closely spaced prints suggesting an individual either standing with feet astride or who slowed before walking forward with an increasing step length (Fig. 2). Typical stride lengths vary from 460 to 785 mm, with step lengths in the range of 431 to 536 mm in the direction of travel; foot angles vary, but are typically parallel to the direction of travel, diverging by 1° to 24°. Using the average of the last three strides (average = 876 mm) and a hip height of 860 mm estimated from foot length (258 mm, the average of the four clearest prints of the FUT1 trail), we estimate a velocity of approximately 0.63 m/s (*14*). This is a slow speed consistent with someone beginning to walk from a standing (or slowed) position and the variability in step lengths may attest to the challenges of walking on an uneven muddy surface already marred by a range of animal prints. Although precise foot-to-stature ratios are unknown for early Pleistocene hominins, we use as a rough estimate ones developed for Australian Aborigines (*18*) and Kenyan Dassenach (*14*), on the basis that their statures are adapted for a semiarid environment. Using this relationship, we estimate the average height of the individuals from the prints on the upper surface to be 1.75 ± 0.26 m and 1.76 ± 0.26 m for those on the lower surface, excluding the potential subadult (print FUI1) which gives a height of 0.92 ± 0.13 m (table S2). The large stature and mass estimates derived from the Ileret prints compare well with those of *Homo ergaster/erectus* on the basis of postcranial remains and are significantly larger than postcrania-based stature and mass estimates for *Paranthropus boisei* and *Homo habilis* (table S3) (*19–21*), suggesting that the prints at FwJj14E were made by *Homo ergaster/erectus* individuals.

Behrensmeyer and Laporte (*22*) reported hominin footprints at site GaJi10 (latitude 3°44′15″N, longitude 36°55′48″E), 45 km to the south of FwJj14E, in 1981. The footprint surface occurs below a prominent tuff, sampled and correlated here to the Akait Tuff, dated to 1.435

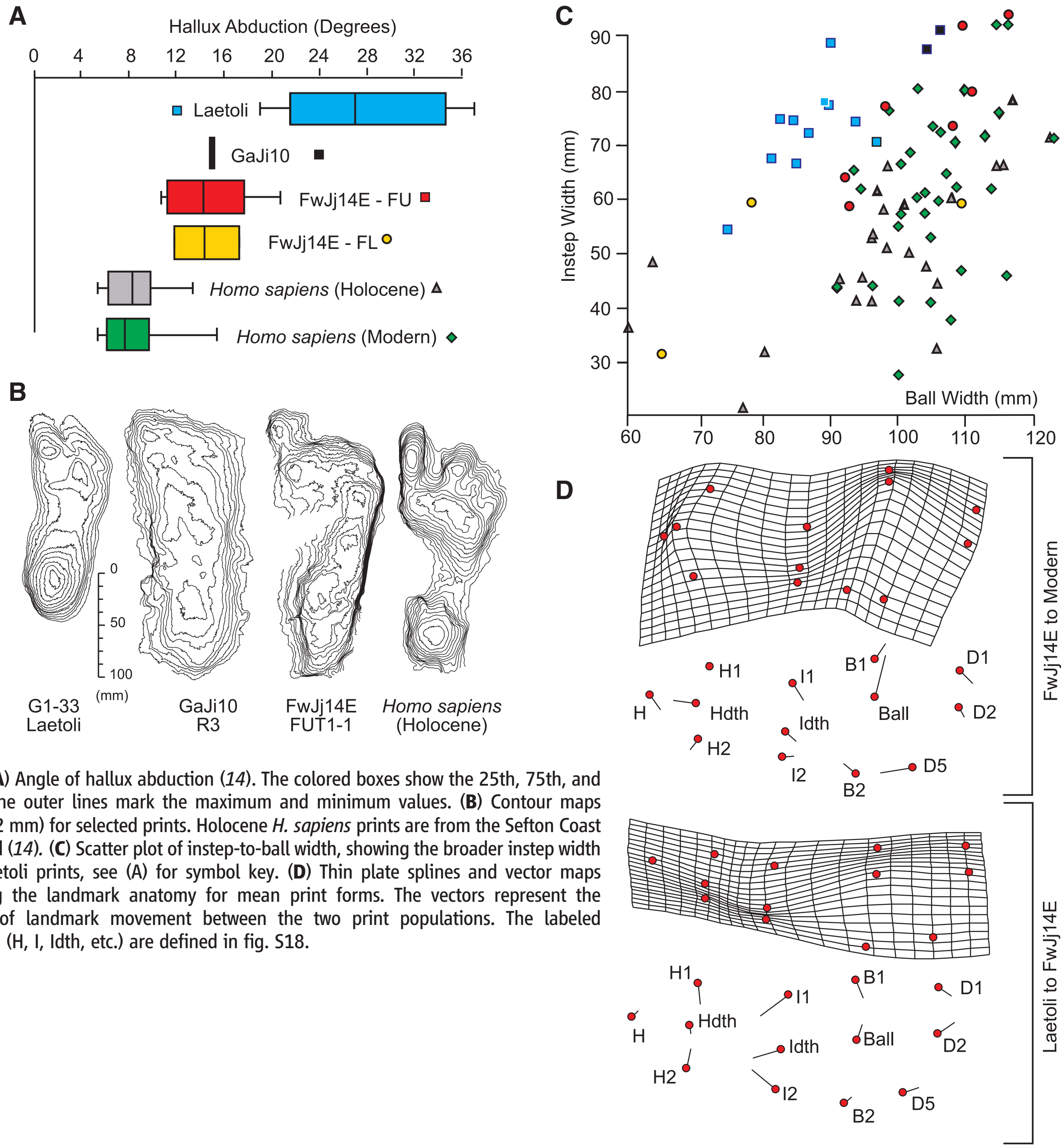


**Fig. 4.** (**A**) Angle of hallux abduction (*14*). The colored boxes show the 25th, 75th, and median; the outer lines mark the maximum and minimum values. (**B**) Contour maps (interval, 2 mm) for selected prints. Holocene *H. sapiens* prints are from the Sefton Coast in England (*14*). (**C**) Scatter plot of instep-to-ball width, showing the broader instep width of the Laetoli prints, see (A) for symbol key. (**D**) Thin plate splines and vector maps comparing the landmark anatomy for mean print forms. The vectors represent the direction of landmark movement between the two print populations. The labeled landmarks (H, I, Idth, etc.) are defined in fig. S18.

Ma (*14*). Re-excavation of these prints uncovered four of the original seven prints and a new print, and the two best-preserved examples are comparable to those at FwJj14E (*14*).

To further evaluate the morphology of the prints at FwJj14E and compare them objectively with samples of modern human and Laetoli footprints, we digitized 13 landmarks on each footprint scan and used generalized Procrustes analysis to compare their shapes (*14*). Figure 4D shows thin plate splines and landmark vector maps comparing the mean landmark positions (in two dimensions) of the Ileret prints with the modern human and Laetoli prints (*14*). When the prints from the two levels at FwJj14E are compared, both surfaces show similar anatomical differences from the modern prints, with a narrower heel and ball area and a wider instep associated with less pronounced arch elevation (Fig. 4D and fig. S20). When compared to the Laetoli prints, the Ileret prints have a more contracted proximal mid-foot region, including a deeper instep (Fig. 4D), suggesting the presence of a medial longitudinal arch. The location of the narrowest point of the instep also lies farther forward (more distal) in the Laetoli prints than in both the modern and Ileret prints, possibly reflecting differences in foot proportions or a lack of definition of the instep. Discriminant analysis was also used to compare the different print populations (*14*); 8 of the 10 prints at FwJj14E used in the analysis were classified with modern prints and two with Laetoli prints (table S6).

The Ileret footprints show the earliest evidence of a relatively modern human–like foot with an adducted hallux, a medial longitudinal arch, and medial weight transfer before push-off. Although we cannot conclude with certainty what hominin species made the footprints at FwJj14E or GaJi10, these modern human characteristics, in combination with the large size of the prints, are most consistent with the large size and tall stature evident in some *Homo ergaster/erectus* individuals (*19*, *20*). These prints add to the anatomical (*19*, *20*, *23*) and archaeological (*24*, *25*) evidence pointing to a major transition in human evolution with the appearance of hominins with long lower limbs, conferring advantages at a lower energetic cost (*26*), and archaeological indications of activities in a variety of ecological settings and the transport of resources over long distances (*27*). These lines of evidence, together with the earliest evidence of a relatively modern foot anatomy and function, support the hypothesis that this was a hominin with a larger home range related to increasing average body size and enhanced dietary quality (*28*). These factors add to an emerging picture of the paleobiology of *H. ergaster/erectus* that suggests a shift in cultural and biological adaptations relative to earlier hominins.

### References and Notes

1. B. G. Richmond, W. L. Jungers, *Science* **319**, 1662 (2008).
2. N. L. Griffin, B. Wood, in *The Human Foot: A Companion to Clinical Studies*, L. Klenerman, B. Wood, Eds. (Springer-Verlag, London, 2006), pp. 27–79.
3. T. D. White, G. Suwa, *Am. J. Phys. Anthropol.* **72**, 485 (1987).
4. R. H. Tuttle, D. Webb, E. Weidle, M. Baksh, *J. Arch. Sci.* **17**, 347 (1990).
5. J. Charteris, J. J. Wall, J. W. Nottrodt, *Am. J. Phys. Anthropol.* **58**, 133 (1982).
6. D. J. Meldrum, in *From Biped to Strider: The Emergence of Modern Human Walking, Running, and Resource Transport*, D. J. Meldrum, C. E. Hilton, Eds. (Kluwer Academic/Plenum Publishers, New York, 2004), pp. 63–83.
7. J. T. Stern Jr., R. L. Susman, *Am. J. Phys. Anthropol.* **60**, 279 (1983).
8. J. R. L. Allen, *Philos. Trans. R. Soc. London Ser. B* **352**, 481 (1997).
9. H. Elftman, J. Manter, *Am. J. Phys. Anthropol.* **20**, 69 (1935).
10. E. Vereecke, K. D'Aout, D. De Clercq, L. Van Elsacker, P. Aerts, *Am. J. Phys. Anthropol.* **120**, 373 (2003).
11. E. Morag, P. R. Cavanagh, *J. Biomech.* **32**, 359 (1999).
12. C. S. Feibel, F. H. Brown, I. McDougall, *Am. J. Phys. Anthropol.* **78**, 595 (1989).
13. M. D. Leakey, R. L. Hay, *Nature* **278**, 317 (1979).
14. Materials and methods, additional illustrations, statistical analysis, and discussion of tuff geochemistry at FwJj14E are available as supporting material on *Science* Online.
15. J. W. K. Harris, D. G. Braun, J. T. McCoy, B. L. Pobiner, M. J. Rogers, *J. Hum. Evol.* **42**, A15 (2002).
16. F. H. Brown, B. Haileab, I. McDougall, *J. Geol. Soc. London* **163**, 185 (2006).
17. M. D. Sockol, D. A. Raichlen, H. Pontzer, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 12265 (2007).
18. S. Webb, M. L. Cupper, R. Robins, *J. Hum. Evol.* **50**, 405 (2006).
19. C. B. Ruff, A. Walker, in *The Nariokotome* Homo erectus *Skeleton*, A. C. Walker, R. E. F. Leakey, Eds. (Harvard Univ. Press, Cambridge, MA, 1993), pp. 234–265.
20. H. M. McHenry, K. Coffing, *Annu. Rev. Anthropol.* **29**, 125 (2000).
21. F. Spoor *et al.*, *Nature* **448**, 688 (2007).
22. A. K. Behrensmeyer, L. F. Laporte, *Nature* **289**, 167 (1981).
23. B. G. Richmond, L. C. Aiello, B. A. Wood, *J. Hum. Evol.* **43**, 529 (2002).
24. M. J. Rogers, J. W. K. Harris, C. S. Feibel, *J. Hum. Evol.* **27**, 139 (1994).
25. B. L. Pobiner, M. J. Rogers, C. M. Monahan, J. W. K. Harris, *J. Hum. Evol.* **55**, 103 (2008).
26. H. Pontzer, *J. Exp. Biol.* **210**, 1752 (2007).
27. D. R. Braun, J. W. K. Harris, D. N. Maina, *Archaeometry* 10.1111/j.1475-4754.2008.00399 (2008).
28. S. C. Antón, W. R. Leonard, M. L. Robertson, *J. Hum. Evol.* **43**, 773 (2002).
29. This project is part of the Koobi Fora Research and Training Program, which is jointly run by the National Museum of Kenya and Rutgers University and supported by the Holt Family Foundation, M. Weiss, B. Mortenson, G. Noland, C. Chip, and the Wenner-Gren Foundation for Anthropological Research. B. Wood provided valuable comments on a draft of the paper, and the dating was undertaken by F. Brown. The field assistance of H. Manley, S. A. Morse, and M. J. Steele is acknowledged.

**Supporting Online Material**

www.sciencemag.org/cgi/content/full/323/5918/1197/DC1
Methods
SOM Text
Figs. S1 to S22
Tables S1 to S6
References and Notes

5 November 2008; accepted 19 January 2009
10.1126/science.1168132

---

# RNA Polymerase IV Functions in Paramutation in *Zea mays*

**Karl F. Erhard Jr., Jennifer L. Stonaker,* Susan E. Parkinson,* Jana P. Lim, Christopher J. Hale, Jay B. Hollick†**

**Plants have distinct RNA polymerase complexes (Pol IV and Pol V) with largely unknown roles in maintaining small RNA–associated gene silencing. Curiously, the eudicot *Arabidopsis thaliana* is not affected when either function is lost. By use of mutation selection and positional cloning, we showed that the largest subunit of the presumed maize Pol IV is involved in paramutation, an inherited epigenetic change facilitated by an interaction between two alleles, as well as normal maize development. Bioinformatics analyses and nuclear run-on transcription assays indicate that Pol IV does not engage in the efficient RNA synthesis typical of the three major eukaryotic DNA-dependent RNA polymerases. These results indicate that Pol IV employs abnormal RNA polymerase activities to achieve genome-wide silencing and that its absence affects both maize development and heritable epigenetic changes.**

In maize, mouse, and other eukaryotes, paramutation refers to a process by which heritable changes in gene regulation are facilitated by interactions between alleles on homologous chromosomes (*1*). As typically described, alleles conferring relatively high gene action invariably change to a repressed expression state when heterozygous with specific alleles or allelic states (*1*). Operationally, paramutation violates the first law of Mendelian inheritance that alleles segregate unchanged from a heterozygote and thus has important implications for normal genome function and evolution (*2*, *3*), though few examples have proved experimentally tractable.

In maize, paramutations occurring at the *Pl1-Rhoades* (*Pl1-Rh*) allele of the *purple plant 1* (*pl1*) locus involve at least four genes (*4*–*6*), two of which appear to be part of a small interfering RNA (siRNA) heterochromatin pathway (*7*–*9*). Repressed expression states of *Pl1-Rh* resulting from

---

Department of Plant and Microbial Biology, 111 Koshland Hall, University of California, Berkeley, CA 94720-3102, USA.

*These authors contributed equally to this work.

†To whom correspondence should be addressed. E-mail: hollick@nature.berkeley.edu

paramutation are heritably maintained by Required to Maintain Repression1 (RMR1), a previously undiscovered Sucrose Nonfermenting 2 (SNF2)–like adenosine triphosphatase, and Mediator of Paramutation1 (MOP1), a putative RNA-dependent RNA polymerase (RDR) (*4*, *8*, *9*). RMR1 and MOP1 appear to be similar to proteins involved in an RNA-directed DNA methylation (RdDM) pathway in the eudicot *Arabidopsis thaliana* (*7*). RMR1 is most similar to *Arabidopsis* CLASSY1 (CLSY1) and DEFECTIVE IN RNA-DIRECTED METHYLATION1 (DRD1) (*5*, *7*), whereas MOP1 appears orthologous with *Arabidopsis* RDR2 (*8*, *9*). In *Arabidopsis*, RDR2- and DICER-LIKE3-dependent 24-nucleotide (nt) siRNAs associate with a specific argonaute protein (AGO4) and are thought to guide, via sequence homology, the deposition of de novo cytosine methylation marks associated with transcriptionally repressed heterochromatin (*10*). The majority of genomic loci targeted by RdDM, and related small RNA (sRNA)–based silencing pathways in other eukaryotes, are transposable elements and other repetitive sequences (*10*, *11*), which suggests an interplay between paramutation-induced repression and the means by which eukaryotic genomes cope with potentially deleterious repetitive features.

A terminal fragment of a CACTA-like DNA transposon related to the *doppia* subfamily (*12*) resides immediately 5′ of the *Pl1-Rh* coding region and is targeted by an RdDM-type pathway (*7*). Plants homozygous for *rmr1*, *mop1*, or *rmr6* mutations have hypomethylated cytosines across this *doppia* fragment relative to heterozygous siblings (*7*), and RMR1 is also required for accumulation of *doppia*-like small (~24 nt) RNA species (*7*). However, the functional connection between *doppia* cytosine methylation and *Pl1-Rh* paramutation remains unclear because no correlations between cytosine methylation levels and *Pl1-Rh* expression levels in nonmutant plants have been observed (*7*). In contrast, *rmr6* mutant plants have increases in *Pl1-Rh* transcription rates that are correlated with 5′ hypomethylation in all cytosine contexts within, 5′ of, and immediately 3′ of the *doppia* fragment (*6*, *7*).

Many of the *Arabidopsis* sRNA-associated repression mechanisms, including RdDM, are hypothesized to be dependent on the actions of two presumed DNA-dependent RNA polymerase complexes, Pol IV (Pol IVa) and Pol V (Pol IVb) (*10*, *13*). All plants appear to have a conserved Pol IV largest subunit (RPD1) derived from an ancient duplication of the Pol II largest subunit (RPB1), and all flowering plants appear to have an RPD1 duplicate called RPE1 (*14*). The *Arabidopsis* RPD1 (AtNRPD1a) and RPE1 (AtNRPD1b) form two functionally distinct multisubunit complexes (Pol IV and Pol V, respectively) that share a single subunit, AtNRPD2 (*13*). Although both polymerase forms appear to be related to Pol II, neither the ability to produce an RNA transcript nor the exact nature of any potential nucleic acid template is known for the Pol IV complex (*13*), whereas recent evidence does indicate the involvement of Pol V in generating low-abundance RNAs (*15*). Additionally, knockout mutants show that both Pol IV and Pol V are nonessential for normal plant development in *Arabidopsis* (*13*).

Because *rmr6* mutants are defective for both paramutation and proper maize development (*6*, *16*) (fig. S1), we set out to identify the gene product. We mapped *rmr6* to a ~450-kb interval on maize chromosome 1L, which is syntenic with a region on rice chromosome 8 (fig. S2A) and contains six contiguous maize-specific gene models that are absent from the syntenic rice interval (fig. S2B) (*17*). Four of these models coalesced to a single, putative, 18-exon gene of ~65 kb (Fig. 1A) with the potential to encode a 1444–amino acid polypeptide with similarity to NRPD1a, the largest subunit of *Arabidopsis* Pol IV (Fig. 1B) (*18*). Because Pol IV acts in an RdDM-type pathway to maintain epigenetic repression of genomic targets (*13*), DNA from this candidate gene was amplified and sequenced in homozygous *rmr6* mutant plants of four different ethyl metanesulfonate–derived *rmr6* alleles (*17*). These efforts identified single transition–type lesions within the predicted coding regions of the *NRPD1a*-like candidate gene (Fig. 1B). Three mutations (*rmr6-1*, *rmr6-7*, and *rmr6-14*) create premature nonsense codons, and one (*rmr6-8*) encodes a nonconservative amino acid substitution at a cysteine residue that is highly conserved among predicted RPD1 sequences (Fig. 1B and fig. S3). These molecular lesions strongly indicated that *rmr6* encodes a Pol IV largest subunit.

Additional molecular profiles also supported the assignment of RMR6 as a functional RPD1 ortholog. *Arabidopsis* showed a 14- to 15-fold reduction of 21- to 24-nt siRNAs in *nrpd1a*/*nrpd1b* double mutants relative to wild-type plants, and the accumulation of ~94% of these RNAs was RPD1-dependent (*19*, *20*). We fractionated sRNAs from immature cobs and noted by ethidium bromide staining intensity that the ~24-nt RNA species was reduced an average of 82% (±5% SEM; *n* = 3 biological replicates) in *rmr6-1* mutants, whereas the ~21-nt RNAs were not affected (Fig. 1C). In *Arabidopsis*, the accumulation of 21-nt RNAs representing mature microRNAs (miRNAs) and *trans*-acting siRNAs (tasiRNAs) also appeared to be unaffected in Pol IV and Pol V mutants (*7*). Because Pol IV–dependent siRNAs are largely representative of transposon-like sequences in *Arabidopsis* (*19*, *20*), and because of the effect *rmr6* mutations have on cytosine methylation patterns 5′ of the *Pl1-Rh* coding region nearby and inside the proximal *doppia*-like transposon sequence (*7*), we specifically tested the effect of *rmr6* mutations on the accumulation of *doppia*-like sRNAs. Northern blots showed that *doppia*-related sRNAs representing both strands were present in nonmutant plants but were undetectable in *rmr6-1* mutants (Fig. 1D and fig. S5). Given the prevalence of *doppia* elements throughout the genome and the sequence arrangement within the terminal inverted repeats (*12*), it is unknown whether these sRNAs emanate from one or more elements or from any particular orientation. These molecular profiles, combined with the null-mutant phenotypes, indicate that *rmr6* encodes a *Zea mays* Pol IV largest subunit, hereafter referred to as ZmRPD1.

All maize *rmr* loci were identified in mutant screens for plants unable to maintain paramutation-induced repression of alleles encoding color factors (for example, *Pl1-Rh*) (*5*, *6*); however, *rmr6* mutations also affect development (*16*). ZmRPD1

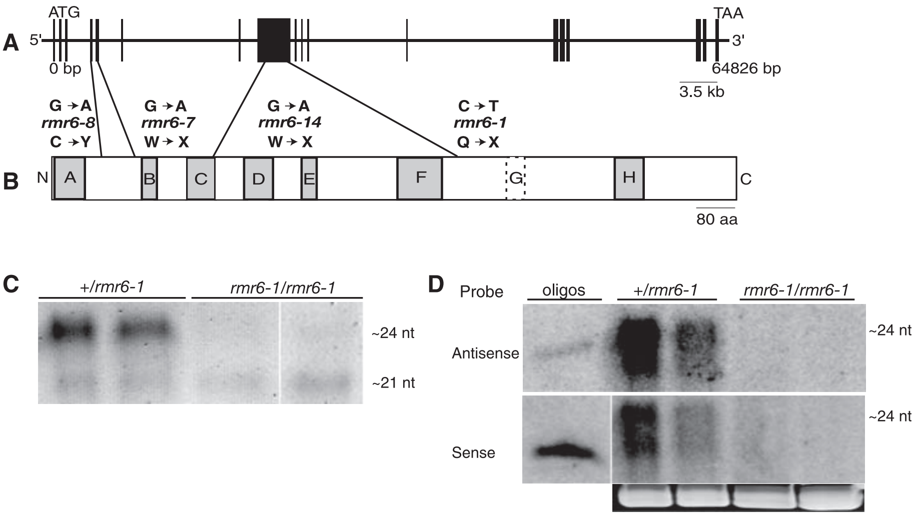


**Fig. 1.** (**A**) Diagram of the *NRPD1a*-like candidate gene model with lines connecting locations of transition-type lesions to resulting codon changes in the translated polypeptide represented in (B). (**B**) Peptide regions with >20% sequence similarity to conserved domains (A to H) of Pol II largest subunits (*22*) are shown in gray boxes, and the annotated domain G region (<20% sequence similarity) is identified with a dashed box. (**C**) Enriched sRNAs isolated from immature cobs of *rmr6-1* homozygotes and heterozygous siblings. (**D**) Northern blot hybridization of enriched sRNAs from immature cobs with radiolabled *doppia* sequences. Stained transfer RNA bands (tRNA) below the blots are loading controls.

controls sex-determination fates by delimiting expression patterns of *silkless1* (*16*), a gene responsible for protecting nascent ovules from a programmed-cell-death pathway (*21*). However, a reverse transcriptase polymerase chain reaction profile of *rmr6* (fig. S6) showed that it is expressed in all rapidly dividing tissues throughout plant development, including the developing ear in which *silkless1* action is required for ovule development (*21*). Lateral meristems are also often derepressed at nodal regions in *rmr6* mutants, giving rise to ectopic organs (fig. S1) (*16*). Thus, Pol IV represents a molecular link between the regulation of developmentally important genes in maize and paramutation (*6*, *16*).

Rice has undergone a duplication of the genes encoding RPD1 and RPE1 (*14*). However, we found only a single intact copy of an RPD1-encoding gene in all other available plant genomes, including *Sorghum bicolor*, a close relative of maize, the basal grass *Brachypodium distachyon*, and in the draft genome of maize, strongly indicating that the duplication of the RPD1-encoding gene is rice-specific. Furthermore, Southern blot analysis confirmed that *rmr6* is singularly distinct in the maize genome (fig. S7), and in silico searches for *rmr6*-related DNA sequences in the regions syntenic with the rice RPD1-encoding duplicates were negative. The position of *rmr6* within the maize B73 genome is also specific in relation to the *S. bicolor* genome.

A MAFFT (http://align.bmr.kyushu-u.ac.jp/mafft/online/server) based protein alignment was used to construct a parsimony tree (Fig. 2A) that agrees with the proposed evolutionary relationship of RPD1 to RPE1 and RPB1 (*14*) and confirms the conservation of catalytic center residues within RPD1 sequences (Fig. 2B) (*13*, *18*). However, our alignment did not agree with previously published comparisons (*18*) over a ~250–amino acid region of low sequence identity near the C terminus of the proteins, which includes domain G, one of the eight normally highly conserved RNA polymerase domains (*22*). We tested different alignment methods [MAFFT, T-COFFEE (www.ebi.ac.uk/Tools/t-coffee/index.html), and MUSCLE

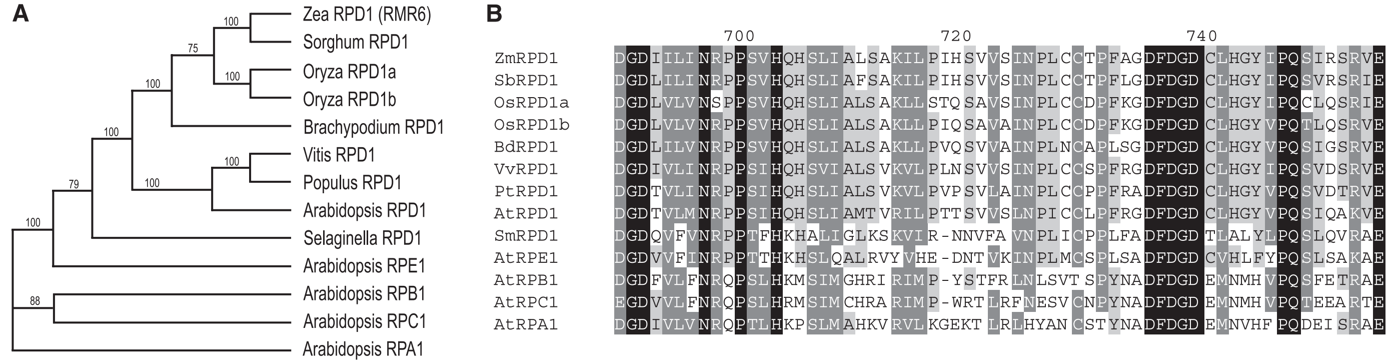


**Fig. 2.** Phylogenomic analysis of RPD1. (**A**) Parsimony tree with bootstrap values generated from MAFFT alignment of full-length polymerase protein sequences from diverse plant species. (**B**) Excerpt from alignment used to generate the phylogenetic tree showing conserved DNA-dependent RNA polymerase domain D (*33*).

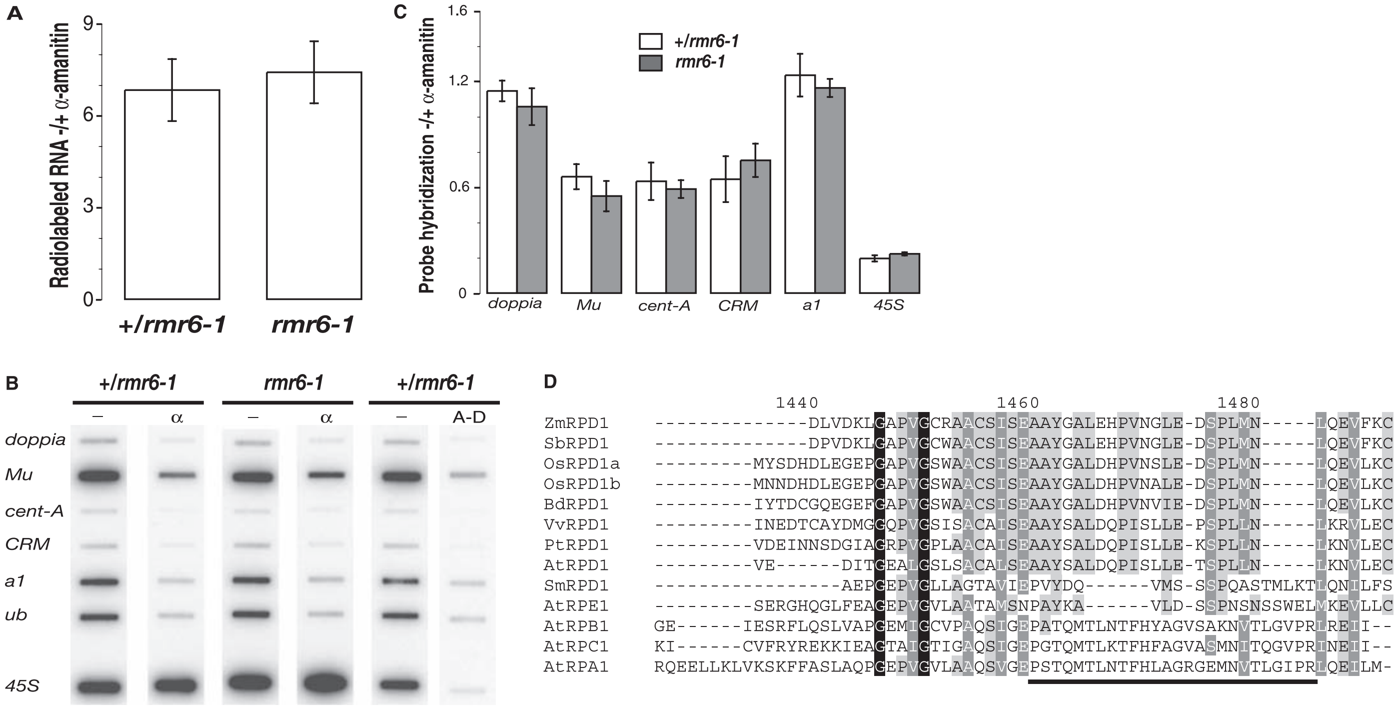


**Fig. 3.** (**A**) Average ratios (± SEM) of bulk $^{32}$P-UTP incorporation (–/+ α-amanitin) into RNA with isolated nuclei from husks of *rmr6-1* homozygotes and *+/rmr6-1* heterozygous siblings ($n$ = 3 biological replicates). (**B**) Representative exposure of hybridized labeled run-on RNAs isolated from seedling nuclei of indicated genotype that were untreated (–) or treated with α-amanitin (α) or actinomycin-D (A-D). (**C**) Average ratios (± SEM) of quantified hybridization signal for run-on RNAs (–/+ α-amanitin) by densitometry for *rmr6-1* homozygotes (gray bars) and heterozygous siblings (open bars) ($n$ = 4 biological replicates for all probes except *doppia* and *a1*, for which $n$ = 3 biological replicates). Hybridization signals were standardized to ubiquitin before calculation of –/+ α-amanitin ratio. (**D**) MAFFT alignment of ZmRPD1 domain G with region of distinct RPD1 sequence underlined.

(www.ebi.ac.uk/Tools/muscle/index.html)] which identified an RPD1 region that contains a deeply conserved backbone relative to all other polymerase largest subunits but which also contained different RPD1 sequences. We annotated this region as RPD1 domain G (Fig. 1B). In most DNA-dependent RNA polymerases, domain G contains a highly conserved motif known as the trigger loop (TL) that is both an interaction site of the potent inhibitor α-amanitin and a mediator of nucleotide selectivity during rapid transcript synthesis in yeast (*23*).

*Arabidopsis* Pol IV does not appear to have any in vitro activity on isolated DNA substrates, and yet it is required for the in vivo accumulation of specific siRNA species (*13*). Cytological studies indicate that Pol IV acts upstream in the RdDM pathway in *Arabidopsis* (*24*), where it is hypothesized to generate single-stranded RNA transcripts of an aberrant nature from its target loci (*13*). If the maize Pol IV is a functional polymerase, Pol IV–dependent primary transcription should be lost in *rmr6* mutants.

We performed nuclear run-on assays with nuclei isolated from husks of homozygous *rmr6-1* mutants and heterozygous siblings in the presence or absence of α-amanitin. We detected no significant Pol IV–derived α-amanitin–insensitive transcription, as measured by comparing the ratio of $\alpha^{32}$P–uridine 5′-triphosphate (UTP) incorporated into run-on transcripts (–/+ α-amanitin) isolated from nuclei of nonmutant and homozygous *rmr6-1* plants (Fig. 3A). Additionally, hybridization of labeled run-on RNAs from both α-amanitin–treated and –untreated nuclei from seedlings to slot blots with strand-specific RNA probes for a suite of maize sequence did not detect ZmRPD1-dependent transcripts (Fig. 3, B and C). The RNA probes used were homologous to potential ZmRPD1 templates represented by both the *doppia* and *Mutator* DNA transposons and the *CRM* and *Cent-A* retroelements (*25*). Although these repetitive sequences that are presumed targets of RdDM were proportionately less affected by α-amanitin treatment than the presumed Pol II–dependent transcript of the pigment biosynthesis factor A1 (Fig. 3C), they did not show the α-amanitin insensitivity of the Pol I–derived 45*S* precursor transcript (Fig. 3C) (*26*), nor did they show differences in abundance between *rmr6* mutants and heterozygous wild-type siblings. Additionally, actinomycin-D treatment (Fig. 3B), which should abolish any DNA-dependent RNA polymerase activity (*27*), confirmed that all transcripts assayed were products of RNA synthesized from DNA templates.

Any Pol IV synthetic action should be α-amanitin–insensitive because the known α-amanitin interaction sites in RPB1 (bridge-helix and multiple TL residues) (*23*, *28*) are divergent in RPD1 (31% identity/50% similarity across the bridge-helix and 11% identity/41% similarity across the TL) (Fig. 3D and fig. S3). Additionally, relative to comparisons between RPB1 and RPD1, the bridge helix and domain G of yeast Rpa1 and Rpc1 are more highly conserved to that of yeast Rpb1 (*18*), yet the cognate Pol I and Pol III complexes are both α-amanitin–insensitive (*29*). If maize Pol IV was enzymatically competent to synthesize RNA from either a DNA or nascent RNA template (*13*), we expected to detect α-amanitin–insensitive transcription, especially in light of the high-repeat content (~70%) of the maize genome (*30*). Our results indicate that either Pol IV does not act as a RNA polymerase or that its relative amount of RNA synthesis is too low to detect by our comparisons. Extrapolations (*17*) of the data indicate that any potential ZmRPD1-dependent RNAs represent no more than ~5% of the total labeled RNA in wild-type nuclei.

The reason for inefficient ZmRPD1 RNA polymerase activity may be tied to its altered domain G (Fig. 3D), which contains the TL in Pol II. Both α-amanitin treatment and a specific TL mutation decrease the synthetic rate (approximately 35-fold and 10-fold, respectively) of *Saccharomyces cereviseae* Pol II and increase the misincorporation incidence of both ribonucleotide and 2′-deoxyribonucleotides in in vitro assays (*23*). The altered domain G within the plant-specific RPD1 and RPE1 proteins may therefore be of considerable importance regarding any Pol IV enzymology.

