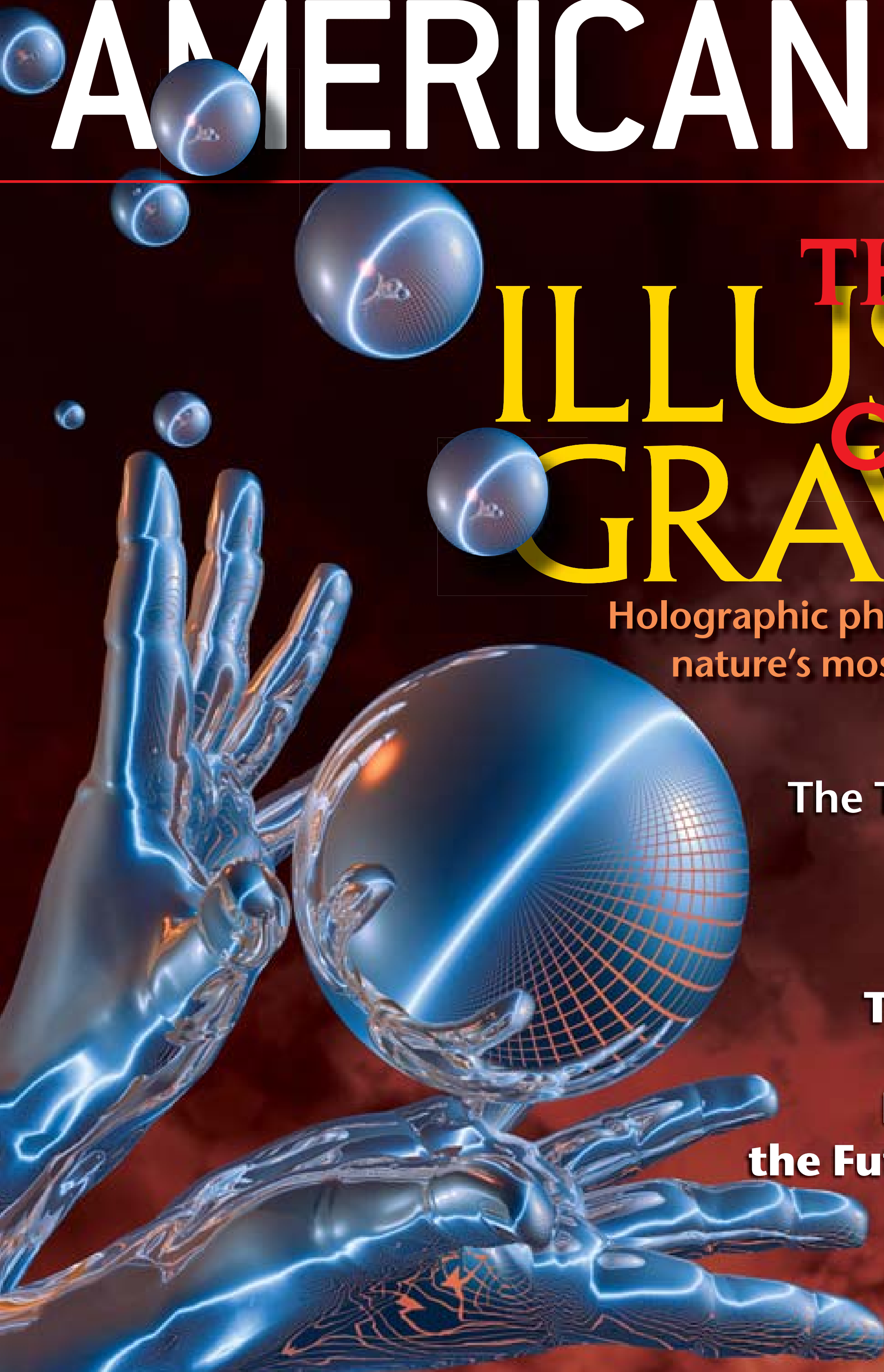
SPECIAL REPORT: PREPARING FOR PANDEMIC FLU
IS ANYONE READY FOR THIS GLOBAL KILLER?

SCIENTIFIC AMERICAN

PANSPERMIA:
Martian Cells Could Have Reached Earth

NOVEMBER 2005
WWW.SCIAM.COM



THE ILLUSION OF GRAVITY

Holographic physics might explain nature's most baffling force

The Two-Billion-Year-Old **Nuclear Reactor**

The First Drug from **Transgenic Animals**

Nanotech Wires and **the Future of Computing**

november 2005 contents

SCIENTIFIC AMERICAN Volume 293 Number 5

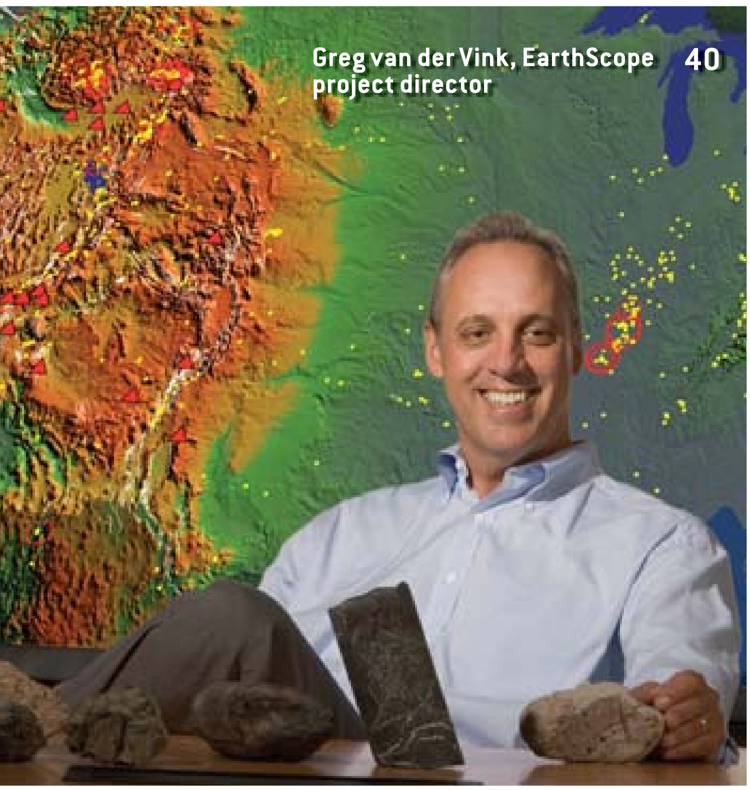
features

SCIENTIFIC AMERICAN Volume 293 Number 5

departments

18

116



Greg van der Vink, EarthScope project director 40

columns

Cover image by Phil Saunders, Space Channel Ltd.; photograph at left by Kay Chernush.

Scientific American (ISSN 0036-8733), published monthly by Scientific American, Inc., 415 Madison Avenue, New York, N.Y. 10017-1111. Copyright © 2005 by Scientific American, Inc. All rights reserved. No part of this issue may be reproduced by any mechanical, photographic or electronic process, or in the form of a phonographic recording, nor may it be stored in a retrieval system, transmitted or otherwise copied for public or private use without written permission of the publisher. Periodicals postage paid at New York, N.Y., and at additional mailing offices. Canada Post International Publications Mail (Canadian Distribution) Sales Agreement No. 40012504. Canadian BN No. 127387652RT; QST No. Q1015332537. Publication Mail Agreement #40012504. Return undeliverable mail to Scientific American, P.O. Box 819, Stn Main, Markham, ON L3P 8A2. Subscription rates: one year $34.97, Canada $49 USD, International $55 USD. Postmaster: Send address changes to Scientific American, Box 3187, Harlan, Iowa 51537. **Reprints available:** write Reprint Department, Scientific American, Inc., 415 Madison Avenue, New York, N.Y. 10017-1111; (212) 451-8877; fax: (212) 355-0408. **Subscription inquiries: U.S. and Canada (800) 333-1199; other (515) 248-7684. Send e-mail to sacust@sciam.com** Printed in U.S.A.

Preparing for the Worst

Hindsight is very often 20–20, but sometimes foresight is, too. Mark Fischetti's article "Drowning New Orleans" in the October 2001 *Scientific American* all too accurately depicted the devastation that an inevitable strong hurricane would bring to that city, as have articles in many other publications since that time. Those predictions sprang from years of published scientific analyses. Any official who claims to have been surprised by the tragic events that unfolded in New Orleans after Katrina simply wasn't paying attention.

LONG WAIT in Wheaton, Ill., for a limited supply of flu vaccine.

Realistically, even with this knowledge, a poor city in one of the nation's poorest states could never have done enough by itself to prevent Katrina's devastation or to cope with it afterward. "Catastrophic disasters are best defined in that they totally outstrip local and state resources, which is why the federal government needs to play a role," then FEMA director Joe Allbaugh told the New Orleans *Times-Picayune* in 2002. More's the pity, then, that in advance of the catastrophe the federal government severely underfunded the Army Corps of Engineers project to fortify New Orleans's levees. Rebuilding the city is sure to be considerably more expensive.

If federal priority-setters do not develop a healthier respect for the scale of damage that nature can and often will inflict, we will continue to be unprepared for another crisis now looming. "Preparing for a Pandemic," our special report starting on page 44, explains why scientists around the world are sounding an alarm and how governments and health authorities have begun belatedly to address the threat.

The leading candidate to cause the next flu pandemic is an avian virus dubbed H5N1 that has the potential to be even more lethal than the infamous 1918 pandemic strain. If it sickened a third of the population and killed just 5 percent of those, the death toll in the U.S. alone would exceed 10 times the pre-Katrina population of New Orleans. The U.S. pandemic plan that was due in October offers states much advice on how to prepare but little else with which to defend against an exceptionally infectious and virulent flu strain. Vaccines would take months to manufacture using archaic technology, and antiviral drugs that work against H5N1 would be in short supply. Both problems are fundamentally caused by a feeble market for these products in nonpandemic years.

That is why flu specialists have urged the U.S. government for at least a decade to prepare for a pandemic by taking the control of seasonal flu more seriously. Promoting yearly flu vaccinations for all age groups would be a good start, and mandatory vaccinations for schoolchildren are worth discussing: either measure could dramatically reduce seasonal transmission and thereby avert tens of thousands of flu-related hospitalizations and deaths annually. The policy would leverage a small government investment by giving vaccine makers incentive to improve their technology and expand manufacturing capacity. Similarly, if the U.S. would encourage doctors to use in-office diagnostic tests such as those that Japanese physicians commonly employ to diagnose flu on the spot and start treatments early, the rising demand for antiviral drugs would give pharmaceutical firms incentive to make more of them.

Flu season comes every year as reliably as hurricane season. If we shore up our defenses against both, we will be in a much stronger position when the "big ones" hit.

THE EDITORS editors@sciam.com

SCOTT OLSON *Getty Images*

How to Contact Us

EDITORIAL

For Letters to the Editors:
Letters to the Editors
Scientific American
415 Madison Ave.
New York, NY 10017-1111
or
editors@sciam.com
Please include your name and mailing address, and cite the article and the issue in which it appeared. Letters may be edited for length and clarity. We regret that we cannot answer all correspondence.

For general inquiries:
Scientific American
415 Madison Ave.
New York, NY 10017-1111
212-451-8200
fax: 212-755-1976
or
editors@sciam.com

SUBSCRIPTIONS

For new subscriptions, renewals, gifts, payments, and changes of address:
U.S. and Canada
800-333-1199
Outside North America
515-248-7684
or
www.sciam.com
or
Scientific American
Box 3187
Harlan, IA 51537

REPRINTS

To order reprints of articles:
Reprint Department
Scientific American
415 Madison Ave.
New York, NY 10017-1111
212-451-8877
fax: 212-355-0408
reprints@sciam.com

PERMISSIONS

For permission to copy or reuse material from *SA*:
www.sciam.com/permissions
or
212-451-8546 for procedures
or
Permissions Department
Scientific American
415 Madison Ave.
New York, NY 10017-1111
Please allow three to six weeks for processing.

ADVERTISING

www.sciam.com has electronic contact information for sales representatives of Scientific American in all regions of the U.S. and in other countries.

New York
Scientific American
415 Madison Ave.
New York, NY 10017-1111
212-451-8893
fax: 212-754-1138

Los Angeles
310-234-2699
fax: 310-234-2670

San Francisco
415-403-9030
fax: 415-403-9033

Detroit
Karen Teegarden & Associates
248-642-1773
fax: 248-642-6138

Midwest
Derr Media Group
847-615-1921
fax: 847-735-1457

Southeast
Publicitas North America, Inc.
404-262-9218
fax: 404-262-3746

Southwest
Publicitas North America, Inc.
972-386-6186
fax: 972-233-9819

Direct Response
Special Aditions Advertising, LLC
914-461-3269
fax: 914-461-3433

Australia
IMR Pty Ltd.
Tel: +612-8850-2220
Fax: +612-8850-0454

Belgium
Publicitas Media S.A.
+32-(0)2-639-8420
fax: +32-(0)2-639-8430

Canada
Derr Media Group
847-615-1921
fax: 847-735-1457

France and Switzerland
PEM-PEMA
+33-1-46-37-2117
fax: +33-1-47-38-6329

Germany
Publicitas Germany GmbH
+49-211-862-092-0
fax: +49-211-862-092-21

Hong Kong
Hutton Media Limited
+852-2528-9135
fax: +852-2528-9281

India
Convergence Media
+91-22-2414-4808
fax: +91-22-2414-5594

Japan
Pacific Business, Inc.
+813-3661-6138
fax: +813-3661-6139

Korea
Biscom, Inc.
+822-739-7840
fax: +822-732-3662

Middle East
Peter Smith Media & Marketing
+44-140-484-1321
fax: +44-140-484-1320

Scandinavia and Finland
Publicitas Nordic AB
+46-8-442-7050
fax: +46-8-442-7059

U.K.
The Powers Turner Group
+44-207-592-8331
fax: +44-207-630-9922

On the Web

WWW.SCIAM.COM

UPDATED EVERY WEEKDAY

Visit **www.sciam.com/ontheweb** to find these recent additions to the site:

Chimp Genome—and First Fossils—Unveiled

YERKES NATIONAL PRIMATE RESEARCH CENTER

Many animals, ranging from the rat to the puffer fish, have had their genomes sequenced, and now humankind's closest living relative has joined the group. The recent publication of a draft sequence of the chimpanzee genome provides the most detailed look yet at the similarities, and differences, between humans and chimps.