We have shown that run-on transcription assays can be used as a powerful tool to study alternative polymerase function in plants. Although the biochemical function of RPD1 remains unclear, our results indicate that it acts as a component of a largely dysfunctional polymerase. Paradoxically, Pol IV is required for accumulation of the majority of ~24-nt sRNAs, yet it provides no detectable RNA synthesis for genomic regions represented by those sRNAs in either maize (Fig. 3, B and C) or *Arabidopsis* (*15*). Pol II appears to be the primary, if not exclusive, polymerase for the repetitive features surveyed by our run-on transcription assays, including hypermethylated and repressed transposons such as *doppia* (*7*). These results, together with our previous analyses of ZmRPD1 function in facilitating paramutation (*6*), predict models of Pol IV action that can now be tested by combining biochemical and genetic approaches. Maize plants, which provide large amounts of specific tissues, represent an ideal system for further analysis of alternative polymerase action.

Although it remains unknown how ZmRPD1 is recruited to specific genomic features, its requirement for the accumulation of most 24-nt sRNAs indicates that it operates at repetitive features. Recruitment to transposon sequences proximal to Pol II templates could interfere with Pol II–dependent RNA synthesis, resulting in the production of abnormal Pol II transcripts. Local competition by RPD1 for Pol II holoenzyme subunits may interfere with Pol II function or cotranscriptional RNA processing; this idea is supported by the conservation of domain-specific RPD1 peptide sequences that are necessary for Rpb1 interaction with other Pol II subunits between yeast and maize (*18*). Alternatively, Pol IV may directly compete for Pol II templates and produce small amounts of RNA targeted for degradation pathways and sRNA amplification. Our analyses, combined with the recent characterization of TL function in yeast Pol II (*23*), indicate that such Pol IV RNAs might contain improper nucleotides that serve as a molecular tag for particular RNA degradation pathways. Such models involving abnormal RNA polymerase activities could apply to examples of paramutation outside the plant kingdom.

Our results point to a largely unknown facet of RNA metabolism by which homologous chromosomes can interact to affect gene regulation and heritable epigenetic change. Given that *Rmr6* is required for both paramutation (*6*) and normal maize development (*16*), its identification as ZmRPD1 illustrates that Pol IV plays a broader role in the biology of domesticated maize than in the eudicot *Arabidopsis*. This finding establishes a long-anticipated link between paramutation, developmental gene control, and heterochromatin function (*31*). The highly repetitive and heterochromatin-rich maize genome may influence the chromatin environment (and thus transcription) of a higher percentage of genes than in *Arabidopsis*, which has a smaller and less-repetitive genome. Identification of other genes influenced by *rmr6* mutations thus promises new avenues for understanding these fundamental differences in plant development strategies and the role of paramutation in normal genome function.

*Note added in proof:* Mass spectrometry indicates that 7 of 12 Pol II subunits copurify with *Arabidopsis* RPD1 (*32*).

### References and Notes

1. V. L. Chandler, *Cell* **128**, 641 (2007).
2. J. L. Kermicle, M. Alleman, *Development* **108** (suppl.), 9 (1990).
3. R. A. Brink, *Am. Nat.* **98**, 193 (1964).
4. J. E. Dorweiler *et al*., *Plant Cell* **12**, 2101 (2000).
5. J. B. Hollick, V. L. Chandler, *Genetics* **157**, 369 (2001).
6. J. B. Hollick, J. L. Kermicle, S. E. Parkinson, *Genetics* **171**, 725 (2005).
7. C. J. Hale, J. L. Stonaker, S. M. Gross, J. B. Hollick, *PLoS Biol.* **5**, 2156 (2007).
8. M. R. Woodhouse, M. Freeling, D. Lisch, *PLoS Biol.* **4**, e339 (2006).
9. M. Alleman *et al*., *Nature* **442**, 295 (2006).
10. B. Huettel *et al*., *Biochim. Biophys. Acta* **1769**, 358 (2007).
11. M. Zaratiegui, D. V. Irvine, R. A. Martienssen, *Cell* **128**, 763 (2007).
12. K. C. Cone *et al*., *Plant Cell* **5**, 1807 (1993).
13. C. S. Pikaard, J. R. Haag, T. Ream, A. T. Wierzbicki, *Trends Plant Sci.* **13**, 390 (2008).
14. J. Luo, B. D. Hall, *J. Mol. Evol.* **64**, 101 (2007).
15. A. T. Wierzbicki, J. R. Haag, C. S. Pikaard, *Cell* **135**, 635 (2008).
16. S. E. Parkinson, S. M. Gross, J. B. Hollick, *Dev. Biol.* **308**, 462 (2007).
17. Materials and methods are available as supporting material on *Science* Online.
18. A. J. Herr, M. B. Jensen, T. Dalmay, D. C. Baulcombe, *Science* **308**, 118 (2005).
19. X. Zhang, I. R. Henderson, C. Lu, P. J. Green, S. E. Jacobsen, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 4536 (2007).
20. R. A. Mosher, F. Schwach, D. Studholme, D. C. Baulcombe, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 3145 (2008).
21. A. Calderon-Urrea, S. L. Dellaporta, *Development* **126**, 435 (1999).
22. R. S. Jokerst, J. R. Weeks, W. A. Zehring, A. L. Greenleaf, *Mol. Gen. Genet.* **215**, 266 (1989).

23. C. D. Kaplan, K.-M. Larsson, R. D. Kornberg, *Mol. Cell* **30**, 547 (2008).
24. O. Pontes *et al.*, *Cell* **126**, 79 (2006).
25. R. J. Mroczek, R. K. Dawe, *Genetics* **165**, 809 (2003).
26. M. D. McMullen, B. Hunter, R. L. Phillips, I. Rubenstein, *Nucleic Acids Res.* **14**, 4953 (1986).
27. H. M. Sobell, *Proc. Natl. Acad. Sci. U.S.A.* **82**, 5328 (1985).
28. D. A. Bushnell, P. Cramer, R. D. Kornberg, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 1218 (2002).
29. T. J. Lindell, F. Weinberg, P. W. Morris, R. G. Roeder, W. J. Rutter, *Science* **170**, 447 (1970).
30. R. Liu *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 114 (2007).
31. B. McClintock, *Cold Spring Harbor Symp. Quant. Biol.* **XVI**, 13 (1951).
32. T. S. Ream *et al.*, *Mol. Cell* **33**, 1 (2009).
33. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.
34. We thank X. Parisky for technical assistance; N. Bolduc for providing nucleic acid samples; S. McCormick and D. Zilberman for comments; and the maize, *Sorghum bicolor*, and *Brachypodium* genome-sequencing projects for providing sequence information before publication. *Brachypodium* and *Sorghum bicolor* sequences were produced by the U.S. Department of Energy Joint Genome Institute (www.jgi.doe.gov/). cDNA sequences for *Rmr6-A619*, *rmr6-1*, *rmr6-7*, *rmr6-8*, and *rmr6-14* are archived under GenBank accession numbers FJ426107 to FJ426111, respectively. Germ plasm containing mutations in the *rmr6* gene is covered by U.S. patent 07264970 awarded to the Regents of the University of California. Supported by the National Research Initiative of the U.S. Department of Agriculture Cooperative State Research, Education and Extension Service (grants 99-35301-7753, 2001-35301-10641, and 2006-35304-17399) and NSF (grant MCB-0419909).

**Supporting Online Material**
www.sciencemag.org/cgi/content/full/323/5918/1201/DC1
Materials and Methods
Figs. S1 to S7
Tables S1 to S4
References

11 August 2008; accepted 16 December 2008
10.1126/science.1164508

---

# Mutations in the *FUS/TLS* Gene on Chromosome 16 Cause Familial Amyotrophic Lateral Sclerosis

**T. J. Kwiatkowski Jr.,[1]* D. A. Bosco,[1,2] A. L. LeClerc,[1,2] E. Tamrazian,[1] C. R. Vanderburg,[3] C. Russ,[1,4] A. Davis,[1] J. Gilchrist,[5] E. J. Kasarskis,[6] T. Munsat,[7]† P. Valdmanis,[8] G. A. Rouleau,[8] B. A. Hosler,[1] P. Cortelli,[9] P. J. de Jong,[10] Y. Yoshinaga,[10] J. L. Haines,[11] M. A. Pericak-Vance,[12] J. Yan,[13] N. Ticozzi,[1,2,14] T. Siddique,[13] D. McKenna-Yasek,[1] P. C. Sapp,[1,15] H. R. Horvitz,[15] J. E. Landers,[1,2] R. H. Brown Jr.[1,2]***

Amyotrophic lateral sclerosis (ALS) is a fatal degenerative motor neuron disorder. Ten percent of cases are inherited; most involve unidentified genes. We report here 13 mutations in the *fused in sarcoma/translated in liposarcoma* (*FUS/TLS*) gene on chromosome 16 that were specific for familial ALS. The FUS/TLS protein binds to RNA, functions in diverse processes, and is normally located predominantly in the nucleus. In contrast, the mutant forms of *FUS/TLS* accumulated in the cytoplasm of neurons, a pathology that is similar to that of the gene *TAR DNA-binding protein 43* (*TDP43*), whose mutations also cause ALS. Neuronal cytoplasmic protein aggregation and defective RNA metabolism thus appear to be common pathogenic mechanisms involved in ALS and possibly in other neurodegenerative disorders.

Amyotrophic lateral sclerosis (ALS) is a progressive, uniformly fatal, age-dependent degenerative disorder of motor neurons. Its incidence (0.6 to 2.6 per 100,000 humans) peaks in the sixth decade of life (*1*). Death from ALS is typically 2 to 5 years after onset and is usually a consequence of respiratory paralysis. Familial cases account for about 10% of ALS. Of these, about 20% are caused by mutations in the *superoxide dismutase 1* (*SOD1*) gene (*2*). A small number of both familial and apparently sporadic cases are caused by mutations in various other genes, including *TAR DNA-binding protein 43* (*TDP43*) (*3*–*5*), although in the majority of familial ALS (FALS) cases the causative gene is unknown.

In a family of Cape Verdean origin (F577) (fig. S1A), four members developed distinctive ALS with onset in the proximal upper extremities and spreading to the lower extremities, but not the bulbar region. The maternal grandparents of the proband were first cousins; the family originates from a small island of roughly 6000 inhabitants, raising the possibility that the inheritance pattern is recessive. To pursue this, we conducted loss-of-heterozygosity (LOH) mapping and identified a major LOH cluster within a previously reported chromosome 16 ALS locus. Five smaller regions of homozygosity were observed in other chromosomal regions (table S1).

The major LOH cluster spanned approximately 4 Mb and contained 56 candidate ALS genes. Genomic sequencing of these genes revealed a sequence variant in the index ALS cases in exon 15 of the *fused in sarcoma/translated in liposarcoma* (*FUS/TLS*) gene (*6*). This variant, a base pair C1551G missense mutation (*7*), substituting glutamine for histidine at amino acid residue 517 (Table 1), was heterozygously present in asymptomatic individuals and homozygously present in four individuals with FALS and three individuals who were asymptomatic; one was just entering the age of risk for ALS, whereas the other two were below that age. The variant was not detected in 1446 control DNA samples from North America; a single heterozygote was observed in 66 DNA samples (132 chromosomes) from Cape Verde.

We next fully sequenced all of the 15 *FUS/TLS* exons in two families genetically linked to chromosome 16 (F55, fig. S1B, and NUFMS9900, fig. S1C), (*8*). In family 55 (F55), we detected a missense mutation (C1561G) that substituted glycine for arginine at residue 521; this was coinherited with ALS (in five of five affected members for whom DNA was available but in no unaffected individuals). Incomplete penetrance is evident in F55: two mutation carriers lived past the average age of onset without developing ALS. In family NUFMS9900, a G1553A mutation substituting lysine for arginine at residue 518 was present in 10 of 10 available affected members. These variants were not detected in 1446 control subjects. Haplotype analysis of six single-nucleotide polymorphisms (SNPs) across the *FUS/TLS* locus does not exclude a common founder among apparently unrelated R521G pedigrees (*9*).

We sequenced all 15 exons in 81 other unrelated FALS cases and 293 sporadic ALS (SALS) DNA samples, and an additional 209 ALS families were screened for mutations in exon 15. Overall, in 17 different FALS families, 13 different mutations were detected, including 10 in exon

---

[1] Department of Neurology, Massachusetts General Hospital, 114 16th Street, Charlestown, MA 02129, USA. [2] Department of Neurology, University of Massachusetts School of Medicine, Worcester, MA 01655, USA. [3] Massachusetts General Institute for Neurodegeneration, Massachusetts General Hospital, 114 16th Street, Charlestown, MA 02129, USA. [4] Broad Institute, Massachusetts Institute of Technology (MIT), Cambridge, MA 02141, USA. [5] Department of Neurology, Rhode Island Hospital, Providence, RI 02192, USA. [6] Department of Neurology, Veterans Affairs Medical Center, University of Kentucky, Lexington, KY 40511, USA. [7] Department of Neurology, Tufts New England Medical Center, Tufts University, Boston, MA 02111, USA. [8] Department of Neurology, Centre de Recherche du Centre Hospitalier de l'Université de Montreal, University of Montreal, Quebec H2L 4M1, Canada. [9] Clinica Neurologica, Departimento di Scienze Neurologiche, University de Bologna, Via Ugo Foscola, 7,40123 Bologna, Italy. [10] Childrens' Hospital Oakland Research Institute, Oakland, CA 94609, USA. [11] Center for Human Genetics Research, Molecular Physiology and Biophyscis, Vanderbilt University, Nashville, TN 37240, USA. [12] Miami Institute of Human Genetics, Miami, FL 33136, USA [13] Departments of Neurology and Molecular and Cellular Biology, Northwestern Feinberg School of Medicine, Chicago, IL 60611, USA. [14] Department of Neurology, University of Milan Medical School, Dino Ferrari Center, Istituto Auxologico Italiano, Milan 20149, Italy. [15] Department of Biology, Howard Hughes Medical Institute, MIT, Cambridge, MA 02139, USA.

*To whom correspondence should be addressed. E-mail: tkwiatkowski@partners.org (T.J.K.); robert.brown@umassmed.edu (R.H.B.)

†Retired.

15, two in exon 5, and one in exon 6 (Table 1 and Fig. 1). The most common mutations were R521C and R521G, each present in 3 of the 17 families. No mutations were found in the SALS DNA set. None of the exon 15 variants was observed in 1446 control individuals sequenced. All of the mutated residues in exon 15 of *FUS/TLS* were highly conserved (fig. S1D); each of the 5 arginines in exon 15 is mutated (Fig. 1).

Autopsy tissue from a single patient from F55 (R521G mutation) showed loss of motor neurons in the anterior horn of the spinal cord and the hypoglossal nucleus. Myelin pallor in the anterior corticospinal tracts, macrophages surrounding shrunken Betz cells in the motor cortex (neuronophagia), and increased lipofuscin staining in neurons were also noted. The distribution of FUS/TLS was assessed using immunostaining of frozen brain and spinal cord sections from this affected individual. Both control and patient sections revealed immunostaining of FUS/TLS within the nuclei (of neurons and of nonneuronal cells). However, in F55 patient sections there was additional prominent cytoplasmic staining (Fig. 2). Further immunostaining revealed diffuse ubiquitin positivity in nuclei in the patient's tissue but not control tissue (fig. S2), suggesting that one or more nuclear proteins are misfolded.

Thus, the R521G missense mutation in FUS/TLS led to aberrant trafficking with subsequent cytoplasmic retention of the mutant protein. To evaluate this, we analyzed the subcellular distribution of wild type (WT), R521G (F55), or H517Q (F577) FUS/TLS fused to green fluorescent protein (GFP) 24 hours after transient transfection in N2A and SKNAS cells. With immunofluorescence, the R521G (F55) mutant *FUS/TLS*–GFP transfections showed dense cytoplasmic staining (Fig. 3A). The ratio of the fraction of cells with exclusively nuclear staining versus cells with combined cytoplasmic and nuclear staining was approximately 0.80/0.20 for the WT protein. In contrast, the ratio for R521G was roughly 0.50/0.50. These ratios were similar in the N2A and SKNAS cells. Immunofluorescence did not detect cytoplasmic retention in cells transfected with the H517Q-GFP (F577)

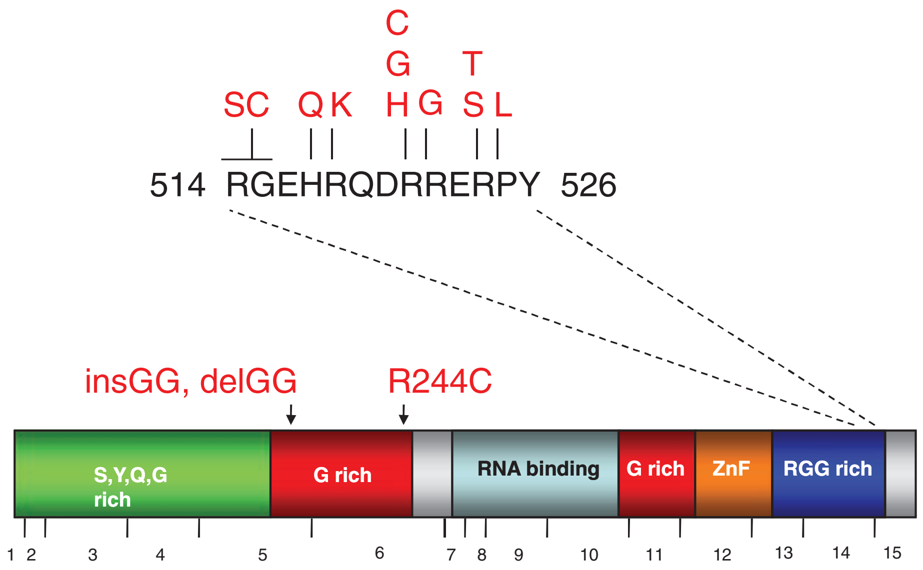


**Fig. 1.** Positions of the FUS/TLS mutations superimposed on the exon and domain organization of the gene. FUS/TLS is encoded by 15 exons that span a genomic region of 11.6 kb. S,Y,Q,G–rich denotes a region rich in the amino acids serine, tyrosine, glutamine, and glycine; G-rich and RGG-rich regions are enriched in either glycine or the arginine-glycine-glycine motif, respectively (*11*). Further details for each mutation are in Table 1.

**Table 1.** *FUS/TLS* mutations in ALS cases, with onset and disease duration data. Base numbering begins with the start codon; amino acid numbering begins with the methionine start codon.

| ID | Mutation | | Exon | Controls | Pattern | Age onset (years) | | Duration (months) | |
|---|---|---|---|---|---|---|---|---|---|
| | Amino acid | Base pair | | | | Mean ± SD | Patients | Mean ± SD | Patients |
| Index cases | | | | | | | | | |
| F577 | H517Q | C1551G* | 15 | 1446 | AR | 45.0 ± 3.6 | 4 | 168.0 | 1 |
| F55 | R521G | C1561G | 15 | 1446 | AD | 39.6 ± 13.3 | 13 | 26.6 ± 16.5 | 13 |
| Other cases | | | | | | | | | |
| F213 | insGG | insGAGGTG523 | 5 | 176 | AD | 65 | 1 | 31 | 1 |
| MTL 10 | delGG | delGAGGTG523 | 5 | 176 | AD | 43.3 ± 6.4 | 3 | 147.0 ± 55.2 | 2 |
| MTL 7 | R244C | C730T | 6 | 231 | AD | 54.0 ± 8.5 | 2 | 28.5 ± 10.6 | 2 |
| F360 | R514S,<br>G515C | G1542T,<br>G1543T† | 15 | 1446 | AD | 32.5 ± 3.5 | 2 | 36.0 | 1 |
| NUFMS9900 | R518K | G1553A | 15 | 1446 | AD | 40.3 ± 6.6 | 12 | 26.7 ± 14.4 | 9 |
| F072 | R521C | C1561T | 15 | 1446 | AD | 35.0 ± 14.8 | 3 | 26.0 ± 8.2 | 3 |
| F080 | R521C | C1561T | 15 | 1446 | AD | 54.0 ± 2.8 | 2 | 14.0 ± 2.8 | 2 |
| F085 | R521C | C1561T | 15 | 1446 | AD | 39.7 ± 10.8 | 6 | 31.2 ± 17.7 | 5 |
| F002 | R521G | C1561G | 15 | 1446 | AD | 60.7 ± 10.7 | 3 | 28.3 ± 17.9 | 3 |
| F136 | R521G | C1561G | 15 | 1446 | AD | 37.5 ± 5.0 | 2 | 24.0 | 1 |
| F067 | R521H | G1562A | 15 | 1446 | AD | 57.7 ± 9.0 | 3 | 54.1 ± 26.2 | 3 |
| F287 | R522G | A1564G | 15 | 1446 | AD | 28.5 ± 14.8 | 2 | 25.1 ± 15.6 | 2 |
| F597 | R524T | G1571C | 15 | 1446 | AD | 61.6 ± 4.8 | 2 | 32.9 ± 18.2 | 2 |
| F346 | R524S | G1572C | 15 | 1446 | AD | 34.0 | 1 | 39.0 | 1 |
| F568 | P525L | C1574T | 15 | 1446 | AD | 22.0 | 1 | 6.0 | 1 |
| Average | | | | | All | 46.3 ± 14.7 | 62 | 33.1 ± 25.3 | 52 |
| | | | | | AD | 46.4 ± 15.1 | 58 | 30.6 ± 17.3 | 51 |

*1 heterozygote C1551G (and no homozygotes) was detected in 66 Cape Verdean controls (132 chromosomes). †These mutations are in cis.

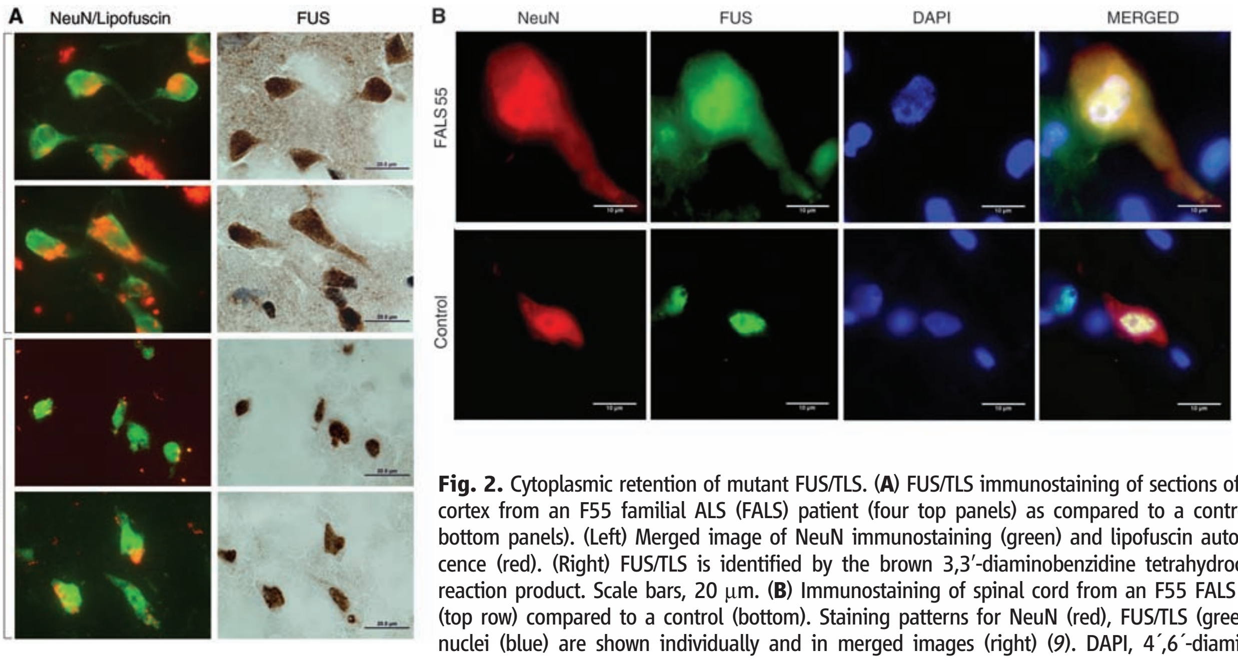


**Fig. 2.** Cytoplasmic retention of mutant FUS/TLS. (**A**) FUS/TLS immunostaining of sections of frontal cortex from an F55 familial ALS (FALS) patient (four top panels) as compared to a control (four bottom panels). (Left) Merged image of NeuN immunostaining (green) and lipofuscin autofluorescence (red). (Right) FUS/TLS is identified by the brown 3,3′-diaminobenzidine tetrahydrochloride reaction product. Scale bars, 20 μm. (**B**) Immunostaining of spinal cord from an F55 FALS patient (top row) compared to a control (bottom). Staining patterns for NeuN (red), FUS/TLS (green), and nuclei (blue) are shown individually and in merged images (right) (*9*). DAPI, 4´,6´-diamidino-2-phenylindole. Scale bars, 10 μm.

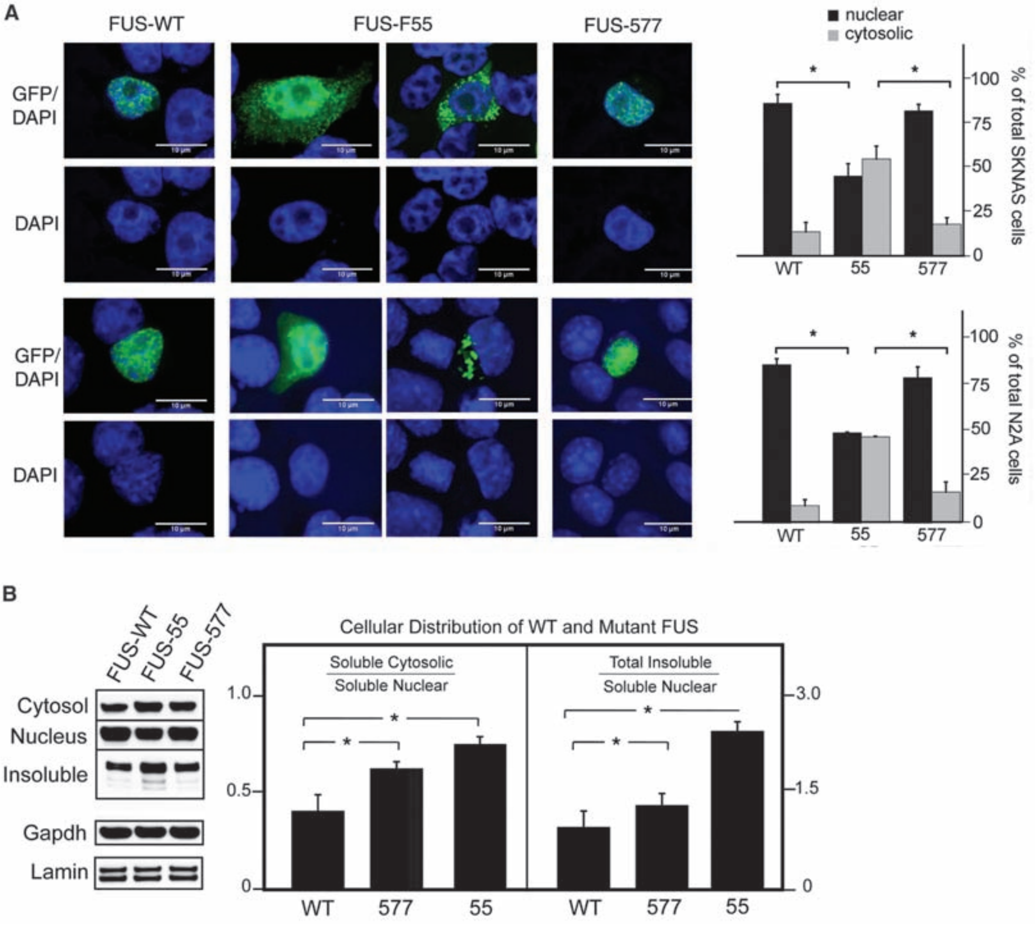


**Fig. 3.** Mislocalization of mutant FUS/TLS. (**A**) SKNAS (top) or N2A (bottom) cells transfected with WT or mutant (F55 = R521G or F577 = H517Q) recombinant FUS/TLS–GFP fusion protein, counterstained with DAPI. Top row, merged images; lower row, DAPI only. The percentage of cells observed with nuclear-only (dark bars) and any cytosolic (light bars) FUS/TLS staining is indicated in the bar graphs at right. The error bars indicate SD (from total number of cells counted on three coverslips). Asterisks indicate statistically significant differences between FUS constructs [$P < 0.0001$; analysis of variance (ANOVA) followed by Holms test for multiple comparisons]. Scale bars, 10 μm. (**B**) Cell fractionation studies. SKNAS cells transfected with WT or mutant (F55 = R521G or F577 = H517Q) recombinant FUS/TLS–GFP fusion protein were harvested and fractionated at 24 hours for analysis by immunoblotting using an antibody to GFP, with binding quantified by chemiluminescence. Lamin and glyceraldehyde-3-phosphate dehydrogenase (GAPDH) loading controls are shown below. Densitometric ratios are at right. The error bars indicate SD from three triplicate immunoblot measurements. Asterisks indicate statistically significant differences between WT and mutant FUS constructs ($P < 0.0001$, ANOVA followed by Holms test for multiple comparisons).

mutant. Subcellular localization of FUS/TLS was additionally studied by compartmental fractionation of SKNAS cells transfected with WT, R521G, or H517Q FUS/TLS–GFP fusion proteins. Immunoblotting of fractions followed by immunostaining with an antibody to GFP demonstrated a substantially higher ratio of soluble cytosolic to soluble nuclear FUS/TLS for both mutants (Fig. 3B). Additionally, a higher ratio of total insoluble to soluble nuclear FUS/TLS protein was also seen for both mutants, although it is more pronounced for the R521G mutant; this reflected both an increase in total insoluble FUS and a decrease in soluble nuclear FUS (fig S3).

The major defined RNA-interacting domains of FUS/TLS are located in the mid-region of the protein, from amino acids 280 to 370, encoded by exons 9 to 11 (*10*, *11*); sequences of target RNA domains recognized by FUS/TLS have been reported. To show that the FALS-associated FUS/TLS mutations detected here do not alter the RNA-binding domain of FUS, we performed in vitro RNA-binding experiments with recombinant histidine (His)–tagged mutant and WT FUS/TLS proteins and RNA 24-nucleotide oligomer containing GGUG motifs and known to bind FUS/TLS (*10*). Binding of the RNA oligomers was similar for mutant and WT FUS/TLS protein (fig. S4).

FUS/TLS is a nucleoprotein that functions in DNA and RNA metabolism (*12*–*15*). It has also been implicated in tumorigenesis (*6*, *16*, *17*) and RNA metabolism. *FUS/TLS* knockout mice show perinatal mortality (*18*) or male sterility and radiation sensitivity (*19*). FUS/TLS–deficient neurons show decreased spine arborization with abnormal morphology. In hippocampal neuronal slice cultures, the protein is found in RNA granules that are transported to dendritic spines for local RNA translation in response to metabotropic glutamate receptor (mGluR5) stimulation (*20*).

We detected 13 FUS/TLS mutations in patients with FALS but none in patients with SALS. We estimate that FUS/TLS mutations are detected in about 5% of FALS; this is comparable to the frequency of *TDP43* gene mutations in ALS but less than that for *SOD1* (mutated in ~20% of FALS cases). The FUS/TLS mutations described here led to cytoplasmic retention and apparent aggregation of FUS/TLS. This is reminiscent of several models of the pathogenesis of FALS that are mediated by the aggregation of mutant superoxide dismutase (*21*) and the mislocalization in ALS of both mutant and WT TDP43 (*4*, *22*). FUS/TLS has also been reported to be a major nuclear aggregate–interacting protein in a model of Huntington's disease (*23*). Genes implicated in other motor neuron diseases also involve aspects of DNA and RNA metabolism [table S5 in (*24*)]; understanding the convergent pathophysiologies of these genetic variants will provide insights into new targets for therapies for the motor neuron diseases.

### References and Notes

1. L. M. Nelson, *Clin. Neurosci.* **3**, 327 (1995).
2. D. R. Rosen, *Nature* **364**, 362 (1993).
3. E. Kabashi *et al.*, *Nat. Genet.* **40**, 572 (2008).
4. J. Sreedharan *et al.*, *Science* **319**, 1668 (2008).
5. V. M. Van Deerlin *et al.*, *Lancet Neurol.* **7**, 409 (2008).
6. A. Crozat, P. Aman, N. Mandahl, D. Ron, *Nature* **363**, 640 (1993).
7. In the mutants, other amino acids were substituted at certain locations; for example, R182Q indicates that arginine at position 182 was replaced by glutamine. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.
8. P. Sapp *et al.*, *Am. J. Hum. Genet.* **73**, 397 (2003).
9. Materials and methods are available as supporting material on *Science* Online.
10. A. Lerga *et al.*, *J. Biol. Chem.* **276**, 6807 (2001).
11. F. Morohoshi *et al.*, *Gene* **221**, 191 (1998).
12. H. Baechtold *et al.*, *J. Biol. Chem.* **274**, 34337 (1999).
13. P. Bertrand, A. T. Akhmedov, F. Delacote, A. Durrbach, B. S. Lopez, *Oncogene* **18**, 4515 (1999).
14. L. Yang, L. J. Embree, S. Tsai, D. D. Hickstein, *J. Biol. Chem.* **273**, 27761 (1998).
15. X. Wang *et al.*, *Nature* **454**, 126 (2008).
16. T. H. Rabbitts, A. Forster, R. Larson, P. Nathan, *Nat. Genet.* **4**, 175 (1993).
17. W. J. Law, K. L. Cann, G. G. Hicks, *Brief. Funct. Genomics Proteomics* **5**, 8 (2006).
18. G. G. Hicks *et al.*, *Nat. Genet.* **24**, 175 (2000).
19. M. Kuroda *et al.*, *EMBO J.* **19**, 453 (2000).
20. R. Fujii *et al.*, *Curr. Biol.* **15**, 587 (2005).
21. S. Boillee, C. Vande Velde, D. W. Cleveland, *Neuron* **52**, 39 (2006).
22. M. J. Winton *et al.*, *J. Biol. Chem.* **283**, 13302 (2008).
23. H. Doi *et al.*, *J. Biol. Chem.* **283**, 6489 (2008).
24. C. Simpson *et al.*, *Hum. Mol. Genet.* **18** (no. 3), 472 (2008).

25. We gratefully acknowledge I. Carr (University of Leeds, UK) for support with AutoSNPa and IBDfinder software, A. Storey for assistance with sequencing, C. LeClerc for genealogical investigations, and D. Crowe for administrative assistance. This work was supported by NIH grants NS050557 (R.H.B. and T.K.) and NS050641 (T.S., R.H.B., J.L. H., and M.P.-V.). R.H.B. also receives support from the Angel Fund, the ALS Therapy Alliance, the ALS Association Project ALS, the Al-Athel ALS Research Foundation, and the Pierre L. de Bourgknecht ALS Research Foundation. T.S. also receives support from The Les Turner ALS Foundation, Vena E. Schaff ALS Research Fund, Harold Post Research Professorship, Herbert and Florence C. Wenske Foundation, Ralph and Marian Falk Medical Research Trust, The David C. Asselin M.D. Memorial Fund, Les Turner ALS Foundation/Herbert C. Wenske Foundation Professorship, Help America Foundation and the ALS Therapy Alliance, Inc. H.R.H. is an investigator of and was supported by the Howard Hughes Medical Institute. R.H.B. is a cofounder of AviTx Inc., which targets development of ALS therapies. R.H.B. and T.J.K. have applied for a patent covering FUS mutations in ALS. We dedicate this report to the memories of Jimmy and Christopher Kennedy, Sharon Timlin, and Ginny Delvecchio.

**Supporting Online Material**

www.sciencemag.org/cgi/content/full/323/5918/1205/DC1
Materials and Methods
Figs. S1 to S4
Table S1
References

17 September 2008; accepted 6 January 2009
10.1126/science.1166066

---

# Mutations in FUS, an RNA Processing Protein, Cause Familial Amyotrophic Lateral Sclerosis Type 6

Caroline Vance,[1]* Boris Rogelj,[1]* Tibor Hortobágyi,[1]* Kurt J. De Vos,[2]*
Agnes Lumi Nishimura,[1] Jemeen Sreedharan,[1] Xun Hu,[1] Bradley Smith,[1]
Deborah Ruddy,[1] Paul Wright,[1] Jeban Ganesalingam,[1] Kelly L. Williams,[3] Vineeta Tripathi,[1]
Safa Al-Saraj,[1] Ammar Al-Chalabi,[1] P. Nigel Leigh,[1] Ian P. Blair,[3,5] Garth Nicholson,[3,4,5]
Jackie de Belleroche,[6] Jean-Marc Gallo,[1] Christopher C. Miller,[1,2] Christopher E. Shaw[1]†

Amyotrophic lateral sclerosis (ALS) is a fatal neurodegenerative disease that is familial in 10% of cases. We have identified a missense mutation in the gene encoding fused in sarcoma (FUS) in a British kindred, linked to ALS6. In a survey of 197 familial ALS index cases, we identified two further missense mutations in eight families. Postmortem analysis of three cases with FUS mutations showed FUS-immunoreactive cytoplasmic inclusions and predominantly lower motor neuron degeneration. Cellular expression studies revealed aberrant localization of mutant FUS protein. FUS is involved in the regulation of transcription and RNA splicing and transport, and it has functional homology to another ALS gene, *TARDBP*, which suggests that a common mechanism may underlie motor neuron degeneration.

Amyotrophic lateral sclerosis (ALS) causes progressive muscular weakness due to the degeneration of motor neurons in the brain and spinal cord. The average age at onset is 60 years, and annual incidence is 1 to 2 per 100,000. Death due to respiratory failure occurs on average 3 years after symptom onset (*1*). Autosomal dominant familial ALS (FALS) is clinically and pathologically indistinguishable from sporadic disease (SALS) and accounts for ~10% of cases (*2*). Three genes have been confidently linked to classical FALS: *SOD1*, encoding CuZn superoxide dismutase (SOD1) (*ALS1* OMIM 105400) (*3*); *ANG*, encoding angiogenin (*4*–*6*); and *TARDP*, encoding TAR DNA binding protein TDP-43 (*ALS10* OMIM 612069) (*7*). SOD1 mutations are detected in 20% of FALS and 5% of SALS cases (*3*, *8*). Mice transgenic for mutant human SOD1 develop selective motor neuron degeneration due to a toxic gain of function (*9*) that is not cell autonomous

(*10*) and is associated with multimeric detergent-resistant cytoplasmic mutant SOD1 aggregates (*11*). Ubiquitinated TDP-43 inclusions are detected in the perikaryon in the motor neurons of ~90% of ALS cases and in the cortical neurons of a subset of frontotemporal lobar dementia (FTLD) cases (*12*, *13*). Fourteen missense TDP-43 mutations have now been associated with ALS, and all but one reside in the C-terminal domain (*7*, *14*–*16*).

In addition, individual FALS kindreds have been linked to chromosomes 18q (*ALS3*, OMIM 606640) (*17*) and 20p (ALS7, OMIM 608031) (*18*), whereas multiple kindreds with ALS and FTLD have been linked to chromosomes 9q (FTD/ALS1 OMIM 105550) (*19*) and 9p (ALSFTD2, OMIM 611454) (*20*–*22*). We previously linked one large multigenerational British kindred (F1) to a 42-Mb region on chromosome 16 with a multipoint lod score (logarithm of the odds ratio for linkage) of 3.85 (ALS6 OMIM 608030) (*23*). As two more individuals developed ALS, we repeated the genome-wide scan using the 10K single-nucleotide polymorphism (SNP) arrays. This confirmed linkage to chromosome 16 with a conserved haplotype containing more than 400 genes (fig. S1). Using a candidate gene approach, we sequenced 279 exons from 32 genes and expressed sequence tags (ESTs) and 10 noncoding RNAs (tables S1 and S2). After our detection of mutations in *TARDBP*, we prioritized six genes containing similar domains within the linked region. A single base-pair change was identified in exon 15 of *FUS* (also known as translocation in liposarcoma or *TLS*) in the proband from the F1 kindred (fig. S2A). This change is predicted to result in an arginine to cysteine substitution at position 521 (R521C), which is in a highly conserved region at the very C terminus of the protein (fig. S2, B and C). The 1561 C>T (R521C) mutation was present in the DNA of all six affected members of F1 and cosegregated with the linked haplotype (Fig. 1A).

To determine the frequency of FUS mutations in FALS, we screened 197 index FALS cases in which *SOD1*, *VAPB*, *ANG*, *Dynactin*, *CHMP2B*, and *TARDBP* mutations had been excluded. We identified the same R521C mutation in a further four families, including three affected individuals from one kindred (ALS53) (Fig. 1B) and three more index FALS cases. In addition, we found a different base-pair change that is predicted to result in an arginine to histidine substitution, also at position 521 (1562 G>A, R521H) (fig. S2, A to C). The R521H mutation was identified in two families where it segregates with disease (F1015 and ALS156) (Fig. 1, C and D) and one index case. A third mutation was identified in exon 14 that is predicted to result in an arginine to glycine substitution at position 514 (1540 A>G, R514G) (fig. S2, A to C). This was detected in two affected members of the pedigree F1083 (Fig. 1E). No mutations were detected in exons 1 to 13 in the other FALS cases. Exons 14 and 15 of *FUS* were sequenced in 400 age, sex, and ethnically matched controls, and no mutations were detected.

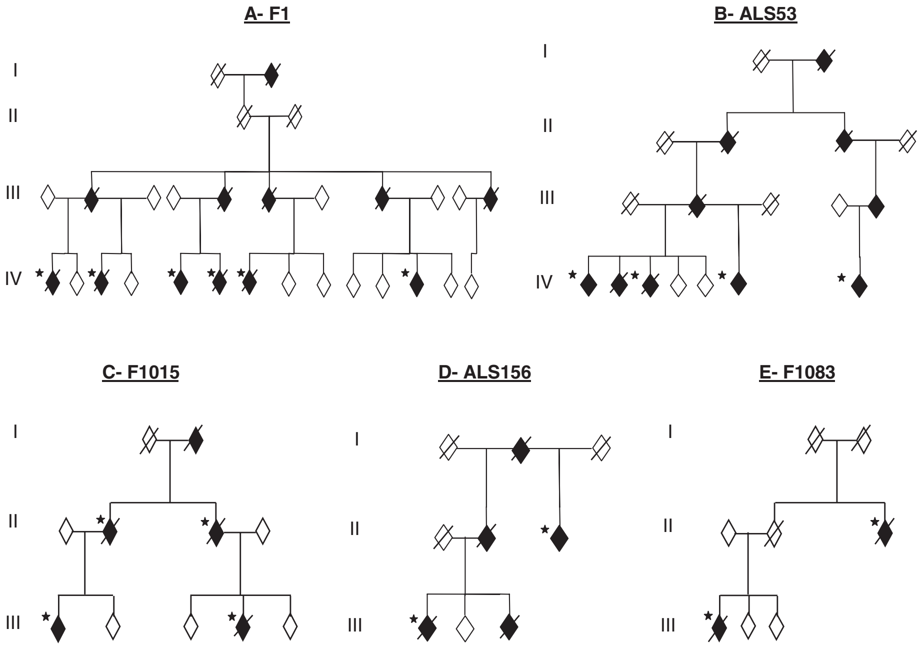


**Fig. 1.** FUS mutations segregate with disease in five dominant ALS kindreds. Affected individuals are indicated by black symbols and unaffected individuals by open symbols. Slashed symbols indicate deceased individuals. Asterisks indicates an affected individual for whom DNA was available. All affected individuals tested carried their families' mutation. Gender, birth order, and the mutation status of unaffected individuals have been omitted for reasons of confidentiality.

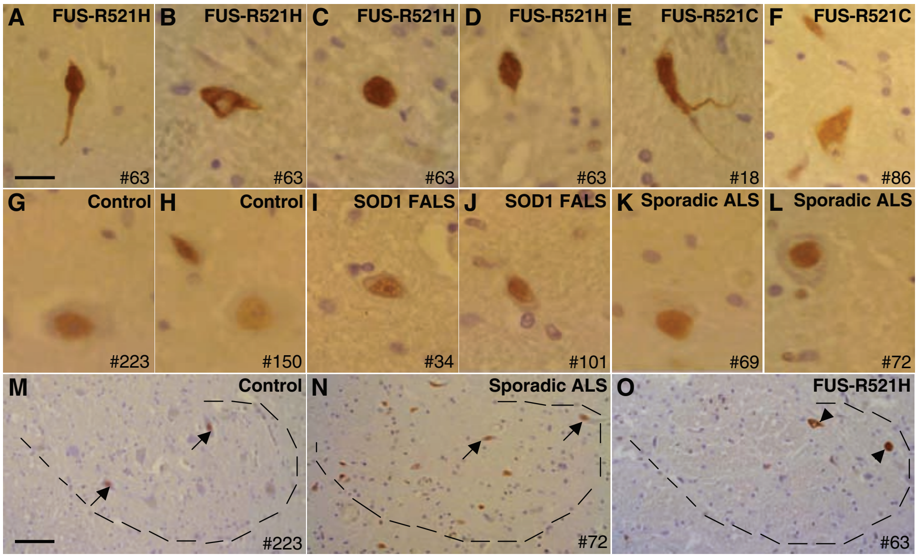


**Fig. 2.** Patients with *FUS* mutations develop cytoplasmic FUS immunoreactive inclusions in lower motor neurons. Antibody to FUS immunolabels inclusions within the anterior horn (dotted line) of the spinal cord in patients with FUS mutations (**A** to **F** and arrowheads in **O**). Staining in controls (**G**, **H**, and **M**), mutant SOD1 FALS (**I** and **J**), and SALS (**K**, **L**, and **N**) demonstrate diffuse nuclear staining (arrows) with variable intensity without inclusions. Scale bars, 12 μm [(A) to (L)], 50 μm [(M) to (O)].

---

[1]Department of Clinical Neuroscience, King's College London, Medical Research Council (MRC) Centre for Neurodegeneration Research, Institute of Psychiatry, London SE5 8AF, UK. [2]Department of Neuroscience, King's College London, MRC Centre for Neurodegeneration Research, Institute of Psychiatry, London SE5 8AF, UK. [3]Northcott Neuroscience Laboratory, Australian and New Zealand Army Corps (ANZAC) Research Institute, Concord, NSW 2139, Australia. [4]Molecular Medicine Laboratory, Concord Hospital, Concord, NSW 2139, Australia. [5]Faculty of Medicine, University of Sydney, Sydney, NSW 2139, Australia. [6]Division of Neurosciences and Mental Health, Faculty of Medicine, Imperial College London, Hammersmith Hospital, London W12 0NN, UK.

*These authors contributed equally to this work.

†To whom correspondence should be addressed. E-mail: chris.shaw@iop.kcl.ac.uk

The frequency of FUS mutations in British and Australian SOD1-negative cases is ~4% (equivalent to 3% of all FALS cases).

Sufficient information to define a clinical phenotype was available in 20 individuals with FUS mutations. There was an even gender distribution (9 male, 11 female), the average age at onset was 44.5 years ($n$ = 20), and average survival was 33 months ($n$ = 18). The site of onset was cervical (10), lumbar (5), and bulbar (3). No affected individual developed cognitive deficits. Three cases with the FUS$_{\text{R521C}}$ and FUS$_{\text{R521H}}$ mutations underwent detailed neuropathological examination. All revealed severe lower motor neuron loss in the spinal cord (Fig. 2O) and to a lesser degree in the brain stem, whereas dorsal horn neurons appeared unaffected. There was evidence of mild myelin loss in the dorsal columns, but only one case had major pallor of the corticospinal tracts. There was mild to moderate upper motor neuron loss in the motor cortex. Ubiquitin and p62-immunopositive skeinlike cytoplasmic inclusions in anterior horn motor neurons of the spinal cord, which are hallmarks of classical ALS, were very rare, and TDP-43–positive inclusions within the cytoplasm and nucleus of neurons and glial cells were absent. Antibody to FUS labeled large globular and elongated cytoplasmic inclusions in spinal cord motor neurons and dystrophic neurites in all three ALS cases with FUS mutations (Fig. 2, A to F, and O). These were absent in normal control individuals (Fig. 2, G, H, and M), SOD1 mutant ALS cases (Fig. 2, I and J), and sporadic ALS cases (Fig. 2, K, L, and N).

To assess the functional importance of these mutations, we transiently expressed N-terminally hemagglutinin (HA)–tagged wild-type FUS (FUS$_{\text{WT}}$) and two mutants, (FUS$_{\text{R521C}}$ and FUS$_{\text{R521H}}$) in CV-1 and N2A cell lines and rat cortical neurons. Staining for endogenous FUS in nontransfected cells indicated that the protein is predominantly localized to the nucleus. Transfected cells show an increase in cytoplasmic localization of the FUS$_{\text{R521C}}$ and FUS$_{\text{R521H}}$ protein when compared to the wild type, both with immunofluorscent labeling and by immunoblot (Fig. 3).

FUS is a ubiquitously expressed, predominantly nuclear, protein that is involved in DNA repair and the regulation of transcription, RNA splicing, and export to the cytoplasm. The N terminus has a transcriptional activation domain which, after chromosomal translocation, produces fusion proteins that can cause Ewing's sarcoma and acute myeloid leukemia (*24*). The C terminus contains RNA recognition motif (RRM), Arg-Gly-Gly (RGG) repeat rich, and zinc finger domains involved in RNA processing. FUS binds to the motor proteins kinesin (KIF5) (*25*) and myosin-Va (*26*) and is involved in mRNA transport; it is notable that defects in axonal transport are a pathological feature of ALS (*27*).

Mutations in the C-terminal domain of TDP-43 have been linked to ALS (*7*, *14*–*16*) and are associated with mislocalization and inclusion formation, providing an almost exact parallel with FUS mutations. In the case of TDP-43 proteinopathies, it is not clear whether disease is due to a toxic gain, or loss, of function. Although toxicity might be a consequence of misfolded proteins overwhelming the protein surveillance machinery, the loss of TDP-43 from the nucleus in neurons with inclusions provides circumstantial evidence that nuclear RNA processing may be impaired. Equally, the presence of cytoplasmic aggregates might lead to the sequestration of proteins and/or RNAs in the cytoplasm, causing selective toxicity to cortical and motor neurons. The absence of TDP-43 inclusions in mutant FUS ALS cases implies that the FUS disease pathway is independent of TDP-43 aggregation. An exploration of how these two functionally related proteins cause neurodegeneration should provide new insights into the mechanisms underlying ALS pathogenesis.

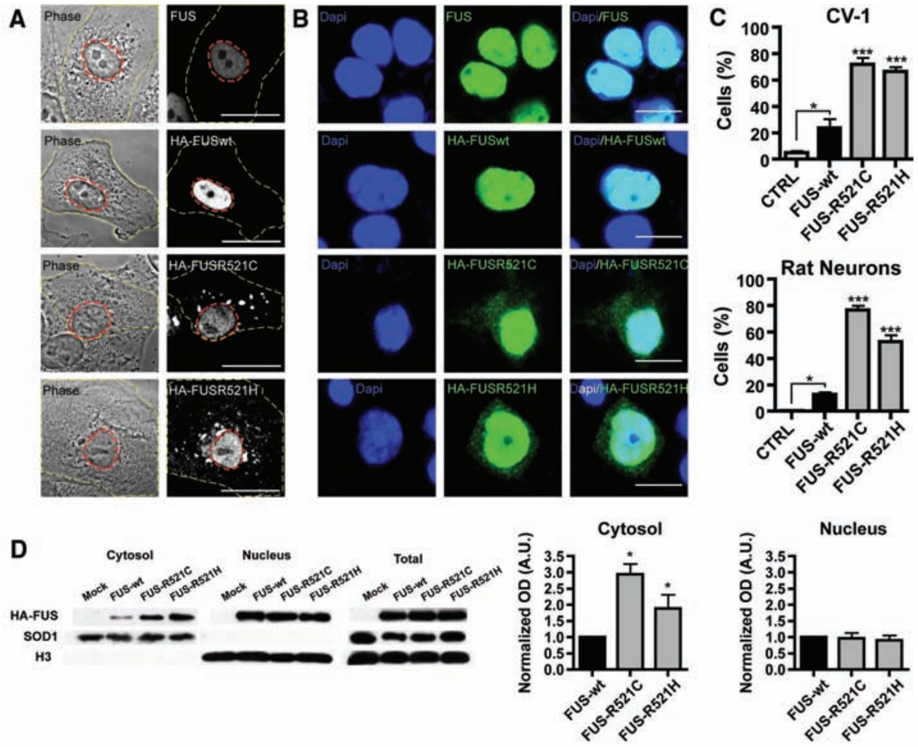


**Fig. 3.** FUS mutations cause subcellular mislocalization. (**A** and **B**) The subcellular localization of endogenous FUS, transfected wild-type (HA-FUS$_{\text{WT}}$), and FUS mutants (HA-FUS$_{\text{R521C}}$, HA-FUS$_{\text{R521H}}$) was determined in CV-1 cells (A) and rat cortical neurons (B) by immunofluorescence. The plasma membrane (A, yellow dash) and nucleus (A, red dash) of representative CV-1 cells are outlined on phase contrast (A, left panel) and corresponding FUS immunofluorescence images (A, right panel). Nuclear diamidino-2-phenylindole (DAPI) staining (B, left panel), FUS immunofluorescence (B, middle panel), and overlay (B, right panel) of representative cortical neurons are shown. Scale bars, 25 μm (A), 10 μm (B). (**C**) The percentage of CV-1 cells (upper panel) and cortical neurons (lower panel) showing FUS staining in the cytoplasm was significantly increased for mutant forms of FUS compared to wild-type and endogenous. Statistical significance was determined by one-way analysis of variance ($P < 0.0001$ for CV1 and neurons) followed by Bonferroni's Multiple Comparison Test (*, $P < 0.05$; ***, $P < 0.001$). (**D**) FUS mutants show a significant increase in the cytosolic fraction. N2A cells transfected with FUS$_{\text{WT}}$, FUS$_{\text{R521C}}$, and FUS$_{\text{R521H}}$ were separated into cytosolic and nuclear fractions, and the amount of transfected FUS was determined by densitometry of antibody to HA immunoblots (mean ± SEM, $n$ = 3). The purity of fractions was determined with antibody to SOD1 (cytosol) and antibody to histone H3 (nucleus), and the transfection efficiencies were compared in total cell lysates. Statistical significance was determined by Wilcoxon Signed Rank Test (*, $P < 0.05$).

### References and Notes

1. C. E. Shaw, A. al-Chalabi, N. Leigh, *Curr. Neurol. Neurosci. Rep.* **1**, 69 (2001).
2. C. E. Shaw *et al.*, *Neurology* **49**, 1612 (1997).
3. D. R. Rosen, *Nature* **364**, 362 (1993).
4. F. L. Conforti *et al.*, *Neuromuscul. Disord.* **18**, 68 (2008).
5. M. J. Greenway *et al.*, *Nat. Genet.* **38**, 411 (2006).
6. A. Paubel *et al.*, *Arch. Neurol.* **65**, 1333 (2008).
7. J. Sreedharan *et al.*, *Science* **319**, 1668 (2008).
8. C. E. Shaw *et al.*, *Ann. Neurol.* **43**, 390 (1998).
9. M. E. Gurney, H. Pu, A. Y. Chiu, *Science* **264**, 1772 (1994).
10. S. Boillee *et al.*, *Science* **312**, 1389 (2006).
11. J. Wang, G. Xu, D. R. Borchelt, *Neurobiol. Dis.* **9**, 139 (2002).
12. I. R. Mackenzie *et al.*, *Ann. Neurol.* **61**, 427 (2007).
13. M. Neumann *et al.*, *Science* **314**, 130 (2006).
14. M. A. Gitcho *et al.*, *Ann. Neurol.* **63**, 535 (2008).
15. E. Kabashi *et al.*, *Nat. Genet.* **40**, 572 (2008).
16. V. M. Van Deerlin *et al.*, *Lancet Neurol.* **7**, 409 (2008).
17. C. K. Hand *et al.*, *Am. J. Hum. Genet.* **70**, 251 (2001).
18. P. C. Sapp *et al.*, *Am. J. Hum. Genet.* **73**, 397 (2003).
19. B. A. Hosler *et al.*, *JAMA* **284**, 1664 (2000).
20. M. Morita *et al.*, *Neurology* **66**, 839 (2006).
21. P. N. Valdmanis *et al.*, *Arch. Neurol.* **64**, 240 (2007).

22. C. Vance *et al.*, *Brain* **129**, 868 (2006).
23. D. M. Ruddy *et al.*, *Am. J. Hum. Genet.* **73**, 390 (2003).
24. W. J. Law, K. L. Cann, G. G. Hicks, *Brief. Funct. Genomics Proteomics* **5**, 8 (2006).
25. Y. Kanai, N. Dohmae, N. Hirokawa, *Neuron* **43**, 513 (2004).
26. A. Yoshimura *et al.*, *Curr. Biol.* **16**, 2345 (2006).
27. K. J. De Vos, A. J. Grierson, S. Ackerley, C. C. Miller, *Annu. Rev. Neurosci.* **31**, 151 (2008).

28. This publication is dedicated to the patients and families who have contributed to this project and to B. Coote and C. Cecere for blood sample collection. Postmortem tissues were provided by MRC Neurodegenerative Diseases Brain Bank. This work was supported in the United Kingdom by grants from the American ALS Association, the Middlemass family, Lady Edith Wolfson Trust, Motor Neurone Disease Association UK, The Wellcome Trust, European Union (APOPIS consortium, contract LSHM-CT-2003-503330, and NeuroNE Consortium), National Institute for Health Research Biomedical Research Centre for Mental Health, The South London and Maudsley National Health Service Foundation Trust, Medical Research Council UK, a Jack Cigman grant from King's College Hospital Charity, The Heaton-Ellis Trust, and The Psychiatry Research Trust of the Institute of Psychiatry. In Australia the work was supported by the Australian National Health and Medical Research Council (CDA 511941) and a Peter Stearne grant from the Motor Neuron Disease Research Institute of Australia.

**Supporting Online Material**
www.sciencemag.org/cgi/content/full/323/5918/1208/DC1
Materials and Methods
Figs. S1 and S2
Tables S1 to S4
References

15 September 2008; accepted 16 December 2008
10.1126/science.1165942

---

# Synchronous Hyperactivity and Intercellular Calcium Waves in Astrocytes in Alzheimer Mice

**Kishore V. Kuchibhotla,[1,2] Carli R. Lattarulo,[1] Bradley T. Hyman,[1] Brian J. Bacskai[1]***

**Although senile plaques focally disrupt neuronal health, the functional response of astrocytes to Alzheimer's disease pathology is unknown. Using multiphoton fluorescence lifetime imaging microscopy in vivo, we quantitatively imaged astrocytic calcium homeostasis in a mouse model of Alzheimer's disease. Resting calcium was globally elevated in the astrocytic network, but was independent of proximity to individual plaques. Time-lapse imaging revealed that calcium transients in astrocytes were more frequent, synchronously coordinated across long distances, and uncoupled from neuronal activity. Furthermore, rare intercellular calcium waves were observed, but only in mice with amyloid-β plaques, originating near plaques and spreading radially at least 200 micrometers. Thus, although neurotoxicity is observed near amyloid-β deposits, there exists a more general astrocyte-based network response to focal pathology.**

Growing evidence supports the hypothesis that in Alzheimer's disease (AD), synapses fail and dendritic spines are lost in the amyloid-β (Aβ) plaque microenvironment through a combination of changes in synaptic drive, calcium overload, and the activation of calcium-dependent degenerative processes (*1*–*4*). Neurons, however, make up only part of the brain's volume, with astrocytes making up the bulk of the remainder. Astrocytes form a structurally interconnected network that, in vitro, exhibit distinct long-distance signaling properties that might be revealed in vivo only after pathological trauma. The idea that neural network dysfunction and degeneration also fully mediate the memory loss in AD does not reflect the growing in vivo evidence that astrocytes play an important role in cortical circuit function (*5*–*7*). In AD, pathological studies of human cases and mouse models have shown that astrocytes surround plaques and might play a critical role in Aβ deposition and clearance (*8*–*10*). Given the profound impact of Aβ deposition on nearby neuronal calcium homeostasis and synaptic function, it is reasonable to hypothesize that astrocyte networks would also be perturbed and might contribute to cortical dysfunction (*11*). We sought to test whether senile plaque deposition would similarly affect astrocyte calcium homeostasis or dynamic signaling in vivo in a mouse model of AD.

To answer these questions, we used multiphoton fluorescent lifetime imaging microscopy (FLIM) to measure resting calcium levels in astrocytes of live mice with cortical plaques (*12*). We multiplexed the fluorescent properties of a small-molecule calcium dye [Oregon-Green BAPTA-1 (OGB)] in the same experimental model and for the same group of cells (Fig. 1A); we used OGB both as a relative indicator of astrocytic activity (intensity) and as a quantitative measure of steady-state intracellular calcium concentration ([Ca]i) (lifetime). We used mice that express mutant human Aβ precursor protein (APP) and mutant presenilin 1 (PS1) (APPswe:PS1ΔE9) in neurons. These mutations lead to an increase in Aβ production and plaque deposition beginning at ~4.5 months of age (*13*, *14*). In mice with plaques, resting [Ca]i in astrocytes was higher than in wild-type animals (Fig. 1, B to F). The resting [Ca]i of astrocytes in wild-type mice was 81 ± 3 nM, whereas in transgenic mice the resting [Ca]i was 149 ± 6 nM ($P < 0.05$). We confirmed that the surrounding neuropil signal minimally contaminated the astrocyte [Ca]i (fig. S2). We next mapped the spatial distribution of astrocytes versus Aβ plaque location in three dimensions (fig. S3). There was no effect of plaque proximity on resting [Ca]i in individual astrocytes (Fig. 1G; $n > 25$ cells per bin). However, plaques deposit rapidly in these mice, so that very few astrocytes were farther than 100 μm away from a plaque.

Using time-lapse calcium imaging of OGB intensity, we measured spontaneous astrocytic activity. Astrocytes in adult APP/PS1 mice (6 to 8 months old) with cortical plaques exhibited a significant increase in spontaneous activity (Fig. 2 and movies S1 and S2). 27.9 ± 6.0% of all astrocytes were active in APP/PS1 mice when compared with the relatively rare spontaneous events seen in wild-type mice (8.1 ± 2.3%; $P < 0.05$) (Fig. 2B). APP/PS1 mutant mice did not show evidence of altered spontaneous activity before they developed senile plaques (at 3 to 3.5 months old) (Fig. 2B). Furthermore, the amplitude $\Delta F/F$ of the calcium signals was significantly higher in APP/PS1 mice ($\Delta F/F$, 33.6 ± 1.1%) when compared with that of wild-type mice ($\Delta F/F$, 23.2 ± 0.8%, $P < 0.001$) (fig. S4). This change in astrocytic function was also independent of plaque proximity (Fig. 2C and fig. S4), again implicating a global astrocytic response to plaque deposition. It was possible that our measurements of baseline resting calcium using FLIM (Fig. 1) reflected this increased astrocyte spontaneous activity rather than a change in baseline resting calcium. To test this, we coregistered FLIM data with spontaneous activity data to allow single-cell identification across imaging modalities. Spontaneously active cells did not have significantly different resting calcium than nonactive cells (Fig. 1H, $P = 0.811$).

In this mouse model, astrocytes do not produce Aβ, which suggests a non–cell-autonomous mechanism behind the increased activity. Hyperactive neurons in the plaque penumbra (*2*) might be responsible by inducing increased activity in functionally connected astrocytes. To test this, we blocked neuronal activity with the sodium-channel blocker tetrodotoxin (TTX) (Fig. 2, D to F, and fig. S5). Under these conditions, astrocytic activity persisted (Fig. 2F). Thus, the increase in astrocyte activity does not derive from neuronal hyperactivity near Aβ deposits, which supports the hypothesis that Aβ interacts directly with the astrocyte network.

Astrocytes play a local role in supporting neuronal activity and in mirroring neuronal responses

---

[1] Massachusetts General Hospital, Department of Neurology/Alzheimer's Disease Research Laboratory, 114 16th Street, Charlestown, MA 02129, USA. [2] Program in Biophysics, Harvard University, Cambridge, MA 02138, USA.

*To whom correspondence should be addressed. E-mail: bbacskai@partners.org

to sensory stimuli (*6*). There is little in vivo evidence supporting the hypothesis that astrocytes form long-range signaling networks (*7*, *15*). To test this in the AD mouse model, we measured the correlation in activity for all active astrocyte cell pairs in a given imaging field. Two cells that were synchronously active would have a robust central peak in the cross-correlogram. There was a sizable increase in the central peak of the correlogram for the transgenic astrocyte cell pairs (Fig. 3A and fig. S6).

We next compared the average correlation coefficient as a function of the pairwise distance of the two cells (Fig. 3B) and modeled the relationship with distance as an exponential function. In wild-type mice, activity was correlated only when cells were within 50 µm of each other ($P <$ 0.05, Kruskal-Wallis with Tukey-Kramer post hoc test). In transgenic mice, there was increased correlated activity as compared with that of the wild type ($P < 0.001$, Mann-Whitney *U* test) that persisted at distances beyond 50 µm ($P < 0.01$, Kruskal-Wallis with Tukey-Kramer post hoc test). Even at cell-to-cell distances of up to 200 µm, there was a significant elevation in correlated activity as compared with that of the wild type. This can be further seen when comparing the exponential fits: the decay constant was 2.4 times longer in transgenic (–0.020) than in wild-type (–0.008) mice.

Thus, astrocytes in mice with cortical plaques exhibit functional coordination of their intracellular calcium signals at long distances. Indeed, we observed intercellular calcium waves (ICWs) in six out of eight transgenic mice but never in wild-type mice (Fig. 4; $n$ = 8 transgenic and 5 nontransgenic mice; $P < 0.05$, Fisher exact test). These waves traveled 196 ± 41 µm at a speed of 22.7 ± 6.4 µm/s, similar to that observed in cells

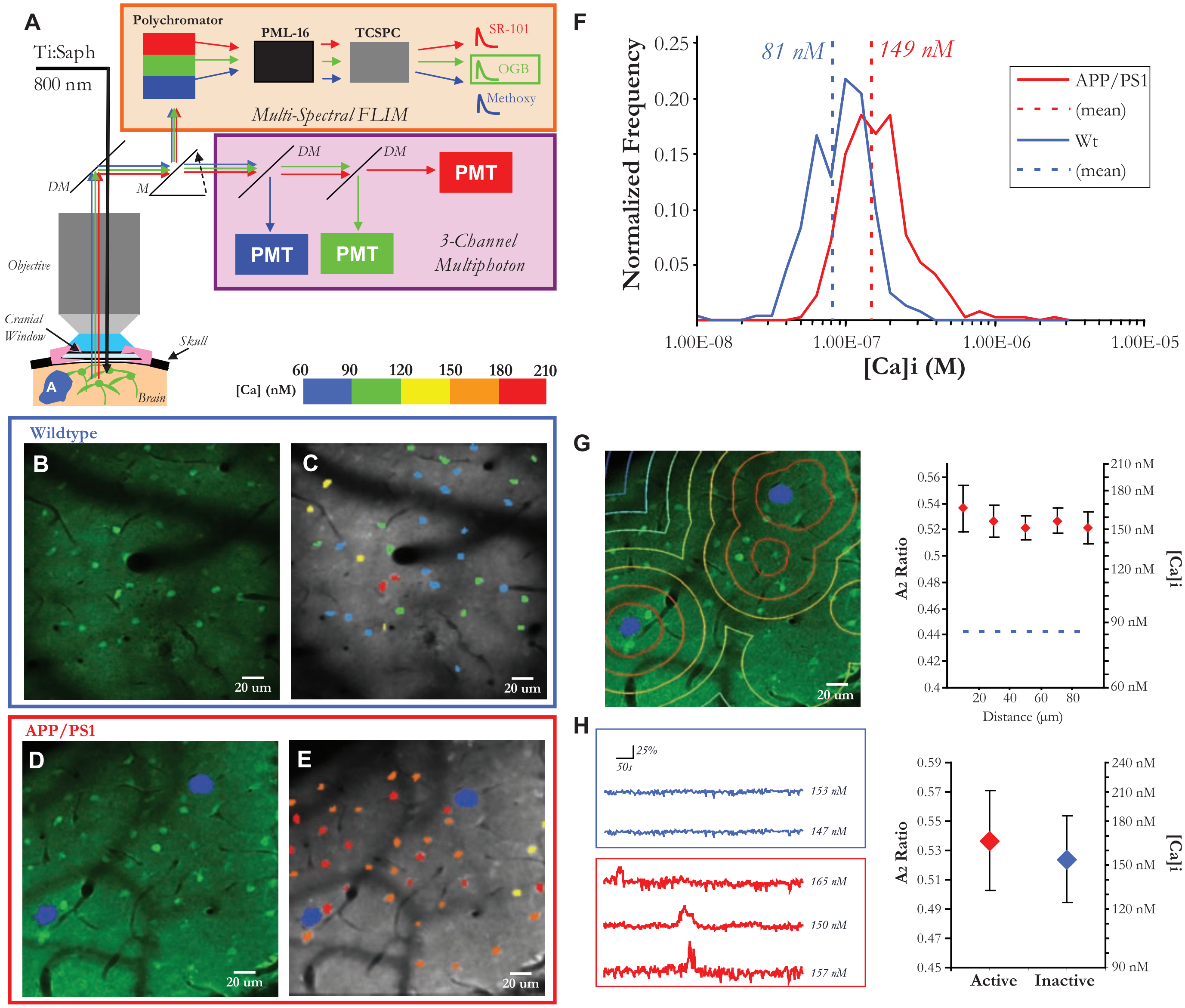


**Fig. 1.** Resting calcium is globally elevated in astrocytic networks. (**A**) Multiphoton laser illumination simultaneously excited methoxy-XO4 (blue, Aβ), OGB (green, neurons and astrocytes), and SR-101 (red, astrocytes) through a cranial window. The resulting fluorescence emission was sent to either (*1*) a three-channel intensity-based photomultiplier tube (PMT) module or (*2*) a 16-channel multispectral FLIM detector. A single-photon counter (TCSPC) recorded fluorescence lifetime data. (**B** to **E**) Fluorescence decay curves were fit with a calcium-bound lifetime (2359 ps) and an unbound calcium lifetime (569 ps) for each pixel. The pixel data were averaged to obtain single-cell calcium levels, depicted with a calibrated color bar [(C) and (E)]. (**F**) Astrocytes in APP/PS1 mice with cortical plaques [(D) and (E), blue] exhibited significantly higher levels of [Ca]i than those in wild-type mice [(B) and (C)] {$P < 0.05$, Student's *t* test, $n = 241$ cells in 3 mice [wild type (Wt)], $n =$ 364 cells in 3 mice [transgenic (Tg)]}. (**G**) Astrocyte resting [Ca]i did not depend on proximity to a plaque ($P = 0.9194$, Kruskal-Wallis test, $n > 25$ cells for each distance group). (**H**) There was no difference in resting calcium between cells that were active versus inactive ($P = 0.811$, Student's *t* test, $n =$ 209 cells in three mice).

in culture (*16*). The "initiator" astrocyte was located 24.8 ± 7.8 μm from the nearest senile plaque, a proximity that has been previously associated with increased synapto- and neurotoxicity (*3*, *4*). This was significantly closer than the average distance of an astrocyte to the nearest plaque (Fig. 4D; active = 52.5 ± 2.3 μm, inactive = 52.3 ± 1.2 μm, $P < 0.05$). Calcium transients during an ICW ($\Delta F/F$, 45.2 ± 2.4%) were higher in magnitude than a typical spontaneous event in transgenic mice (Fig. 4E; $\Delta F/F$, 30.1 ± 1.5%, $P < 0.01$), suggesting that this signal was different from that spread by nonwave activity.

Astrocytes, in normal conditions, can have highly localized and independent responses (*6*).