Flexible "E-skin" Could Endow Robots with Humanlike Sense of Touch

Robots with emotions are a staple of science fiction and a holy grail for AI researchers. But machines with another kind of feeling—a sense of touch—might be a more attainable goal for the near future. To that end, the results of a new study represent significant progress. Scientists have developed a pliable artificial skin that can sense pressure and temperature.

Americans and Chinese Differ in Their Worldview—Literally

A study of Chinese and American students has determined that the two groups looked at scenes in photographs in distinct ways. The findings indicate that previously observed cultural differences in judgment and memory between East Asians and North Americans derive from differences in what they actually see.

Ask the Experts

What causes ice cream headaches?

Robert Smith, professor emeritus and founding director of the department of family medicine at the University of Cincinnati and the founder of the University of Cincinnati Headache Center, provides an explanation.

Scientific American Digital

12-year archive with more than 150 issues

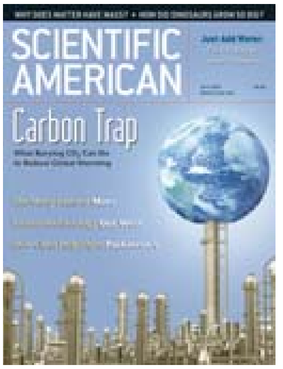
Visit www.sciamdigital.com to subscribe

Save $5 (use code WEB5)

Offer ends November 30, 2005

Letters

EDITORS@SCIAM.COM



IN THE JULY ISSUE our special report in collaboration with the *Financial Times* on "The Future of Stem Cells" drew letters expressing views on all facets of the debate. (If you did not see our limited print edition of the report, please visit www.sciam.com/stemcells)

It also elicited a well-versed observation on the use of embryonic stem cells from Kevin B. Keating of Los Angeles: "A stem cell's a miracle, you see, / It could extend the life of you—or me. / Only week-old embryos are needed for stems, / The problem is—we once were them," which prompted *Scientific American* editor in chief and sometimes rhymer John Rennie to compose a counter-verse: "IVF embryos have potential for life, / But when destined to end from a cold dish's strife, / Why not harvest their stem cells for future folk's good? / It's a waste and a sin if we don't when we could."

STEM SELLS

In regard to "The Future of Stem Cells" (special report with the *Financial Times*), I would like to point out the error of considering umbilical cord blood stem cells in the same category as adult stem cells, as well as the failure to address the financial distinctions between embryonic and cord blood stem cells.

It has been known that cord blood, in addition to being able to produce hematopoietic cells, contains stem cells called "beginning cells" or "Berashis cells," which can produce tissue beyond the hematopoietic system. Cord blood stem cells have produced promising results when tested in preclinical animal models of nearly all human diseases that embryonic stem cells have been postulated to help. Regardless, there remains slow development in cord blood research.

Given that the *Financial Times* coedited this special report, the important economic distinction between embryonic and cord blood stem cells should also have been noted. Embryonic stem cells have the potential to be patented (similar to hybrid corn), whereas only cord blood expansion techniques can be patented, not the cells derived from umbilical cord blood. This economic factor may have played a significant role in the slow research development of cord blood cells.

Norman Ende
UMDNJ–New Jersey Medical School
Newark, N.J.

LIVE LONGER AND PROSPER

Michael Shermer notes in "Hope Springs Eternal" [Skeptic] that progress is "unpredictable," and that is a common view. I've studied technology trends for 25 years and find that certain measurements of information-based processes, such as the cost of sequencing a base pair of DNA or the spatial and temporal resolution of brain scanning, are remarkably predictable. In computing, the doubly exponential trend line goes back a century, long before Gordon Moore of Moore's Law was born.

We see other examples in science of predictable results that arise from the interactions of a vast number of unpredictable elements (for instance, in thermodynamics). In technology, the ongoing exponential trend is not dependent on a single paradigm (such as shrinking transistors on an integrated circuit) but shifts among paradigms. These exponential trends are not limited to electronics: the amount of sequenced genetic data continues to double every year.

Shermer fails to fully appreciate the implications of this. He is thinking linearly, applying the current pace of progress and not appreciating the growing power of the tools of scientific inquiry.

He also states that "Ray and Terry's Longevity Program" consists of 250 daily nutritional supplements. Our program is personalized; if someone has no serious health issues, 10 or fewer are sufficient. The goal of nutrition, supplements, exercise and stress management is primarily to avoid premature death, not to exceed what Shermer calls the "maximum human life span of 120 years." Extending longevity, however, is not a new idea; two centuries ago life expectancy was only 37 years. Biotechnology will extend it enough for many of us to benefit from the nanotechnology revolution, which will enable us to go beyond the limitations of biology.

Ray Kurzweil
Boston

SHERMER REPLIES: I am, by nature, an optimist. I developed my skepticism over decades of bubble-bursting encounters between my dreamy-eyed meliorism and steel-jacketed realism. I hope Kurzweil's projections are right, but I remain unconvinced of his analogy between computer hardware and human software when it comes to Moore's Law. We should remember Wirth's Law, described by Nicklaus Wirth of the Swiss Federal Institute of Technology, which says that "software gets slower faster than hardware gets faster." This is one of the reasons that robots with human-scale artificial intelligence have yet to be produced and do not appear to be anywhere on the horizon. Moore's Law does not apply to computer software, nor does it apply to human biology. As for human life expectancy, better science, technology, medicine, nutrition, exercise and public health have combined to double the average of 37, but that's quite different from increasing the maximum life span of 120 years, which remains locked in genetic stone. When that barrier falls, then we can think about multicentury life spans. But until then, eat, drink, and be merry, for the end is nigh …

SPACE CENTS

I mostly agree with your criticism in "Lost in Space" [SA Perspectives] of the shifting budget priorities at NASA. But a critical point is always overlooked in any discussion of budgets involving millions or billions of dollars: the out-of-pocket cost to the median American wage earner for the entire NASA budget is probably less than an annual subscription to *Scientific American*. Relatively, the cost to the taxpayer for the Voyager and Landsat programs, then, amounts to couch-cushion change.

I firmly believe that if you folks were to include this perspective when reporting government funding for any scientific endeavor, you'd probably get a lot more taxpayers to think, "Really? That's all it costs me?" and (one would hope) urge their congressional representatives to support increased funding.

Steve Warren
Waldoboro, Me.

SPACE SHUTTLE *DISCOVERY:* No free launch.

PERMEATING REALITY

The box on "Properties of the Elusive Higgs" in "The Mysteries of Mass," by Gordon Kane, shows that the lowest energy of the electromagnetic field is zero, whereas the lowest energy of the Higgs field is a nonzero value below zero.

When I flip on a light, I add energy and raise the ambient electromagnetic field strength. Can I add energy to the Higgs field to raise its strength up to zero? That would be a useful thing to do to a spaceship because a zero-mass spaceship would be cheaper to accelerate.

Jim Soldini
Palm Bay, Fla.

This article is thought-provoking, but there are some confusing statements. Why does the Higgs field have two low-energy values, with neither one existing at the zero field strength, as illustrated in the section on permeating reality? Does such an energy state with two minima exist in any other physical system?

Ed Hammer
Mayfield Village, Ohio

KANE REPLIES: Soldini asks if one could add energy to raise the Higgs field energy to zero. That was the natural state of the universe shortly after the big bang. We know of no way to re-create that state. We probably wouldn't want to, because the world would be very different—if quarks and electrons were massless, the neutron would be lighter than the proton, so there would be no hydrogen atom, and any heavier atoms would be huge, so chemistry would be effectively nonexistent.

Hammer wonders why the Higgs field has an energy level below zero. This can be explained by the supersymmetric Standard Model. Basically, in the calculation of the energy, certain terms that are positive for most fields turn out to be negative for the Higgs. That leads to the multiple minima below zero energy. Energy curves with two or more minima occur in a wide range of other physical systems, including permanent magnets.

ERRATA In the box "Environmental Culprits" ["New Movement in Parkinson's," by Andres M. Lozano and Suneil K. Kalia], J. Timothy Greenamyre's affiliation was given as Emory University. He is on the faculty of the University of Pittsburgh School of Medicine.

In "The California Gambit," by W. Wayt Gibbs [stem cell special report], the gubernatorial recall election that Arnold Schwarzenegger won was incorrectly given as November 2003. It occurred in October 2003. The proposition that allotted funds for stem cell research was passed in November 2004.

CLARIFICATION In "Simulating Ancient Societies," Timothy A. Kohler, George J. Gumerman and Robert G. Reynolds omitted the fact that Alan Swedlund of the University of Massachusetts Amherst modeled the characteristics of the households in the Long House Valley simulation.

NASA

150, 100 & 150 Years Ago

FROM *SCIENTIFIC AMERICAN*

Before Tectonics ▪ Before Missiles ▪ Before Film

NOVEMBER 1955

FOUNDERING THEORY—"The size and peculiar shape of the Pacific trenches stir our sense of wonder. What implacable forces could have caused such large-scale distortions of the sea floor? And what is the significance of the fact that they lie along the Pacific 'ring of fire'—the zone of active volcanoes that encircles the vast ocean? Speculating from what we know, we may imagine that forces deep within the earth cause a foundering of the sea floor, forming a V-shaped trench. The depth stabilizes at about 35,000 feet, but crustal material, including sediments, may continue to be dragged downward into the earth. This is suggested by the fact that the deepest trenches contain virtually no sediments, although they are natural sediment traps."

NOVEMBER 1905

WIND POWER—"For several years the Danish government has experimented with windmills, to ascertain the relative amount of electrical power that can be generated. In this country similar experimental tests have been tried, and although the instances are not numerous, the data furnished indicate a useful future for this form of prime mover. This is particularly true of the agricultural regions of the West, where innumerable windmills have been constructed in the past ten years for irrigating purposes."

THE CARIBS—"The Indian [*see photograph*] stands in such fear of the Venezuelan and his government that he frequently prefers to follow the smaller waterways of the Guiana region or take overland trips through the virgin forest rather than use the broad highway of the Orinoco River that is his rightful heritage from countless ancestors. This disappearance of the Indian has greatly impeded the gathering of rubber, tonca beans, and other natural products. Since immigration is not encouraged and continuous revolutions have scattered or killed the settlers of European and mixed descent, it would seem that the country is steadily retrograding."

CARIB INDIANS of the Lower Orinoco, Venezuela, 1905

TORPEDO MISS—"The Whitehead torpedo has exercised a greater controlling influence upon naval construction and tactics than perhaps any other single weapon of naval warfare. However, it cannot be denied that the torpedo has, at times, been greatly overrated. Indeed, the experience of the recent war seems to prove that only under exceptional and very favorable conditions can the torpedo get in its blow. In the fleet engagements on the high seas it seems to have exercised very little, if any, influence upon battle formations. Consequently, we think it unlikely that torpedo tubes will be fitted into future warships."

NOVEMBER 1855

YELLOW FEVER—"At a recent meeting of the New York Academy of Medicine, Dr. Stowe, a distinguished surgeon of New Orleans, gave some valuable information respecting the terrible disease of yellow fever. In his opinion, it is the same everywhere, unmodified by topographical causes or changes of climate. Many attempts have been made to discover its cause, but, like cholera, it escapes observation. Warm climate is essential. Moisture seems not essential. New Orleans has daily showers at certain seasons, yet without disease. This year it was very dry, and the sugar cane died for want of moisture, and all were suffocated by dust when the disease first appeared."

TALBOT'S SUCCESS—"Photography is the general name now applied to sun painting on paper and glass, as being different from the daguerreotype, which is produced on metallic plates. The inventor of photographs is Fox Talbot, of England, who secured patents in Britain and America, but has thrown them open to the public. Photography is destined apparently to supersede the art of Daguerre. In France, the splendid display of photographs in the Great Exhibition of Industry, and the limited number of pictures on metallic plates, affords conclusive proof that, with the French artists, the daguerreotype is becoming obsolete."

news SCAN

FLOOD CONTROL

Protecting against the Next Katrina

WETLANDS MITIGATE FLOODING, BUT ARE THEY TOO DAMAGED IN THE GULF? BY MARK FISCHETTI

In the aftermath of Hurricane Katrina's devastation, the nation has vowed to rebuild New Orleans and Gulf Coast communities while improving protection against raging storms. But before engineers redesign a single levee, they must consider a fundamental question: Can the Mississippi Delta be restored as a lush, hardy buffer that can absorb surges and rising seas? Or is it too far gone, necessitating a 300-mile wall to hold back the Gulf?

VINCENT LAFORET *AP Photo/Pool*

DROWNING CITY: Battered by the August 29 arrival of Hurricane Katrina, levees protecting New Orleans crumbled, such as this one along the Inner Harbor Navigational Canal. Tattered wetlands around the city offered little protection against stormy seas.

Researchers have known for at least five decades that wetlands help to stop storm surges from crashing inland. But for a century, the U.S. Army Corps of Engineers leveed the Mississippi River to its mouth to stop annual floods. That spared New Orleans but starved the wetlands south and east of the city of the sediment, nutrients and freshwater they need to thrive. The levees also cut off sediment flow that builds barrier islands ringing the delta.

In 1998 scientists and engineers proposed a $14-billion master plan, which Congress never funded. Called Coast 2050, it detailed strategies to revive the delta and control flooding [see "Drowning New Orleans," by Mark Fischetti; Scientific American, October 2001]. Seven years and a bad hurricane later, the wetlands and barrier islands are so much more tattered that traditional restoration techniques may no longer suffice. "I raised this issue at a Corps meeting several years ago already," notes S. Jeffress Williams of the U.S. Geological Survey, who has studied the coast for 20 years. "It was not well received. But considering the rate of wetland loss, land subsidence, sea-level rise, and increasing frequency and severity of storms, the question should be on the table."

Coast 2050–style measures carry some

DUTCH LESSON: The Oosterschelde, a storm surge barrier in the Netherlands, allows the North Sea to mix with freshwater, thus helping to keep wetlands intact.

F. DAMM zefa/Corbis

significant complications. For one, massive superlevees such as those being raised in Osaka, Japan, would be needed and would overrun city streets and private land. The alternative would be to connect the barrier islands and outer marshes with large levees, dams and floodgates, creating a continuous rim around the delta. After a 1953 North Sea storm sent a surge into the Netherlands that killed 2,000 people, the Dutch government agreed to that very kind of network, which today safeguards 400 miles of coast. Joop Weijers, a senior engineer at the Dutch Ministry of Transport, Public Works and Water Management, which oversees the installation, known as Delta Works, says that "building the whole system right now would cost $15 billion to $16 billion."