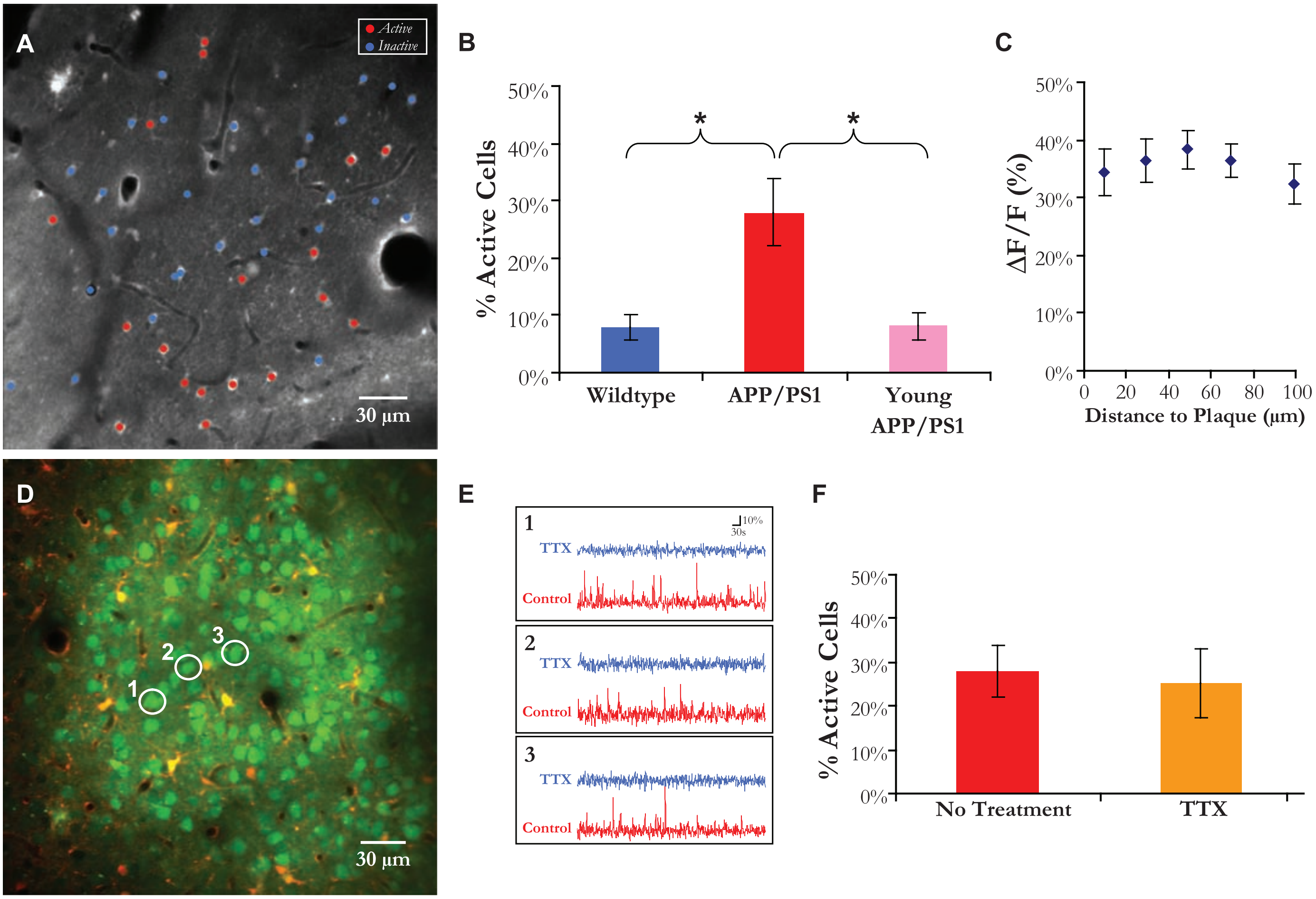


**Fig. 2.** Neuron-independent increase in spontaneous activity throughout the astrocytic network. (**A**) Cell-activity map overlaid on a multiphoton image of astrocytes. (**B**) There were more spontaneously active cells in mice with cortical plaques (*$P < 0.05$, Kruskal-Wallis with Tukey-Kramer post hoc test, $n = 15$ mice: 8 APP/PS1 with plaques, 4 nontransgenic; and 3 APP/PS1 before plaque deposition). (**C**) The amplitude of the [Ca]i transients ($n = 160$ astrocytes) did not depend on proximity to plaques ($P = 0.9178$, Kruskal-Wallis test). (**D**) Multiphoton image of neurons and astrocytes in an APP/PS1 transgenic mouse with cortical plaques. Three neurons are highlighted to show their spontaneous activity traces before and after application of 1 μM TTX in (**E**). (**F**) There was no reduction in the percentage of active astrocytes in the presence of TTX [27.9 ± 6.0% in control ($n = 8$ mice, 1241 astrocytes) versus 25.4 ± 7.8% under TTX ($n = 4$ mice, 818 astrocytes), $P = 0.8081$, Mann-Whitney *U* test].

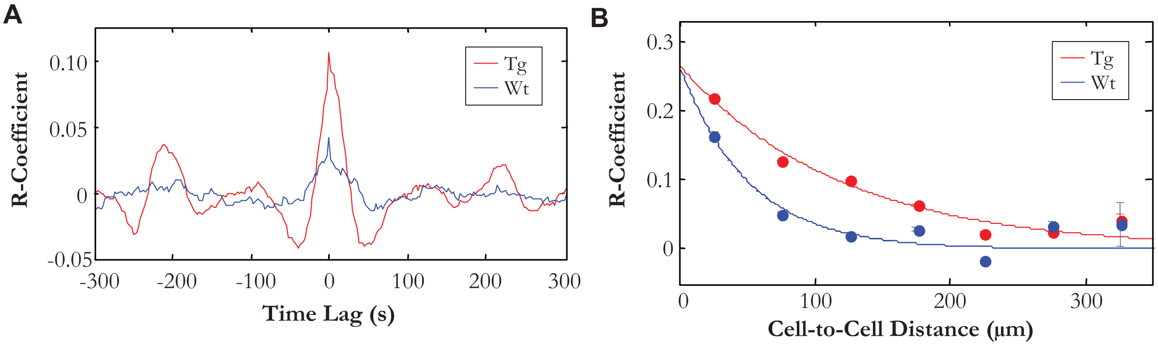


**Fig. 3.** Spatiotemporal synchrony of astrocytic calcium signaling in APP/PS1 mice. (**A**) The mean cross-correlogram for all active cell pairs (excluding autocorrelations) in transgenic and nontransgenic mice. In APP/PS1 mice there was an increase in the probability that two cells had coordinated activity ($n = 3$ mice for APP/PS1 and wild type, $n = 1257$ cell pairs in transgenic and $n = 471$ cell pairs in wild type). (**B**) Cell-pair distance (*x* axis) versus correlation coefficient (*y* axis). Data were fit to a mono-exponential decay curve. In transgenic mice (red), cell pairs exhibited significantly correlated activity at distances up to 200 μm ($P < 0.01$, Kruskal-Wallis with Tukey-Kramer post hoc test), whereas in wild-type mice (blue) cell pairs were not correlated at intercellular distances greater than 50 μm ($P = 0.65$, Kruskal-Wallis with Tukey-Kramer post hoc test).

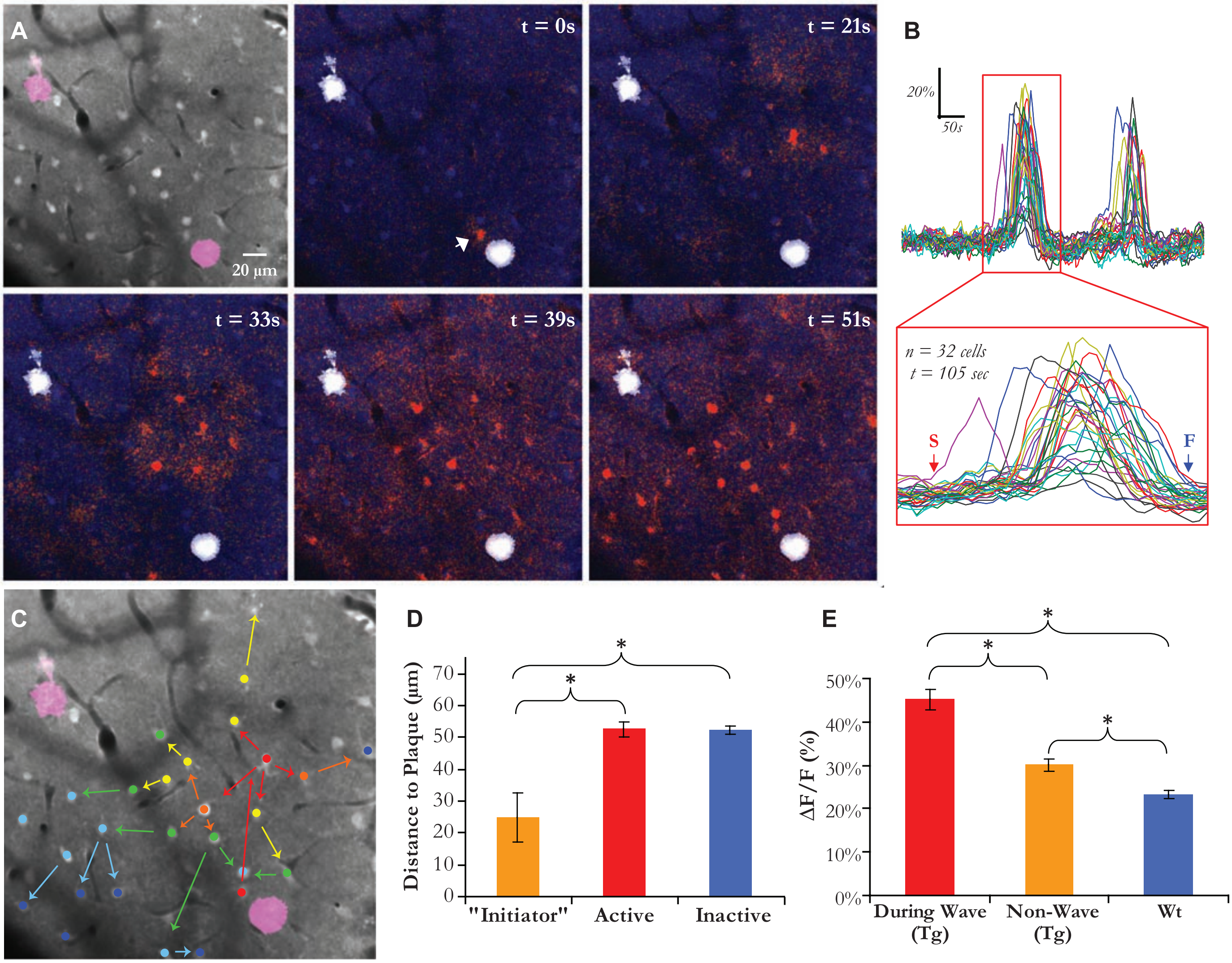


**Fig. 4.** ICWs in mice with cortical plaques. (**A**) A time-lapse Δ*F*/*F* image filmstrip and (**B**) cell-resolved time course of calcium activity in an APP/PS1 transgenic mouse. The white arrow in the second panel points to the initiator astrocyte, and the sequential panels show the propagation of the wave. (**C**) A summary of (A) in which color denotes temporal sequence and arrows denote spatial propagation. (**D**) Astrocytes that initiate ICWs were closer to senile plaques than the average active or inactive astrocyte ($*P < 0.05$, Kruskal-Wallis with Tukey-Kramer post hoc test). (**E**) There was a significant increase in the amplitude of the calcium signal during a wave [$*P < 0.001$, analysis of variance with Tukey-Kramer post hoc test, $n = 47$ cells in transgenic mice (wave), $n = 156$ cells in transgenic mice (nonwave), $n = 168$ in wild-type mice].

In the AD model, astrocytes exhibited network-wide elevations in resting calcium and increased network-level functional activity. Thus, it appears that astrocytes may represent functionally adaptive cells that play distinct roles in health versus disease. The astrocytic network amplified the effects of focal amyloid deposition across a larger cortical network landscape, perhaps contributing to the global alterations in cortical function and possibly the memory disorders seen in AD. One mechanism underlining this amplification effect is the propagation of ICWs. The observed ICWs typically began in the local plaque microenvironment, suggesting that plaques or plaque-associated bioactive species might induce these powerful calcium waves that travel across the cortex. Since the initial discovery of ICWs nearly two decades ago (*17*), their existence in culture and in acute slices has never been confirmed in vivo (*16*, *18*). Their propagation, through gap junctions (*19*) or via adenosine 5′-triphosphate (*20*, *21*), has been postulated to signal the existence of a pathological insult (*16*, *18*). Our data support this hypothesis because plaque deposition has been linked extensively to neuronal trauma (*1*, *4*, *22*).

The increased astrocyte activity cannot be explained by a simple coupling mechanism with neuronal activity. First, neuronal calcium homeostasis is most severely impaired near senile plaques (*4*), whereas here resting calcium was globally elevated in astrocytes. Second, neurons exhibit a pronounced hyperactivity near plaques (*2*), whereas here astrocytes were more active both near and far from plaques. Third, abolishing neuronal activity had no measurable effect on astrocytic calcium oscillations. Thus, whereas senile plaque deposition induces local synapto- and neurotoxicity, the same Aβ deposits might catalyze astrocytic intra- and intercellular signaling events. This idea is supported in part by evidence that cultured astrocytes exhibit elevated calcium upon application of Aβ but are surprisingly resistant to cell death (unlike neurons) (*23*). An important question that remains unanswered is whether modulating astrocytic calcium signaling, via genetic (*24*) or pharmacological manipulations, will result in dynamic changes in amyloid accumulation (*10*) or alter neuronal network activity associated with behavioral impairments.

### References and Notes

1. M. Meyer-Luehmann *et al.*, *Nature* **451**, 720 (2008).
2. M. A. Busche *et al.*, *Science* **321**, 1686 (2008).
3. T. L. Spires *et al.*, *J. Neurosci.* **25**, 7278 (2005).

4. K. V. Kuchibhotla *et al.*, *Neuron* **59**, 214 (2008).
5. X. Wang *et al.*, *Nat. Neurosci.* **9**, 816 (2006).
6. J. Schummers, H. Yu, M. Sur, *Science* **320**, 1638 (2008).
7. H. Hirase, L. Qian, P. Bartho, G. Buzsaki, *PLoS Biol.* **2**, E96 (2004).
8. N. J. Maragakis, J. D. Rothstein, *Nat. Clin. Pract. Neurol.* **2**, 679 (2006).
9. J. Wegiel *et al.*, *Neurobiol. Aging* **22**, 49 (2001).
10. T. Wyss-Coray *et al.*, *Nat. Med.* **9**, 453 (2003).
11. T. Takano, X. Han, R. Deane, B. Zlokovic, M. Nedergaard, *Ann. N. Y. Acad. Sci.* **1097**, 40 (2007).
12. C. D. Wilms, H. Schmidt, J. Eilers, *Cell Calcium* **40**, 73 (2006).
13. M. Garcia-Alloza *et al.*, *Neurobiol. Dis.* **24**, 516 (2006).
14. D. R. Borchelt *et al.*, *Neuron* **19**, 939 (1997).
15. C. Agulhon *et al.*, *Neuron* **59**, 932 (2008).
16. E. Scemes, C. Giaume, *Glia* **54**, 716 (2006).
17. A. H. Cornell-Bell, S. M. Finkbeiner, M. S. Cooper, S. J. Smith, *Science* **247**, 470 (1990).
18. T. A. Fiacco, K. D. McCarthy, *Glia* **54**, 676 (2006).
19. S. Boitano, E. R. Dirksen, M. J. Sanderson, *Science* **258**, 292 (1992).
20. M. L. Cotrina, J. H. Lin, J. C. Lopez-Garcia, C. C. Naus, M. Nedergaard, *J. Neurosci.* **20**, 2835 (2000).
21. P. B. Guthrie *et al.*, *J. Neurosci.* **19**, 520 (1999).
22. M. Garcia-Alloza, S. A. Dodwell, M. Meyer-Luehmann, B. T. Hyman, B. J. Bacskai, *J. Neuropathol. Exp. Neurol.* **65**, 1082 (2006).
23. A. Y. Abramov, L. Canevari, M. R. Duchen, *J. Neurosci.* **23**, 5088 (2003).
24. J. Petravicz, T. A. Fiacco, K. D. McCarthy, *J. Neurosci.* **28**, 4967 (2008).

25. We thank K. O. Ohki and R. C. Reid for assistance with the multicell bolus loading technique. This work was supported by NIH grants EB000768 (B.J.B.), AG08487 (B.T.H.), and NS580752 (K.V.K.).

**Supporting Online Material**
www.sciencemag.org/cgi/content/full/323/5918/1211/DC1
Materials and Methods
Figs. S1 to S6
Movies S1 and S2
References

27 November 2008; accepted 16 January 2009
10.1126/science.1169096

---

# Meropenem-Clavulanate Is Effective Against Extensively Drug-Resistant *Mycobacterium tuberculosis*

**Jean-Emmanuel Hugonnet,[1] Lee W. Tremblay,[1] Helena I. Boshoff,[2] Clifton E. Barry 3rd,[2] John S. Blanchard[1]***

**β-lactam antibiotics are ineffective against *Mycobacterium tuberculosis*, being rapidly hydrolyzed by the chromosomally encoded *blaC* gene product. The carbapenem class of β-lactams are very poor substrates for BlaC, allowing us to determine the three-dimensional structure of the covalent BlaC-meropenem covalent complex at 1.8 angstrom resolution. When meropenem was combined with the β-lactamase inhibitor clavulanate, potent activity against laboratory strains of *M. tuberculosis* was observed [minimum inhibitory concentration ($\text{MIC}_{\text{meropenem}}$) less than 1 microgram per milliliter], and sterilization of aerobically grown cultures was observed within 14 days. In addition, this combination exhibited inhibitory activity against anaerobically grown cultures that mimic the "persistent" state and inhibited the growth of 13 extensively drug-resistant strains of *M. tuberculosis* at the same levels seen for drug-susceptible strains. Meropenem and clavulanate are Food and Drug Administration–approved drugs and could potentially be used to treat patients with currently untreatable disease.**

Tuberculosis is perhaps the most persistent human disease caused by an infectious bacterium, *Mycobacterium tuberculosis*. The death toll remains extremely high, despite the introduction of modern multidrug chemotherapy in the 1960s, with between 1.6 and 2 million fatalities annually. An increasing percentage of human clinical isolates are drug-resistant or multidrug-resistant strains that threaten the ability to treat the disease (*1*). The continued use of multidrug therapy has caused an even more dire problem: strains of *M. tuberculosis* resistant to all first-, second-, and third-line agents. In a recent study from South Africa, 54 of 54 patients infected with such highly resistant strains died with a mean survival time from diagnosis of 16 days (*2*).

Since the discovery of penicillin in 1929 (*3*), the β-lactam class of antibiotics has included some of the most clinically important antibacterial agents. The development of broad-spectrum derivatives of penicillin, such as the cephalosporins and olivanic acid (*4*), coupled with their low inherent toxicity have made them the drugs of choice for the treatment of both Gram-negative and Gram-positive bacterial infections. This class, however, has never provided a compound useful in the treatment of tuberculosis, and β-lactams are only rarely used in the treatment of this disease. One important reason for the lack of efficacy was found in the genome sequence of *M. tuberculosis*, which contains a single, highly active, chromosomally encoded class A (Ambler) β-lactamase (*5*). Recently a genetic knockout of the *blaC*-encoded β-lactamase showed that strains lacking this enzyme were more sensitive to β-lactams (*6*). This suggested that the chemical recapitulation of the genetic knockout could similarly resensitize the organism to existing β-lactam antibiotics.

We recently cloned and expressed the *M. tuberculosis blaC* gene and reported a detailed enzymatic characterization (*7*). BlaC exhibits an exceptionally broad substrate specificity, hydrolyzing penicillins at nearly the diffusion-limited rate, all classes of cephalosporins, and, unexpectedly for a class A extended-spectrum β-lactamase, imipenem and meropenem, both carbapenems. Equally unexpected, the enzyme was only transiently inhibited by the β-lactamase inhibitors sulbactam and tazobactam, penicillanic acid sulfones with potent inhibitory activity against other class A β-lactamases. However, clavulanic acid is the only Food and Drug Administration (FDA)–approved β-lactamase inhibitor that irreversibly inhibits BlaC, suggesting that clavulanic acid may recapitulate the genetic knockout, rendering *M. tuberculosis* susceptible to β-lactam antibiotics.

We have previously shown that meropenem was an extremely slow substrate for *M. tuberculosis* BlaC, being hydrolyzed five orders of magnitude slower than ampicillin. A more detailed investigation of the kinetics of meropenem hydrolysis under near stoichiometric enzyme concentrations revealed a steady-state burst with a magnitude dependent on the concentration of BlaC (Fig. 1A). The reaction of meropenem with the enzyme to form the acyl-enzyme intermediate (acylation half-reaction) is fast relative to hydrolysis of the substrate (deacylation). Extrapolation of the final, linear rate to zero time revealed that enzyme acylation was stoichiometric with meropenem. At a single catalytic enzyme concentration, the linear rates yielded plots typical of Michaelis-Menten kinetics (Fig. 1B), with Michaelis constant $K_m = 3.4 \pm 0.7\ \mu\text{M}$ and turnover number $k_{cat} = 0.08 \pm 0.01\ \text{min}^{-1}$. Because of its extremely slow turnover rate, we investigated the possibility that meropenem could act as an inhibitor of BlaC and whether it was possible to trap the covalently acylated form of the enzyme. Meropenem acts as a slow, tight-binding inhibitor of the hydrolysis of the chromogenic β-lactam nitrocefin by BlaC. The time courses of nitrocefin hydrolysis are nonlinear in the presence of meropenem, and an analysis of these data yielded an inhibition constant ($K_i$) value of $16 \pm 2\ \mu\text{M}$ and a $K_i{}^*$ value of $1.1 \pm 0.8\ \mu\text{M}$ (fig. S1, A and B). The ability of meropenem to act as an inhibitor of BlaC in addition to being a very poor substrate for BlaC added to its potential as an active partner with clavulanate.

The rapid acylation and slow deacylation of BlaC by meropenem suggested that we could observe the covalently bound species by Fourier transform ion cyclotron resonance (FTICR) mass spectrometry. A freshly prepared solution of

---

[1] Department of Biochemistry, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, NY 10461, USA. [2] Tuberculosis Research Section, Laboratory of Clinical Infectious Diseases, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD 20892, USA.

*To whom correspondence should be addressed. E-mail: blanchar@aecom.yu.edu

BlaC and meropenem displayed a peak corresponding to the mass of the covalently acylated BlaC-meropenem complex [charge/mass ($m/z$) = 29,167.5] and a second peak whose mass corresponded to the mass of the covalently acylated BlaC-meropenem complex –44 ($m/z$ = 29,123.6) (Fig. 1C). After 7 min of incubation, both these peaks decreased in intensity with the corresponding appearance of the free enzyme. Small-molecule mass spectrometry revealed the presence of two species, one with the expected mass for hydrolyzed meropenem ($m/z$ = 402) and another with a mass 44 mass units smaller ($m/z$ = 358, fig. S2). Hydrolysis of meropenem in 1 N NaOH followed by mass spectrometry revealed only the presence of hydrolyzed meropenem. Together, these experiments suggest that, after β-lactam ring opening, the covalently bound meropenem partitions between direct hydrolysis and enzyme-catalyzed decomposition of the C6 hydroxyethyl substituent, yielding acetaldehyde ($m/z$ = 44). The proposed chemical mechanism is discussed below.

The mass spectrometry results suggested that soaking of crystals of BlaC with meropenem followed by vitrification would trap the meropenem complex at the active site. Crystals of BlaC were prepared as previously described (*8*). Crystals soaked for 90 min with 50 mM meropenem containing 20% glycerol were vitrified and analyzed by x-ray diffraction at the Brookhaven National Laboratory Synchrotron Radiation Source. Crystals were present in the same space group as we previously reported for the BlaC-clavulanate complex, and diffraction data to 1.8 Å resolution were used to solve the structure (final model: $R_{\text{work}}$ = 0.152 and $R_{\text{free}}$ = 0.192, table S1) with molecular replacement methods and the structure of the BlaC-clavulanate complex (*9*). Clear electron density that was contiguous with the β-hydroxyl side chain of $\text{Ser}^{70}$ (Ambler numbering is used throughout for residue identification) was observed in the active site (Fig. 2, A and B). The carbonyl oxygen of the covalent enzyme-meropenem ester is oriented to interact with the main chain amide of $\text{Ser}^{70}$ and $\text{Thr}^{253}$, residues comprising the “oxyanion” hole. When compared to the structure of the BlaC-clavulanate complex, a number of active site rearrangements are evident. Most notably, the ε-amino group of $\text{Lys}^{73}$ and the $\text{Ser}^{130}$ side chain hydroxyl no longer hydrogen-bond to the $\text{Ser}^{70}$ ester oxygen. The $\text{Lys}^{73}$ ε-amino group and the carboxyl side chain of $\text{Glu}^{166}$ interact with the meropenem C8 hydroxyl group, whereas the $\text{Ser}^{130}$ side chain hydroxyl group hydrogen bonds to the nitrogen of the pyrroline ring (fig. S3). The pyrroline ring shows a collinear relationship of C6, N1, C2, and C3, requiring that the double bond originally present between C2 and C3 has isomerized, as first proposed by Knowles (*10*). The structure clearly shows that the thioether sulfur atom at the C3 position is in the *S* configuration requiring protonation of the *re* face of the C2-C3 double bond. We thus propose a chemical mechanism for meropenem hydrolysis by BlaC as shown in Fig. 2C. The electron density beyond the thioether sulfur atom is weak and discontinuous, so the modeled configuration of the terminal pyrazole ring is uncertain at this time. Whether double bond isomerization is concerted with β-lactam ring opening is also unclear, although we draw it as concerted (Fig. 2C).

The combination of clavulanate and amoxicillin has previously been shown to inhibit the growth of *M. tuberculosis* strains (*11*). Because amoxicillin is one of the best substrates of BlaC, complete inactivation of BlaC would be required to maintain inhibitory concentrations of the antibiotic. More recently, the addition of clavulanate to susceptible and multidrug-resistant strains of *M. tuberculosis* has been shown to potentiate the effects of all classes of β-lactams (*12*, *13*). We determined the minimum inhibitory concentration (MIC) values of *M. tuberculosis* H37Rv in 7H9 medium at 37°C by penicillins, cephalosporins, and carbapenems in the absence and presence of 2.5 $\mu\text{g ml}^{-1}$ clavulanate. The addition of clavulanate had only a modest effect on the MIC values of ampicillin and amoxicillin but a significant effect on the MIC values of cephalothin, imipenem, and meropenem (table S2). On the basis of the low MIC value of meropenem in the presence of clavulanate (0.32 $\mu\text{g ml}^{-1}$) and its low rate of hydrolysis by BlaC, we selected this carbapenem for detailed analysis. When various combinations of meropenem and clavulanate were added daily for 5 consecutive days to cultures of *M. tuberculosis* Erdman under aerobic growth conditions, the colony-forming units per milliliter ($\text{CFU ml}^{-1}$) dropped rapidly until complete sterilization was obtained after 9 to 12 days (Fig. 3A). Although the number and identity of the cell wall cross-linking transpeptidase targets of meropenem in *M. tuberculosis* are not known (see below), it is clear that a combination of clavulanate and meropenem rapidly sterilizes actively growing aerobic cultures of *M. tuberculosis*.

An important problem in tuberculosis therapy is the phenotypic drug resistance of populations of organisms that are in a nonreplicative state, termed “persistence” (*14*, *15*). The *M. tuberculosis* L,D-transpeptidase has recently been reported to be a target for carbapenems, and this enzyme is thought to be expressed as the organism enters into the persistent state, with corresponding changes in the nature of peptidoglycan cross-linking (*16*). There are several in vitro models of this state, but the most widely used is the Wayne model (*17*), where organisms that are grown in sealed tubes enter into a viable but nonreplicative (NRP) state after 2 weeks because of consumption of available oxygen. Combinations of clavulanate and meropenem were tested for their ability to sterilize organisms in this state. Drug combinations were added to NRP2 cultures within an anaerobic chamber, and cellular viability was assessed 1 week and 2 weeks later by measuring intracellular adenosine triphosphate (ATP) concentrations as well as CFUs. All clavulanate–β-lactam combinations were effective in reducing viability, but the decrease was more pronounced with the two carbapenems, imipenem and meropenem, than with amoxicillin and cefuroxime (fig. S4). In the case of the clavulanate-meropenem combination, we observed more than a log kill over a 2-week exposure time, comparable to metronidazole, which is shown as a control compound in Fig. 3B (*12*).

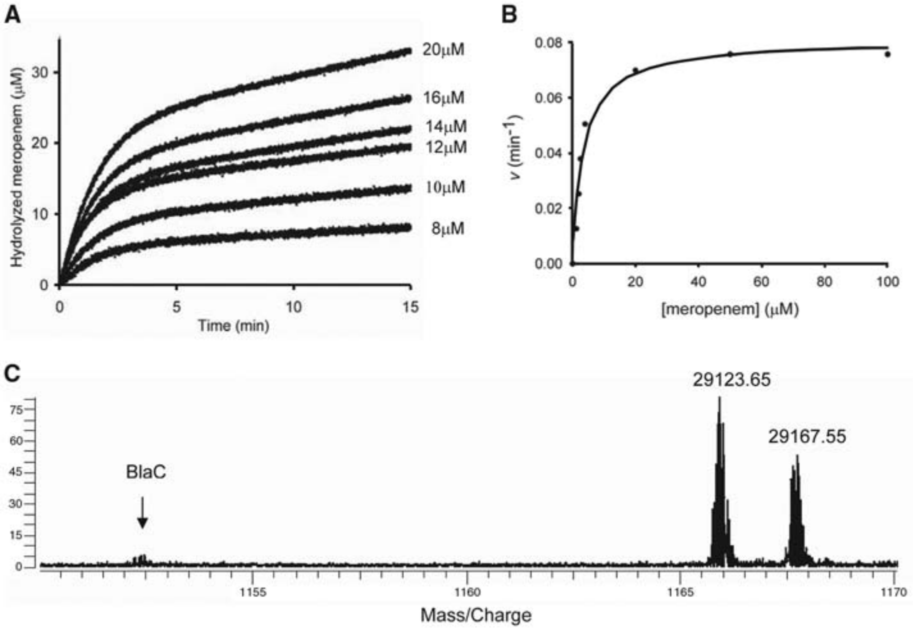


**Fig. 1.** Kinetics of BlaC with meropenem. (**A**) Time courses of meropenem hydrolysis with various concentrations of BlaC. Enzyme concentrations are reported on the right. (**B**) Michaelis-Menten kinetics of BlaC with meropenem at single enzyme concentration (0.8 μM). (**C**) Mass spectra of enzyme-meropenem species. The 25+ charge state ions are shown. The mass reported for each peak was calculated as described in (*20*) from the two ions with $m/z$ values of 1165.946 and 1167.702.

The combination of clavulanate with β-lactams, especially meropenem, was also tested for the ability to inhibit the growth of extensively drug-resistant (XDR) clinical strains of *M. tuberculosis*. Thirteen clinical isolates exhibiting the XDR phenotype were tested (*18*). Clavulanate was used at a concentration of 2.5 μg $ml^{-1}$, and the MIC values of these strains for meropenem were determined. The susceptibility of these strains was experimentally indistinguishable from that determined for H37Rv and the Erdman strain, that is, ≤1 μg $ml^{-1}$ (Table 1). In contrast, substantial variability in the MIC values to ampicillin, amoxicillin, cephalothin, and imipenem was observed for these same strains (table S2). The clavulanate-meropenem combination is thus equally effective against both susceptible and XDR strains.

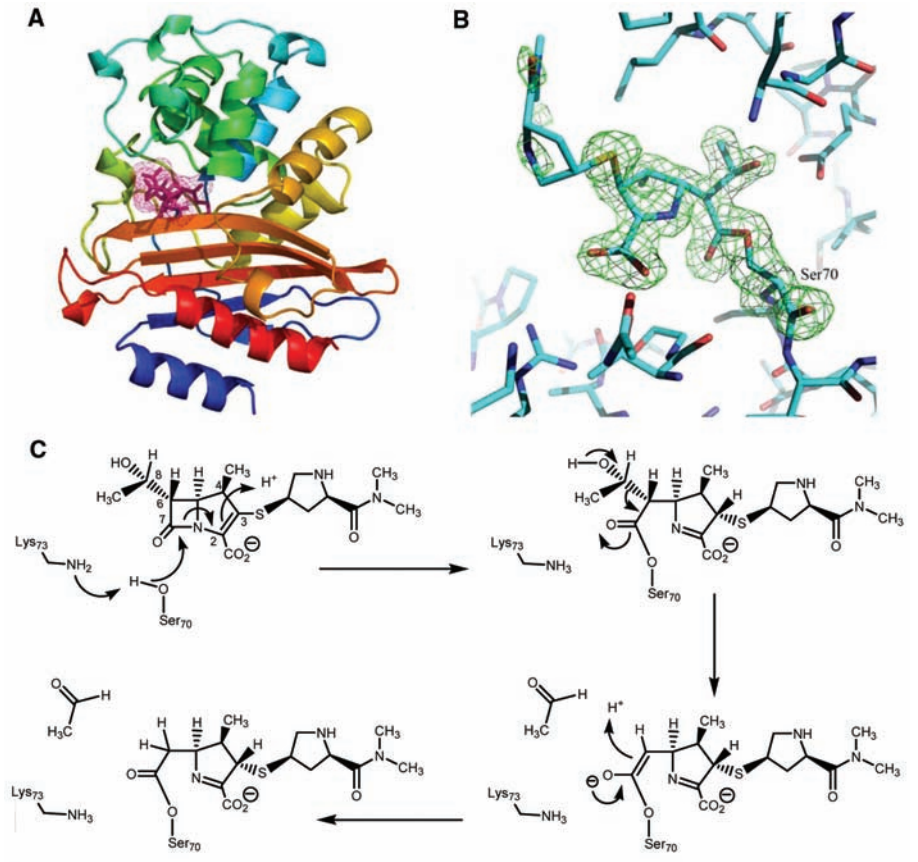


**Fig. 2.** (**A**) Overall structure of BlaC displayed in rainbow from N term (blue) to the C term (red), with the meropenem adduct displayed as a surface mesh. (**B**) $F_o$–$F_c$ omit density (green) contoured at 4.0 σ surrounds the covalent meropenem adduct formed at the Ambler active-site residue $Ser^{70}$. Structure figures were produced using Pymol (www.pymol.org). (**C**) Proposed chemical mechanism for the BlaC-catalyzed reaction with meropenem.

These structural and mechanistic studies of carbapenem interactions with the BlaC β-lactamase have revealed properties of specific β-lactam antibiotics that can be exploited in the treatment of tuberculosis, including the treatment of multidrug- and extensively drug-resistant strains. The structure of the meropenem-inactivated acyl-enzyme, in combination with our mechanistic proposal for its hydrolysis and the structure of the clavulanate-BlaC complex, provides the information necessary to design improved tuberculosis-specific β-lactams that could form longer-lived acyl-enzyme intermediates. Among currently approved β-lactams, however, meropenem is superior on the basis of its poor activity as a substrate for BlaC, ability to transiently inhibit BlaC, and activity against nonreplicating organisms. This activity provides experimental evidence that peptidoglycan remodeling occurs in *M. tuberculosis* in the nonreplicating state, which may be an important determinant of clinical efficacy.

Ten years ago, a report on the early bactericidal activity of amoxicillin-clavulanate in patients with tuberculosis appeared (*19*), but no additional reports have appeared since then. Our studies reveal that clinical strain-to-strain variability is observed with combinations of clavulanate with penicillins, cephalosporins, and imipenem but not with meropenem. The synergism of the clavulanate-meropenem combination and the uniform activity against drug-susceptible, laboratory, and XDR clinical strains suggest this combination could be useful in the treatment of

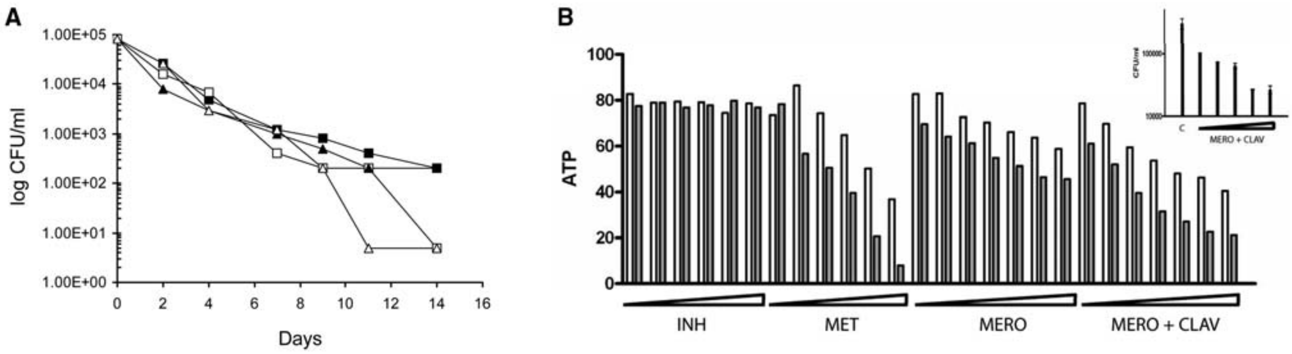


**Fig. 3.** Killing curves of *M. tuberculosis* after exposure to β-lactams and clavulanate. (**A**) Aerobic growth using the microdilution method. Meropenem and clavulanate were added at 2 μg $ml^{-1}$ + 1 μg $ml^{-1}$ (■), 2 μg $ml^{-1}$ + 2 μg $ml^{-1}$ (□), 4 μg $ml^{-1}$ + 1 μg $ml^{-1}$ (▲), and 4 μg $ml^{-1}$ + 2 μg $ml^{-1}$ (Δ), respectively, for 5 consecutive days. (**B**) Meropenem is cidal for nonreplicating anaerobic *M. tuberculosis*. Hypoxically adapted *M. tuberculosis* H37Rv was treated under anaerobic conditions with twofold dilutions of meropenem (0.19 to 12.5 μg $ml^{-1}$) in the presence or absence of 2.5 μg $ml^{-1}$ clavulanate. Isoniazid (0.16 to 1.0 μg $ml^{-1}$) and metronidazole (4.6 to 73 mM) served as negative and positive controls, respectively. Survival was determined by measurement of ATP amounts in surviving bacteria during aerobic outgrowth of 100-fold diluted cells at either 1 week (white bars) or 2 weeks (shaded bars) of treatment or by enumeration of CFUs (inset) after 2 weeks of compound exposure.

**Table 1.** MIC values for $\beta$-lactams in the presence of 2.5 $\mu$g ml$^{-1}$ clavulanic acid. The XDR strains were a subset of those previously reported (*18*).

| Strain | $\beta$-lactam | MIC value ($\mu$g ml$^{-1}$) |
|---|---|---|
| Erdman | Meropenem | 0.5 |
| H37Rv | Amoxicillin | >10 |
| H37Rv | Ampicillin | 5.0 |
| H37Rv | Cefotaxime | 1.25 |
| H37Rv | Cephalothin | 0.94 |
| H37Rv | Imipenem | 0.16 |
| H37Rv | Meropenem | 0.32 |
| XDR-1 | Meropenem | 0.94 |
| XDR-2 | Meropenem | 0.625 |
| XDR-3 | Meropenem | 0.625 |
| XDR-4 | Meropenem | 0.625 |
| XDR-5 | Meropenem | 0.625 |
| XDR-6 | Meropenem | 0.625 |
| XDR-7 | Meropenem | 0.625 |
| XDR-8 | Meropenem | 0.94 |
| XDR-9 | Meropenem | 1.25 |
| XDR-10 | Meropenem | 0.47 |
| XDR-11 | Meropenem | 0.23 |
| XDR-12 | Meropenem | 0.625 |
| XDR-13 | Meropenem | 0.32 |

tuberculosis. Both clavulanate and meropenem are FDA-approved drugs, and both clavulanate and meropenem are sufficiently free of side effects to be approved for pediatric use in children over 3 months old.

### References and Notes

1. H. I. Boshoff, C. E. Barry 3rd, *Nat. Rev. Microbiol.* **3**, 70 (2005).
2. N. R. Gandhi *et al.*, *Lancet* **368**, 1575 (2006).
3. A. Fleming, *Br. J. Exp. Pathol.* **10**, 226 (1929).
4. K. Bush, S. Mobashery, *Adv. Exp. Med. Biol.* **456**, 71 (1998).
5. S. T. Cole *et al.*, *Nature* **393**, 537 (1998).
6. A. R. Flores, L. M. Parsons, M. S. Pavelka Jr., *Microbiology* **151**, 521 (2005).
7. J.-E. Hugonnet, J. S. Blanchard, *Biochemistry* **46**, 11998 (2007).
8. F. Wang, C. Cassidy, J. C. Sacchettini, *Antimicrob. Agents Chemother.* **50**, 2762 (2006).
9. L. W. Tremblay, J.-E. Hugonnet, J. S. Blanchard, *Biochemistry* **47**, 5312 (2008).
10. C. J. Easton, J. R. Knowles, *Biochemistry* **21**, 2857 (1982).
11. M. H. Cynamon, G. S. Palmer, *Antimicrob. Agents Chemother.* **24**, 429 (1983).
12. C. Segura, M. Salvado, I. Collado, J. Chaves, A. Coira, *Antimicrob. Agents Chemother.* **42**, 1524 (1998).
13. I. Dincer, A. Ergin, T. Kocagoz, *Int. J. Antimicrob. Agents* **23**, 408 (2004).
14. J. E. Gomez, J. D. McKinney, *Tuberculosis (Edinburgh)* **84**, 29 (2004).
15. Y. Zhang, *Front. Biosci.* **9**, 1136 (2004).
16. M. Lavollay *et al.*, *J. Bacteriol.* **190**, 4360 (2008).
17. L. G. Wayne, L. G. Hayes, *Infect. Immun.* **64**, 2062 (1996).
18. C. Y. Jeon *et al.*, *Clin. Infect. Dis.* **46**, 42 (2008).
19. H. F. Chambers, T. Kocagoz, T. Sipit, J. Turner, P. C. Hopewell, *Clin. Infect. Dis.* **26**, 874 (1998).
20. Materials and methods are available as supporting material on *Science* Online.
21. The authors wish to thank J. Chan and E. Russel for help in the Wayne model studies and H. Xiao for assistance in mass spectrometry. This work was supported partially by the NIH (AI33696 to J.S.B.); in part by the Intramural Research Program of the NIH, National Institute of Allergy and Infectious Diseases; and in part by a grant from the Bill and Melinda Gates Foundation and the Wellcome Trust through the Grand Challenges in Global Health Initiative. A provisional U.S. patent application was filed on 27 May 2008 related to this work. Structure coordinates have been deposited as Protein Data Bank identification code 3DWZ.

**Supporting Online Material**

www.sciencemag.org/cgi/content/full/323/5918/1215/DC1
Materials and Methods
Figs. S1 to S4
Tables S1 and S2
References

21 October 2008; accepted 6 January 2009
10.1126/science.1167498

---

# Analysis of *Drosophila* Segmentation Network Identifies a JNK Pathway Factor Overexpressed in Kidney Cancer

**Jiang Liu,[1,2]* Murad Ghanim,[1,2]*† Lei Xue,[3]* Christopher D. Brown,[1,2] Ivan Iossifov,[1,4]‡ Cesar Angeletti,[5] Sujun Hua,[1,2] Nicolas Nègre,[1,2] Michael Ludwig,[1,2,6] Thomas Stricker,[1,2,7] Hikmat A. Al-Ahmadie,[7] Maria Tretiakova,[7] Robert L. Camp,[5] Montse Perera-Alberto,[8] David L. Rimm,[5] Tian Xu,[3] Andrey Rzhetsky,[1,4] Kevin P. White[1,2,6]§**

We constructed a large-scale functional network model in *Drosophila melanogaster* built around two key transcription factors involved in the process of embryonic segmentation. Analysis of the model allowed the identification of a new role for the ubiquitin E3 ligase complex factor SPOP. In *Drosophila*, the gene encoding SPOP is a target of segmentation transcription factors. *Drosophila* SPOP mediates degradation of the Jun kinase phosphatase Puckered, thereby inducing tumor necrosis factor (TNF)/Eiger–dependent apoptosis. In humans, we found that SPOP plays a conserved role in TNF-mediated JNK signaling and was highly expressed in 99% of clear cell renal cell carcinomas (RCCs), the most prevalent form of kidney cancer. SPOP expression distinguished histological subtypes of RCC and facilitated identification of clear cell RCC as the primary tumor for metastatic lesions.

Over the last three decades, extensive molecular and genetic analyses have characterized the identity of and interactions between components of the *Drosophila* segmentation process (*1*). Maternal factors distributed in gradients along the anterior-posterior (A-P) axis activate zygotic transcription of gap genes, which encode transcription factors that activate sets of pair-rule genes including the homeobox transcription factors Even-skipped (Eve) and Fushi tarazu (Ftz). These pair-rule proteins then directly regulate segment polarity genes that determine the internal A-P orientation of each segment. Many of the human homologs of these genes and their downstream targets play critical roles in human diseases, especially cancers (*2*, *3*). In an effort to extract new information from the *Drosophila* segmentation network, as well as to mine this network for previously unknown disease-related genes, we built a large-scale predictive network model around Ftz and Eve.

We analyzed gene expression changes between individual wild-type embryos and embryos with null mutations in *ftz* and *eve* (*1*) collected during a developmental time course from 2 hours until 7 hours after egg laying (AEL). By focusing on the effects of Ftz and Eve 2 to 3 hours AEL (early zygotic expression), we found 1310 genes differentially expressed between the *ftz* mutant and wild type, and 1074 genes differentially expressed between the *eve* mutant and wild type (false discovery rate < 0.001; tables S1 and S2).

Using antibodies specific for Ftz or Eve, we performed chromatin immunoprecipitation (ChIP) and mapped genome-wide transcription factor binding in cellular blastoderm embryos 2 hours AEL on custom-designed high-density DNA microarrays (*4*). We found 1286 Ftz- and

1499 Eve-bound probes (intensity $P < 0.0001$ and Z score > 1.96, see supplemental methods; 21 probes on both lists map within 500 base pairs). We analyzed several methods for probe-to-target gene mapping to maximize the overlap between the differentially expressed and ChIP target gene sets (see supplemental methods and fig. S1). The greatest such enrichment was obtained by designating genes within 1 kb of a binding site as targets. At this threshold, we identified 969 Ftz ChIP-chip target genes and 932 Eve ChIP-chip target genes (overlap 175 genes; tables S3 and S4).

Genes both differentially expressed and targeted by ChIP-chip binding site mapping were considered as putative direct target genes. We thus identified 137 Ftz direct target genes (Fig. 1A) and 98 Eve direct target genes (fig. S2; overlap 9 genes). Figure 1A (right) shows the locations of binding sites at Ftz or Eve direct target genes. Analysis of direct target gene annotations indicates that 39 genes (21%) regulate transcription and 74 genes (40%) are involved in developmental processes (Fig. 1A, center); both annotation classes were significantly enriched compared to the 9.6% and 18% of *Drosophila* genes annotated as transcriptional or developmental regulators, respectively ($P = 1.05 \times 10^{-6}$ and $P = 1.81 \times 10^{-12}$; hypergeometric test). A complete target list can be found in tables S5 and S6.

To extend our Ftz-Eve network model beyond direct transcriptional regulation, we included automated literature-mining methods to capture published interactions of target genes (*5*). We then integrated yeast two-hybrid–based protein-protein interaction data (*6*) into our model by connecting protein interactions between existing components in the network. To limit the size of the network, we extended the protein-protein interaction only one degree from the direct targets of Ftz or Eve. The resulting Ftz-Eve regulatory network model included 4084 genes/proteins and 6648 interactions between them (fig. S3).

To confirm parts of the network model topology, we examined several genes that are expressed in segmental patterns (*7*) and validated a limited set of interactions by genetic and biochemical testing of simple predictions from our network model (fig. S4). Analysis of the Eve-Ftz network identified 150 different genes as direct targets of Eve or Ftz that also have unambiguous human homologs. From this gene set, we identified a top candidate, *CG9924* or *roadkill* (*rdx*), which ranks first in network betweenness-centrality and thus constitutes a major network hub (*8*) (see supplemental methods and table S8). The *rdx* gene encodes a BTB domain protein that has been shown to regulate Cubitus interruptus (Ci) degradation in the Hedgehog pathway (*9*, *10*). This product of the *rdx* gene is 79% identical to the human protein SPOP, and these proteins appear to be orthologs (fig. S5) (*9*, *10*); we refer to the *rdx* gene product(s) as *Drosophila* SPOP (D-SPOP).

Our network model indicates that the *D-SPOP* gene is a direct target of Ftz at 2 to 3 hours AEL and that the D-SPOP protein interacts with the Jun kinase phosphatase Puckered (Puc) (Fig. 1B). RNA in situ hybridizations for *D-SPOP* mRNA in *ftz* mutant embryos confirmed that *ftz* is indeed required for *D-SPOP* expression in parasegments that normally express Ftz (Fig. 1C). We did not observe significant misexpression of *D-SPOP* mRNA in *eve* mutant embryos at 2 to 3 hours AEL (fig. S6), suggesting that the Ftz effects on *D-SPOP* mRNA expression occur before the Eve effect. We found that the D-SPOP protein segmental expression pattern was completely lost in *eve* mutant embryos at 6 to 7 hours AEL (Fig. 1D), similar to the expression pattern observed for the well-characterized Ftz and Eve target gene *engrailed* (*11*, *12*). Previous studies also indicate that D-SPOP is regulated by Hedgehog (Hh) later in development, indicating another layer of D-SPOP regulation by the segment polarity system

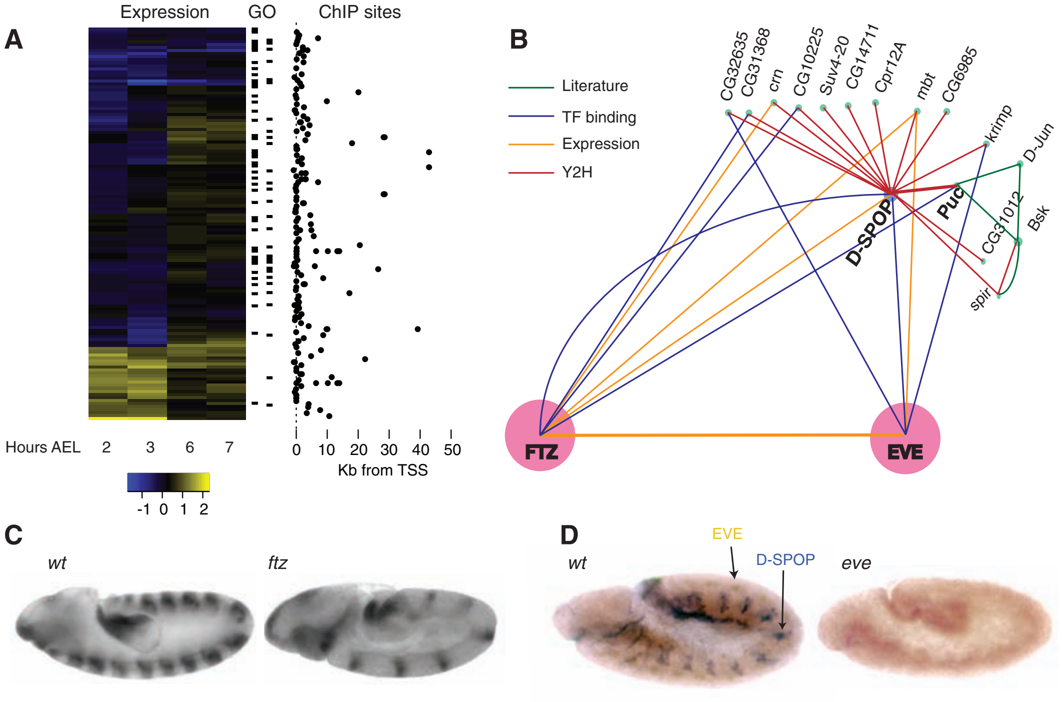


**Fig. 1.** *Drosophila* segmentation network. (**A**) Identification of direct targets of Ftz. (Left) Heat map depicts $\log_2$-fold change in gene expression of mutant versus wild type. Columns represent time points in hours after egg laying (AEL). Rows depict individual genes, sorted by hierarchical clustering. (Center) Genes involved in regulation of development (left tick marks) or transcriptional regulation (right tick marks) from gene ontology (GO) annotations. (Right) Locations of binding sites relative to the transcription start site of each gene. (**B**) Detailed D-SPOP subnetwork showing that D-SPOP is regulated by FTZ and EVE and interacts with puc. (**C**) RNA in situ hybridization of D-SPOP in wild-type and *ftz* mutant backgrounds. Fourteen and 7 stripes of expression were observed in wild-type and *ftz* embryos, respectively. (**D**) Staining with antibodies specific for D-SPOP and EVE in wild-type and *eve* mutant backgrounds. D-SPOP [BCIP-NBT (bromochloroindolyl phosphate–nitro blue tetrazolium), blue] was stained in segmental grooves in wild type, but all expression was lost in *eve* mutant embryos. Eve (diaminobenzidine, brown) was stained in neurons.

---

[1] Institute for Genomics and Systems Biology, The University of Chicago and Argonne National Laboratory, Chicago, IL 60637, USA. [2] Department of Human Genetics, The University of Chicago, Chicago, IL 60637, USA. [3] Howard Hughes Medical Institute, Department of Genetics, Yale University School of Medicine, New Haven, CT 06519, USA. [4] Department of Medicine, The University of Chicago, Chicago, IL 60637, USA. [5] Department of Pathology, Yale University School of Medicine, New Haven, CT 06520, USA. [6] Department of Ecology and Evolution, The University of Chicago, Chicago, IL 60637, USA. [7] Department of Pathology, The University of Chicago, Chicago, IL 60637, USA. [8] Department of Anatomy, La Laguna University, La Laguna, 38320 Tenerife, Spain.

*These authors contributed equally to this work.

†Present address: Department of Entomology, The Volcani Center, Bet Dagan, Israel.

‡Present address: Cold Spring Harbor Laboratory, One Bungtown Road, P.O. Box 100, Cold Spring Harbor, NY 11742, USA.

§To whom correspondence should be addressed. E-mail: kpwhite@uchicago.edu

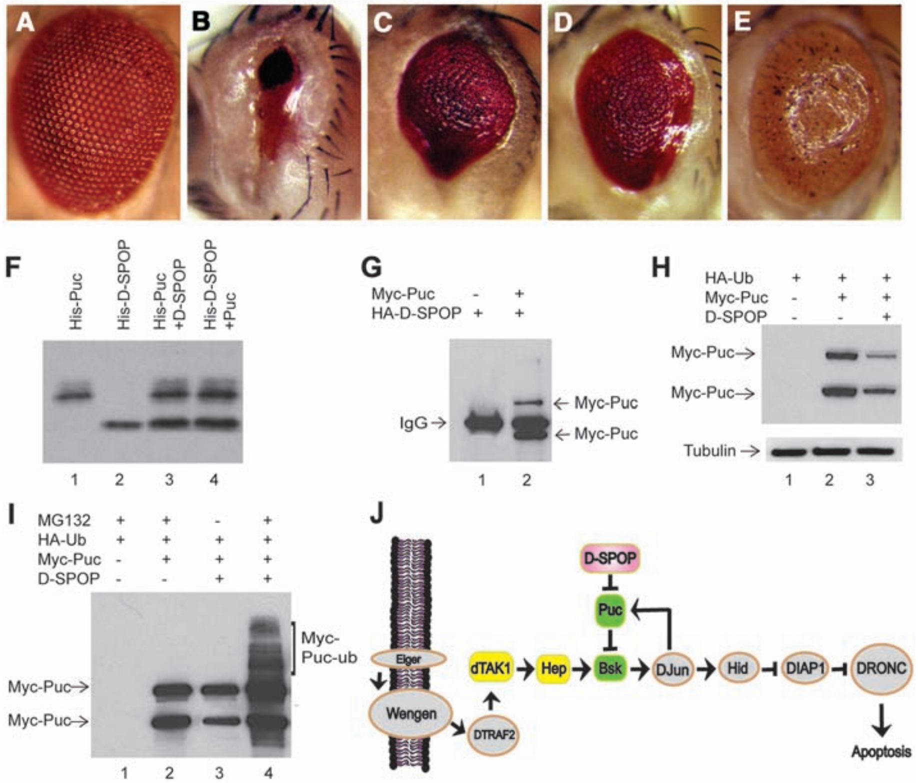


**Fig. 2.** D-SPOP promotes puc ubiquitination and degradation. (**A**) Light micrographs of *Drosophila* adult eyes for wild-type (*GMR-Gal4*/+). (**B**) Egr triggered cell death and produced a small eye phenotype (*GMR-Gal4; UAS-Egr*/+). (**C**) Deleting one copy of *D-SPOP* (*GMR-Gal4; UAS-Egr*/+; *D-SPOP*$\Delta$*6*/+) suppressed the phenotype of (*GMR-Gal4; UAS-Egr*/+). (**D**) Coexpression of *D-SPOP-RNAi* (*GMR-Gal4; UAS-Egr*/+; *UAS-D-SPOP-RNAi*) suppressed the phenotype of (*GMR-Gal4; UAS-Egr*/+). (**E**) Expression of D-SPOP (*GMR-Gal4; UAS-D-SPOP*/+) produced rough eyes with slightly reduced size. (**F**) D-SPOP interacts with Puc in an in vitro affinity assay. Proteins were translated in vitro, purified on a nickel-ion ($Ni^{2+}$) column, and detected by Western blot (see supplementary methods). Lane 1, His-Puc; lane 2, His-D-SPOP; lane 3, untagged D-SPOP copurifies with His-Puc; lane 4, untagged Puc copurifies with His-D-SPOP. (**G**) D-SPOP interacts with Puc in an in vivo immunopreciptation (IP) assay. IPs were carried out with an antibody to hemagglutin (anti-HA) using cell lysates from S2 cells expressing the indicated fusion proteins, followed by immunoblot with anti-Myc epitope. Lane 1, IP of HA-D-SPOP control transfection shows only the immunoglobulin G (IgG) band; lane 2, IP of HA-D-SPOP in cells cotransfected with Myc-Puc, which is detected as two bands. (**H**) Puc degradation is promoted by SPOP. Western blots were performed with anti-Myc to detect Myc-Puc fusion protein. Lane 1, HA-ubiquitin (HA-Ub) transfection negative control; lane 2, Myc-Puc transfection positive control; lane 3, Myc-Puc and D-SPOP cotransfection shows reduced Myc-Puc abundance relative to lane 2. Tubulin is detected as a loading control. (**I**) In vivo ubiquitination of Puc is promoted by D-SPOP. Myc-Puc was detected by immunoblot with anti-Myc epitope. Lane 1, HA-Ub expression and MG132 treatment as a negative control; lane 2, HA-Ub and Myc-Puc coexpression with MG132 treatment but without D-SPOP; lane 3, Myc-Puc abundance decreased with coexpression of HA-Ub and D-SPOP but without addition of MG132; lane 4, Myc-Puc polyubiquitination in the presence of HA-Ub, D-SPOP, and MG132. (**J**) TNF/Eiger-induced apoptosis pathway. D-SPOP (pink) is downstream of dTAK1 (JNKKK, yellow) and Hep (JNKK, yellow), and upstream of Bsk (DJNK, green) and Puc (MKP, green).

(*9*). Together, these data strongly indicate that *D-SPOP* expression is downstream of the pair-rule genes in the segmentation hierarchy.

Knockdown of *D-SPOP* mRNA expression by RNA interference (RNAi) and P-element insertion mutagenesis of the *D-SPOP* gene resulted in severe and consistent disruption of both the peripheral and the central nervous system (CNS) (fig. S7). Such phenotypes are recapitulated by mutating *ftz* or *eve* and are likely due to mid-embryonic functions of D-SPOP when Ftz and Eve become active in the CNS (*13*, *14*). Furthermore, it was recently demonstrated that the *Drosophila* Eiger/TNF (tumor necrosis factor) pathway regulates embryonic neuroblast division (*15*). Thus, we hypothesized that the function of D-SPOP in nervous tissue development may result from its interaction with Puc, which mediates a feedback loop by negatively regulating *basket* (*Drosophila* JNK) in the *Drosophila* Eiger/TNF pathway (*16*) (Fig. 1B).

In *Drosophila*, ectopic expression of Eiger in neuronal cells in the developing eye induces apoptosis through the JNK pathway, resulting in a reduced adult eye size (Fig. 2, A and B) (*17*). Deletion of one wild-type copy of *D-SPOP* or RNAi knockdown of *D-SPOP* mRNA partially suppresses the eye phenotype of Eiger expression (Fig. 2, C and D). Additionally, ectopic expression of D-SPOP in the developing eyes produces a small and rough eye phenotype (Fig. 2E). Analysis of genetic interactions between the genes encoding D-SPOP and other members of the Eiger-JNK pathway (fig. S8) indicates that D-SPOP is acting downstream of dTAK1 (JNKKK) and Hep (JNKK) and upstream of Bsk (JNK) and Puc. Our experiments therefore indicate that D-SPOP functions as an essential positive regulator for Eiger-triggered apoptosis, consistent with the interaction between D-SPOP and Puc predicted in the Ftz-Eve network model.

A physical interaction between D-SPOP and Puc was confirmed by both in vitro pull-down and in vivo immunoprecipitation assays (Fig. 2, F and G). D-SPOP contains two conserved domains, a MATH domain and a BTB/POZ domain (*18*). MEL-26, the *Caenorhabditis elegans* ortholog of human SPOP, was first identified as a BTB protein that serves as an adaptor of Cul3-based ubiquitin ligase (*18*). Recently, human SPOP has been shown to mediate ubiquitination of death domain–associated protein (Daxx) (*19*), the Polycomb group protein BMI-1, the histone variant MacroH2A (*20*), and the transcription factor Gli (*10*). We found that Puc protein abundance was reduced when coexpressed with D-SPOP in S2 cells (Fig. 2H). Furthermore, D-SPOP promoted Puc ubiquitination in S2 cells treated with the proteasome inhibitor MG132 (Fig. 2I). Together, these results indicate that D-SPOP induces apoptosis in the Eiger/TNF pathway by mediating Puc ubiquitination and degradation (Fig. 2J).

Homologs of several Ftz and Eve targets have been shown to be involved in human cancers (*21*); a large body of experimental and clinical data indicates that defects in ubiquitin signaling pathways have roles in the genesis of different tumor types (*22*); and JNK activation is required for cellular transformation induced by RAS, an oncogene mutated in 30% of human cancers (*23*). To determine whether human SPOP's role in modulating TNF-stimulated JNK signaling is conserved, we treated human embryonic kidney 293 (HEK293) cells overexpressing SPOP with TNF-α, then analyzed the abudance of phosphorylated JNK (P-JNK) and phosphorylated c-Jun (P-c-Jun). Consistent with its role in *Drosophila* as an activator of the pathway, overexpression of SPOP increases the amount of P-JNK and P-c-Jun, indicating conservation of its function in modulating the JNK pathway (Fig. 3A).

To test whether SPOP is associated with human cancers, we screened SPOP protein abundance with tissue microarrays that contained 20 tumors from each of 18 different organs. We found that 85% of renal cell carcinomas (RCCs) showed high expression of SPOP, whereas normal kidney tissue was uniformly negative (Fig. 3B and Table 1). To further investigate the potential

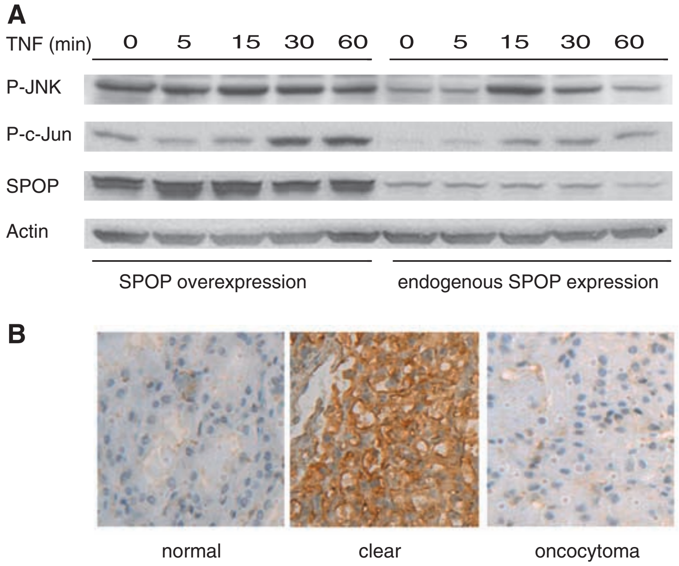


**Fig. 3.** Function of SPOP in the mammalian TNF pathway and overexpression in RCC. (**A**) Overexpression of SPOP increases the abundance of P-JNK and P-c-Jun. HEK293 cells were transfected with SPOP and then treated with TNF (50 ng/ml) at 0, 5, 15, 30, 60 min, followed by immunoblotting. In control treatments, HEK293 cells were transfected with empty pcDNA3.1 vector, and thus only express endogenous amounts of SPOP. (**B**) SPOP expression in RCC tissue. Tissue images are from normal kidney, oncocytoma, and clear cell renal carcinoma. A SPOP-specific monoclonal antibody (SPOP-5G) was used to detect SPOP expression in tissue (diaminobenzidine, brown staining).

**Table 1.** Tissue microarray screening for SPOP expression in 18 cancer types from different organs. SPOP is highly expressed in RCC as compared to normal renal tissue. Endometrial carcinoma and germ-cell tumors also display higher SPOP staining relative to comparable normal tissues.

| Tissue type | Normal tissue | Cancer tissue |
|---|---|---|
| Kidney/renal | 5 negative | 17 positive; 3 negative |
| Uterus/endometrial | 5 negative | 10 positive; 4 negative |
| Testis/germ cell | 5 weak | 18 positive; 2 negative |
| Soft tissue/sarcoma | 5 negative | 20 weak or negative |
| Lung | 5 weak | 20 weak |
| Bladder | 5 weak or negative | 20 weak or negative |
| Breast | 5 weak or negative | 20 weak or negative |
| Prostate | 5 negative | 1 positive; 19 negative |
| Pancreas | 5 negative | 20 negative |
| Lymphoma | 5 negative | 19 weak or negative |
| Thyroid | 5 negative | 20 weak or negative |
| Colon | 5 negative | 20 negative |
| Ovary | 5 weak or negative | 20 weak or negative |
| Skin/melanoma | 5 weak or negative | 20 weak or negative |
| Liver | 5 positive | 7 positive; 13 negative |
| Skin/squamous | 5 weak or negative | 20 weak or negative |
| Brain | 5 weak or negative | 20 weak or negative |
| Stomach/gastric | 5 weak or negative | 20 weak or negative |

**Table 2.** SPOP expression in RCCs. RCC tissue sections were analyzed by staining with SPOP-specific monoclonal antibody (SPOP-5G). Patient samples are classified into different categories depending on cell type.

| Renal cell type | Positive | Negative |
|---|---|---|
| Clear cell | 179 | 1 |
| Chromophobe | 6 | 1 |
| Papillary | 6 | 21 |
| Oncocytic | 2* | 28 |
| Rare type | 6 | 8 |
| Total tumor | 199 | 59 |
| Normal tissue | 0 | 295 |

*Both cases are weak staining.

**Table 3.** SPOP expression in metastatic lesions where RCCs were the primary tumors. Metastatic tissues were analyzed by staining with SPOP-specific monoclonal antibody (SPOP-5G). Patient samples are classified into different categories depending on cell type of the primary tumor.

| Primary RCC type | Positive | Negative |
|---|---|---|
| Clear cell | 71 | 2 |
| Chromophobe | 3 | 1 |
| Papillary | 1 | 2 |
| Rare type | 3 | 4 |

of SPOP as a marker, we designed a large tissue array containing more than 300 RCC samples. Of the tumor samples, 77% were positive for SPOP staining; normal kidney samples were all negative (Table 2).

RCC is a heterogeneous group of tumors with distinct histological subtypes, including clear cell, papillary, chromophobe, and other rare subtypes in addition to oncocytoma, which is a benign solid renal tumor (*24*). Most RCCs (up to 75%) are of the clear cell type and can be subtyped by hematoxylin and eosin staining morphology, but diagnostic difficulties arise when clear cell RCCs display morphologic features that overlap with those of other RCC subtypes and nonrenal tumors (*25*–*27*). At present, a panel of immunohistochemical markers is used to differentiate the major subtypes of RCC in difficult cases (*26*, *27*). Unfortunately, these panels lack a specific and sensitive marker that is positive in clear cell RCC (*26*, *27*). Recently, carbonic anhydrase IX has been proposed as a positive marker for clear cell RCC, but it is positive in other RCCs and several other tumor types as well (*28*–*30*). Patient tumor samples in our studies were classified into different types according to the recent World Health Organization classification system. We found that 99% of the clear cell RCCs and 86% of the chromophobe RCCs showed positive staining for SPOP, but only 22% of papillary-type RCCs were SPOP positive. Four out of 31 papillary RCCs from the general pathology reports were shown to be misdiagnosed as clear cell RCCs when the tumor biopsies were reanalyzed by urological pathologists. All four of these misdiagnosed RCCs have papillary architecture and were subsequently shown to stain positive for SPOP. Our tissue array also included benign oncocytomas, which can mimic RCC both clinically and pathologically, in turn potentially subjecting patients to unnecessary surgeries and additional morbidities. Only 6% of oncocytomas showed weak positive staining. These results indicate that SPOP is a highly sensitive and specific diagnostic biomarker for clear cell RCC and can help distinguish histological subtypes of RCC.

Up to 30% of RCC patients present with metastases; half of the rest will develop metastases later in their course, 90% of which are clear cell RCCs. Accordingly, we further screened for SPOP staining in confirmed metastases from RCC and found that 97% of them were positive (Table 3), indicating that SPOP may be a useful biomarker to identify clear cell RCC as the site of the primary tumors in cases of metastases of unknown origin. Together, our results demonstrate that novel functions for conserved molecules can readily be extracted from data mining of large-scale networks in *Drosophila* and provide a strategy for rapid identification of factors that may have clinical relevance as biomarkers or drug targets for human diseases.

### References and Notes

1. A. Nasiadka, B. H. Dietrich, H. M. Krause, in *Advances in Developmental Biology and Biochemistry*, M. DePamphilis, Ed. (Elsevier, Amsterdam, 2002), vol. 12, p. 155.
2. S. A. Hahn *et al.*, *Science* **271**, 350 (1996).