A front wall against the Gulf could condemn to death the wetlands behind it, by changing the tidal mixing patterns of saltwater and freshwater. Weijers admits that in the first decades of Delta Works marshland was lost. To prevent that, the Dutch engineers changed strategy; instead of installing solid dams, they erected a long series of sluices whose huge doors remain open year-round, allowing the sea in. The doors close only when storms approach. Other retrofits are creating more natural water flows. As a result, the loss of wetlands has stopped.

Insight can also be gained from storm surge work going on in Venice, Italy, says Rafael Bras, a civil and environmental engineer at the Massachusetts Institute of Technology. The venerable city will rely on mobile floodgates in its lagoon that lie flat on the seabed under normal conditions and rise only during extremely high tides. Further lessons, he notes, are coming from current efforts to restore natural freshwater flows throughout the Florida Everglades and from the successful restoration of 15,000 acres of wetlands in south San Francisco Bay.

At the same time, "we need a new vision to revive southeastern Louisiana," says Denise Reed, a marsh specialist at the University of New Orleans. Even if a Mississippi version of Delta Works defends wetlands, certain Coast 2050–style projects will still be needed to negate damage done. The region produces one third of the country's seafood and provides wintering for 70 percent of the nation's migratory waterfowl, yet at least 25 square miles are still disappearing every year. And Katrina may have accelerated the destruction. Although investigations are just beginning, Reed expects that Katrina's surge pushed saltwater much farther into the marsh than normal. "That salt has likely baked into the soil," Reed says, "and won't readily wash out," which could kill wide areas of marsh grass from the roots up.

Lessons from New Orleans could save other cities already at sea level, such as New York and Miami. Global-warming models indicate that sea level will climb one to three feet this century and increase the frequency and intensity of storms. "Even if you believe that sea level and storm activity are rising because of natural cycles," Williams says, "these cycles run for 25 to 30 years, and we are only eight years into the current one. And the cycle will return." The USGS has recently finished fine-tuning a formula, called the Coastal Vulnerability Index, which can predict how at risk an open coastline is to high seas.

Despite the urgency, experts say the U.S. should not rush to erect new protection around New Orleans. "Politicians want to act now," Weijers remarks, but "you need a long-term strategy." His counsel? "You must learn how to work with nature instead of defying it."

NEED TO KNOW: SINKING CITIES

Sucking out groundwater for human use faster than rainfall can recharge it dries soils and reduces the pore pressure from below, allowing them to compact and sink. As a result, San Jose, Calif., at the southern tip of San Francisco Bay, is becoming a bowl like New Orleans, with some sections already slumped several feet below sea level.

Extracting oil and natural gas also deflates land, in part because it draws deep groundwater with it. Houston is sinking even faster than New Orleans. Long Beach, Calif., on the Gulf of Santa Catalina, dropped 15 to 20 feet from the 1950s to the 1970s before extraction was slowed; some neighborhoods now lie seven feet below the sea.

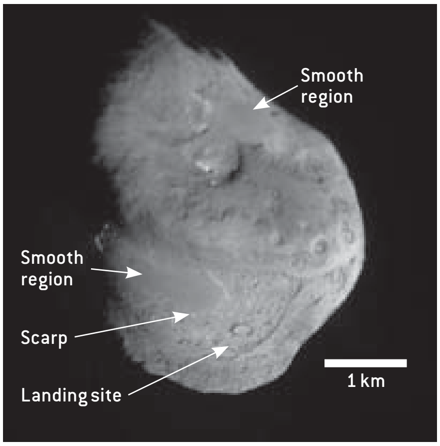
For the nation's highs and lows: http://erg.usgs.gov/isb/pubs/booklets/elvadist/elvadist.html

SPACE

Comet Dust Bunny

TEMPEL 1 PROVES TO BE A BALL OF FLUFF BY GEORGE MUSSER

In the course of the Space Age, planets have gone from astronomical objects—wandering specks in the night sky—to geologic ones: full-fledged worlds you could imagine yourself walking on. In the 1990s asteroids made the same transition. And now it is comets' turn, as exemplified by July's (deliberate) crash of the Deep Impact probe into the nucleus of Comet Tempel 1. In September researchers announced a batch of findings at an American Astronomical Society meeting in Cambridge, England.



COMET TEMPEL 1'S unexpectedly textured surface has dozens of suspected impact craters.

The impact threw up a cone of dusty debris some 500 meters high; it extended right from the surface, indicating that the cometary material put up no significant resistance to the projectile. The flight path of debris particles revealed the strength of the comet's gravity, hence its density—on average, about half that of water. The body must be riddled with voids. So much fine dust flew out that it could not have been created by the impact itself but rather must have already been sitting on the surface. A lack of big pieces suggests that the comet has no outer crust.

Along with other data, these results indicate that Tempel 1 is not a compacted snowball, as many had conjectured, but a loose, powdery fluff ball—an agglomerate of primordial dust that came together at low speeds and gently clung like dust bunnies under a bed. If other comets are like this, landing on one—as the European Space Agency's Rosetta mission plans to do in 2014—could be tricky.

Despite its flimsiness, Tempel 1 has an almost planetlike surface, covered with what appear to be impact craters (the first ever observed on a comet), cliffs, and distinct layers. The composition hints at an unexpectedly complex history of chemical reactions. Humans have always found comets mysterious, and seeing them up close has done nothing to change that.

NASA/JPL/UNIVERSITY OF MARYLAND

NEURO-BIOLOGY

Baby to Brain

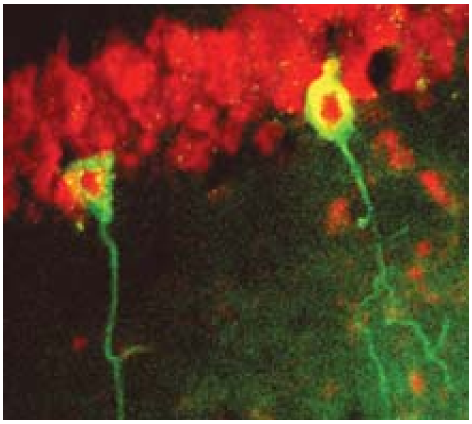
THERAPY CLUES FROM FETAL CELLS THAT ENTER MOM'S BRAIN BY CHARLES Q. CHOI

Mothers could literally always have their kids on their minds. Researchers find that in mice, cells from fetuses can migrate into a mother's brain and apparently develop into nervous system cells.

The discovery comes from Gavin S. Dawe of the National University of Singapore and Zhi-Cheng Xiao of Singapore General Hospital, along with their colleagues from China and Japan. They were looking to design therapies for stroke or diseases such as Alzheimer's. Scientists have known for years that fetal cells can enter a mother's blood; in humans, they may remain there at least 27 years after birth. Like stem cells, they can become many other kinds of cells and in theory might help repair damaged organs.

The neurobiologists bred normal female mice with males genetically modified to uniformly express a green fluorescent protein. They found green fetal cells in the mothers' brains. "In some regions of some mothers' brains, there are as many as one in 1,000 to sometimes even 10 in 1,000 cells of fetal origin," Xiao reports.

The fetal cells transformed into what seem like neurons, astrocytes (which help to feed neurons), oligodendrocytes (which insulate neurons) and macrophages (which ingest germs and damaged cells). Moreover, after the scientists chemically injured the mouse brains, nearly six times as many fetal cells made their way to damaged areas than elsewhere, suggesting the cells could be re-



FETAL CELLS (*green*) can make their way to a mother's brain. Neuronal nuclei are stained red.

"FETAL MICROCHIMERISM IN THE MATERNAL MOUSE BRAIN," BY XIAO-WEI TAN ET AL. IN *STEM CELLS* ONLINE; AUGUST 10, 2005

sponding to molecular distress signals released by the brain.

Just how the fetal cells make it through the capillaries separating the brain from the blood system is not known—the cells of the vessels are densely packed, preventing most compounds from crossing the barrier. The researchers speculate that biomolecules such as proteins or sugars adorning fetal cell surfaces interact with the blood-brain barrier and allow the cells to wriggle past. The team feels confident that fetal cells can also pass to the brains of males and nonpregnant females, given little evidence of major differences between their blood-brain barriers and those of pregnant females, Dawe says. The scientists hope next to show that the fetal cells become functional neurons.

The finding, published online August 10 by *Stem Cells,* gives fresh hope in treating brain disorders. Because of the blood-brain barrier, transplant therapies for the brain normally evoke thoughts of drilling into the skull. Identifying the molecules typical of fetal cells that enter the brain and become nervous system cells could help find similar cells from sources other than fetuses, such as umbilical cord blood. Such research could lead to noninvasive cell transplants for the brain requiring only intravenous injections. Any cells used for therapies would be matched to patients as closely as possible to avoid triggering immune disease. It remains uncertain whether injected cells meant for the brain could end up grafting somewhere else, "but we don't know yet if that happening would even be a problem," Dawe says.

The investigators are also now looking to see if the passage of fetal cells to the brain occurs in humans as readily as it does in mice. They plan on looking at postmortem brain tissue from mothers of boys. Signs of a Y chromosome would confirm the effect in humans. It would also, Xiao points out, raise the issue of "whether there are any behavioral or psychological implications."

Charles Q. Choi is a frequent contributor.

PREGNANT PROTECTION?

Fetal cells may circulate inside a mother's body to protect her health, according to one theory. Indeed, the ability of fetal cells to enter the brain preferentially in response to injury adds substance to that theory, put forward by medical geneticist Diana W. Bianchi of Tufts University, who first discovered that fetal cells can persist for decades in mothers. "If their mothers die, the babies are in deep trouble, so there's a speculated evolutionary advantage here," she explains.

Other evidence contradicts the theory: some studies have linked the development of disease on very rare occasions with fetal cells that have engrafted themselves onto mothers' tissues. Bianchi feels, however, that these grafted fetal cells are not causing such disease but are responding in order to help.

STANDARDS

Wait a Second

IS IT TIME TO DECOUPLE TIME FROM THE EARTH'S SPIN? BY WENDY M. GROSSMAN

The arrival of the new year will take a bit longer this winter. The reason: bringing the calendar back into synchronization with the earth's rotation requires the addition of an extra second. Leap seconds are not new: 22 have been added since 1972, on June 30 or December 31. But under proposals to be discussed at a November meeting of the radiocommunication sector of the International Telecommunication Union—the United Nations agency in charge of broadcasting official time—this leap second might be the last. Astronomers, historians and others are upset about the prospect of losing the leap second and have initiated a spirited debate to keep time tied to the earth's rotation.

How long can people spend arguing about a second? "It's been going on for about five years," says Judah Levine, a physicist in the time and frequency division at the National Institute of Standards and Technology (NIST) in Boulder, Colo. Word of the proposed abolition became public in July, thanks to an e-mail tip to timekeeping officials by Daniel Gambis, head of the Earth Orientation Center of the Frankfurt, Germany–

COUNTING THE HOURS

The most common timescale based on the earth's rotation is Coordinated Universal Time (or UTC, an acronym derived from the French version), which Britons still call Greenwich Mean Time. International Atomic Time (TAI) is based on seconds ticked off an atomic clock. No single entity keeps time for the world; the task falls to several different authorities. The three most important are the U.S. Naval Observatory, the National Institute of Standards and Technology (NIST) and Germany's equivalent to NIST, Physikalisch-Technische Bundesanstalt, often referred to more simply as PTB.

based International Earth Rotation and Reference Systems Service (IERS), which decides when to add leap seconds.

Leap seconds are needed because the earth's spin is slowing down, gradually and unevenly. The rotational changes arise because of tidal forces exerted by the moon and inertial effects related to the liquid outer core sloshing around

NO TIME OUT: Official timekeeping may depend on atoms, not day-night cycles.

IMAGES.COM/CORBIS

and to the cycle of evaporation, in which water at the equator gets deposited at the poles as ice that melts seasonally.

The present system is a compromise between taking advantage of the most accurate timepieces—that is, atomic clocks—and respecting traditional timekeeping via the sun's position. To people who want to end leap seconds, Levine explains, "the really fundamental quantity is not time but frequency—and frequency comes from quantum mechanics; it is a property of atoms. And what these folks really want is for time to represent frequency in a smooth, continuous way." Levine does not speak officially for NIST, but he is the person who, on December 1, will formally alert authorities to add the leap second at the end of the month.

The primary advocates of a frequency standard are people behind modern telecommunications systems. GPS, for example, does not adjust for leap seconds—a problem made apparent in 2003, when the length of time since the last leap second caused the clocks in some GPS receivers to malfunction so that they gave the time as 62:28:15.

To illustrate the issue posed by leap seconds, Levine points to navigation. "Every time there's a leap second, the thing that's moving continues to move, but the clock stops," he explains. "So the people who deal with physical processes do not want leap seconds." Tom O'Brian, chief of NIST's time and frequency division, adds that leap seconds occur unpredictably. The IERS did not decide on this year's leap second, the first in seven years, until July 2005. To keep clocks from drifting too far from the day-night cycle, abolitionists would presumably need to add, say, several leap minutes every few hundred years.

The existing compromise system, Levine notes, also sows confusion. For one, the leap second occurs in the middle of the day in Asia and Australia, causing a time hiccup during stock trading. For another, the more timescales there are, the easier it is for a programmer to make an error in calculations.

Astronomers are deeply dismayed at the prospect, which would decouple time from the earth's rotation, much as going off the gold standard did for currency. "We need to know where the earth is pointing so we can point a radio" for spacecraft communications, says Mike Hapgood, the geophysical secretary of Britain's Royal Astronomical Society. Modifying equipment, he argues, could become expensive.

Not every astronomer thinks that adjustments will be arduous. Cornell University's Donald Campbell, former director of the center that runs the Arecibo radiotelescope, states that rewriting the software to adjust for the difference between atomic time and universal time is not a big deal.

Even if the cost turns out to be negligible, a sense of history and tradition may not be. After all, for thousands of years humans have tracked time by the position of the sun, not by the oscillation of the atom.

Wendy M. Grossman writes about information technology from London.

ASTRONOMY

Red Star Rising

SMALL, COOL STARS MAY BE HOT SPOTS FOR LIFE BY MARK ALPERT

As every comic-book fan knows, Superman was born on the planet Krypton, which orbited a red star. Scientists are now learning that the Superman legend may contain a kernel of truth: the best places to find life in our galaxy could be on planets that circle the small but common stars known as red dwarfs.