3. R. L. Johnson *et al.*, *Science* **272**, 1668 (1996).
4. V. Stolc *et al.*, *Science* **306**, 655 (2004).
5. A. Rzhetsky *et al.*, *J. Biomed. Inform.* **37**, 43 (2004).
6. L. Giot *et al.*, *Science* **302**, 1727 (2003).
7. P. Tomancak *et al.*, *Genome Biol.* **8**, R145 (2007).
8. H. Jeong, S. P. Mason, A. L. Barabasi, Z. N. Oltvai, *Nature* **411**, 41 (2001).
9. D. Kent, E. W. Bush, J. E. Hooper, *Development* **133**, 2001 (2006).
10. Q. Zhang *et al.*, *Dev. Cell* **10**, 719 (2006).
11. K. Harding, C. Rushlow, H. J. Doyle, T. Hoey, M. Levine, *Science* **233**, 953 (1986).
12. J. B. Jaynes, M. Fujioka, *Dev. Biol.* **269**, 609 (2004).
13. C. Q. Doe, Y. Hiromi, W. J. Gehring, C. S. Goodman, *Science* **239**, 170 (1988).
14. J. Broadus *et al.*, *Mech. Dev.* **53**, 393 (1995).
15. H. Wang, Y. Cai, W. Chia, X. Yang, *EMBO J.* **25**, 5783 (2006).
16. E. Martin-Blanco *et al.*, *Genes Dev.* **12**, 557 (1998).
17. T. Igaki *et al.*, *EMBO J.* **21**, 3009 (2002).
18. L. Xu *et al.*, *Nature* **425**, 316 (2003).
19. J. E. Kwon *et al.*, *J. Biol. Chem.* **281**, 12664 (2006).
20. I. Hernandez-Munoz *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 7635 (2005).
21. C. J. Potter, G. S. Turenchalk, T. Xu, *Trends Genet.* **16**, 33 (2000).
22. D. Hoeller, C. M. Hecker, I. Dikic, *Nat. Rev. Cancer* **6**, 776 (2006).
23. C. Weiss, D. Bohmann, *Cell Cycle* **3**, 111 (2004).
24. H. T. Cohen, F. J. McGovern, *N. Engl. J. Med.* **353**, 2477 (2005).
25. J. Rosai, *Rosai and Ackerman's Surgical Pathology* (Mosby, New York, ed. 9, 2004).
26. B. F. Skinnider, M. B. Amin, *Semin. Diagn. Pathol.* **22**, 51 (2005).
27. M. Zhou, A. Roma, C. Magi-Galluzzi, *Clin. Lab. Med.* **25**, 247 (2005).
28. H. A. Al-Ahmadie *et al.*, *Am. J. Surg. Pathol.* **32**, 377 (2008).
29. B. C. Leibovich *et al.*, *J. Clin. Oncol.* **25**, 4757 (2007).
30. C. Potter, A. L. Harris, *Cell Cycle* **3**, 164 (2004).
31. We thank J. Jiang, M. Van Lohuizen, C. Chung, D. McEwen for providing expression vectors. Microarray data described in this paper have been deposited in the NCBI Gene Expression Omnibus (GEO) under accession code GSE14086 (expression data) and GSE14289 (ChIP data). M.G. was supported by Vaadia-BARD Postdoctoral Fellowship Award No. FI-315-2001 from BARD, The United States–Israel Binational Agricultural Research and Development Fund. C.D.B was supported by a Lilly Life Science Research Fellowship. This work was supported by grants from the W. M. Keck Foundation, the Arnold and Mabel Beckman Foundation, and the Searle Funds at The Chicago Community Trust from the Chicago Biomedical Consortium to K.P.W.

**Supporting Online Material**
www.sciencemag.org/cgi/content/full/1157669/DC1
Materials and Methods
Figs. S1 to S8
Tables S1 to S8
References

12 March 2008; accepted 14 January 2009
Published online 22 January 2009;
10.1126/science.1157669
Include this information when citing this paper.

---

# In Bad Taste: Evidence for the Oral Origins of Moral Disgust

**H. A. Chapman,[1]* D. A. Kim,[1] J. M. Susskind,[1] A. K. Anderson[1,2]***

In common parlance, moral transgressions "leave a bad taste in the mouth." This metaphor implies a link between moral disgust and more primitive forms of disgust related to toxicity and disease, yet convincing evidence for this relationship is still lacking. We tested directly the primitive oral origins of moral disgust by searching for similarity in the facial motor activity evoked by gustatory distaste (elicited by unpleasant tastes), basic disgust (elicited by photographs of contaminants), and moral disgust (elicited by unfair treatment in an economic game). We found that all three states evoked activation of the levator labii muscle region of the face, characteristic of an oral-nasal rejection response. These results suggest that immorality elicits the same disgust as disease vectors and bad tastes.

Although rationalist theories of moral psychology have long emphasized the role of conscious reasoning in morality (*1*, *2*), recent empirical (*3*–*5*) and theoretical (*6*, *7*) work suggests that emotion may also play a key role in moral judgment. These newer theories make the claim that moral cognition relies primarily on phylogenetically older affective systems, rather than on more recently evolved higher cognitive functions (*6*, *7*). For example, it has been proposed that the violation of moral norms might evoke a kind of moral revulsion or disgust in victims or onlookers (*8*–*10*). Disgust is a somewhat surprising candidate for a moral emotion, given its hypothesized origins in the very concrete, nonsocial, and straightforwardly adaptive functions of rejecting toxic or contaminated food and avoiding disease (*8*). In the moral domain, this rejection impulse might have been co-opted to promote withdrawal from transgressors, or even from the thought of committing a transgression. If the primitive motivational system of disgust is indeed activated by abstract moral transgressions, it would provide strong support for the idea that the human moral sense is built from evolutionarily ancient precursors.

The notion that moral transgressions might evoke the same disgust as potential toxins and disease agents has not gone unchallenged, however. Some have argued that just as a "thirst" for knowledge does not denote a desire to drink, moral "disgust" may reflect not the engagement of more primitive forms of disgust but merely the use of a compelling metaphor for socially offensive behavior (*11*, *12*). As well, prominent theories of disgust have proposed that although moral disgust may be related to contamination-based disgust (typically evoked by potential disease vectors), it is distinct from the most primitive forms of disgust related to the ingestion of potential toxins, having differentiated from the ancient oral distaste response rooted in chemical sensory rejection (*13*). Thus, the "bad taste" of moral disgust may serve as an abstract metaphor rather than reflect a concrete origin in oral distaste.

The evidence that does exist for the specific involvement of disgust in morality is also problematic. Moral transgressions elicit negative emotions (*9*), and induction of negative emotions such as disgust heightens sensitivity to moral transgressions (*5*). However, these studies do not specifically implicate disgust versus other negative emotions such as anger, nor do they demonstrate that moral "disgust" arises from oral disgust. As well, verbal reports of "disgust" in response to moral transgressions are suspect, because the word "disgusting" is used in colloquial English to describe angering or irritating situations (*14*). Thus, verbal self-report measures of subjective experience alone are not diagnostic of disgust. With respect to neural data, moral transgressions sometimes activate the insula (*10*, *15*), which has also been associated with oral disgust (*16*, *17*). However, many affective and cognitive functions besides disgust are associated with activation of the insula, including anger (*18*), anxiety (*19*), general viscerosensory awareness (*19*), and uncertainty (*20*).

The aim of the current series of studies was to provide a more powerful and direct test of the alleged involvement of disgust in morality, and hence of the notion that moral cognition calls on a phylogenetically older motivational system originating in the rejection of hazardous food. We tested the relationship among simple chemosensory distaste, basic forms of contamination-related disgust, and moral disgust by examining subjective experiential reports and objective facial motor activity associated with these states. Recent work supports Darwin's thesis (*21*) that the configuration of emotional facial expressions has evolved from a functional role in regulating sensory intake (*22*). These ancestral configurations may later have proven useful as social signals, assuming a new function without needing to change their basic form (*21*, *22*). Consequently, if moral disgust really is born from the same emotion involved in rejection of hazardous foods, then there should be continuity in facial actions between moral and oral disgust, despite the former being far removed from its purported origin in food rejection. Because moral disgust might result in subtle overt facial movements reflecting the residual engagement of a primitive oral disgust motor program, we recorded facial motor activity with electromyography (EMG), which enables greater sensitivity in detection relative to visual scoring techniques (*23*).

Our first goal was to collect objective measurements of the basic disgust expression with which to compare the moral disgust expression. The most

---

[1] Department of Psychology, University of Toronto, Toronto, Ontario M5S 3G3, Canada. [2] Rotman Research Institute, Baycrest Centre for Geriatric Care, Toronto, Ontario M6A 2E1, Canada.

*To whom correspondence should be addressed. E-mail: hanah@aclab.ca (H.A.C.); anderson@psych.utoronto.ca (A.K.A.)

primitive manifestation of disgust is thought to be distaste, a motivational response to the ingestion of unpleasant-tasting substances, many of which are harmful or toxic (*8*). However, little is known about spontaneous expressions elicited by distaste in adults and their underlying facial motor activity. Therefore, in the first experiment, we recorded facial EMG data while participants drank small samples of unpleasant-tasting bitter, salty, and sour liquids (*24*). A sweet solution of approximately equivalent subjective intensity to the unpleasant solutions was used as a control for nonspecific effects of gustatory stimulation; water was used as a neutral control.

We focused on measuring the activation of the levator labii muscle region of the face, which raises the upper lip and wrinkles the nose (*23*). These movements are thought to be characteristic of the facial expression of disgust (*25*, *26*) and may aid in the function of oral-nasal rejection of aversive chemosensory stimuli (*22*). Consistent with the proposed origin of disgust in distaste, drinking the unpleasant solutions resulted in activation of the levator labii region relative to drinking water or the sweet solution [repeated-measures analysis of variance (ANOVA): $F(2,52) = 8.07$, $P < 0.01$] (Fig. 1A). More specifically, levator labii region activity was greater for unpleasant solutions relative to the sweet solution [paired-samples *t* test: $t(26) = 2.89$, $P < 0.01$]. Levator labii region activity did not reflect a nonspecific response to intense tastes, as the pleasant sweet solution did not evoke significant activity relative to neutral water (paired-samples *t* test: $t < 1$). After each taste trial, participants rated the subjective valence of the preceding sample. These ratings of unpleasantness were highly correlated with the strength of levator labii region activity evoked (Fig. 1B, linear Pearson $r = 0.77$, $P < 0.001$; quadratic $r = 0.93$, $P < 0.001$).

To more accurately visualize the distaste response and the source of the activity in the levator labii region, we recorded on digital video an additional group of participants during ingestion of the solutions. A computerized facial appearance model was constructed to uncover the underlying action tendencies associated with the distaste response. The model represented facial expressions as vectors corresponding to variations from the average face common to face exemplars and then exaggerated these action tendencies by a factor of 2. Figure 1C (right panel) shows the model's depiction of a canonical distaste expression. Consistent with our finding of increased activity in the levator labii region, the upper lip raise and nose wrinkle are apparent, bearing striking similarity in facial actions to the putative disgust expression (*25*). Thus, tasting unpleasant liquids results in facial actions that are the precursors of more elaborated forms of disgust, and these actions may contribute to the adaptive sensory regulatory function of defending the senses against aversive and potentially harmful chemosensory stimuli (*22*).

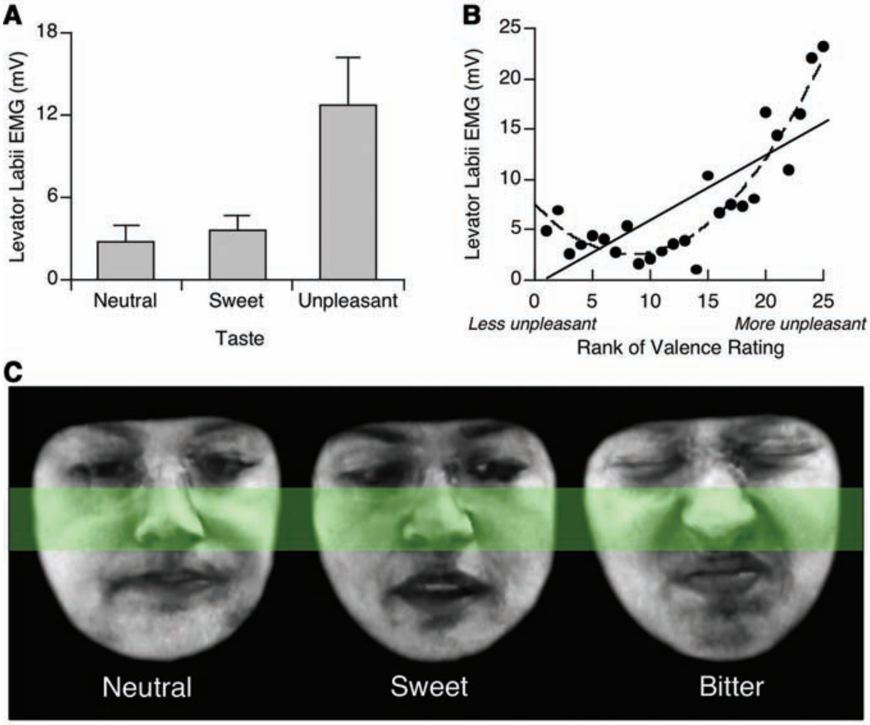


**Fig. 1.** (**A**) Mean levator labii region EMG response evoked by ingestion of neutral, sweet, and unpleasant liquids ($N = 27$). Error bars are +1 SEM [within-subjects (*31*)]. (**B**) Correlation between valence ratings and levator labii region EMG response. For each participant, valence ratings and the corresponding EMG responses for all trials were rank-ordered by decreasing valence. The EMG responses at each rank were then averaged across participants. Points on the plot show this average EMG response by rank; higher rank indicates greater unpleasantness. Linear (solid line) and quadratic (dashed line) fits are shown. (**C**) Appearance model–generated average facial expressions of the five most expressive individuals (from a total sample of 20) tasting neutral, sweet, and bitter solutions. The upper lip and nose areas are highlighted to show the action of the levator labii muscle (upper lip raise and nose wrinkle) across conditions.

We next moved beyond simple gustatory stimuli to examine the facial movements associated with more conceptual, but still relatively concrete, forms of disgust. We recorded EMG data from the levator labii region while participants viewed photographs of uncleanliness and contamination-related disgust stimuli, including feces, injuries, insects, etc. Sad photographs of equivalent negativity were used as a control for nonspecific effects of negative emotional arousal, and neutral photographs were also presented. All photographs were selected from the International Affective Picture System (*24*). Only the disgusting photographs resulted in increased activation of the levator labii region [repeated-measures ANOVA: $F(2,34) = 8.58$, $P < 0.001$] (Fig. 2A). Disgusting photographs resulted in significantly greater levator labii region activity than sad photographs [paired-samples *t* test: $t(17) = 3.71$, $P < 0.01$], whereas sad photographs did not differ from neutral photographs (paired-samples *t* test: $t < 1$). After viewing each photograph, participants rated their feelings of disgust and sadness. Subjective ratings of disgust were significantly correlated with activation of the levator labii region: The stronger the self-reported experience of disgust in response to a photograph, the more levator labii region activity was evoked (Fig. 2B; linear Pearson $r = 0.52$, $P < 0.001$). A more significant quadratic trend (Pearson $r = 0.80$, $P < 0.001$) suggests that the levator labii region may be most responsive to strong levels of disgust. In contrast to the disgust ratings, sadness ratings did not predict levator labii region activity (Fig. 2C; Pearson $r = 0.19$, $P = 0.15$). Because negative emotional arousal associated with increasing sadness did not correlate with activation of the levator labii region, facial motor activity in this area is not a general response to aversive experience (*25*, *26*). These results indicate that more abstract and complex—but still nonsocial or nonmoral—forms of disgust evoke facial motor activity that is very similar to that evoked by unpleasant tastes.

Having determined that both the primitive distaste response and more complex forms of disgust evoke levator labii region activity that is proportional to the degree of disgust or distaste experienced, we next examined whether the same pattern of results would hold for moral transgressions. Given that fairness is a cornerstone of human morality and sociality (*27*), we examined the facial motor activity associated with violations of the norm of fairness. We used the Ultimatum Game as a model of unfairness in social interactions. In our version of the Ultimatum Game, two players split \$10: The proposer suggests how the money should be split (an "offer"), which the responder can accept or reject. If the responder accepts, the money is split as proposed; if he or she rejects the offer, neither player receives anything. While undergoing EMG recording, participants played 20 rounds of the Ultimatum Game in the role of responder, one with each of 10 human proposers (confederates) and 10 with a computer partner. Participants treated offers from humans and computers almost equivalently, in terms of both behavior and emotional response (*24*), so the data presented below are collapsed across proposer identity. All offers were actually generated by a computer algorithm so as to control the number and size of offers made, which ranged from "fair" (proposing a \$5:\$5 split between proposer and responder) to very "unfair" (proposing a \$9:\$1 split). The EMG signal from the period when the proposer's offer was displayed was used to analyze facial activity associated with varying levels of unfairness.

In addition to measuring facial activity, we assessed subjective experience using a nonverbal self-report method that bypasses linguistic emotion labels, so as to avoid the common confusion between disgust and anger. At the end of each Ultimatum Game round, participants reported their experience by rating how well their feelings about the preceding offer were represented by photographs of seven different canonical emotional facial expressions [disgust, anger, contempt, fear, sadness, surprise, and happiness (*28*)]. To confirm that this self-report method separates disgust and its chemical sensory origins from other emotions, an independent group of observers matched the seven expressions to a variety of written emotion labels. The disgust expression was selected as the best match for disgust-relevant labels such as "tasting something bad" and "smelling something bad" in 85% of responses, whereas the anger expression was chosen in only 4% of responses [$\chi^2(1) = 35.4$, $P < 0.001$] and the contempt expression was never chosen. By contrast, the disgust expression was judged to portray anger-relevant labels such as "frustration" and being "pissed off" in only 12% of responses (table S1).

Participants accepted all fair (\$5:\$5) offers, with rejection rates increasing significantly as offers became increasingly unequal [fig. S1; repeated-measures ANOVA: $F(1,15) = 46.7$, $P < 0.001$], suggesting a motivation to punish unfair proposers even at personal financial cost. Of the seven emotions measured, four tracked the unfairness of offers: Disgust, anger, and sadness endorsement increased, whereas happiness endorsement decreased, as the offers became increasingly unfair [Fig. 3A; repeated-measures ANOVA, main effect of offer: $F(9,135) = 25.2$, $P < 0.001$]. Endorsement of contempt, surprise, and fear did not vary with the unfairness of offers [repeated-measures ANOVA, main effect of offer: $F(3,45) = 1.36$, $P = 0.27$; fig. S2]. In addition to reporting increasing disgust with increasing offer unfairness [focused contrast: $F(1,135) = 64.8$, $P < 0.001$], unfair offers evoked disgust to a greater degree than both anger [focused contrast: $F(1,135) = 25.0$, $P < 0.001$] and sadness [focused contrast: $F(1,135) = 25.0$, $P < 0.001$]. In other words, when participants received unfair offers, they judged their experience as most similar to tasting or smelling something bad.

To provide a visualization of participants' ratings of their internal feeling states, we used the emotion endorsements for \$9:\$1 offers to modify the expression photographs used in the self-report task. Using our facial appearance model, we weighted the vector representations of each expression such that the intensity of each expression matched its endorsement strength. Participants endorsed strong disgust, moderate anger, and mild sadness (Fig. 3C). Only disgust expressions showed levator labii activation (highlighted in green). Moreover, a comparison of the disgust expression to distaste and the six other emotion prototypes showed a significantly positive correlation only with distaste (Pearson $r = 0.58$, $P < 0.001$; fig. S3), which suggests that the emotion endorsed in response to unfair offers was most similar in appearance to that displayed by participants consuming unpleasant tastes in our earlier experiment.

Because the emotional response to unfairness was not characterized by disgust alone, we examined what the total self-reported emotional response to unfairness might look like. We created a blend of the disgust, anger, and sadness expressions that participants rated during the self-report task, weighted according to the strength of emotion endorsement for \$9:\$1 offers. The far right panel of Fig. 3C shows the resulting model of the facial response to unfairness, reflecting a complex blend of multiple emotions. We note that the presence of disgust is more subtle, likely because the presence of other emotions dilutes its appearance.

Confirming the subjective reports of disgust, levator labii region EMG was significantly affected by the type of offer presented [repeated-measures ANOVA: $F(3,45) = 3.51$, $P < 0.05$]. Specifically, the increase in self-reported disgust was paralleled by a parametric increase in levator labii region activity as offers became more unfair [Fig. 3B; linear contrast: $F(1,45) = 6.34$, $P < 0.05$]. Focused contrasts revealed that levator labii region activity was greater for \$9:\$1 offers (which were most often rejected; mean of 73% rejection) relative to \$5:\$5 and \$7:\$3 offers (which were most often accepted; mean of 10% rejection) [$F(1,45) = 9.32$, $P < 0.004$]. The decrease in levator labii region activity for \$7:\$3 offers relative to \$5:\$5 was not statistically reliable [focused contrast: $F(1,45) = 1.11$, $P = 0.29$].

The association between self-reported disgust and levator labii activity received further support from a significant correlation between disgust experience and the strength of levator labii region activity. Offers rated as more disgusting were associated with more activation of the levator labii region (Fig. 3D; Pearson $r = 0.61$, $P < 0.01$). Note that the period of EMG analysis preceded the viewing of facial expressions during self-report to ensure the independence of these measures. Although anger and sadness endorsement also increased with increasing unfairness, these ratings did not correlate with levator labii region activity (Fig. 3, E and F; anger, Pearson $r = 0.14$, $P > 0.5$; sadness, Pearson $r = 0.052$, $P > 0.8$). Contempt, another emotion that has been theoretically linked to immorality (*9*), also did not correlate with activation of the levator labii region (Pearson $r = 0.26$, $P > 0.2$). Levator labii region activity was thus specifically related to feelings of tasting or smelling something bad.

In sum, participants showed both subjective (self-report) and objective (facial motor) signs of disgust that were proportional to the degree of unfairness they experienced. These results bear a strong resemblance to the findings of the first two experiments, suggesting that moral transgressions trigger facial motor activity that is also evoked by distasteful and basic disgust stimuli, even though the "bad taste" left by immorality is abstract rather

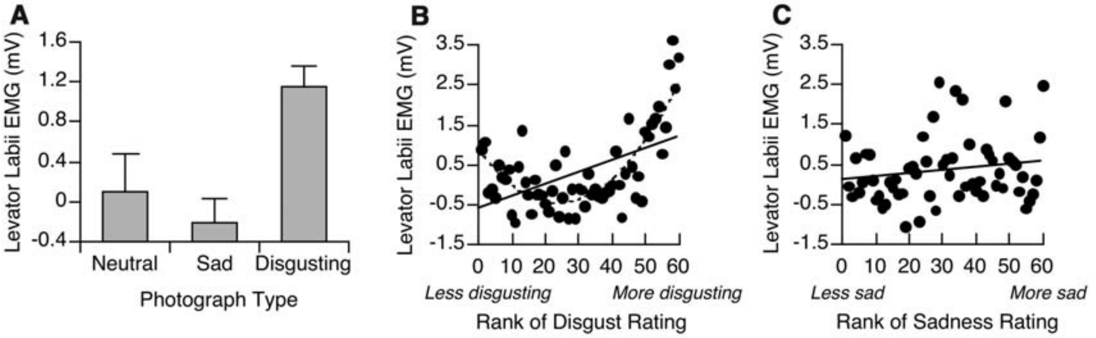


**Fig. 2.** (**A**) Mean levator labii region EMG response evoked by viewing neutral, sad, and disgusting photographs ($N = 18$). Error bars are +1 SEM [within-subjects (*31*)]. (**B**) Correlation between disgust ratings and levator labii region EMG response. Disgust ratings and the paired EMG responses for all photographs were rank-ordered for each participant in order of increasing disgust. The EMG responses at each rank were then averaged across participants. Points on the graph show this average EMG response by rank. Higher rank indicates a more disgusting photograph. Linear (solid line) and quadratic (dashed line) fits are shown. (**C**) Correlation between sadness ratings and levator labii region EMG response, showing linear fit. Correlation was calculated as in (B).

than literal. These data provide direct evidence for Darwin's notions regarding the primitive origins of facial expressions (*21*), as well as their exaptation into the social or moral domain (*22*).

Our results support the idea that moral transgressions evoke disgust as well as other negative emotions. However, the importance of these feelings is not yet clear: What effect do they have on decision-making, if any? One possibility is that negative emotions accompany unfairness but are irrelevant to the decision to reject unfair offers. However, contrary to this interpretation, we found that self-reported disgust was strongly correlated with the decision to reject unfair offers: The more disgust a participant reported, the more likely he or she was to reject a given unfair offer (Pearson $r = 0.70$, $P < 0.05$). In contrast to disgust, self-reports of anger were moderately correlated with rejection (Pearson $r = 0.58$, $P = 0.078$), whereas sadness was not significantly correlated with rejection (Pearson $r = 0.34$, $P = 0.37$). Thus, in addition to being experientially most salient, feelings of disgust were the strongest predictor of decision-making. Providing further support for an association between disgust and the behavioral response to offers, activation of the levator labii region was correlated with the tendency to reject unfair offers: Unfair offers that were associated with stronger levator labii region activity during offer presentation were more likely to be rejected (Pearson $r = 0.71$, $P < 0.05$). These results suggest that the spoiling of economically rational behavior—rejection of unfair offers (i.e., receiving no money versus some)—is strongly associated with subjective and objective measures of disgust. Although these correlations cannot establish a causal relationship between decision-making and emotionality, they do suggest that emotions evoked in response to unfairness—in particular disgust—are relevant to subsequent behavioral choices.

Taken together, our results provide direct evidence of the primitive oral origins of moral disgust. A facial motor action program evoked by aversive chemical sensory stimulation extends to other concrete forms of disgust related to cleanliness and contamination and is also triggered when the everyday moral code of fairness is violated. Furthermore, subjective feelings of tasting or smelling something bad were evoked in response to unfairness, and, in parallel with disgust-related facial motor activity, predicted increasing rejection of unfair offers. The disgust evoked by moral transgressions thus appears to be similar to that evoked by bad tastes and potential disease agents.

These results are consistent with the idea that in humans, the rejection impulse characteristic of distaste may have been co-opted and expanded to reject offensive stimuli in the social domain (*8*). Although some theories have proposed that moral disgust is reserved for transgressions that are conceptually related to notions of moral contamination or purity, with anger and contempt being the more likely response to violations of individual rights and community norms (*9*), our data suggest that moral disgust may in fact be triggered by a wider range of offenses. The role of disgust in active rejection and distancing could explain why

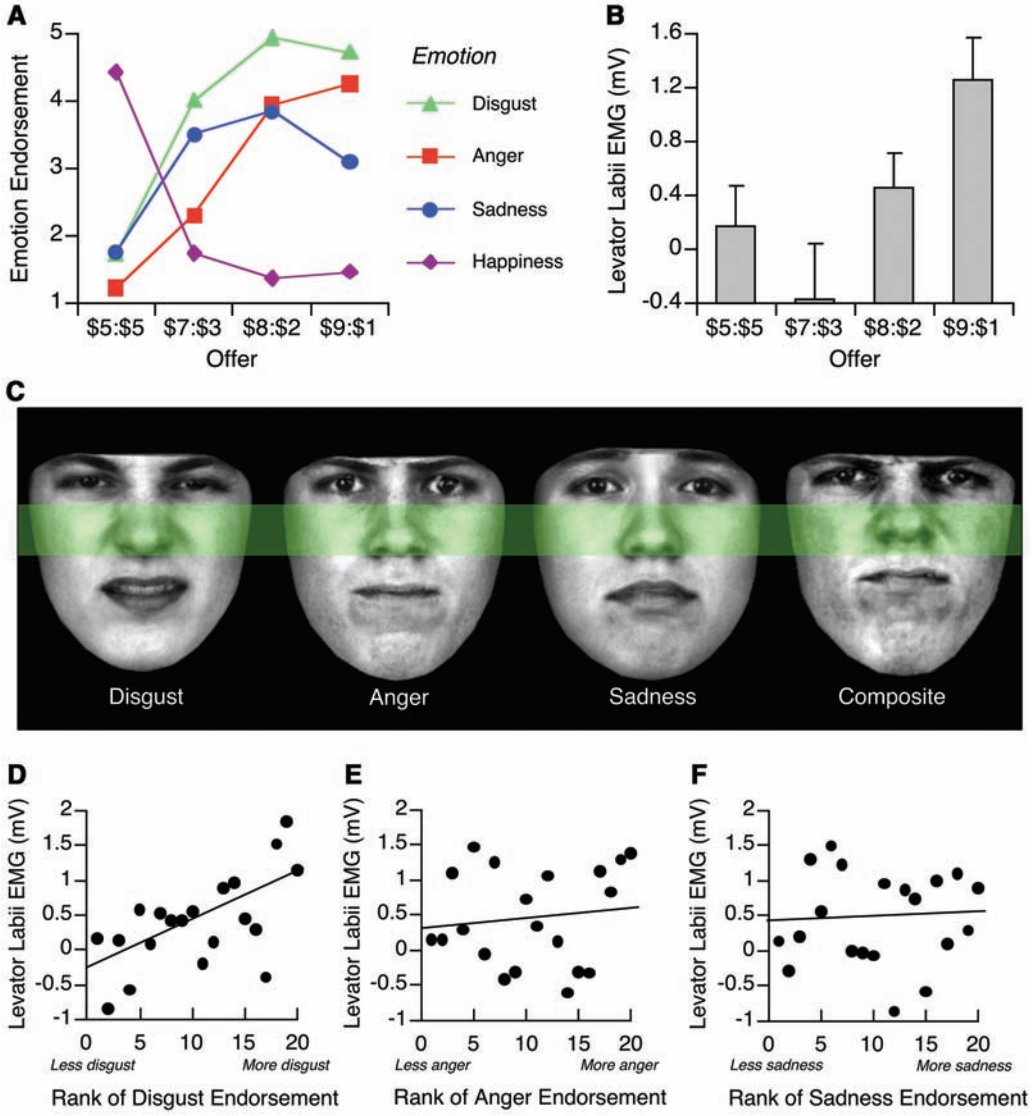


**Fig. 3.** (**A**) Mean self-reported emotion in response to different offers in the Ultimatum Game ($N = 16$). Only emotions that varied with offer type are shown. Higher numbers indicate greater endorsement. (**B**) Mean levator labii region EMG response evoked by the different offers. Error bars are +1 SEM [within-subjects (*31*)]. (**C**) Negative emotions endorsed for $9:$1 offers. Disgust, anger, and sadness expression photographs used in the self-report task were modified using a computer model of facial appearance by weighting the expression intensity by the strength of emotion endorsement; the far right panel shows a composite of these expressions. The upper lip and nose areas are highlighted, showing the action of the levator labii muscle (upper lip raise and nose wrinkle) in the disgust expression and composite. (**D**) Correlation between disgust ratings and levator labii region EMG response. For each participant, disgust ratings and the corresponding EMG responses for all trials were rank-ordered by increasing disgust. The EMG responses at each rank were then averaged across participants. Points on the plot show this average EMG response by rank. Higher rank indicates greater disgust endorsement. Linear fit is shown. (**E**) Correlation between anger ratings and levator labii region EMG response, calculated as in (D). (**F**) Correlation between sadness ratings and levator labii region EMG response, calculated as in (D).

immorality evokes this emotion in addition to others such as anger: Whereas anger is associated with approach motivation (*29*), disgust may motivate vigorous withdrawal (*8*). Thus, unfair offers may be received like a plate of spoiled food. This turning away or rejection of unfair actions may also extend to later avoidance of transgressors.

The ability to detect and avoid toxins appears to be very ancient: Sea anemones, which evolved about 500 million years ago, evert their gastrovascular cavities in response to being fed a bitter substance (*30*). That a system with the ancient and critical adaptive function of rejecting toxic foods should be brought to bear in the moral sphere speaks to the vital importance of regulating social behavior for human beings. Although the stimulus triggers for this rejection mechanism may have shifted far from their chemical sensory origins to the moral domain, the basic behavioral program of oral rejection appears to have been conserved. Thus, the metaphorical "bad taste" left by moral transgressions may genuinely have its origins in oral distaste.

### References and Notes

1. E. Turiel, M. Killen, C. C. Helwig, in *The Emergence of Morality in Young Children*, J. Kagan, S. Lamb, Eds. (Univ. of Chicago Press, Chicago, 1987), pp. 155–243.
2. L. Kohlberg, in *Handbook of Socialization Theory and Research*, D. A. Goslin, Ed. (Rand McNally, Chicago, 1969), pp. 347–380.
3. J. D. Greene, R. B. Sommerville, L. E. Nystrom, J. M. Darley, J. D. Cohen, *Science* **293**, 2105 (2001).
4. M. Koenigs *et al.*, *Nature* **446**, 908 (2007).
5. S. Schnall, J. Haidt, G. L. Clore, A. H. Jordan, *Pers. Soc. Psychol. Bull.* **34**, 1096 (2008).
6. J. Haidt, *Science* **316**, 998 (2007).
7. J. Greene, J. Haidt, *Trends Cogn. Sci.* **6**, 517 (2002).
8. P. Rozin, J. Haidt, C. McCauley, in *Handbook of Emotions*, M. Lewis, J. M. Haviland-Jones, Eds. (Guilford, New York, 2000), pp. 637–653.
9. P. Rozin, L. Lowery, S. Imada, J. Haidt, *J. Pers. Soc. Psychol.* **76**, 574 (1999).
10. A. G. Sanfey, J. K. Rilling, J. A. Aronson, L. E. Nystrom, J. D. Cohen, *Science* **300**, 1755 (2003).
11. P. Bloom, *Descartes' Baby: How the Science of Child Development Explains What Makes Us Human* (Basic Books, New York, 2004).
12. D. Jones, *Nature* **447**, 768 (2007).
13. P. Rozin, L. Lowery, R. Ebert, *J. Pers. Soc. Psychol.* **66**, 870 (1994).
14. R. L. Nabi, *Cogn. Emotion* **16**, 695 (2002).
15. J. Moll *et al.*, *Cogn. Behav. Neurol.* **18**, 68 (2005).
16. M. L. Phillips *et al.*, *Nature* **389**, 495 (1997).
17. A. J. Calder, J. Keane, F. Manes, N. Antoun, A. W. Young, *Nat. Neurosci.* **3**, 1077 (2000).
18. A. R. Damasio *et al.*, *Nat. Neurosci.* **3**, 1049 (2000).
19. H. D. Critchley, S. Wiens, P. Rotshtein, A. Ohman, R. J. Dolan, *Nat. Neurosci.* **7**, 189 (2004).
20. A. Simmons, S. C. Matthews, M. P. Paulus, M. B. Stein, *Neurosci. Lett.* **430**, 92 (2008).
21. C. Darwin, *The Expression of the Emotions in Man and Animals* (HarperCollins, London, 1872/1998).
22. J. M. Susskind *et al.*, *Nat. Neurosci.* **11**, 843 (2008).
23. L. G. Tassinary, J. T. Cacioppo, in *Handbook of Psychophysiology*, J. T. Cacioppo, L. G. Tassinary, G. G. Berntson, Eds. (Cambridge Univ. Press, Cambridge, 2000), pp. 163–198.
24. See supporting material on *Science* Online.
25. P. Ekman, W. Friesen, J. C. Hager, *Facial Action Coding System* (Research Nexus, Salt Lake City, 2002).
26. S. R. Vrana, *Psychophysiology* **30**, 279 (1993).
27. J. P. Henrich, *Foundations of Human Sociality: Economic Experiments and Ethnographic Evidence from Fifteen Small-Scale Societies* (Oxford Univ. Press, Oxford, 2004).
28. D. Matsumoto, P. Ekman, *Japanese and Caucasian Facial Expressions of Emotion (JACFEE)* [Slides] (Intercultural and Emotion Research Laboratory, Department of Psychology, San Francisco State University, 1988).
29. E. Harmon-Jones, *Pers. Individ. Differ.* **35**, 995 (2003).
30. J. Garcia, W. G. Hankins, in *Olfaction and Taste V*, D. A. Denton, J. P. Coghlan, Eds. (Academic Press, New York, 1975), pp. 39–45.
31. G. R. Loftus, M. E. Masson, *Psychon. Bull. Rev.* **1**, 476 (1994).
32. Supported by the Natural Sciences and Engineering Research Council of Canada and by the Canada Research Chairs Program (A.K.A.). We thank G. Cosgrove, S. Couto, R. Landy, A. Meyers, J. Robinson, and M. Sutrisno for assistance in collecting and processing the data, and W. Grabski for technical assistance.

**Supporting Online Material**

www.sciencemag.org/cgi/content/full/323/5918/1222/DC1

Materials and Methods

SOM Text

Figs. S1 to S3

Table S1

References

5 September 2008; accepted 23 December 2008
10.1126/science.1165565

---

# Blue or Red? Exploring the Effect of Color on Cognitive Task Performances

**Ravi Mehta and Rui (Juliet) Zhu***

**Existing research reports inconsistent findings with regard to the effect of color on cognitive task performances. Some research suggests that blue or green leads to better performances than red; other studies record the opposite. Current work reconciles this discrepancy. We demonstrate that red (versus blue) color induces primarily an avoidance (versus approach) motivation (study 1, *n* = 69) and that red enhances performance on a detail-oriented task, whereas blue enhances performance on a creative task (studies 2 and 3, *n* = 208 and 118). Further, we replicate these results in the domains of product design (study 4, *n* = 42) and persuasive message evaluation (study 5, *n* = 161) and show that these effects occur outside of individuals' consciousness (study 6, *n* = 68). We also provide process evidence suggesting that the activation of alternative motivations mediates the effect of color on cognitive task performances.**

Color is a fundamental aspect of human perception, and its effects on cognition and behavior have intrigued generations of researchers. Although a large amount of research has been done in this domain, the psychological processes through which color operates have not been explored fully. As a result, the field has observed certain conflicting results. One inconsistency, which is the focus of this report, concerns the effect of color on cognitive task performance. Most research examining this topic has focused on two of the three primary colors—red versus blue (or green). Some have proposed that red enhances cognitive task performance as compared with blue or green (*1*, *2*); others have shown exactly the opposite (*3*, *4*).

This report details our effort to understand the theory behind the psychological process through which color affects cognitive task performances. Based on our theorizing, we are able to reconcile the above-described inconsistency. We demonstrate that red and blue activate different motivations and consequently enhance performances on different types of cognitive tasks. In line with most of the extant research, we limit our research to the two primary colors, red and blue.

Color theorists believe that color influences cognition and behavior through learned associations (*3*). When people repeatedly encounter situations where different colors are accompanied by particular experiences and/or concepts, they form specific associations to colors. Red and blue have been shown to have different associations within the cognitive domain. Red is often associated with dangers and mistakes [e.g., errors that are circled with a red ink pen, stop signs, and warnings (*3*)]. Claims have been made linking the color red to the highest level of hazard and also the highest level of compliance (*5*, *6*). In contrast, blue is often associated with openness, peace, and tranquility [e.g., ocean and sky (*7*)]. A word association test confirmed that people indeed generate these different associations to red versus blue color in the cognitive task domain (*8*, *9*).

We propose that these different associations related to red versus blue color can induce alternative motivations. Specifically, red, because of its association with dangers and mistakes, should activate an avoidance motivation, which has been shown to make people more vigilant and risk-averse (*10*–*12*). Thus, red, compared with blue, should enhance performance on detail-oriented tasks (i.e., tasks that require focused, careful attention). In contrast, because blue is usually associated with openness, peace, and tranquility, it is likely to activate an approach motivation, because these associations signal a benign environment that encourages people to use innovative as opposed to "tried-and-true" problem-solving strategies (*13*). Indeed, an approach motivation has been shown to make people behave in a more

---

Sauder School of Business, University of British Columbia, 2053 Main Mall, Vancouver, BC V6T 1Z2, Canada.

*To whom correspondence should be addressed. E-mail: juliet.zhu@sauder.ubc.ca

explorative, risky manner (*10*, *11*). Thus, blue versus red should enhance performance on creative tasks.

We report six studies (*14*) that offered systematic support to our hypotheses. Most studies were conducted on computers, and color was manipulated through the background screen color. Color is usually defined along three dimensions: hue (the pigment of the color, e.g., blue, red, etc.), chroma (saturation of color), and value (degree of darkness or lightness of the color) (*15*). In order to reduce confounds and to be consistent with prior research, we manipulated only hue (i.e., red versus blue) and kept chroma and value constant. For computer-based studies, we employed the HSL (hue-saturation-lightness) scheme (red: hue = 0, saturation = 240, lightness = 120; blue: hue = 160, saturation = 240, lightness = 120) (*8*) (table S1). To provide a baseline for comparison, we also included a neutral condition in some studies where computer background color was set to be white.

Study 1 tested our hypothesis that red color will induce primarily an avoidance motivation, whereas blue will activate an approach motivation. Sixty-nine participants were randomly assigned to the red, blue, or neutral background color condition and completed a computer-based study that consisted of two tasks. In the first task, participants solved a series of 12 anagrams, with three of them having target words related to avoidance motivation (e.g., prevent), another three having target words related to approach motivation (e.g., adventure) (*16*, *17*), and the remaining six that were unrelated to either motivation (e.g., computer) (*8*). The response times for each type of correctly solved anagrams were averaged to create three reaction time (RT) indices—avoidance, approach, and neutral RT indices. Faster reaction time to approach-related (or avoidance-related) anagrams would imply a stronger activation of an approach (or avoidance) motivation (*18*). As anticipated (Fig. 1), for approach-related anagrams, those in the blue condition [10.93 ± 5.51 s (mean ± SD)] responded faster than those in the red [18.53 ± 12.25 s; $t(66) = 2.81$, $P < 0.01$; Cohen's $d = 0.81$] or neutral condition [17.50 ± 9.17 s; $t(66) = -2.29$, $P < 0.03$; Cohen's $d = 0.91$]. For avoidance-related anagrams, the reverse pattern appeared, such that those in the red condition (10.40 ± 3.64 s) responded faster than those in the blue [20.39 ± 14.73 s; $t(66) = -3.21$, $P < 0.01$; Cohen's $d = 0.96$] or neutral condition [19.14 ± 11.48 s; $t(66) = -2.67$, $P < 0.01$; Cohen's $d = 1.1$]. For neutral anagrams, however, no effect of color was observed (red, 10.56 ± 5.48 s; blue, 12.64 ± 7.53 s; neutral, 11.58 ± 4.37 s; $t < 1$).

In the second task, participants read descriptions of three pairs of brands and reported their preferences along a scale from 1 (prefer brand A) to 7 (prefer brand B). Within each pair, one brand highlighted a negative outcome people try to avoid, whereas the other brand highlighted a positive outcome people try to approach. For example, one pair featured two brands of toothpastes, with brand A being particularly good for cavity prevention (avoidance focused), and brand B being particularly good for tooth whitening (approach focused) (*8*). Across three pairs, we found that those in the blue color condition (4.03 ± 1.55) indicated greater preference for brands that were approach-oriented (highlighting positive benefits) than those in the red [2.79 ± 1.65; $t(66) = -2.80$, $P < 0.01$; Cohen's $d = 0.79$] and the neutral condition [3.05 ± 1.43; $t(66) = 2.08$, $P < 0.05$; Cohen's $d = 0.67$]. Thus, this study demonstrated that within a cognitive task domain, red (versus blue) can activate an avoidance (versus approach) motivation. A post hoc study ruled out mood as an alternative explanation (*8*).

The next two studies tested whether red (versus blue), because of its activation of avoidance (versus approach) motivation, enhances performance on a detail-oriented (versus a creative) task. Study 2 ($n = 208$) contained two tasks, a detail-oriented and a creative task. A set of participants completed the detailed-oriented task (i.e., a memory exercise) presented on computers with red, blue, or neutral background color. They studied a list of 36 words for 2 min and were asked to recall as many words as they could after a 20-min delay. Three measures confirmed that red indeed enhanced performance on this memory task. Those in the red condition (15.89 ± 5.90) recalled more correct items than those in the blue condition [12.31 ± 5.48; $t(100) = 2.50$, $P < 0.02$; Cohen's $d = 0.64$] (Fig. 2). Furthermore, blue led to more false recalls (0.86 ± 1.29) than red [0.34 ± 0.64; $t(100) = -2.42$, $P < 0.02$; Cohen's $d = 0.52$] or neutral [0.38 ± 0.55; $t(100) = 2.21$, $P < 0.03$; Cohen's $d = 0.48$] condition. These two measures have been shown to reflect people's attention to details (*19*, *20*). However, color manipulation did not affect the total number of items recalled ($P > 0.11$) (*8*).

Another set of participants completed a creative task where they were asked to generate as many creative uses for a brick as they could think of within 1 min (*21*). Consistent with prior research (*21*), each participant's responses were coded into three categories: (i) total number of uses generated, (ii) mean creativity score as rated by a panel of judges, and (iii) total number of creative uses. Participants in the three color conditions produced equal number of uses in total ($F < 1$; red, 4.83 ± 2.31; blue, 4.67 ± 2.62; neutral, 4.94 ± 1.68). However, the quality of these uses differed by color conditions. Those in the blue condition (3.97 ± 0.99) demonstrated a higher mean creativity score than those in the red [3.39 ± 0.97; $t(102) = -2.81$, $P < 0.01$; Cohen's $d = 0.6$] or neutral color condition [3.50 ± 0.63; $t(102) = 2.23$, $P < 0.03$; Cohen's $d = 0.57$] (Fig. 3). Similarly, those in the blue (1.64 ± 1.46) condition produced more creative uses than those in the red [0.86 ± 0.97; $t(102) = -2.93$, $P < 0.01$; Cohen's $d = 0.64$] or neutral condition [0.91 ± 0.83; $t(102) = 2.70$, $P < 0.01$; Cohen's $d = 0.62$] (*8*). Findings from this study suggest that, although color did not affect the amount of processing, as shown in the equal recall level for the memory task and comparable uses generated for the brick task, it affected the quality of responses, i.e., red led to superior performances on detail-oriented tasks and blue, on creative tasks.

Study 3 ($n = 118$) tested the generalizability of results observed in study 2 by using two different tasks. Moreover, it aimed to demonstrate that the activation of avoidance or approach motivations is the underlying force that drives our results. As in study 2, this study was computer-based, and color was manipulated using the computer background screen color. The detail-oriented task in this study was a proofreading task (*22*). Participants examined five sets of items, with each set containing a pair of names or addresses, which were either identical or slightly different (*8*). Participants' task was to judge whether items within each pair were identical or not. To assess whether color-induced motivations drive our expected effects, we also asked participants to answer three questions concerning the extent to which they focused on accuracy (mistake-avoidance motivation) versus speed (approach motivation). Results revealed that red color condition (4.33 ± 0.77) led to more correct responses than blue [3.53 ± 0.80;

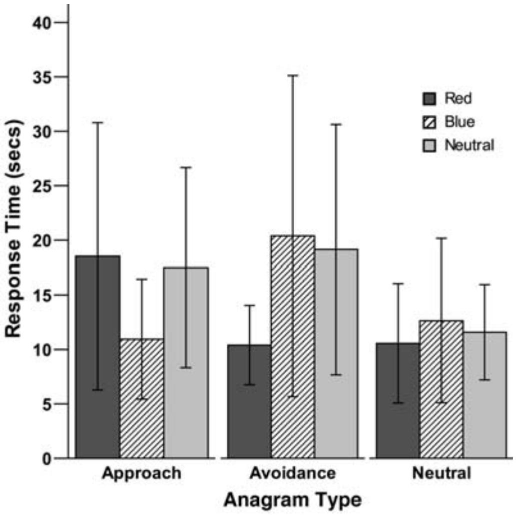


**Fig. 1.** Participants' response times to approach-related, avoidance-related, and neutral anagrams under red, blue, and neutral color conditions (study 1). $F_{2,66} = 8.79$, $P < 0.001$. Error bars, ±1.00 SD.

$t(51) = 2.49$, $P < 0.02$; Cohen's $d = 1.05$] or neutral [$3.68 \pm 1.20$; $t(51) = 2.07$, $P < 0.05$; Cohen's $d = 0.66$] condition ($F_{2,51} = 3.56$, $P < 0.04$). Further, mediation analysis (*23*) revealed that approach versus avoidance motivations were indeed the driving force for the observed effect.

The Remote Associates Test (RAT), which is widely used as a test of creative thinking (*24*), was used as the creative task. Each RAT item consists of three or four stimulus words (e.g., "Shelf," "Read," and "End") that are in some way related to a fourth or fifth unreported word (e.g., "Book"). Participants were presented with five RAT items and were asked to determine what the target words were. As predicted, those in the blue condition ($4.00 \pm 0.74$) produced more correct answers (thus exhibiting higher creativity) than those in the red [$3.45 \pm 0.89$; $t(61) = -2.35$, $P < 0.03$; Cohen's $d = 0.69$] or neutral [$3.38 \pm 0.67$; $t(61) = 2.67$, $P < 0.01$; Cohen's $d = 0.9$] condition ($F_{2,61} = 4.33$, $P < 0.02$). Mediation analysis again confirmed that the alternative motivations activated by color drive the observed effect (*8*).

Study 4 aimed to further extend the previous studies by using a single task that could examine both people's creativity level and their attention to details. For this purpose, participants were presented with a sheet of paper with drawings of 20 different parts (fig. S1). Participants ($n = 42$) were required to use any five parts and draw a design of a toy a child (age 5 to 11) could use to play with (*8*, *25*). Unlike previous studies, the color manipulation in this study was done by presenting the 20 parts either in red or blue color. Twelve judges evaluated black-and-white copies of each design on two dimensions, one assessing the originality and novelty (reflecting creativity level) and the other assessing the practicality and appropriateness (reflecting attention to details) of the design. The correlation between these two dimensions was 0.29 ($P = 0.08$) and interjudge reliability was 0.75 for originality and 0.83 for appropriateness. As expected, toys designed by those in the red color condition were judged to be more practical and appropriate ($3.47 \pm 0.79$) than those in the blue condition ($2.95 \pm 0.88$; $F_{1,40} = 4.16$, $P < 0.05$; Cohen's $d = 0.64$), but were judged less original and novel ($2.94 \pm 0.55$) than those in the blue condition ($3.37 \pm 0.76$; $F_{1,40} = 4.46$, $P < 0.05$; Cohen's $d = 0.67$). A set of anagrams similar to those described in study 1 were included in this study to test whether avoidance or approach motivations were the driving force for the effect. Results revealed that those in the red versus blue condition responded faster to the avoidance-related anagrams (which indicated an activation of an avoidance motivation) and, subsequently, exhibited a higher score on practicality and appropriateness. In contrast, those in the blue versus red condition responded faster to the approach-related anagrams (which indicated an activation of an approach motivation) and, subsequently, exhibited a higher score on originality and novelty.

Study 5 tested our theorizing in yet another domain, namely persuasive message evaluation. Participants ($n = 161$) evaluated one of two versions of an advertisement for a camera on a computer screen with the background color set to be either red or blue. The two versions of the advertisement were identical except for the visuals (*26*). In one version, the visuals represented specific product details of the camera (e.g., lens) and thus fitted a detail-oriented processing style (*27*). We expected that red, which enhances attention to details, would lead to higher persuasion for this version. The other version included visuals that represented rather remotely related associations (e.g., a road sign, a dining table in a restaurant, and a map), which would require creative thinking to connect all these images to a camera-related theme, e.g., travel (fig. S2). Thus, we expected that blue, which appears to enhance creative cognition, would lead to more persuasion for this version (*27*). Participants evaluated one of the ads on three seven-point items assessing its appeal, favorability, and effectiveness. For the red color computer background screen, participants formed more favorable evaluations when the ad included visuals representing specific product details ($4.69 \pm 1.26$) as opposed to remotely related associations ($4.11 \pm 1.28$; $F_{1,150} = 3.41$, $P < 0.07$; Cohen's $d = 0.42$). In contrast, when the background color was blue, the reverse occurred, i.e., more favorable evaluations emerged when the ad contained visuals representing remotely related associations ($4.41 \pm 1.47$) versus specific product details ($3.60 \pm 1.59$; $F_{1,150} = 6.01$, $P < 0.02$; Cohen's $d = 0.56$) (*8*).

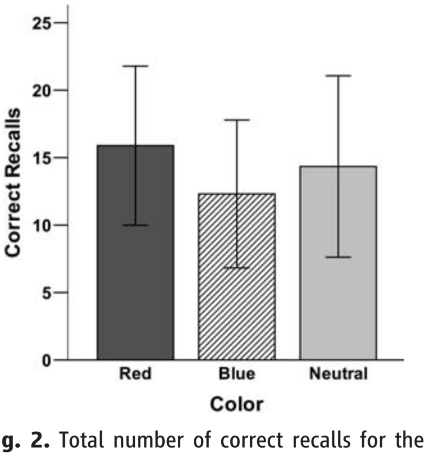


**Fig. 2.** Total number of correct recalls for the memory task (study 2). $F_{2,100} = 3.15$, $P < 0.05$. Error bars, ±1.00 SD.

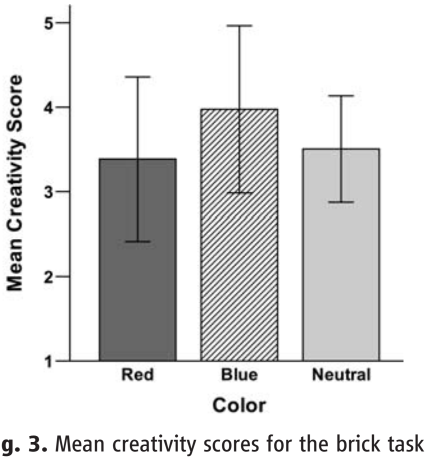


**Fig. 3.** Mean creativity scores for the brick task (study 2). $F_{2,102} = 4.43$, $P < 0.02$. Error bars, ±1.00 SD.

In the last study, we tested whether people are aware of the differential effects of red versus blue color. No color manipulation was done for this study, and all the instructions and the focal task were presented in black color font with white background screen color on computers. Participants ($n = 68$) were told that one of these tasks they would complete requires detailed, careful, and systematic processing of information, and it could be presented to them with either a red or a blue background color. Participants' task was to select one color that they thought would enhance their performance on that task. A sample of the red and blue colors was presented (fig. S3). On the next screen, participants were told that another task in this study would require creative, imaginative, and outside-of-the-box thinking and were asked to select one of the two colors that they thought might enhance their performance on the creative task.

The data revealed that significantly more participants chose the blue (66%) versus red (34%) color when the task was described to be creative [$\chi^2\ (1) = 7.12$, $P < 0.01$]. However, interestingly, the same pattern of results emerged when the task was described to be detail-oriented, i.e., more people thought the blue (74%) versus red (26%) background color would enhance their performance even on the detail-oriented task [$\chi^2\ (1) = 15.06$, $P < 0.001$] (*8*). These results are consistent with the general belief that people have an overall preference for blue versus red color, although we found that red can be beneficial when the focal task requires detailed attention.

From a series of six studies, using various tasks covering a number of different domains, we demonstrate that red (versus blue) can activate an avoidance (versus approach) motivation and subsequently can enhance performance on detail-oriented (versus creative) cognitive tasks. This research thus offers a reconciliation of the conflicting results reported in the extant literature and advances current research on the effect of color on cognition and behavior [e.g., (*3*)]. More important, our findings offer a wide range of implications for daily human life. What wall color do we pick for an educational facility? What color enhances persuasion in a consumption context? What color enhances creativity in a new product design process? Results from this research suggest that, depending on the nature of the task, different colors might be beneficial. If the task on hand requires people's vigilant attention (e.g., memorizing important information or understanding the side effects of a new drug), then red (or another color that activates an avoidance motivation) might be particularly appropriate. However, if the task calls for creativity and imagination (e.g., designing an art shop, or a new product idea brainstorming session), then blue (or another color that activates an approach motivation) would be more beneficial.

### References and Notes

1. N. Kwallek, C. M. Lewis, *Appl. Ergon.* **21**, 275 (1990).
2. N. J. Stone, *J. Environ. Psychol.* **23**, 63 (2003).
3. A. J. Elliot, M. A. Maier, A. C. Moller, R. Friedman, J. Meinhardt, *J. Exp. Psychol. Gen.* **136**, 154 (2007).
4. S. Soldat, R. C. Sinclair, M. M. Mark, *Soc. Cogn.* **15**, 55 (1997).
5. C. C. Braun, N. C. Silver, *Ergonomics* **38**, 2207 (1995).
6. D. J. Williams, J. M. Noyes, *Theor. Issues Ergon. Sci.* **8**, 1 (2007).
7. N. Kaya, H. H. Epps, *Coll. Stud. J.* **38**, 396 (2004).
8. Materials and methods are available as supporting material on *Science* Online.
9. We recognize that red and blue can also imply other associations. For example, red might connote excitement or femininity, whereas blue might suggest sadness. Although we acknowledge these different associations, we simply suggest that in the cognitive task domain, red is predominately associated with dangers and mistakes and blue with openness and peace. In addition, we note that the same color may have different associations across cultures (*28*). All our studies were run in a North American university. Thus, future research should test whether our results can be generalized to other cultures.
10. R. S. Friedman, J. Förster, *J. Exp. Soc. Psychol.* **38**, 41 (2002).
11. R. S. Friedman, J. Förster, *J. Pers. Soc. Psychol.* **88**, 263 (2005).
12. S. Koch, R. W. Holland, A. van Knippenberg, *Cognition* **109**, 133 (2008).
13. N. Schwarz, in *Handbook of Motivation and Cognition*, E. T. Higgins, R. M. Sorrentino, Eds. (Guilford Press, New York, 1990), pp. 527–561.
14. All studies were run at the University of British Columbia, where students participated in exchange for either course credit (studies 1, 2, 4, 5, and 6) or money (color association study, study 3). The majority of the participants in our studies were raised in Canada, mainland China, and Hong Kong (in decreasing order of prevalence). The participants' ages ranged between 17 and 39 years.
15. E. Thompson, A. Palacios, F. J. Varela, *Behav. Brain Sci.* **15**, 1 (1992).
16. E. T. Higgins, *Am. Psychol.* **52**, 1280 (1997).
17. E. T. Higgins, in *Advances in Experimental Social Psychology*, M. P. Zanna, Ed. (Academic Press, New York, 1998), pp. 1–46.
18. A. J. Elliot, M. V. Coviogton, *Educ. Psychol. Rev.* **13**, 73 (2001).
19. F. I. Craik, J. M. McDowd, *J. Exp. Psychol. Learn. Mem. Cogn.* **13**, 474 (1987).
20. E. Crowe, E. T. Higgins, *Organ. Behav. Hum. Decis. Process.* **69**, 117 (1997).
21. R. S. Friedman, J. Förster, *J. Pers. Soc. Psychol.* **81**, 1001 (2001).
22. S. Kaplan, *J. Environ. Psychol.* **15**, 169 (1995).
23. R. M. Baron, D. A. Kenny, *J. Pers. Soc. Psychol.* **51**, 1173 (1986).
24. V. Griskevicius, R. B. Cialdini, D. T. Kenrick, *J. Pers. Soc. Psychol.* **91**, 63 (2006).
25. C. P. Moreau, D. W. Dahl, *J. Consum. Res.* **32**, 13 (2005).
26. R. J. Zhu, J. Meyers-Levy, *J. Consum. Res.* **34**, 89 (2007).
27. E. T. Higgins, L. C. Idson, A. L. Freitas, S. Spiegel, D. C. Molden, *J. Pers. Soc. Psychol.* **84**, 1140 (2003).
28. M. M. Aslam, *J. Mark. Commun.* **12**, 15 (2006).

29. This work benefited from financial support from the Social Sciences and Humanities Research Council to R.J.Z. We thank J. Meyers-Levy and K. D. Vohs for very helpful comments on earlier drafts of this paper. Research assistance from P. Behmardi, S. Ho, and S. Park is greatly appreciated.

**Supporting Online Material**

www.sciencemag.org/cgi/content/full/1169144/DC1
Materials and Methods
Figs. S1 to S3
Tables S1 to S3
References

28 November 2008; accepted 5 January 2009
Published online 5 February 2009;
10.1126/science.1169144
Include this information when citing this paper.

---

# Self-Sustained Replication of an RNA Enzyme

**Tracey A. Lincoln and Gerald F. Joyce***

**An RNA enzyme that catalyzes the RNA-templated joining of RNA was converted to a format whereby two enzymes catalyze each other's synthesis from a total of four oligonucleotide substrates. These cross-replicating RNA enzymes undergo self-sustained exponential amplification in the absence of proteins or other biological materials. Amplification occurs with a doubling time of about 1 hour and can be continued indefinitely. Populations of various cross-replicating enzymes were constructed and allowed to compete for a common pool of substrates, during which recombinant replicators arose and grew to dominate the population. These replicating RNA enzymes can serve as an experimental model of a genetic system. Many such model systems could be constructed, allowing different selective outcomes to be related to the underlying properties of the genetic system.**

A long-standing research goal has been to devise a nonbiological system that undergoes replication in a self-sustained manner, brought about by enzymatic machinery that is part of the system being replicated. One way to realize this goal, inspired by the notion of primitive RNA-based life, would be for an RNA enzyme to catalyze the replication of RNA molecules, including the RNA enzyme itself (*1*–*4*). This has now been achieved in a cross-catalytic system involving two RNA enzymes that catalyze each other's synthesis from a total of four component substrates.

The "R3C" RNA enzyme is an RNA ligase that binds two oligonucleotide substrates through Watson-Crick pairing and catalyzes nucleophilic attack of the 3′-hydroxyl of one substrate on the 5′-triphosphate of the other, forming a 3′,5′-phosphodiester and releasing inorganic pyrophosphate (*5*). The R3C ligase was configured to self-replicate by joining two RNA molecules to produce another copy of itself (*6*). This process was inefficient because the substrates formed a nonproductive complex that limited the extent of exponential growth, with a doubling time of about 17 hours and no more than two successive doublings.

The R3C ligase subsequently was converted to a cross-catalytic format (Fig. 1A), whereby a plus-strand RNA enzyme (E) catalyzes the joining of two substrates (A′ and B′) to form a minus-strand enzyme (E′), which in turn catalyzes the joining of two substrates (A and B) to form a new plus-strand enzyme (*7*, *8*). This too was inefficient because of the formation of nonproductive complexes and the slow underlying rate of the two enzymes. The enzymes E and E′ operate with a rate constant of only ~0.03 $\text{min}^{-1}$ and a maximum extent of only 10 to 20% (*9*). These rates are about 10 times slower than that of the parental R3C ligase (*5*), and when the two cross-catalytic reactions are carried out within a common mixture the rates are even slower (*7*).

The catalytic properties of the cross-replicating RNA enzymes were improved by the use of in vitro evolution, optimizing the two component reactions in parallel and seeking solutions that would apply to both reactions when conducted in the cross-catalytic format (*9*). The 5′-triphosphate–bearing substrate was joined to the enzyme via a hairpin loop (B′ to E and B to E′), and nucleotides within both the enzyme and the separate 3′-hydroxyl–bearing substrate (A′ and A) were randomized at a frequency of 12% per position. The two resulting populations of molecules were subjected to six rounds of stringent in vitro selection, selecting for their ability to react in progressively shorter times, ranging from 2 hours to 10 ms. Mutagenic polymerase chain reaction was performed after the third round to maintain diversity in the population. After the sixth round, individuals were cloned from both populations and sequenced. There was substantial sequence variability among the clones, but all contained mutations just upstream from the ligation junction that resulted in a G•U wobble pair at this position.

The G•U pair was installed in both enzymes and both 3′-hydroxyl–bearing substrates (Fig. 1B). In the trimolecular reaction (with two separate substrates), the optimized enzymes E and E′ exhibited a rate constant of 1.3 and 0.3 $\text{min}^{-1}$ with a maximum extent of 92 and 88%, respectively. The optimized enzymes underwent robust exponential amplification at a constant temperature of 42°C, with more than 25-fold amplification after 5 hours, followed by a leveling off as the supply of substrates became depleted (Fig. 2A). The data fit well to the logistic growth equation $[\mathrm{E}]_t = a/(1 + be^{-ct})$, where $[\mathrm{E}]_t$ is the concentration of E (or E′) at time $t$, $a$ is the maximum extent of growth, $b$ is the degree of sigmoidicity, and $c$ is the exponential growth

---

Department of Chemistry, Department of Molecular Biology, and the Skaggs Institute for Chemical Biology, The Scripps Research Institute, La Jolla, CA 92037, USA.

*To whom correspondence should be addressed. E-mail: gjoyce@scripps.edu

rate. For the enzymes E and E′, the exponential growth rate was 0.92 and 1.05 hour$^{-1}$, respectively.

Exponential growth can be continued indefinitely in a serial transfer experiment in which a portion of a completed reaction mixture is transferred to a new reaction vessel that contains a fresh supply of substrates. Six successive reactions were carried out in this fashion, each 5 hours in duration and transferring 4% of the material from one reaction mixture to the next. The first mixture contained 0.1 μM E and 0.1 μM E′, but all subsequent mixtures contained only those enzymes that were carried over in the transfer. Exponential growth was maintained throughout 30 hours total of incubation, with an overall amplification of greater than $10^8$-fold for each of the two enzymes (Fig. 2B).

It is possible to construct variants of the cross-replicating RNA enzymes that differ in the regions of Watson-Crick pairing between the cross-catalytic partners without markedly affecting replication efficiency. These regions are located at the 5′ and 3′ ends of the enzyme (Fig. 1B). Four nucleotide positions at both the 5′ and 3′ ends were varied, adopting the rule that each region contains one G•C and three A•U pairs so that there would be no substantial differences in base-pairing stability. Of the 32 possible pairs of complementary sequences for each region, 12 were chosen as a set of designated pairings (Fig. 1C). Each pairing was associated with a particular sequence within the catalytic core of both members of a cross-replicating pair. Twelve pairs of cross-replicating enzymes were synthesized, as well as the 48 substrates (12 each of A, A′, B, and B′) necessary to support their exponential amplification. Each replicator was individually tested and demonstrated varying levels of catalytic activity and varying rates of exponential growth (fig. S1). The pair shown in Fig. 1B (now designated E1 and E1′) had the fastest rate of exponential growth, achieving about 20-fold amplification after 5 hours. The various cross-replicating enzymes shown in Fig. 1C had the following rank order of replication efficiency: E1, E10, E5, E4, E6, E3, E12, E7, E9, E8, E2, E11. The top five replicators all achieved more than 10-fold amplification after 5 hours, and all except E11 achieved at least fivefold amplification after 5 hours.

A serial-transfer experiment was initiated with a 0.1 μM concentration each of E1 to E4 and E1′ to E4′ and a 5.0 μM concentration of each of the 16 corresponding substrates. Sixteen successive transfers were carried out over 70 hours, transferring 5% of the material from one reaction mixture to the

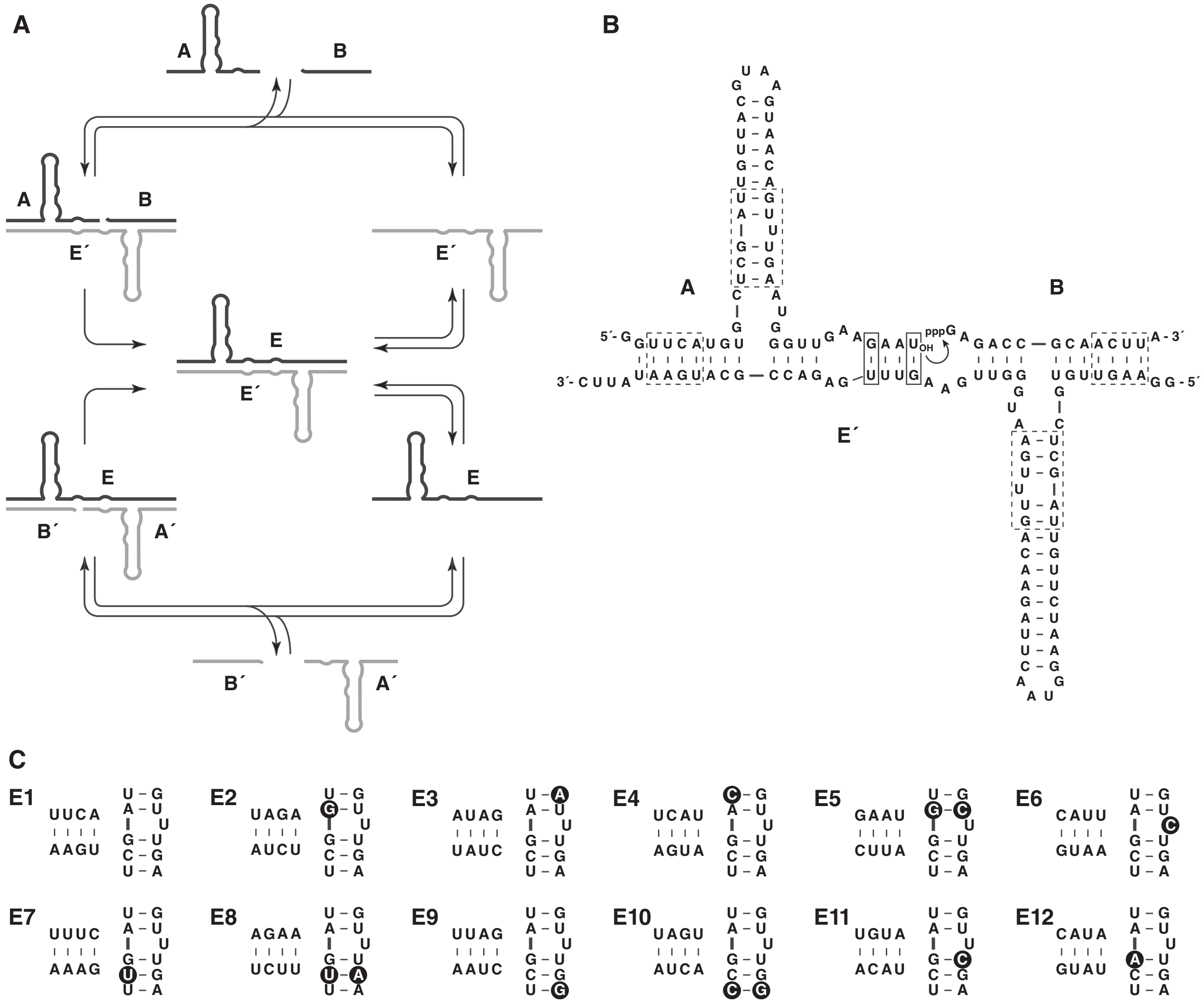


**Fig. 1.** Cross-replicating RNA enzymes. (**A**) The enzyme E′ (gray) catalyzes ligation of substrates A and B (black) to form the enzyme E, whereas E catalyzes ligation of A′ and B′ to form E′. The two enzymes dissociate to provide copies that can catalyze another reaction. (**B**) Sequence and secondary structure of the complex formed between the enzyme and its two substrates (E′, A, and B are shown; E, A′, and B′ are the reciprocal). The curved arrow indicates the site of ligation. Solid boxes indicate critical wobble pairs that provide enhanced catalytic activity. Dashed boxes indicate paired regions and catalytic nucleotides that were altered to construct various cross replicators. (**C**) Variable portion of 12 different E enzymes. The corresponding E′ enzymes have a complementary sequence in the paired region and the same sequence of catalytic nucleotides (alterations relative to the E1 enzyme are highlighted).

next (fig. S2A). Individuals were cloned from the population after the final reaction and sequenced. Among 25 clones (sequencing E′ only), there was no dominant replicator (fig. S2B). E1′, E2′, E3′, and E4′ all were represented, as well as 17 clones that were the result of recombination between a particular A′ substrate and one of the three B′ substrates other than its original partner (or similarly for A and B). Recombination occurs when an enzyme binds and ligates a mismatched substrate. In principle, any A could become joined to any B or B′, and any A′ could become joined to any B′ or B, resulting in 64 possible enzymes. The set of replicators were designed so that cognate substrates have a binding advantage of several kilocalories per mole as compared with noncognate substrates (fig. S2C), but once a mismatched substrate is bound and ligated, it forms a recombinant enzyme that also can cross-replicate. Recombinants can give rise to other recombinants, as well as revert back to nonrecombinants. Based on relative binding affinities, there are expected to be preferred pathways for mutation, primarily involving substitution among certain A′ or among certain B components (fig. S2D).

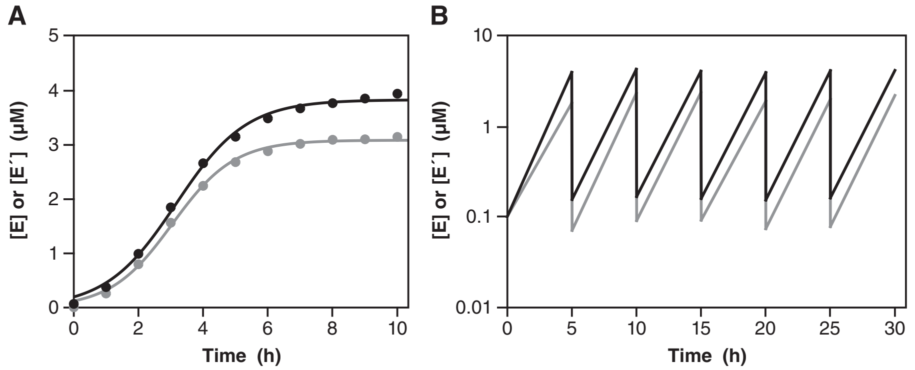


**Fig. 2.** Self-sustained amplification of cross-replicating RNA enzymes. (**A**) The yield of both E (black curve) and E′ (gray curve) increased exponentially before leveling off as the supply of substrates became exhausted. (**B**) Amplification was sustained by performance of a serial transfer experiment, allowing approximately 25-fold amplification before transferring 4% of the mixture to a new reaction vessel that contained a fresh supply of substrates. The concentrations of E and E′ were measured at the end of each incubation.