Last June astronomers reported the discovery of Krypton's real-life counterpart, a planet orbiting a red dwarf called Gliese 876, about 15 light-years from our sun. Although researchers had previously identified two other planets orbiting Gliese 876, the detection of the third body made headlines because it is so much like Earth. Nearly all the planets found outside our solar system to date are gas giants comparable in size to Jupiter or Neptune. The newly discovered world, in contrast, is most likely a rocky body only about twice as large as Earth.

ROCKY PLANET found near red dwarf star Gliese 876 is twice Earth's size (*artist's rendering*).

Red dwarfs, also known as M dwarfs, have less than half the mass of the sun and are hundreds of times dimmer. Scientists had long doubted that life could arise near such faint stars because the "Goldilocks zone"—the area around the star that is neither too hot nor too cold for liquid water—would extend no more than about 40 million kilometers from the M dwarf, less than the distance between the sun and Mercury. At such close quarters, the planet could become tidally locked: one side would always face the star while the other would be turned away. Researchers had previously assumed that the planet would inevitably lose its atmosphere because of freezing on the bitterly cold night side, but computer models have shown that an atmosphere containing modest amounts of carbon dioxide would spread enough heat to the night side to prevent freezing.

M dwarfs are by far the most abundant stars in the Milky Way galaxy, outnumbering sunlike G stars more than 10 to one, so the possibility that they could harbor habitable worlds has excited researchers involved in the search for extraterrestrial intelligence (SETI). In July the SETI Institute in Mountain View, Calif., held a workshop in which scientists debated whether life could emerge on an M-dwarf planet. "No one found any showstoppers to habitability," says Gibor Basri of the University of California, Berkeley. One concern was that because M dwarfs frequently produce flares, the resulting torrents of charged particles could strip the atmosphere off any nearby planet. If the planet had a magnetic field, though, it would deflect the particles from the atmosphere. And even the slow rotation of a tidally locked M-dwarf planet—it spins once for every time it orbits its star—would be enough to generate a magnetic field as long as part of the planet's interior remained molten.

Unfortunately, the new planet found near Gliese 876 is only three million kilometers from the star, so its surface would be far too hot for life as we know it. But the sheer number of M stars in the Milky Way—an estimated 300 billion—has made them a prime target for SETI investigators, who are now building the Allen Telescope Array in northern California to scan the galaxy for radio signals. Besides their abundance, M dwarfs have another advantage over G stars as possible sites for intelligent life: because they shine longer before exhausting their hydrogen fuel and do not brighten with age, they provide a more stable long-term environment for life-bearing planets. One billion years from now, intensifying solar radiation will make Earth uninhabitable, but the galaxy's M dwarfs will burn steadily for hundreds of billions of years. It's enough to make Superman homesick.

TRENT SCHINDLER *National Science Foundation*

COUNTING OUR DIM NEIGHBORS

Before scientists can look for habitable planets or radio signals near red dwarfs, they must overcome the challenge of finding the stars. Because red dwarfs are so faint—even the closest ones cannot be seen with the naked eye—they are seriously underrepresented in the astronomical catalogues. To correct this deficiency, researchers are measuring the parallax of faint red stars to estimate their distance from Earth, which helps to determine whether they are nearby red dwarfs or more luminous red giants that only appear dim because they are farther away. In recent years a team led by Todd J. Henry of Georgia State University has used this method to add about two dozen red dwarfs to the list of stars within 10 parsecs (32.6 light-years) of our solar system.

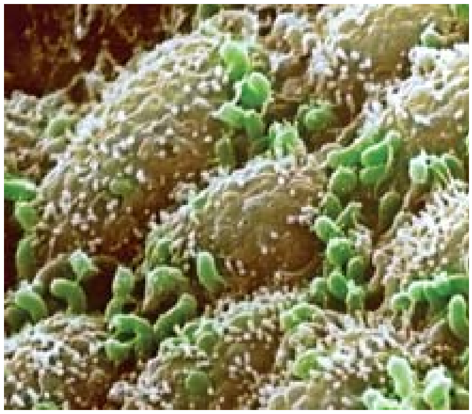
BIOTECH

Bugs and Drugs

GUT BACTERIA COULD DETERMINE HOW WELL MEDICINES WORK BY GUNJAN SINHA

Last year scientists at drug giant Pfizer noticed something peculiar. Rats in a routine study were excreting unusually low levels of a metabolite in their urine called hippuric acid. It was a metabolic oddity that could throw off further laboratory results. So the scientists dug deeper. The rats had been reared at the same facility in Raleigh, N.C., as another group. The metabolite levels should have all been the same. But curiously, the rats in question had been bred in one particular room. Further investigation turned up an unlikely culprit—the rats carried a unique composition of gut microorganisms that had altered their metabolism.

"That was really a surprise," remarks Lora C. Robosky, principal scientist at Pfizer. Robosky is part of a team at the company investigating how spectroscopy and pattern-recognition software could analyze metabolites in body fluids—called metabonomics—to better select drug compounds. Along the way, the technology has revealed a factor often overlooked as a source of individual responses to drugs—gut microflora.



INTESTINAL FORTITUDE: Bacteria (*green*) thrive along the walls of the duodenum. At least 400 species of bacteria live in the gut.

BIOMEDICAL IMAGING UNIT, SOUTHAMPTON GENERAL HOSPITAL *Photo Researchers, Inc.*

Scientists have long speculated that gut bacteria play a role in human health. But the microbes—passed from mother to infant via feeding and physical contact—are uniquely adapted to the gut and are not easily studied outside of it. "Remarkably, the first paper that enumerated human gut microbiota was published just a few months ago," says Jeffrey I. Gordon, director of the Center for Genome Sciences at Washington University in St. Louis. Through genomic sequencing, the paper, by Paul B. Eckburg and his colleagues at Stanford University, estimates at least 400 species in our gut. Each species exists in different strains, multiplying the variation. In humans, microorganisms in the distal intestines may liberate at least 20 percent of calories by breaking down sugars into more digestible forms.

Scientists have only just begun to elucidate how these mysterious bugs influence health. Two years ago, for example, David G. Binion of the Medical College of Wisconsin showed that the sodium butyrate produced by gut microorganisms could inhibit blood vessel growth in the intestine by blocking COX-2—an enzyme implicated in many inflammatory disorders (and the target of drugs such as Vioxx). Other studies have shown that specific strains of *Escherichia coli* can metabolize dimethylarsine, a derivative of arsenic, to produce potentially toxic compounds, which may underlie the carcinogenicity of arsenic in the gut. Bacteria may also free therapeutic compounds in foods such as soy. And people who pack on pounds easily might be suffering from especially thrifty gut bacteria that are contributing to obesity, Gordon and some scientists hypothesize.

Although Pfizer scientists did not investigate their unusual rats' ability to metabolize drugs, the discovery may partly explain why data from presumably identical animals sometimes conflict, Robosky says. What is more, investigators have shown that gut microbes affect drug metabolism, she adds. How significantly, however, remains unknown.

Jeremy K. Nicholson, a pioneer of

CORBIS

RESPONSE TO DRUGS varies among individuals and depends not only on genes but probably also on the type of gut microflora in their bodies.

metabonomics research at Imperial College London, has no doubt that bacteria substantially affect the way the body responds to drugs. "What determines metabolism is largely environmental: how stressed you are, what gut microbes you've got—that turns out to be incredibly important," he argues. For example, many species produce compounds that switch on detoxification enzymes in the liver, and certain microbial metabolites are necessary players in human metabolic pathways.

The microflora may even influence some drugs that turn toxic in a minority of people, such as Vioxx with its cardiovascular complications, Nicholson speculates. But most efforts to identify individual variation focus on genes rather than gene-environment interaction, Nicholson points out, which is likely to be more important.

To take his research further, Nicholson has been collaborating with several companies, including Pfizer and Bristol-Myers Squibb, to develop metabonomic technology that identifies metabolite patterns to predict both a drug compound's toxicity and the biochemical pathways involved. But Nicholson says he and others are working toward a broader ideal: to statistically integrate data from all the "omics" disciplines, such as proteomics and transcriptomics, for a complete picture of a drug's physiological effects.

Gunjan Sinha writes frequently about medicine and biotechnology from Berlin.

SIMPLE SCAN FOR DRUG TOXICITY

The nascent field of metabonomics—the study of the body's metabolites—has begun flexing its muscle. Last year Imperial College London and Pfizer filed a patent on a new metabonomic approach that, by scanning an animal's urine before dosing the animal with certain drugs, can predict toxicity with high accuracy. "Up to 90 percent of toxicity prediction may ultimately be possible based on metabonomic modeling," says Jeremy Nicholson of Imperial College, who helped to develop the work. If the technology pans out for humans as it does for animals, then drugs should become much safer.

BY THE NUMBERS

Fewer Men, More Crime

HOW A FALLING SEX RATIO CAN UNDERMINE A COMMUNITY **BY RODGER DOYLE**

Researchers have long recognized that a large surplus of young men can lead to social disorder, as happened on the frontier in 19th-century America. A counter theory, however, holds that a shortage of men can also be disruptive.

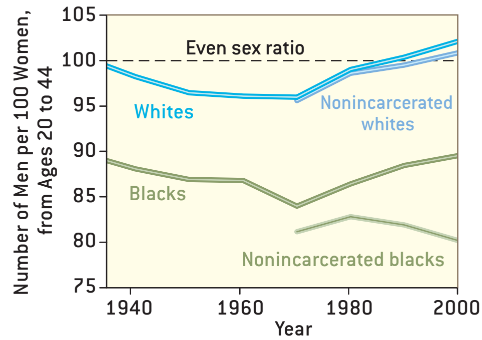
The theory comes from the late Marcia Guttentag of Harvard University and Paul Secord of the University of Houston. They explored the dynamics of marriage and partnership among the African-American community, which has the largest recorded shortage of men of any ethnic subgroup in the world. On purely demographic grounds, they concluded that a low sex ratio—the number of males per 100 females—may make men reluctant to marry or invest in children, particularly if their incomes were inadequate. If they did marry, they were at high risk for divorce or separation. Among the black population, many men could not find work—good-paying factory jobs in Northern cities began disappearing in the 1960s—and some turned to crime. These factors help to destabilize conventional family structure and result in births to unwed mothers, many of them teenagers, who tend to have limited skills in raising children. Their sons tend to be more crime-prone than sons of two-parent families.

The theory gets support from U.S. census data for 153 cities, which show an indirect link between low sex ratios and violent crime. Low sex ratios for blacks were strongly associated with family disruption, which in turn was strongly associated with crime. Cities in which the sex ratio is high had few female-headed households, whereas the opposite was true of cities with low ratios.

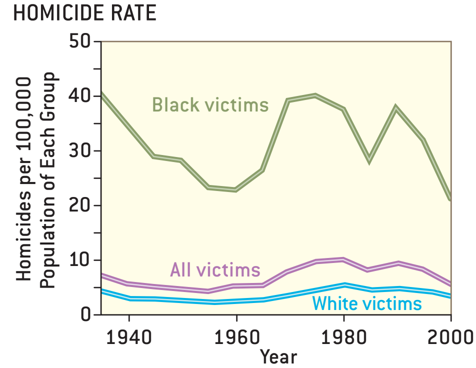
As the charts illustrate, the decline in the black sex ratio accelerated about 1960, which was shortly before violent crime rates among the African-American population began their precipitous rise. (The homicide data, determined by the number of victims, are the most reliable indicators of violent crime available.) Some black men were not available as marriage partners because of service in the armed forces overseas, but far more important was the growing population of black men in prison since 1970. Incarceration effectively reversed an emerging trend toward rising sex ratios among African-Americans aged 20 to 44.

The decline in the crime rate beginning in the early 1990s would seem to weigh against the Guttentag-Secord theory. Possibly, other developments overwhelmed any effect of declining sex ratios and poverty. This period saw significant changes in inner-city neighborhoods, including a reduction in the number of handguns, lower drug and alcohol consumption, better policing, and a tendency for drug dealers to be less violent. According to one of the most heartening explanations, the decline in crime came about mainly because young people in poor neighborhoods developed a revulsion against drug dealing and the spread of AIDS and organized successfully to avoid these hazards.

SEX RATIO



HOMICIDE RATE



Rodger Doyle can be reached at rodgerpdoyle@verizon.net

RODGER DOYLE

FURTHER READING

Too Many Women? The Sex Ratio Question. Marcia Guttentag and Paul F. Secord. Sage Publications, 1983.

The Sex Ratio, Family Disruption, and Rates of Violent Crime: The Paradox of Demographic Structure. Steven F. Messner and Robert J. Sampson in *Social Forces*, Vol. 69, No. 3, pages 693–713; March 1991.

The Improbable Transformation of Inner-City Neighborhoods: Crime, Violence, Drugs, and Youth in the 1990s. Richard Curtis in *Journal of Criminal Law and Criminology*, Vol. 88, No. 4, pages 1233–1276; Summer 1998.

DATA POINTS: ANXIOUS MOMENTS

At some point in their lives, about half of all Americans will to some degree develop a disorder of anxiety, mood, impulse control or substance (abuse or dependency). Sufferers also hesitate in seeking treatment but get help more often now than 10 years ago. The conclusions come from a national survey of 9,282 households.

Percent who will develop a mental disorder during their lives: 46.4

Percent who will develop two: 27.7

Median age of onset for disorders of:

Anxiety: 11

Mood: 30

Impulse control: 11

Substance: 20

Years before treatment is sought for disorders of:

Anxiety: 9 to 23

Mood: 6 to 8

Impulse control: 4 to 13

Substance: 5 to 9

Most prevalent ailments and percent affected at some point in their lives:

Major depressive disorder: 16.6

Alcohol abuse: 13.2

Specific phobia (fear of object or situation): 12.5

Social phobia: 12.1

SOURCE: Archives of General Psychiatry, *June 6*

ATMOSPHERIC SCIENCE

Replenished Ozone

The ozone layer is no longer disappearing. Depletion of stratospheric ozone, which defends the planet from harmful levels of ultraviolet rays, was first detected 25 years ago and stems mostly from industrial pollutants, especially chlorofluorocarbons. Atmospheric scientists analyzed data on the entire ozone layer from satellites and ground stations and found ozone level decline plateaued between 1996 and 2002. The ozone layer has even increased a small amount in parts of the Northern Hemisphere, including much of the area over North America, Europe and Asia. Study co-author Betsy Weatherhead of the University of Colorado at Boulder says that more recent data, through this past August, "absolutely are in agreement with the general conclusions of the paper," which appears in the August 31 *Journal of Geophysical Research.* Overall, however, ozone will remain seriously depleted worldwide, especially at the poles, because ozone-destroying chemicals persist in the atmosphere for decades. *—Charles Q. Choi*

NEUROBIOLOGY

Take Your Breath Away

Those who die in their sleep are often recorded as cases of cardiac failure, but sometimes the real culprit may be sleep apnea, the interruption of breathing during sleep. Researchers injected a cell-specific neurotoxin into a tiny region of the brain stem in rats, the pre-Bötzinger complex, which is essential for breathing. After a few days the rats developed breathing problems, first during REM sleep, then spreading to non-REM sleep and wakefulness. Lapses in breathing rhythm grow more frequent among the elderly and those with neurodegenerative disorders, the researchers note in a *Nature Neuroscience* report, published online August 7. They speculate that depletion of these neurons from disease or old age results in an increasingly severe apnea from which the body finally cannot rouse itself. *—JR Minkel*

PHARMACEUTICALS

No Sleep, No Problem

Exhausted police and hospital workers could someday benefit from a compound observed to perk up sleep-deprived rhesus monkeys. Neurobiologists at the Wake Forest University School of Medicine administered the chemical, called CX717, to 11 monkeys after keeping them awake for 30 to 36 hours (with music, treats, and other stimuli). After receiving the compound, the sleep-deprived animals performed normally at a memory task requiring them to pick out an image matching one they had seen up to 30 seconds before. Scans revealed that the drug reverted their brain activity to that of their normal resting state, the group reports in the August 22 *Public Library of Science, Biology.* Moreover, the chemical boosted the performance of rested monkeys by up to 15 percent. In May, Cortex Pharmaceuticals, the maker of the compound, also reported that the drug improved performance in 16 sleep-deprived male volunteers, but the research is still unpublished. *—JR Minkel*

BETTER THAN COFFEE? A drug seems capable of eliminating the effects of sleep deprivation.

THOMAS HOEPKER *Magnum Photos;* ILLUSTRATION BY MATT COLLINS

GENETICS

Relative Distance

FROM CHIMPAN-A TO CHIMPANZEE: The genome of humans' closest relative has been sequenced, yielding clues about the genetic differences between the species, which are greater than previously thought.

Chimpanzees and humans are more different than previously thought, based on a draft of the chimp genome published in the September 1 *Nature*. Although the human genome differs from our closest relative's by 1.2 percent in terms of single nucleotide changes, the international Chimpanzee Sequencing and Analysis Consortium finds that duplications and rearrangements of larger DNA stretches add another 2.7 percent difference. Seven regions in the human genome that differ from that of chimps bear strong hallmarks of natural selection; for instance, one contains elements regulating a gene implicated in nervous system development and another possessing genes linked with speech. Consortium glycobiologist Ajit Varki and his colleagues also report in the September 9 *Science* the first human-specific protein, which binds to cell-surface sugars and is expressed in brain cells known as microglia. These immune cells are involved in ailments not seen in chimps, such as Alzheimer's, multiple sclerosis and HIV-induced dementia. *—Charles Q. Choi*

ENVIRONMENT

Road Assault on Water

Keeping winter roads clear is turning freshwater salty across the northeastern U.S. Ecologists at the Institute of Ecosystem Studies in Millbrook, N.Y., and their colleagues investigated salinity increases in streams in Baltimore County, the Hudson River Valley and the White Mountains of New Hampshire over the past 30 years. Chloride levels approached up to a quarter the concentration of seawater during the winter. Saltiness was strongly linked to the number of roads, parking lots and other impervious surfaces nearby, which overall in the U.S. cover an area roughly the size of Ohio; that area is expected to expand this decade with one million new homes and 16,000 kilometers of new roads. At this rate, many rural streams in the northeastern U.S. will become toxic to sensitive freshwater life and unfit for human consumption within the next century, the scientists reported online September 12 in the *Proceedings of the National Academy of Sciences USA*. *—Charles Q. Choi*

PAIN

Use Your Illusion

Although placebos may only be sham medicines, the benefits patients believe they receive appear to be more than illusions. Neuroscientists at the Universities of Michigan at Ann Arbor and Maryland injected saltwater into the jaw muscles of healthy young male volunteers to cause pain. The researchers then told them that intravenous saline drips might be painkillers and asked them to rate the intensity of their discomfort on a scale of 0 to 100 every 15 seconds. Brain scans revealed that neural areas linked with responses to pain, stress, rewards and emotion released endorphins, natural analgesics that behave like opiates. The endorphin response matched up in time with reductions in pain intensity and unpleasantness the volunteers reported. The scientists, whose findings appear in the August 24 *Journal of Neuroscience*, plan on researching this effect in women and in patients with chronic pain. *—Charles Q. Choi*

BRIEF POINTS

- The bee's knees: In the first demonstration that insects can learn by watching, bumblebees seeing fellow buzzing foragers on certain flowers were twice as likely to forage on those flowers.

  Biology Letters of the Royal Society online, August 19

- The earth's core appears to be rotating faster than the rest of the planet by up to 0.5 degree a year. The possible cause: the liquid outer core's magnetic field that tugs on the solid iron inner core.

  Science, August 26

- Using alcohol-based hand-sanitizing gels reduces the incidence of gastrointestinal infections by 59 percent.

  Pediatrics, September

- Space rocks could affect local weather: a 10-meter-wide asteroid that burned up over Antarctica left micron-size dust particles in the upper atmosphere—big enough to seed clouds and to cool the air by reflecting sunlight.

  Nature, August 25

GERRY ELLIS *Minden Pictures*

BRAD HINES

Rupert's Resonance

The theory of "morphic resonance" posits that people have a sense of when they are being stared at. What does the research show? By MICHAEL SHERMER

Have you ever noticed how much easier it is to do a newspaper crossword puzzle later in the day? Me neither. But according to Rupert Sheldrake, it is because the collective successes of the morning resonate through the cultural morphic field.

In Sheldrake's theory of morphic resonance, similar forms (morphs, or "fields of information") reverberate and exchange information within a universal life force. "Natural systems, such as termite colonies, or pigeons, or orchid plants, or insulin molecules, inherit a collective memory from all previous things of their kind, however far away they were and however long ago they existed," Sheldrake writes in his 1988 book, *Presence of the Past* (Park Street Press). "Things are as they are because they were as they were." In this book and subsequent ones, Sheldrake, a botanist trained at the University of Cambridge, details the theory.

Morphic resonance, Sheldrake says, is "the idea of mysterious telepathy-type interconnections between organisms and of collective memories within species" and accounts for phantom limbs, how dogs know when their owners are coming home, and how people know when someone is staring at them. "Vision may involve a two-way process, an inward movement of light and an outward projection of mental images," Sheldrake explains. Thousands of trials conducted by anyone who downloaded the experimental protocol from Sheldrake's Web page "have given positive, repeatable, and highly significant results, implying that there is indeed a widespread sensitivity to being stared at from behind."

Skepticism is the default position.

Let us examine this claim more closely. First, science is not normally conducted by strangers who happen on a Web page protocol, so we have no way of knowing if these amateurs controlled for intervening variables and experimenter biases.

Second, psychologists dismiss anecdotal accounts of this sense to a reverse self-fulfilling effect: a person suspects being stared at and turns to check; such head movement catches the eyes of would-be starers, who then turn to look at the staree, who thereby confirms the feeling of being stared at.

Third, in 2000 John Colwell of Middlesex University in London conducted a formal test using Sheldrake's experimental protocol. Twelve volunteers participated in 12 sequences of 20 stare or no-stare trials each and received accuracy feedback for the final nine sessions. Results: subjects could detect being stared at only when accuracy feedback was provided, which Colwell attributed to the subjects learning what was, in fact, a nonrandom presentation of the trials. When University of Hertfordshire psychologist Richard Wiseman also attempted to replicate Sheldrake's research, he found that subjects detected stares at rates no better than chance.

Fourth, confirmation bias (where we look for and find confirmatory evidence for what we already believe) may be at work here. In a special issue in June of the *Journal of Consciousness Studies* devoted to a fierce debate between "Sheldrake and His Critics," I rated the 14 open peer commentaries on Sheldrake's target article (on the sense of being stared at) on a scale of 1 to 5 (critical, mildly critical, neutral, mildly supportive, supportive). Without exception, the 1's, 2's and 3's were all traditional scientists with mainstream affiliations, whereas the 4's and 5's were all affiliated with fringe and proparanormal institutions. (For complete results, see Table 1 in the online version of this column at www.sciam.com)

Fifth, there is an experimenter bias problem. Institute of Noetic Sciences researcher Marilyn Schlitz—a believer in psychic phenomena—collaborated with Wiseman (a skeptic of psi) in replicating Sheldrake's research and discovered that when *they* did the staring Schlitz found statistically significant results, whereas Wiseman found chance results.

Sheldrake responds that skeptics dampen the morphic field, whereas believers enhance it. Of Wiseman, he remarked: "Perhaps his negative expectations consciously or unconsciously influenced the way he looked at the subjects."

Perhaps, but wouldn't that mean that this claim is ultimately nonfalsifiable? If both positive and negative results are interpreted as supporting a theory, how can we test its validity? Skepticism is the default position because the burden of proof is on the believer, not the skeptic. SA

Michael Shermer is publisher of Skeptic *(www.skeptic.com) and author of* The Science of Good and Evil.

Scoping Out the Planet

Greg van der Vink hopes that EarthScope yields unprecedented data about faults and plates. It may do for geoscience what human genome sequencing did for biology By KRISTA WEST

At two years old and $37 million invested, EarthScope is like a young Little League team with the game tied in the second inning—the players are still learning to work together, and the parental units are watching closely. But the game is just getting started.

The passionate, pat-you-on-the-back coach is Greg van der Vink, geophysicist by training, visiting Princeton University professor, and the first project director of the most ambitious earth science project ever attempted. If successful, EarthScope will measure the movement and deformation of the earth below the contiguous U.S. and Alaska with a level of detail and data accessibility never seen in geophysics. The hope is that a clearer understanding of the forces that shape the environment will translate into better assessment of earthquake and volcanic hazards and more precise knowledge of the country's natural resources.

KAY CHERNUSH

GREG VAN DER VINK: AN EARTHLY VISION

- **Project director of EarthScope, an attempt to measure the structure and deformation of the earth underneath the contiguous U.S. and Alaska.**
- **Most cited publication: "Scientifically Illiterate vs. Politically Clueless," in *Science*, about the need for politically literate scientists.**
- **On getting things done at the intersection of science and politics: "Not making a decision is, of course, making a decision."**

When I meet the 48-year-old Greg van der Vink—known as GVDV among colleagues—at a June meeting of the Incorporated Research Institutions for Seismology (IRIS), a consortium of universities and research bodies, I ask for a sports analogy to explain EarthScope. "I wondered if that would come up," he chuckles. The first time van der Vink testified before Congress—in the mid-1980s as a congressional fellow fresh out of graduate school—a senator advised him to use a sports analogy to explain everything in three minutes or less, the idea being that that is about the amount of time you have to make your case. Now a sports analogy is part of the final exam for his Princeton students.

Not surprisingly, van der Vink wants his students to be both scientifically and politically literate. Besides his congressional fellowship, he has been an adviser to Congress and President Bill Clinton on nuclear arms control and treaty verification and a reviewer for the Council on Foreign Affairs to improve how scientific information is used to formulate international policy.

Following his work on Capitol Hill, van der Vink served as the director of IRIS for 14 years, where he was part of the initial planning process for EarthScope. He and his colleagues wanted a way to fill the information gaps in earth science, to look deeper into the earth with better resolution. Today geoscientists rely mostly on data from quakes, gathered only if the limited number of instruments is near the sites—conditions somewhat dependent on chance. Alternatively, they conduct the occasional active source experiment, in which sound waves are created and recorded—conditions somewhat dependent on funding.

Nearly a decade of discussions and workshops, attended by hundreds of geoscientists from more than 50 universities and a dozen state and federal agencies, led to the final EarthScope proposal, which was funded by Congress in 2003. "This will shift the future of earth science," van der Vink declares. "EarthScope will provide a new source of data for the next decade." He adds, "In the future, it may not be a science limited by [the lack of] data."

When it is complete in 2008, EarthScope's web of instruments will provide public data for geoscience, not unlike the way the human genome sequencing effort has for biology. EarthScope will gather information on three scales.

At the smallest scale, or fault level, is the San Andreas Fault Observatory at Depth (SAFOD), a four-kilometer-deep hole that is drilled into the mountains 300 kilometers south of San Francisco. SAFOD, set to be finished in 2007, has drilled down 3,200 meters next to the area that has ruptured seven times since 1857 (including the most recent magnitude 6 Parkfield earthquake in 2004). On August 2 drilling reached the fault zone, where scientists plan to install seismic instruments and take rock and fluid samples.

SEISMOMETER (*green sphere*) with supporting electronics (*black*) is sealed in a tank on a concrete pad to keep the device level. It is one of hundreds used by EarthScope.

At the middle scale is the plate boundary level, where the Plate Boundary Observatory (PBO) will image and characterize the slow deformation of the earth along the western U.S. and Alaska. Roughly 1,000 Global Positioning System (GPS) instruments will be installed in a gridlike pattern from the edge of the Pacific Coast to the Rocky Mountains and from Alaska to Mexico (not including Canada). Mounted on four four-meter-long legs for stability, these instruments will communicate with GPS satellites to measure movement on the surface of the earth. In addition, 175 devices will be clustered along fault zones and magmatic centers to measure strain on the surface and at depth. Crews finished installing instruments on Akutan Volcano in Alaska in early July.

At the largest, continental level is USArray, which consists of three seismic components. The first is "Bigfoot," a mobile grid of 400 seismometers that will stomp across the U.S. over the next 10 years. Within each footprint (which covers about one quarter of the country), a flexible array of 400 portable seismometers will provide dense instrument coverage in geophysically interesting places. To provide reference data for Bigfoot and the flexible array, 43 permanent seismic stations will be added to an existing array maintained by the U.S. Geological Survey. One Bigfoot station resides at the Wishkah Valley School near Aberdeen, Wash., where students now monitor its data collection online.