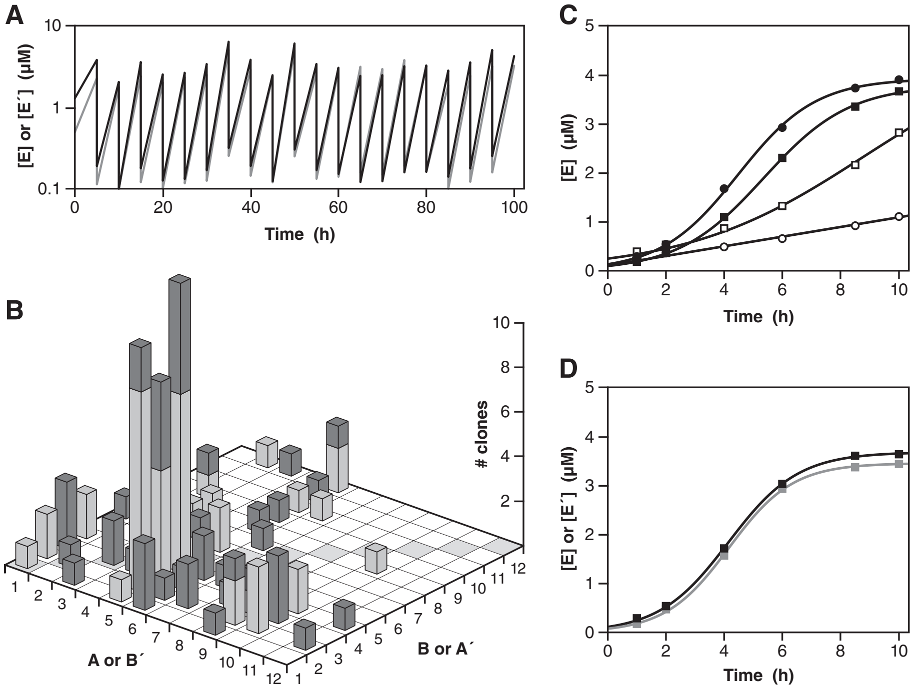


**Fig. 3.** Self-sustained amplification of a population of cross-replicating RNA enzymes, resulting in selection of the fittest replicators. (**A**) Beginning with 12 pairs of cross-replicating RNA enzymes (Fig. 1C), amplification was sustained for 20 successive rounds of ~20-fold amplification and 20-fold dilution. The concentrations of all E (black) and E′ (gray) molecules were measured after each incubation. (**B**) Graphical representation of 50 E and 50 E′ clones (dark and light columns, respectively) that were sequenced after the last incubation. The A and B (or B′ and A′) components of the various enzymes are shown on the horizontal axes, with nonrecombinant enzymes indicated by shaded boxes along the diagonal. The number of clones containing each combination of components is shown on the vertical axis. (**C**) Comparative growth of E1 (circles) and A5B3 (squares) in the presence of either their cognate substrates alone (solid symbols) or all substrates that were present during serial transfer (open symbols). (**D**) Growth of A5B3 (black curve) and B5′A3′ (gray curve) in the presence of the eight substrates (A5, B2, B3, B4, B5′, A2′, A3′, and A4′) that comprise the three most abundant cross-replicating enzymes.

A second serial transfer experiment was initiated with a 0.1 μM concentration each of all 12 pairs of cross-replicating enzymes and a 5.0 μM concentration of each of the 48 corresponding substrates. This mixture allowed 132 possible pairs of recombinant cross-replicating enzymes as well as the 12 pairs of nonrecombinant cross-replicators. Twenty successive reactions were carried out over 100 hours, transferring 5% of the material from one reaction mixture to the next, and achieving an overall amplification of greater than $10^{25}$-fold (Fig. 3A). Of 100 clones isolated after the final reaction (sequencing 50 E and 50 E′), only 7 were nonrecombinants (Fig. 3B). The distribution was highly nonuniform, with sparse representation of molecules containing components A6 to A12 and B5 to B12 (and reciprocal components B6′ to B12′ and A5′ to A12′). The most frequently represented components were A5 and B3 (and reciprocal components B5′ and A3′). The three most abundant recombinants were A5B2, A5B3, and A5B4 (and their cross-replication partners), which together accounted for one third of all clones.

In the presence of their cognate substrates alone, E1 remained the most efficient replicator, but in the presence of all 48 substrates, the most efficient replicator was A5B3 (Fig. 3C). When the A5B3 replicator was provided with a mixture of substrates corresponding to the components of the three most abundant recombinants, its exponential growth rate was the highest measured for any replicator (Fig. 3D). The fitness of a pair of cross-replicating enzymes depends on several factors, including their intrinsic catalytic activity, exponential growth rate with cognate substrates, ability to withstand inhibition by other substrates in the mixture, and net rate of production through mutation among the various cross-replicators. The A5B3 recombinant and its cross-replication partner B5′A3′ have different catalytic cores (Fig. 1C), and both exhibit comparable activity, accounting for their well-balanced rate of production throughout the course of exponential amplification (Fig. 3D). The selective advantage of this cross-replicator appears to derive from its relative resistance to inhibition by other substrates in the mixture (Fig. 3C) and its ability to capitalize on facile mutation among substrates B2, B3, and B4 and among substrates A2′, A3′, and A4′ that comprise the most abundant recombinants (fig. S2D).

Populations of cross-replicating RNA enzymes can serve as a simplified experimental model of a genetic system with, at present, two genetic loci and 12 alleles per locus. It is likely, however, that the number of alleles could be increased by exploiting more than four nucleotide positions at the 5′ and 3′ ends of the enzyme and by relaxing the rule that these nucleotides form one G•C and three A•U pairs. In order to support much greater complexity, it will be necessary to constrain the set of substrates, for example, by using the population of newly formed enzymes to generate a daughter population of substrates (*9*). An important challenge for an artificial RNA-based genetic system is to support a broad range of encoded functions, well beyond replication itself. Ultimately, the system should provide open-ended opportunities for discovering novel function, something that probably has not occurred on Earth since the time of the RNA world but presents an increasingly tangible research opportunity.

### References and Notes

1. F. H. C. Crick, *J. Mol. Biol.* **38**, 367 (1968).
2. J. W. Szostak, D. P. Bartel, P. L. Luisi, *Nature* **409**, 387 (2001).
3. G. F. Joyce, *Nature* **418**, 214 (2002).
4. L. E. Orgel, *Crit. Rev. Biochem. Mol. Biol.* **39**, 99 (2004).
5. J. Rogers, G. F. Joyce, *RNA* **7**, 395 (2001).
6. N. Paul, G. F. Joyce, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 12733 (2002).
7. D.-E. Kim, G. F. Joyce, *Chem. Biol.* **11**, 1505 (2004).
8. K.-S. Kim, S. Oh, S. S. Yea, M.-Y. Yoon, D.-E. Kim, *FEBS Lett.* **582**, 2745 (2008).
9. Materials and methods are available as supporting material on *Science* Online.
10. The authors thank L. Orgel for many stimulating discussions. This research was supported by grants from NASA (NNX07AJ23G) and NIH (R01GM065130) and by the Skaggs Institute for Chemical Biology.

**Supporting Online Material**
www.sciencemag.org/cgi/content/full/1167856/DC1
Materials and Methods
Figs. S1 and S2
References

29 October 2008; accepted 23 December 2008
Published online 8 January 2009;
10.1126/science.1167856
Include this information when citing this paper.

---

# Antagonistic Actions of Msx1 and Osr2 Pattern Mammalian Teeth into a Single Row

**Zunyi Zhang,[1]* Yu Lan,[1]* Yang Chai,[2] Rulang Jiang[1]†**

Mammals have single-rowed dentitions, whereas many nonmammalian vertebrates have teeth in multiple rows. Neither the molecular mechanism regulating iterative tooth initiation nor that restricting mammalian tooth development in one row is known. We found that mice lacking the transcription factor odd-skipped related-2 (Osr2) develop supernumerary teeth lingual to their molars because of expansion of the odontogenic field. Osr2 was expressed in a lingual-to-buccal gradient and restricted expression of bone morphogenetic protein 4 (Bmp4), an essential odontogenic signal, in the developing tooth mesenchyme. Expansion of odontogenic field in *Osr2*-deficient mice required Msx1, a feedback activator of *Bmp4* expression. These findings suggest that the Bmp4-Msx1 pathway propagates mesenchymal activation for sequential tooth induction and that spatial modulation of this pathway provides a mechanism for patterning vertebrate dentition.

Teeth are vertebrate-specific organs, and distinct dentition patterns have played critical roles in vertebrate diversification and specialization (*1*–*3*). In addition to variations in tooth number, size, and shape, many nonmammalian vertebrates have multirowed dentitions, whereas mammals develop teeth in a single row. Studies of tooth development in several fish species showed that multirowed dentitions result from sequential iterative tooth initiation along the mesial-to-distal and labial-to-lingual directions (*3*–*5*). The molecular mechanisms regulating the precise spatiotemporal patterns of sequential tooth initiation are unknown. Development of the single-rowed mammalian dentition likely involves restricting odontogenic field along the buccolingual axis; the mechanism underlying this control is also unknown.

Current understanding of the molecular mechanisms controlling tooth development has come mostly from studies in mice (*1*, *6*). Although supernumerary teeth have been reported in several mutant mouse strains (*7*–*10*), the majority developed within the tooth row from vestigial diastemal tooth buds (*10*, *11*). We recently generated mice lacking the *Osr2* (*odd-skipped related-2*) gene (*12*, *13*) and found that they exhibited supernumerary tooth development lingual to their molar teeth (Fig. 1, A and B, and fig. S1). Histological analyses (*14*) traced initiation of these supernumerary tooth germs to aberrant thickening of oral epithelium lingual to the first molar tooth buds at embryonic day 13.5 (E13.5) (Fig. 1, C and D). By E15.5, as the first molar germs developed to late "cap" stage (*1*), the ectopically thickened oral epithelia in $Osr2^{-/-}$ embryos invaginated and the underlying mesenchyme condensed (Fig. 1E). These ectopic epithelial invaginations resembled cap stage tooth germs by E16.5 (Fig. 1F). Because $Osr2^{-/-}$ mice die shortly after birth resulting from cleft palate (*13*), we transplanted E13.5 mandibular molar tooth germs under renal capsules of adult mice to allow complete tooth morphogenesis (*14*). After 21 days, wild-type and heterozygous molar tooth germs gave rise to two to three mineralized molar teeth, representing the normal molars (Fig. 1G). In contrast, $Osr2^{-/-}$ mutant molar tooth germs gave rise to four to five mineralized teeth (Fig. 1H and fig. S2). These data indicate that a complete odontogenic program was ectopically activated lingual to the molar teeth in $Osr2^{-/-}$ mice.

To gain insight into supernumerary tooth formation in $Osr2^{-/-}$ mice, we examined expression of selected marker genes during early tooth development. *Pitx2* was initially expressed throughout oral epithelium, and its expression selectively maintained in dental epithelium after E11 (*15*). In $Osr2^{-/-}$ embryos, *Pitx2* expression abnormally persisted in oral epithelium lingual to the first molar tooth buds (fig. S3, A and B). By E14.5, strong *Pitx2* expression marked the supernumerary dental placodes and first molar tooth buds (fig. S3, C and D). At E13.5, *sonic hedgehog* (*Shh*) was expressed in the enamel knot of developing molar tooth buds (*16*) (fig. S3E). In $Osr2^{-/-}$ mutants, *Shh* was ectopically expressed in a subset of epithelial cells lingual to the first molar buds (fig. S3F). By E14.5, *Shh* expression was clearly detected in the supernumerary dental placodes in $Osr2^{-/-}$ embryos (fig. S3H). In addition, expression of dental mesenchyme markers *Msx1* and *Lef1* were up-regulated and expanded lingually in $Osr2^{-/-}$ mice (fig. S3, I to L). These data suggest that supernumerary tooth development resulted from lingual expansion of the odontogenic field from the first molar tooth germs.

To understand how Osr2 regulates the odontogenic field, we examined *Osr2* expression during normal tooth development. At E11.5, *Osr2* was strongly expressed in the mesenchyme lingual to the dental lamina in both the maxilla and the mandible (Fig. 2A). *Osr2* was also highly expressed in the proximal mandibular mesenchyme buccal to the dental lamina (Fig. 2A). As tooth buds developed from E12.5 to E14.5, *Osr2* mRNA was expressed in a gradient in the developing tooth mesenchyme, with higher expression lingual and lower expression immediately buccal to the tooth buds (Fig. 2, B to D). Overall, the *Osr2* expression pattern is complementary to that of bone morphogenetic protein 4 (Bmp4) (Fig. 2, E to H), an essential odontogenic signal preferentially expressed on the buccal side in developing molar mesenchyme (*17*–*20*).

The expression pattern and mutant phenotype suggest that Osr2 functions to restrict odontogenic potential in the developing tooth mesenchyme. Consistent with this hypothesis, *Bmp4* expression was up-regulated and expanded into mesenchyme lingual to first molar buds in $Osr2^{-/-}$

---

[1] Center for Oral Biology and Department of Biomedical Genetics, University of Rochester Medical Center, Rochester, NY 14642, USA. [2] Center for Craniofacial Molecular Biology, University of Southern California School of Dentistry, Los Angeles, CA 90033, USA.

*These authors contributed equally to this work.
†To whom correspondence should be addressed. E-mail: Rulang_Jiang@urmc.rochester.edu

embryos by E13.5 compared with its expression in wild-type littermates (Fig. 3, A and B, and fig. S4). Moreover, Smad1 activation was enhanced in the molar tooth germ and expanded lingually to the oral epithelium and mesenchyme in *Osr2*$^{-/-}$ embryos, compared with that in wild-type littermates (Fig. 3, C and D).

To test whether mesenchymal odontogenic field was expanded in *Osr2*$^{-/-}$ mice, we examined the ability of isolated E13.5 mandibular mesenchyme to induce tooth formation in nondental epithelia from E10.5 second branchial arches (*14*). The molar tooth mesenchyme from both wild-type and *Osr2*$^{-/-}$ embryos induced tooth germlike structures in nondental epithelia in recombinant explant cultures (Fig. 3, E and F). Whereas mandibular mesenchyme lingual to wild-type molar tooth germs had no odontogenic activity (Fig. 3G), mandibular mesenchyme lingual to *Osr2*$^{-/-}$ molar tooth germs induced tooth-characteristic changes in nondental epithelia (Fig. 3H). We next cultured the recombinant tissues in vitro followed by transplanting them under renal capsules to allow complete tooth morphogenesis (*14*). Histological analyses showed that the mutant, but not wild-type, mandibular mesenchyme lingual to molar tooth germs induced tooth morphogenesis in nondental epithelium (Fig. 3, E-H), confirming that supernumerary teeth in *Osr2*$^{-/-}$ mice were induced by lingually expanded mesenchymal signals from the molar tooth germs.

We next investigated whether exogenous Bmp4 was sufficient to induce supernumerary tooth formation lingual to the developing first molars. Implantation of Bmp4-soaked beads lingual to the first molar tooth germs of E13.5 wild-type embryos did not cause supernumerary tooth initiation in any of 18 mandibular explants examined (fig. S5), suggesting that supernumerary tooth induction in *Osr2*$^{-/-}$ mutants involved activation of additional mesenchymal odontogenic signals on the lingual side of the first molars.

Activation of odontogenic potential, including *Bmp4* expression, in the dental mesenchyme is mediated by the transcription factor Msx1 (*17*–*20*). Mice lacking *Msx1* exhibited loss of *Bmp4* expression in the dental mesenchyme and molar developmental arrest at the bud stage (*17*, *21*). Because *Msx1* is expressed throughout the early tooth mesenchyme (*20*), the spatially restricted *Bmp4* expression in normal dental mesenchyme and the expansion of *Bmp4* expression in *Osr2*$^{-/-}$ tooth mesenchyme suggest that Osr2 repressed Msx1-mediated activation of odontogenic signals. To test this hypothesis, we examined tooth development in mice carrying mutations in *Osr2* and *Msx1*. In contrast to early tooth developmental arrest in *Msx1*$^{-/-}$ mice, all five *Msx1*$^{-/-}$*Osr2*$^{-/-}$ mutant pups harvested at E18.5 showed first molar development to late bell stage, with well-patterned ameloblast and odontoblast differentiation (Fig. 4, A to C, and fig. S6). Whereas *Bmp4* expression was down-regulated in *Msx1*$^{-/-}$ first molar mesenchyme from E12 to E13.5 (Fig. 4, D and E, and fig. S4), it was partially restored in *Msx1*$^{-/-}$*Osr2*$^{-/-}$ littermates (Fig. 4F). By E14.5, both wild-type and *Msx1*$^{-/-}$*Osr2*$^{-/-}$ first molar germs had developed to the cap stage, with similar patterns of *Bmp4* expression (Fig. 4, G and I), whereas *Msx1*$^{-/-}$ first molar germs remained at the bud stage with little *Bmp4* expression (Fig. 4H). By E15, both wild-type and *Msx1*$^{-/-}$*Osr2*$^{-/-}$ first molar germs had progressed to late cap stage and strongly expressed *Lef1* (Fig. 4, J and L), a downstream target of Bmp4 signaling (*17*), whereas *Msx1*$^{-/-}$ first molar germs remained arrested, with little *Lef1* expression (Fig. 4K). At all stages examined, however, no supernumerary tooth development was detected in *Msx1*$^{-/-}$*Osr2*$^{-/-}$ mutants. Moreover, the pattern of *Osr2* expression during molar tooth development was similar in *Msx1*$^{-/-}$ and wild-type littermates (fig. S7). These data indicate that Msx1 and Osr2 act antagonistically to pattern the tooth morphogenetic field by controlling the expression and spatial distribution of mesenchymal odontogenic signals along the buccolingual axis. Disruption of the balance of this antagonistic interaction may underlie supernumerary tooth formation, such as in *Osr2*$^{-/-}$ mice, and tooth agenesis, such as in *Msx1*$^{-/-}$ mice and in humans with *MSX1* mutations (*21*, *22*).

In *Msx1*$^{-/-}$*Osr2*$^{-/-}$ mutant mice, sufficient amounts of Bmp4 were expressed to drive morphogenesis of the first molar teeth, but no supernumerary tooth initiated. Remarkably, mandibular second molars also failed to develop in *Msx1*$^{-/-}$*Osr2*$^{-/-}$ mutant mice (fig. S8). Kavanagh *et al.* (*23*) recently showed that mouse mandibular first molar tooth germ inhibited second molar development and proposed an inhibitory cascade model, in which initiation of posterior molars depended on a balance between intermolar inhibition and mesenchymal activation, to account for sequential molar initiation in mammals. Similar activator-inhibitor mechanisms have been proposed for periodic dentition patterning in other vertebrates (*3*–*5*, *24*), but

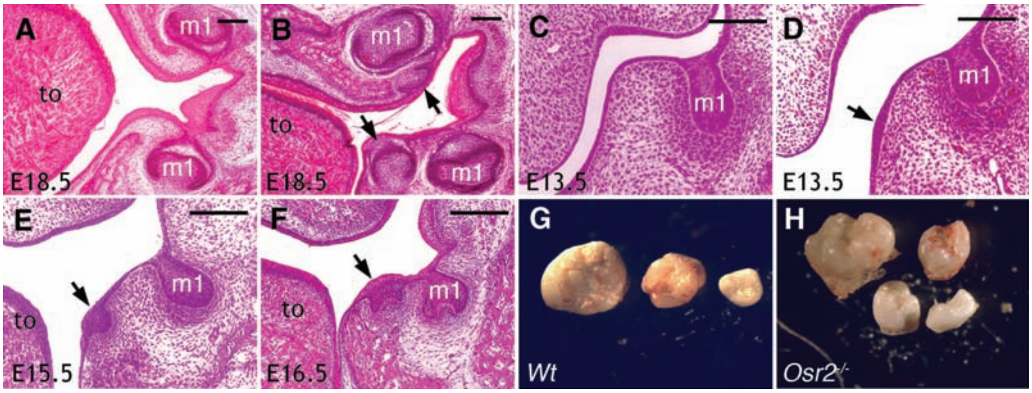


**Fig. 1.** (**A** and **B**) Frontal sections of E18.5 wild-type (A) and *Osr2*$^{-/-}$ (B) littermates. Arrows in (B) point to supernumerary tooth germs. m1, first molar tooth germ, and to, tongue. Scale bars indicate 100 μm. (**C**) Frontal section of wild-type first molar region at E13.5. (**D** to **F**) Frontal sections of *Osr2*$^{-/-}$ mutant first molar regions at E13.5, E15.5, and E16.5. Arrows point to supernumerary tooth germs. (**G** and **H**) Mineralized teeth from renal capsule cultures of E13.5 wild-type (G) and *Osr2*$^{-/-}$ (H) molar tooth germs.

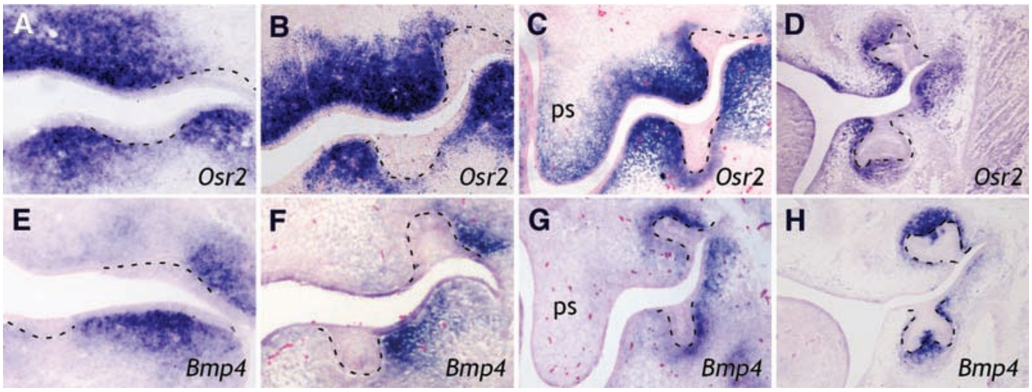


**Fig. 2.** Expression patterns of *Osr2* (**A** to **D**) and *Bmp4* (**E** to **H**) mRNAs along the buccolingual axis of mouse molar tooth germs at E11.5 [(A) and (E)], E12.5 [(B) and (F)], E13.5 [(C) and (G)], and E14.5 [(D) and (H)]. Lingual side is to the left in all panels. Black dashed lines mark the boundary between dental epithelium and mesenchyme. ps, palatal shelf.

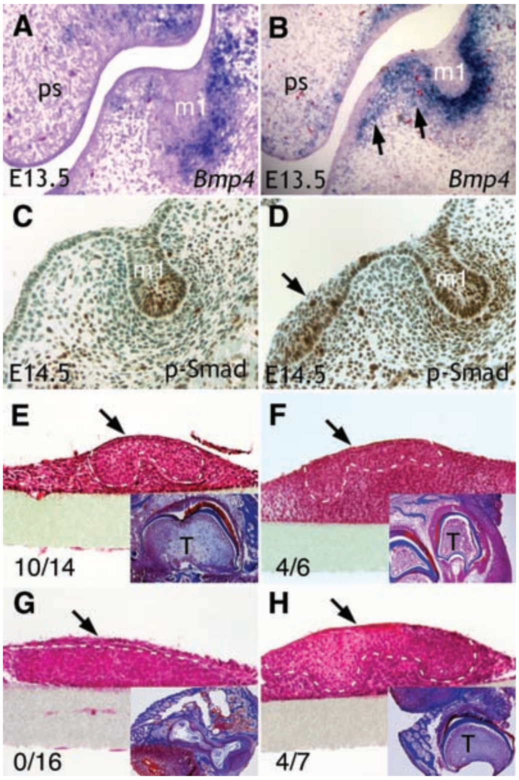


**Fig. 3.** (**A** and **B**) *Bmp4* mRNA expression in E13.5 wild-type (A) and $Osr2^{-/-}$ (B) first molar tooth germs. Arrows in (B) point to lingually expanded *Bmp4* expression. m1, first molar tooth bud; ps, palatal shelf. (**C** and **D**) Increased amounts of phospho-Smad1 accompany the supernumerary dental placode [arrow in (D)] in an E14.5 $Osr2^{-/-}$ embryo compared with wild-type littermate (C). (**E** and **F**) Isolated molar tooth mesenchyme from E13.5 wild-type (E) and $Osr2^{-/-}$ (F) embryos induced tooth formation from E10.5 second branchial arch epithelia. White dashed lines mark the boundary between epithelium and mesenchyme. Insets show sections of renal capsule cultures of corresponding recombinant explants. T marks tooth in renal capsule. Numbers indicate the ratios of corresponding recombinant explants forming teeth in renal capsules. (**G** and **H**) Isolated mesenchyme lingual to E13.5 molar tooth germ of $Osr2^{-/-}$ (H), but not that of wild-type (G), induced tooth formation from E10.5 second branchial arch epithelia.

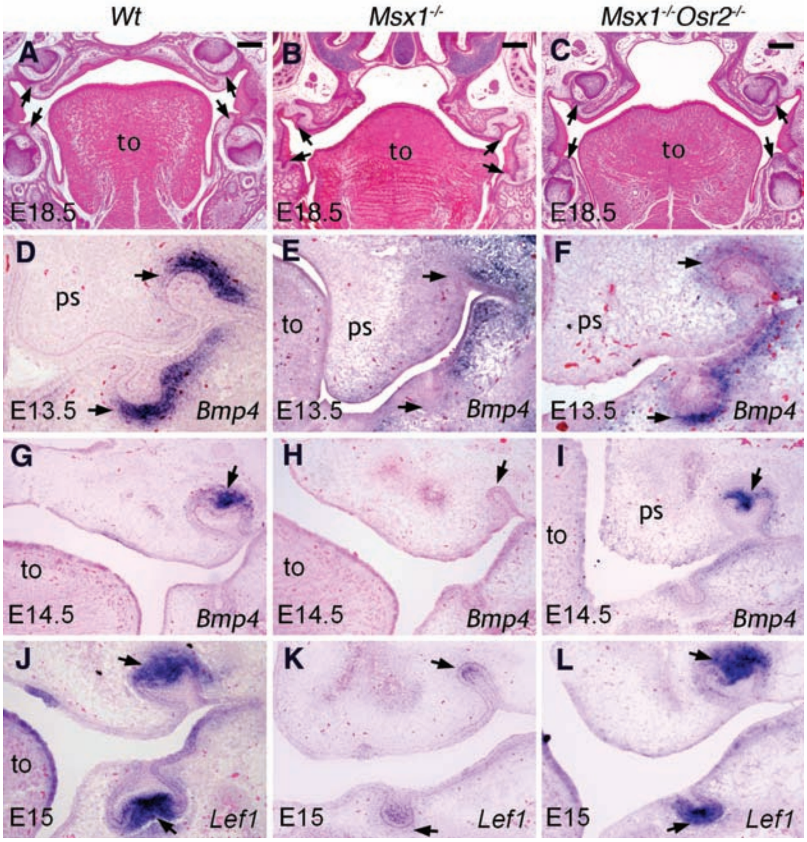


**Fig. 4.** (**A** to **C**) Frontal sections through the first molar tooth germs (arrows) of E18.5 wild-type (A), $Msx^{-/-}$ (B), and $Msx1^{-/-}Osr2^{-/-}$ (C) embryos. ps, palatal shelf; to, tongue. Scale bars, 200 μm. (**D** to **I**) *Bmp4* mRNA expression in the first molar tooth mesenchyme (arrows) in wild-type [(D) and (G)], $Msx1^{-/-}$ [(E) and (H)], and $Msx1^{-/-}Osr2^{-/-}$ [(F) and (I)] embryos at E13.5 [(D) to (F)] and E14.5 [(G) to (I)]. (**J** to **L**) *Lef1* mRNA expression in the first molar tooth mesenchyme (arrows) in E15 wild-type (J), $Msx1^{-/-}$ (K), and $Msx1^{-/-}Osr2^{-/-}$ (L) littermates.

the molecular underpinnings have not been identified. Bmp4 and Msx1 have been shown to regulate each other in a positive feedback loop in the dental mesenchyme (*17*–*19*). Bmp4 also induced expression of ectodin, a secreted Bmp inhibitor whose inactivation caused fusion of first and second molars as well as extra teeth in mice (*9*, *25*). Taken together, these and our finding that Msx1 is required for expansion of the odontogenic field in $Osr2^{-/-}$ mice suggest that the Bmp4-Msx1 pathway is a driving force in the activator-inhibitor network regulating sequential tooth initiation. In mammals, Osr2 suppresses this pathway along the buccolingual axis to restrict molar development to one tooth row. Diversity in dentition patterns in other vertebrates is likely due, at least in part, to evolutionary changes in antagonistic interactions regulating this pathway across the tooth morphogenetic field. In addition, reiterative initiation of other ectodermal organs, such as feather buds and taste papillae, which are also controlled by epithelial-mesenchymal interactions (*6*, *25*, *26*), may be driven by propagation of mesenchymal activators through a similar mechanism.

### References and Notes

1. A. Tucker, P. T. Sharpe, *Nat. Rev. Genet.* **5**, 499 (2004).
2. J. T. Streelman, J. F. Web, R. C. Albertson, T. D. Kocher, *Evol. Dev.* **5**, 600 (2003).
3. M. M. Smith, *Evol. Dev.* **5**, 394 (2003).
4. G. J. Fraser, A. Graham, M. M. Smith, *J. Exp. Zool.* **306B**, 183 (2006).
5. G. J. Fraser, R. F. Bloomquist, J. T. Streelman, *BMC Biol.* **6**, 32 (2008).
6. J. Pispa, I. Thesleff, *Dev. Biol.* **262**, 195 (2003).
7. Q. Zhang *et al.*, *Dev. Dyn.* **227**, 78 (2003).
8. T. Mustonen *et al.*, *Dev. Biol.* **259**, 123 (2003).
9. Y. Kassai *et al.*, *Science* **309**, 2067 (2005).
10. O. D. Klein *et al.*, *Dev. Cell* **11**, 181 (2006).
11. R. Peterkova, H. Lesot, M. Peterka, *J. Exp. Zool.* **306B**, 234 (2006).
12. Y. Lan, P. D. Kingsley, E.-S. Cho, R. Jiang, *Mech. Dev.* **107**, 175 (2001).
13. Y. Lan *et al.*, *Development* **131**, 3207 (2004).
14. Materials and methods are available as supporting material on *Science* Online.
15. M. L. Mucchielli *et al.*, *Dev. Biol.* **189**, 275 (1997).
16. H. R. Dassule, A. P. McMahon, *Dev. Biol.* **202**, 215 (1998).
17. Y. Chen, M. Bei, I. Woo, I. Satokata, R. Mass, *Development* **122**, 3035 (1996).
18. M. Bei, R. Maas, *Development* **125**, 4325 (1998).
19. R. Maas, M. Bei, *Crit. Rev. Oral Biol. Med.* **8**, 4 (1997).
20. S. V. Keranen, T. Aberg, P. Kettunen, I. Thesleff, J. Jernvall, *Dev. Genes Evol.* **208**, 477 (1998).
21. I. Satokata, R. Maas, *Nat. Genet.* **6**, 348 (1994).
22. H. Vastardis, N. Karimbux, S. W. Guthua, J. G. Seidman, C. E. Seidman, *Nat. Genet.* **13**, 417 (1996).
23. K. D. Kavanagh, A. R. Evans, J. Jernvall, *Nature* **449**, 427 (2007).
24. P. M. Kulesa *et al.*, *Theor. Biol.* **180**, 287 (1996).
25. J. Laurikkala, Y. Kassai, L. Pakkasjarvi, I. Thesleff, N. Itoh, *Dev. Biol.* **264**, 91 (2003).
26. H.-S. Jung *et al.*, *Dev. Biol.* **196**, 11 (1998).
27. This work was supported by NIH grants R01DE013681 (R.J.) and T32DE007202 (Z.Z.). We thank K. Maltby for technical assistance and K. Ackerman, C. E. Ovitt, and M. Dixon for comments.

**Supporting Online Material**

www.sciencemag.org/cgi/content/full/323/5918/1232/DC1
Materials and Methods
Figs. S1 to S8

20 October 2008; accepted 8 January 2009
10.1126/science.1167418

## New Products

## Thermocycler

Evaporation of water from the master mix changes the specificity and efficiency of any polymerase chain reaction, so the Mastercycler pro thermocycler offers a new technology to reduce evaporation. The new technology found in the lid features a fluid-filled Teflon-coated membrane. In contrast to traditional solid plates, this innovative lid envelops the samples to guarantee a snug fit to reduce evaporation to a minimum.

Eppendorf
For information 800-645-3050
www.eppendorf.com/pcr

## Gel Protein Recovery System

The GPR-800 is an advanced, microfluidics-based gel protein recovery (GPR) system for the rapid and efficient extraction of proteins from polyacrylamide gels. The system makes use of a proprietary plastic microfluidic chip in a high voltage environment to allow fast, simultaneous recovery in eight parallel microchannels. It is a closed system with minimal dead volume, assuring no introduction of contaminants. The recovered proteins can be easily analyzed by mass spectrometry to determine intact mass or perform top-down proteomics experiments.

Protea Biosciences
For information 877-776-8321
www.proteabio.com

## Illuminator

The Midac Illuminator is a compact Fourier transform infrared module that integrates well with a variety of sample interfaces in virtually any configuration. Inside each Illuminator is an infrared source and interferometer that provide a collimated output beam. The rugged module can be coupled with any sampling accessory equipped with its own infrared detector to create a compact dedicated analyzer. Alternatively, the Illuminator's internal infrared source can be replaced with a detector module for emissions work or for use with an external infrared source. The unit can be powered by a 12-volt battery for use in remote or field applications.

Midac Corp.
For information 714-546-4322
www.midac.com

## Gel Permeation Chromatography

Solvent-enhanced light scattering (SELS) is a new light-scattering technique that can be used in gel permeation chromatography (GPC) applications in which the demands of sample solvent and mobile phase are different. The technique allows users with two different solvents to separately optimize the sample preparation step and the chromatographic conditions to increase the refractive index increment to get better light-scattering responses. Viscotek has published an application note about a fluoropolymer analysis by SELS-GPC in which the fluoropolymer is dissolved in tetrahydrofuran and then this solution is injected into a mobile phase of acetone for analysis.

Viscotek Europe
For information +44-1344-467180
www.viscotek.com

## Platinum Electrodes

Platinum Tweezertrode electrodes are designed to enhance in utero and in vivo transfection applications. These electrodes provide uniform electrical fields and efficient drug and gene delivery. They are available in a variety of sizes (1 mm, 3 mm, 5 mm, 7 mm, and 10 mm) diameters to accommodate a variety of tissue sizes and allow for easier targeting of tissues.

BTX
For information 800-272-5732
www.btxonline.com

## Perfusion System

The Perfusion System for Cell Isolation is designed for harvesting individual cells from the isolated organs of mice, rats, and guinea pigs. It features dual perfusion pathways, in which one circuit delivers perfusate to flush out blood cells while a second circuit delivers an enzyme solution to disintegrate tissue and release individual cells. The system is simple to operate using a single manual switch to change between the perfusion pathways. The dead/mixing volume in the perfusion pathway is less than 100 microliters. A dedicated extension for cardiomyocyte isolation ensures a positive pressure within the heart chamber to prevent bacterial contamination. Adaptations for in situ and in vivo perfusions are available.

Hugo Sachs Elektronik—Harvard Apparatus
For information 800-272-2775
www.harvardapparatus.com

Electronically submit your new product description or product literature information! Go to **www.sciencemag.org/products/newproducts.dtl** for more information.