Van der Vink has helped keep the three components close to schedule, and none has run over budget. Still, EarthScope's future is uncertain. The original plan was to pay for EarthScope in pieces: $200 million over the first five years to build and install instruments (2003 to 2008); $13 million a year for the next 10 years to operate and maintain the facilities, plus an additional $100 million to fund actual science. Congress authorized the first $200 million in 2003 but unexpectedly delayed a decision on bankrolling EarthScope's next decade of operations and maintenance.

"We have no interest in building a facility if there are no funds to use it," says van der Vink, who admits that funding is somewhat out of his hands. "We have a team and structure in place to deliver EarthScope on time and on budget." He remains optimistic that the $40 million he and his colleagues requested in January will come through to kick-start the first year of EarthScope operations and maintenance in 2009.

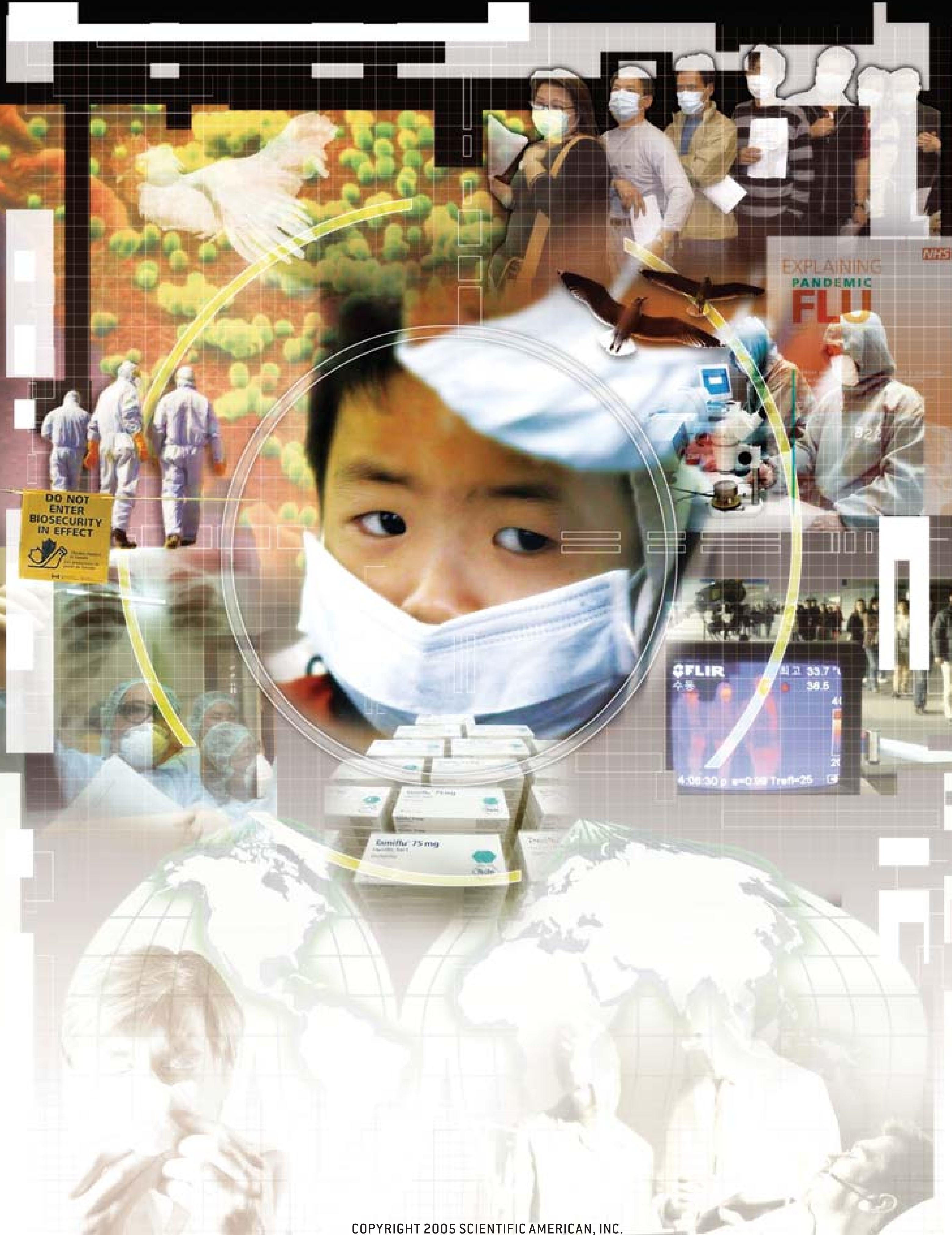
Many are confident in van der Vink's ability to keep the funds flowing. As head of IRIS, van der Vink obtained an extra $50 million through federal and private channels. "Greg is at his best portraying EarthScope's activity and potential on the Hill," says Thorne Lay, the USArray representative on the EarthScope Facilities executive committee.

Although he never set out to work in the world between science and politics, van der Vink confesses that it is a little bit addictive. "Intellectually," he says, "the work [in Congress] is most challenging because you have to find solutions that are not only scientifically accurate but socially, politically and economically acceptable as well."

When he finally complies with my request for a sports analogy describing EarthScope, funding and management challenges are not part of his description. Instead he echoes this multidisciplinary requirement with determination. "EarthScope needs to create the next generation of athletes to compete in the decathlons of the future—not in single events," he says. "The future lies in all-around athletes." SA

Krista West is based in Alaska, the fastest-deforming and most seismically active state in the nation.

EARTHSCOPE


DO NOT
ENTER
BIOSECURITY
IN EFFECT
EXPLAINING
PANDEMIC
FLU
NHS
FLIR
Tamiflu 75 mg

One day a highly contagious and lethal strain of influenza will sweep across all humanity, claiming millions of lives. It may arrive in months or not for years—but the next pandemic is inevitable. **Are we ready?**

Preparing for a Pandemic

By W. Wayt Gibbs and Christine Soares

When the levees collapsed in New Orleans, the faith of Americans in their government's ability to protect them against natural disasters crumbled as well. Michael Chertoff, the secretary of homeland security who led the federal response, called Hurricane Katrina and the flood it spawned an "ultracatastrophe" that "exceeded the foresight of the planners."

But in truth the failure was not a lack of foresight. Federal, state and local authorities had a plan for how governments would respond if a hurricane were to hit New Orleans with 120-mile-per-hour winds, raise a storm surge that overwhelmed levees and water pumps, and strand thousands inside the flooded city. Last year they even practiced it. Yet when Katrina struck, the execution of that plan was abysmal.

The lethargic, poorly coordinated and undersized response raises concerns about how nations would cope with a much larger and more lethal kind of natural disaster that scientists warn will occur, possibly soon: a pandemic of influenza. The threat of a flu pandemic is more ominous, and its parallels to Katrina more apt, than it might first seem. The routine seasonal upsurges of flu and of hurricanes engender a familiarity that easily leads to complacency and inadequate preparations for the "big one" that experts admonish is sure to come.

The most fundamental thing to understand about serious pandemic influenza is that, except at a molecular level, the disease bears little resemblance to the flu that we all get at some time. An influenza pandemic, by definition, occurs only when the influenza virus mutates into something dangerously unfamiliar to our immune systems and yet is able to jump from person to person through a sneeze, cough or touch.

Flu pandemics emerge unpredictably every generation or so, with the last three striking in 1918, 1957 and 1968. They get their start when one of the many influenza strains that constantly circulate in wild and domestic birds evolves into a form that infects us as well. That virus then adapts further or exchanges genes with a flu strain native to humans to produce a novel germ that is highly contagious among people.

Some pandemics are mild. But some are fierce. If the virus replicates much faster than the immune system learns to defend against it, it will cause severe and sometimes fatal illness, resulting in a pestilence that could easily claim more lives in a single year than AIDS has in 25. Epidemiologists have warned that the next pandemic could sicken one in every three people on the planet, hospitalize many of those and kill tens to hundreds of millions. The disease would spare no nation, race or income group. There would be no certain way to avoid infection.

Scientists cannot predict which influenza strain will cause a pandemic or when the next one will break out. They can warn only that another is bound to come and that the conditions now seem ripe, with a fierce strain of avian flu killing people in Asia and infecting birds in a rapid westward lunge toward Europe. That strain, influenza A (H5N1) does not yet pass readily from one person to another. But the virus is evolving, and some of the affected avian species have now begun their winter migrations.

As a sense of urgency grows, governments and health experts are working to bolster four substantial lines of defense against a pandemic: surveillance, vaccines, containment measures and medical treatments. The U.S. plans to release by October a pandemic preparedness plan that surveys the strength of each of these barricades. Some failures are inevitable, but the more robust those preparations are, the less humanity will suffer. The experience of Katrina forces a question: Will authorities be able to keep to their plans even when a large fraction of their own workforce is downed by the flu?

Surveillance: What Is Influenza Up to Now?

Our first defense against a new flu is the ability to see it coming. Three international agencies are coordinating the global effort to track H5N1 and other strains of influenza. The World Health Organization (WHO), with 110 influenza centers in 83 countries, monitors human cases. The World Organization for Animal Health (OIE, formerly the Office International des Épizooties) and the Food and Agriculture Organization (FAO) collect reports on outbreaks in birds and other animals. But even the managers of these surveillance nets acknowledge that they are still too porous and too slow.

Speed is of the essence when dealing with a fast-acting airborne virus such as influenza. Authorities probably have no realistic chance of halting a nascent pandemic unless they can contain it within 30 days [see "Rapid Response," on page 50]. The clock begins ticking the moment that the first victim of a pandemic-capable strain becomes contagious.

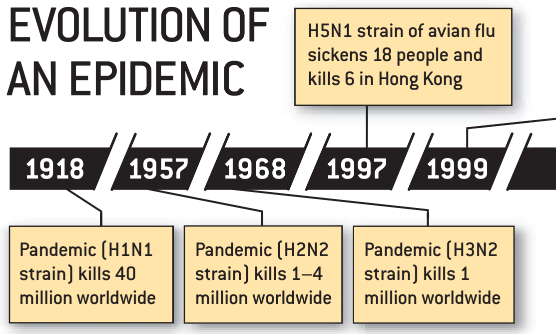
The only way to catch that emergence in time is to monitor constantly the spread of each outbreak and the evolution of the virus's abilities. The WHO assesses both those factors to determine where the world is in the pandemic cycle, which a new guide issued in April divides into six phases.

The self-limiting outbreaks of human H5N1 influenza seen so far bumped the alert level up to phase three, two steps removed from outright pandemic (phase six). Virologists try to obtain samples from every new H5N1 patient to scout for signs that the avian virus is adapting to infect humans more efficiently. It evolves in two ways: gradually through random mutation, and more rapidly as different strains of influenza swap genes inside a single animal or person [*see box on opposite page*].

The U.S. has a sophisticated flu surveillance system that funnels information on hospital visits for influenzalike illness, deaths from respiratory illness and influenza strains seen in public health laboratories to the Centers for Disease Control and Prevention in Atlanta. "But the system is not fast enough to take the isolation or quarantine action needed to manage avian flu," said Julie L. Gerberding, the CDC director, at a February conference. "So we have been broadening our networks of clinicians and veterinarians."

In several dozen cases where travelers to the U.S. from H5N1-affected Asian countries developed severe flulike symptoms, samples were rushed to the CDC, says Alexander Klimov of the CDC's influenza branch. "Within 40 hours of hospitalization we can say whether the patient has H5N1. Within another six hours we can analyze the genetic sequence of the hemagglutinin gene" to estimate the infectiousness of the strain. (The virus uses hemagglutinin to pry its way into cells.) A two-day test then reveals resistance to antiviral drugs, he says.

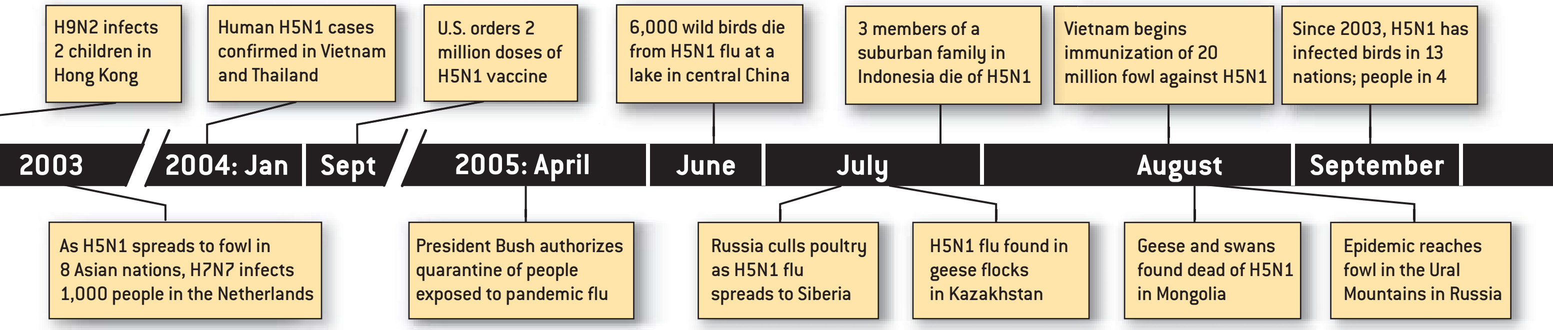
The next pandemic could break out anywhere, including in the U.S. But experts think it is most likely to appear first in Asia, as do most influenza strains that cause routine annual epidemics. Aquatic birds such as ducks and geese are the



Overview/*The Plan to Fight a New Flu*

- Scientists warn that a global epidemic caused by some newly evolved strain of influenza is inevitable and poses an enormous threat to public health.
- The pandemic could occur soon or not for years. H5N1 bird flu has killed more than 60 people in Asia, raising alarms. Even if that outbreak wanes, however, a global surveillance network must remain alert for other threatening strains.
- Flu shots matched to the new virus will arrive too late to prevent or slow the early stages of a pandemic, but rapid response with antiviral drugs might contain an emerging flu strain at its source temporarily, buying time for international preparations.
- Severity of disease will depend on the pandemic strain. In many places, drug supplies and other health resources will be overwhelmed.

PAGE 44: PHOTOILLUSTRATION BY JASON JAROSLAV COOK; NIBSC/PHOTO RESEARCHERS, INC. (*virus SEM*); VINCENT YU *AP Photo* (*line of people*); HOANG DINH NAM/AFP/GETTY IMAGES (*boy with mask* and *doctor with x-ray*); JUAN VRIJDAG *EPA* (*Tamiflu*); BYUN YOUNG-WOOK *EPA* (*monitor of scan*)



natural hosts for influenza, and in Asia many villagers reside cheek by bill with such animals. Surveillance in the region is still spotty, however, despite a slow trickle of assistance from the WHO, the CDC and other organizations.

A recent H5N1 outbreak in Indonesia illustrates both the problems and the progress. In a relatively wealthy suburb of Jakarta, the eight-year-old daughter of a government auditor fell ill in late June. A doctor gave her antibiotics, but her fever worsened, and she was hospitalized on June 28. A week later her father and one-year-old sister were also admitted to the hospital with fever and cough. The infant died on July 9, the father on July 12.

The next day an astute doctor alerted health authorities and sent blood and tissue samples to a U.S. Navy medical research unit in Jakarta. On July 14 the girl died; an internal report shows that on this same day Indonesian technicians in the naval laboratory determined that two of the three family members had H5N1 influenza. The government did not acknowledge this fact until July 22, however, after a WHO lab in Hong Kong definitively isolated the virus.

The health department then readied hospital wards for more flu patients, and I Nyoman Kandun, head of disease control for Indonesia, asked WHO staff to help investigate the outbreak. Had this been the onset of a pandemic, the 30-day containment window would by that time have closed. Kandun called off the investigation two weeks later. "We could not find a clue as to where these people got the infection," he says.

Local custom prohibited autopsies on the three victims. Klaus Stöhr of the WHO Global Influenza Program has complained that the near absence of autopsies on human H5N1 cases leaves many questions unanswered. Which organs does H5N1 infect? Which does it damage most? How strongly does the immune system respond?

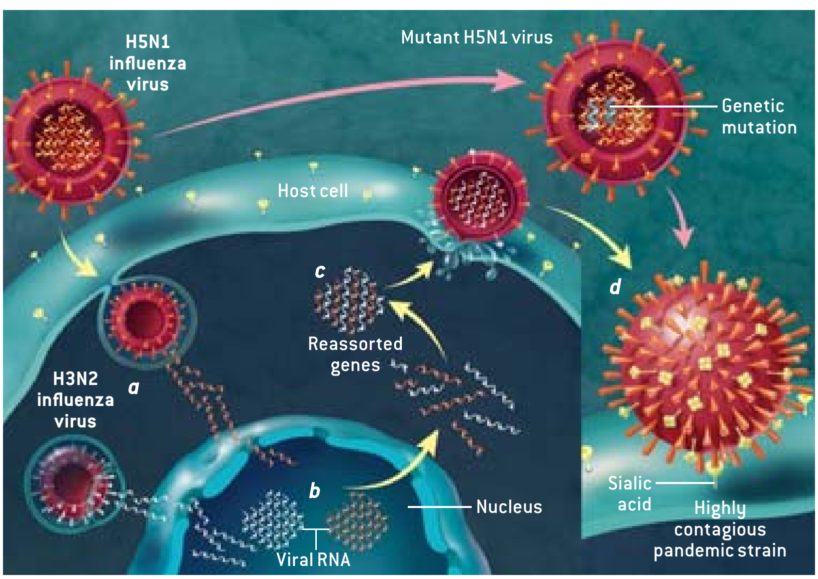
Virologists worry as well that they have too little information about the role of migratory birds in transmitting the disease across borders. In July domestic fowl infected with H5N1 began turning up in Siberia, then Kazakhstan, then Russia. How the birds caught the disease remains a mystery.

Frustrated with the many unanswered questions, Stöhr and other flu scientists have urged the creation of a global task force to supervise pandemic preparations. The OIE in August appealed for more money to support surveillance programs it is setting up with the FAO and the WHO.

"We clearly need to improve our ability to detect the virus," says Bruce G. Gellin, who coordinates U.S. pandemic planning as head of the National Vaccine Program Office at the U.S. Department of Health and Human Services (HHS). "We need to invest in these countries to help them, because doing so helps everybody."

HOW A PANDEMIC STRAIN EMERGES

Avian strains of influenza A, such as H5N1, can evolve via two paths into pandemic-capable virus (able to bind readily to sialic acid on human cells). Genetic mutations and natural selection can render the virus more efficient at entering human cells (*pink path*). Alternatively (*yellow path*), two strains of influenza may infect the same cell (*a*) and release viral RNA, which replicates inside the cell nucleus (*b*). RNA from the two strains can then mix to create a set of "reassorted" genes (*c*) that give rise to a novel and highly contagious pandemic strain.



W. WAYT GIBBS (*timeline*); ALICE Y. CHEN (*illustration*)

Vaccines: Who Will Get Them—and How Quickly?

Pandemics of smallpox and polio once ravaged humanity, but widespread immunization drove those diseases to the brink of extinction. Unfortunately, that strategy will not work against influenza—at least not without a major advance in vaccine technology.

Indeed, if an influenza pandemic arrives soon, vaccines against the emergent strain will be agonizingly slow to arrive and frustratingly short in supply. Biology, economics and complacency all contribute to the problem.

Many influenza strains circulate at once, and each is constantly evolving. "The better the match between the vaccine and the disease virus, the better the immune system can defend against the virus," Gellin explains. So every year manufacturers fashion a new vaccine against the three most threatening strains. Biologists first isolate the virus and then modify it using a process called reverse genetics to make a seed virus. In vaccine factories, robots inject the seed virus into fertilized eggs laid by hens bred under hygienic conditions. The pathogen replicates wildly inside the eggs.

Vaccine for flu shots is made by chemically dissecting the virus and extracting the key proteins, called antigens, that stimulate the human immune system to make the appropriate antibodies. A different kind of vaccine, one inhaled rather than injected, incorporates live virus that has been damaged enough that it can infect but not sicken. The process requires six months to transform viral isolates into initial vials of vaccine.

Because people will have had no prior exposure to a pandemic strain of influenza, everyone will need two doses: a primer and then a booster about four weeks later. So even those first in line for vaccines are unlikely to develop immunity until at least seven or eight months following the start of a pandemic.

And there will undoubtedly be a line. Total worldwide production of flu vaccine amounts to roughly 300 million doses a year. Most of that is made in Europe; only two plants operate in the U.S. Last winter, when contamination shut down a Chiron facility in Britain, Sanofi Pasteur and MedImmune pulled out all stops on their American lines—and produced 61 million doses. The CDC recommends annual flu immunization for high-risk groups that in the U.S. include some 185 million people.

Sanofi now runs its plant at full bore 365 days a year. In July it broke ground for a new facility in Pennsylvania that will double its output—in 2009. Even in the face of an emergency, "it would be very hard to compress that timeline," says James T. Matthews, who sits on Sanofi's pandemic-planning working group. He says it would not be feasible to convert factories for other kinds of vaccines over to make flu shots.

Pascale Wortley of the CDC's National Immunization Program raises another concern. Pandemics typically overlap with the normal flu season, she notes, and flu vaccine plants can make only one strain at a time. Sanofi spokesman Len Lavenda agrees that "we could face a Sophie's choice: whether to stop producing the annual vaccine in order to start producing the pandemic vaccine."

MedImmune aims to scale up production of its inhalable vaccine from about two million doses a year to 40 million doses by 2007. But Gellin cautions that it might be too risky to distribute live vaccine derived from a pandemic strain. There is a small chance, he says, that the virus in the vaccine could exchange genes with a "normal" flu virus in a person and generate an even more dangerous strain of influenza.

Because delays and shortages in producing vaccine against a pandemic are unavoidable, one of the most important functions of national pandemic plans is to push political leaders to decide in advance which groups will be the first to receive vaccine and how the government will enforce its rationing. The U.S. national vaccine advisory committee recommended in July that the first shots to roll off the lines should go to key government leaders, medical caregivers, workers in flu vaccine and drug factories, pregnant women, and those infants, elderly and ill people who are already in the high-priority group for an-

EGG-BASED PROCESS for making flu vaccine imposes bottlenecks that will delay the release of a pandemic vaccine by six months or more. Supplies will fall far short of the demand.

NIAID

nual flu shots. That top tier includes about 46 million Americans.

Among CDC planners, Wortley says, "there is a strong feeling that we ought to say beforehand that the government will purchase some amount of vaccine to guarantee equitable distribution." Australia, Britain, France and other European governments are working out advance contracts with vaccine producers to do just that. The U.S., so far, has not.

In principle, governments could work around these supply difficulties by stockpiling vaccine. They would have to continually update their stocks as new strains of influenza threatened to go global; even doing so, the reserves would probably always be a step or two behind the disease. Nevertheless, Wortley says, "it makes sense to have H5N1 vaccine on hand, because even if it is not an exact match, it probably would afford some amount of protection" if the H5N1 strain evolved to cause a pandemic.

To that end, the U.S. National Institute of Allergy and Infectious Diseases (NIAID) last year distributed an H5N1 seed virus created from a victim in Vietnam by scientists at St. Jude Children's Research Hospital in Memphis. The HHS then placed an order with Sanofi for two million doses of vaccine against that strain. Human trials began in March, and "the preliminary results from the clinical trial indicate that the vaccine would be protective," says NIAID director Anthony S. Fauci. "HHS Secretary Michael Leavitt is trying to negotiate to get up to 20 million doses," he adds. (Leavitt announced in September that HHS had increased its H5N1 vaccine order by $100 million.) According to Gellin, current vaccine producers could contribute at most 15 million to 20 million doses a year to the U.S. stockpile.

Those numbers are probably overoptimistic, however. The trial tested four different concentrations of antigen. A typical annual flu shot has 45 micrograms of protein and covers three strains of influenza. Officials had expected that 30 micrograms of H5N1 antigen—two shots, with 15 micrograms in each—would be enough to induce immunity. But the preliminary trial results suggest that 180 micrograms of antigen are needed to immunize one person.

NEW VACCINE TECHNOLOGIES

Researchers in industry and academia are testing new immunization methods that would stretch the limited supply of vaccines to cover more people. They are also developing technologies that could allow vaccine production to increase rapidly in an emergency.

Technology	Benefits	Readiness	Companies
Intradermal injectors	Delivering flu vaccine into the skin rather than muscle might cut the required dose per shot by a factor of five	Clinical trials show promise, but few nurses and doctors are trained in the procedure	Iomai, GlaxoSmithKline
Adjuvants	Chemical additives called adjuvants can increase the immune response, so that less protein is needed per shot	One such vaccine is licensed in Europe. Others are in active development	Iomai, Chiron, GlaxoSmithKline
Cell-cultured vaccines	Growing influenza virus for vaccine in cell-filled bioreactors, rather than in eggs, would enable faster increases in production if a flu pandemic broke out	Chiron is conducting a large-scale trial in Europe. Sanofi Pasteur and Crucell are developing a process for the U.S.	Chiron, Baxter, Sanofi Pasteur, Crucell, Protein Sciences
DNA vaccines	Gold particles coated with viral DNA could be injected into the skin with a jet of air. Production of DNA vaccines against a new strain could begin in weeks, rather than months. Stockpiles would last years without refrigeration	No DNA vaccine has yet been proved effective in humans. PowderMed expects results from a small-scale trial of an H5N1 DNA vaccine in late 2006	PowderMed, Vical
All-strain vaccines	A vaccine that raises immunity against a viral protein that rarely mutates might thwart every strain of influenza. Stockpiles could then reliably defend against a pandemic	Acambis began developing a vaccine against the M2e antigen this past summer	Acambis

An order for 20 million conventional doses may thus actually yield only enough H5N1 vaccine for about 3.3 million people. The true number could be even lower, because H5 strains grow poorly in eggs, so each batch yields less of the active antigen than usual. This grim picture may brighten, however, when NIAID analyzes the final results from the trial. It may also be possible to extend vaccine supplies with the use of adjuvants (substances added to vaccines to increase the immune response they induce) or new immunization approaches, such as injecting the vaccine into the skin rather than into muscle.

Caching large amounts of prepandemic vaccine, though not impossible, is clearly a challenge. Vaccines expire after a few years. At current production rates, a stockpile would never grow to the 228 million doses needed to cover the three highest priority groups, let alone to the roughly 600 million doses that would be needed to vaccinate everyone in the U.S. Other nations face similar limitations.

The primary reason that capacity is so tight, Matthews explains, is that vaccine makers aim only to meet the demand for annual immunizations when making business decisions. "We really don't see the pandemic itself as a market opportunity," he says.

To raise manufacturers' interest, "we need to offer a number of incentives, ranging from liability insurance to better profit margins to guaranteed purchases," Fauci acknowledges. Long-term solutions, Gellin predicts, may come from new technologies that allow vaccines to be made more efficiently, to be scaled up more rapidly, to be effective at much lower doses and perhaps to work equally well on all strains of influenza.

Rapid Response: Could a Pandemic Be Stopped?

As recently as 1999, WHO had a simple definition for when a flu pandemic began: with confirmation that a new virus was spreading between people in at least one country. Thereafter, stopping the flu's lightning-fast expansion was unthinkable—or so it then seemed. But because of recent advances in the state of disease surveillance and antiviral drugs, the latest version of WHO's guidelines recognizes a period on the cusp of the pandemic when a flu virus ready to burst on the world might instead be intercepted and restrained, if not stamped out.

Computer models and common sense indicate that a containment effort would have to be exceptionally swift and efficient. Flu moves with extraordinary speed because it has such a short incubation period—just two days after infection by the virus, a person may start showing symptoms and shedding virus particles that can infect others. Some people may become infectious a day before their symptoms appear. In contrast, people infected by the SARS coronavirus that emerged from China in 2003 took as long as 10 days to become infectious, giving health workers ample time to trace and isolate their contacts before they, too, could spread the disease.

Contact tracing and isolation alone could never contain flu, public health experts say. But computer-simulation results published in August showed when up to 30 million doses of antiviral drugs and a low-efficacy vaccine were added to the interventions a chance emerged to thwart a potential pandemic.

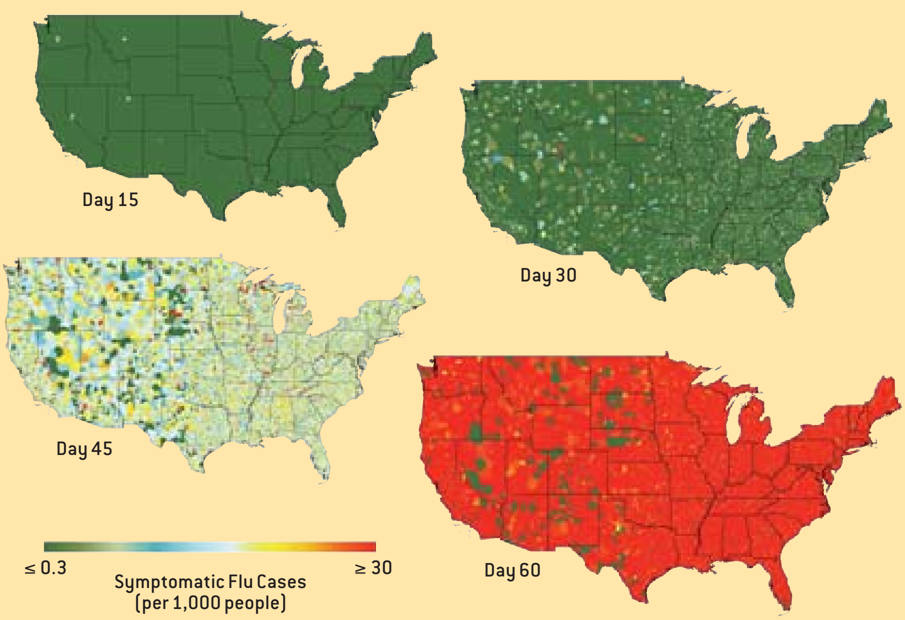
Conditions would have to be nearly ideal. Modeling a population of 85 million based on the demographics and geography of Thailand, Neil M. Ferguson of Imperial College London found that health workers would have at most 30 days from the start of person-to-person viral transmission to deploy antivirals as both treatment and preventives wherever outbreaks were detected.

But even after seeing the model results earlier this year, WHO officials expressed doubt that surveillance in parts of Asia is reliable enough to catch a budding epidemic in time. In practice, confirmation of some human H5N1 cases has taken more than 20 days, WHO flu chief Stöhr warned a gathering of experts in Washington, D.C., this past April. That leaves just a narrow window in which to deliver the drugs to remote areas and dispense them to as many as one million people.

Partial immunity in the population could buy more time, however, according to Ira M. Longini, Jr., of Emory University. He, too, modeled intervention with antivirals in a smaller community based on Thai demographic data, with outcomes similar to Ferguson's. But Longini added scenarios in which people had been vaccinated in advance. He assumed that an existing vaccine, such as the H5N1 prototype version some countries have already developed, would not perfectly match a new variant of the virus, so his model's vaccinees were only 30 percent less likely to be infected. Still, their reduced susceptibility made containing even a highly infectious flu strain possible in simulations. NIAID director Fauci has said that the U.S. and other nations with H5N1 vaccine are still considering whether to direct it toward prevention in the region where a human-adapted version of that virus is most likely to emerge—even if that means less would remain for their own citizens. "If we're smart, we would," Longini says.

Pandemic Flu Hits the U.S.

A simulation created by researchers from Los Alamos National Laboratory and Emory University shows the first wave of a pandemic spreading rapidly with no vaccine or antiviral drugs employed to slow it down. Colors represent the number of symptomatic flu cases per 1,000 people (*see scale*). Starting with 40 infected people on the first day, nationwide cases peak around day 60, and the wave subsides after four months with 33 percent of the population having become sick. The scientists are also modeling potential interventions with drugs and vaccines to learn if travel restrictions, quarantines and other disruptive disease-control strategies could be avoided.



TIMOTHY C. GERMANN, KAI KADAU AND CATHERINE A. MACKEN *Los Alamos National Laboratory;* AND IRA M. LONGINI, JR. *Emory University/Models of Infectious Disease Agent Study, National Institute of General Medical Sciences*

COMMAND CENTER at the U.S. Department of Health and Human Services in Washington, D.C., would be used to track the spread of a flu pandemic. From this location, HHS would coordinate the activities of its divisions, including the CDC and the NIH, and share information with state and federal government agencies such as the Department of Homeland Security.

DEPARTMENT OF HEALTH AND HUMAN SERVICES

Based on patterns of past pandemics, experts expect that once a new strain breaks loose, it will circle the globe in two or three waves, each potentially lasting several months [*see box on opposite page*] but peaking in individual communities about five weeks after its arrival. The waves could be separated by as long as a season: if the first hit in springtime, the second might not begin until late summer or early fall. Because meaningful amounts of vaccine tailored to the pandemic strain will not emerge from factories for some six months, government planners are especially concerned with bracing for the first wave.

Once a pandemic goes global, responses will vary locally as individual countries with differing resources make choices based on political priorities as much as on science. Prophylactic use of antivirals is an option for a handful of countries able to afford drug stockpiles, though not a very practical one. No nation has enough of the drugs at present to protect a significant fraction of its population for months. Moreover, such prolonged use has never been tested and could cause unforeseen problems. For these reasons, the U.K. declared this past July that it would use its pandemic stockpile primarily for treating patients rather than for protecting the uninfected. The U.S., Canada and several other countries are still working out their priorities for who will receive antivirals and when.

For most countries there will be no choice: what the WHO calls nonpharmaceutical interventions will have to be their primary defense. Although the effectiveness of such measures has not been extensively researched, the WHO gathered flu specialists in Geneva in March 2004 to try to determine which actions medical evidence does support. Screening incoming travelers for flu symptoms, for instance, "lacks proven health benefit," the group concluded, although they acknowledged that countries might do it anyway to promote public confidence. Similarly, they were skeptical that public fever screening, fever hotlines or fever clinics would do much to slow the spread of the disease.

The experts recommended surgical masks for flu patients and health workers exposed to those patients. For the healthy, hand washing offers more protection than wearing masks in public, because people can be exposed to the virus at home, at work and by touching contaminated surfaces—including the surface of a mask.

Traditional "social distancing" measures, such as banning public gatherings or shutting down mass transit, will have to be guided by what epidemiologists find once the pandemic is under way. If children are especially susceptible to the virus, for example—as was the case in 1957 and 1968—or if they are found to be an important source of community spread, then governments may consider closing schools.

Treatment: What Can Be Done for the Sick?

If two billion become sick, will 10 million die? Or 100 million? Public health specialists around the world are struggling to quantify the human toll of a future flu pandemic. Casualty estimates vary so widely because until it strikes, no one can be certain whether the next pandemic strain will be mild, like the 1968 virus that some flu researchers call a "wimp"; moderately severe, like the 1957 pandemic strain; or a stone-cold killer, like the "Great Influenza" of 1918.

For now, planners are going by rules of thumb: because no one would have immunity to a new strain, they expect 50 percent of the population to be infected by the virus. Depending on its virulence, between one third and two thirds of those people will become sick, yielding a clinical attack rate of 15 to 35 percent of the whole population. Many governments are therefore trying to prepare for a middle-ground estimate that 25 percent of their entire nation will fall ill.

No government is ready now. In the U.S., where states have primary responsibility for their residents' health, the Trust for America's Health (TFAH) estimates that a "severe" pandemic virus sickening 25 percent of the population could translate into 4.7 million Americans needing hospitalization. The TFAH notes that the country currently

AVIAN FLU PATIENTS at a hospital in Hanoi, Vietnam, this past March were a man (*left*), 21, in critical condition, and his sister, 14. Many of the most severe illnesses and deaths from H5N1 infection have been among previously healthy young adults and children.

has fewer than one million staffed hospital beds.

For frontline health workers, a pandemic's severity will boil down to the sheer number of patients and the types of illness they are suffering. These, in turn, could depend on both inherent properties of the virus and susceptibility of various subpopulations to it, according to Maryland's pandemic planner, Jean Taylor. A so-called mild pandemic, for example, might resemble seasonal flu but with far larger numbers infected.

Ordinarily, those hardest hit by annual flu are people who have complications of chronic diseases, as well as the very young, the very old and others with weak immune systems. The greatest cause of seasonal flu-related deaths is pneumonia brought on by bacteria that invade after flu has depleted the body's defenses, not by the flu virus itself. Modeling a pandemic with similar qualities, Dutch national health agency researchers found that hospitalizations might be reduced by 31 percent merely by vaccinating the usual risk groups against bacterial pneumonia in advance.

In contrast, the 1918 pandemic strain was most lethal to otherwise healthy young adults in their 20s and 30s, in part because their immune systems were so hardy. Researchers studying that virus have discovered that it suppresses early immune responses, such as the body's release of interferon, which normally primes cells to resist attack. At the same time, the virus provokes an extreme immune overreaction known as a cytokine storm, in which signaling molecules called cytokines summon a ferocious assault on the lungs by immune cells.

Doctors facing the same phenomenon in SARS patients tried to quell the storm by administering interferon and cytokine-suppressing corticosteroids. If the devastating cascade could not be stopped in time, one Hong Kong physician reported, the patients' lungs became increasingly inflamed and so choked with dead tissue that pressurized ventilation was needed to get enough oxygen to the bloodstream.

Nothing about the H5N1 virus in its current form offers reason to hope that it would produce a wimpy pandemic, according to Frederick G. Hayden, a University of Virginia virologist who is advising WHO on treating avian flu victims. "Unless this virus changes dramatically in pathogenicity," he asserts, "we will be confronted with a very lethal strain." Many H5N1 casualties have suffered acute pneumonia deep in the lower lungs caused by the virus itself, Hayden says, and in some cases blood tests indicated unusual cytokine activity. But the virus is not always consistent. In some patients, it also seems to multiply in the gut, producing severe diarrhea. And it is believed to have infected the brains of two Vietnamese children who died of encephalitis without any respiratory symptoms.

Antiviral drugs that fight the virus directly are the optimal treatment, but many H5N1 patients have arrived on doctors' doorsteps too late for the drugs to do much good. The version of the strain that has infected most human victims is also resistant to an older class of antivirals called amantadines, possibly as a result of those drugs having been given to poultry in parts of Asia. Laboratory experiments indicate that H5N1 is still susceptible to a newer class of antivirals called neuraminidase inhibitors (NI) that includes two products, oseltamivir and zanamivir, currently on the market under the brand names Tamiflu and Relenza. The former comes in pill form; the latter is a powder delivered by inhaler. To be effective against seasonal flu strains, either drug must be taken within 48 hours of symptoms appearing.

The only formal test of the drugs against H5N1 infection, however, has been in mice. Robert G. Webster of St. Jude Children's Research Hospital reported in July that a mouse equivalent of the normal human dose of two Tamiflu pills a day eventually subdued the virus, but the mice required treatment for eight

OSELTAMIVIR, sold as Tamiflu, is manufactured in a complex multistep process that takes nearly a year. Current stockpile orders will require several years to fill, and generic versions would be difficult to create for an emergency.

EPA (*top*); JUAN VRIJDAG *EPA* (*bottom*)

NEW FLU DRUGS

Today's flu antivirals disable specific proteins on the virus's surface—either M2 (drugs known as amantadines) or neuraminidase (zanamivir and oseltamivir). Some new drugs in development are improved neuraminidase inhibitors. Other novel approaches include blocking the virus's entry into host cells or hobbling its ability to function once inside.

Approach	Drugs	Benefits	Readiness
Inhibition of neuraminidase protein, which the virus uses to detach from one cell and infect another	Peramivir (BioCryst Pharmaceuticals); CS-8958 (Biota/Sankyo)	Neuraminidase inhibitors have fewer side effects and are less likely to provoke viral resistance than the older amantadines. CS-8958 is a long-acting formulation that clings inside lungs for up to a week	Peramivir reached lungs inefficiently in clinical trials of a pill form; trials of intravenous delivery may occur in 2006; initial safety trials are complete on CS-8958
Inhibition of viral attachment to cells	Fludase (NexBio)	Because it blocks the sialic acid receptor that flu viruses use to enter host cells, Fludase should be equally effective on all flu strains	Clinical trials are planned for 2006
Stimulation of RNA interference mechanism	G00101 (Galenea); unnamed (Alnylam Pharmaceuticals)	Uses DNA to activate a built-in defense mechanism in cells, marking viral instructions for destruction. G001498 demonstrated effective against avian H5 and H7 flu viruses in mice	Clinical trials are expected within 18 months
Antisense DNA to block viral genes	Neugene (AVI BioPharma)	Synthetic strands of DNA bind to viral RNA that instructs the host cell to make more virus copies. The strategy should be effective against most strains	Animal testing is scheduled for 2006

days rather than the usual five. The WHO is organizing studies of future H5N1 victims to determine the correct amount for people.

Even at the standard dosage, however, treating 25 percent of the U.S. population would require considerably more Tamiflu, or its equivalent, than the 22 million treatment courses the U.S. Department of Health and Human Services planned to stockpile as of September. An advisory committee has suggested a minimum U.S. stockpile of 40 million treatment courses (400 million pills). Ninety million courses would be enough for a third of the population, and 130 million would allow the drugs to also be used to protect health workers and other essential personnel, the committee concluded.

Hayden hopes that before a pandemic strikes, a third NI called peramivir may be approved for intravenous use in hospitalized flu patients. Long-acting NIs might one day be ideal for stockpiling because a single dose would suffice for treatment or offer a week's worth of prevention.

These additional drugs, like a variety of newer approaches to fighting flu [*see box above*], all have to pass clinical testing before they can be counted on in a pandemic. Researchers would also like to study other treatments that directly modulate immune system responses in flu patients. Health workers will need every weapon they can get if the enemy they face is as deadly as H5N1.

Fatality rates in diagnosed H5N1 victims are running about 50 percent. Even if that fell to 5 percent as the virus traded virulence for transmissibility among people, Hayden warns, "it would still represent a death rate double [that of] 1918, and that's despite modern technologies like antibiotics and ventilators." Expressing the worry of most flu experts at this pivotal moment for public health, he cautions that "we're well behind the curve in terms of having plans in place and having the interventions available."

Never before has the world been able to see a flu pandemic on the horizon or had so many possible tools to minimize its impact once it arrives. Some mysteries do remain as scientists watch the evolution of a potentially pandemic virus for the first time, but the past makes one thing certain: even if the dreaded H5N1 never morphs into a form that can spread easily between people, some other flu virus surely will. The stronger our defenses, the better we will weather the storm when it strikes. "We have only one enemy," CDC director Gerberding has said repeatedly, "and that is complacency." SA

W. Wayt Gibbs is senior writer. Christine Soares is a staff writer and editor.

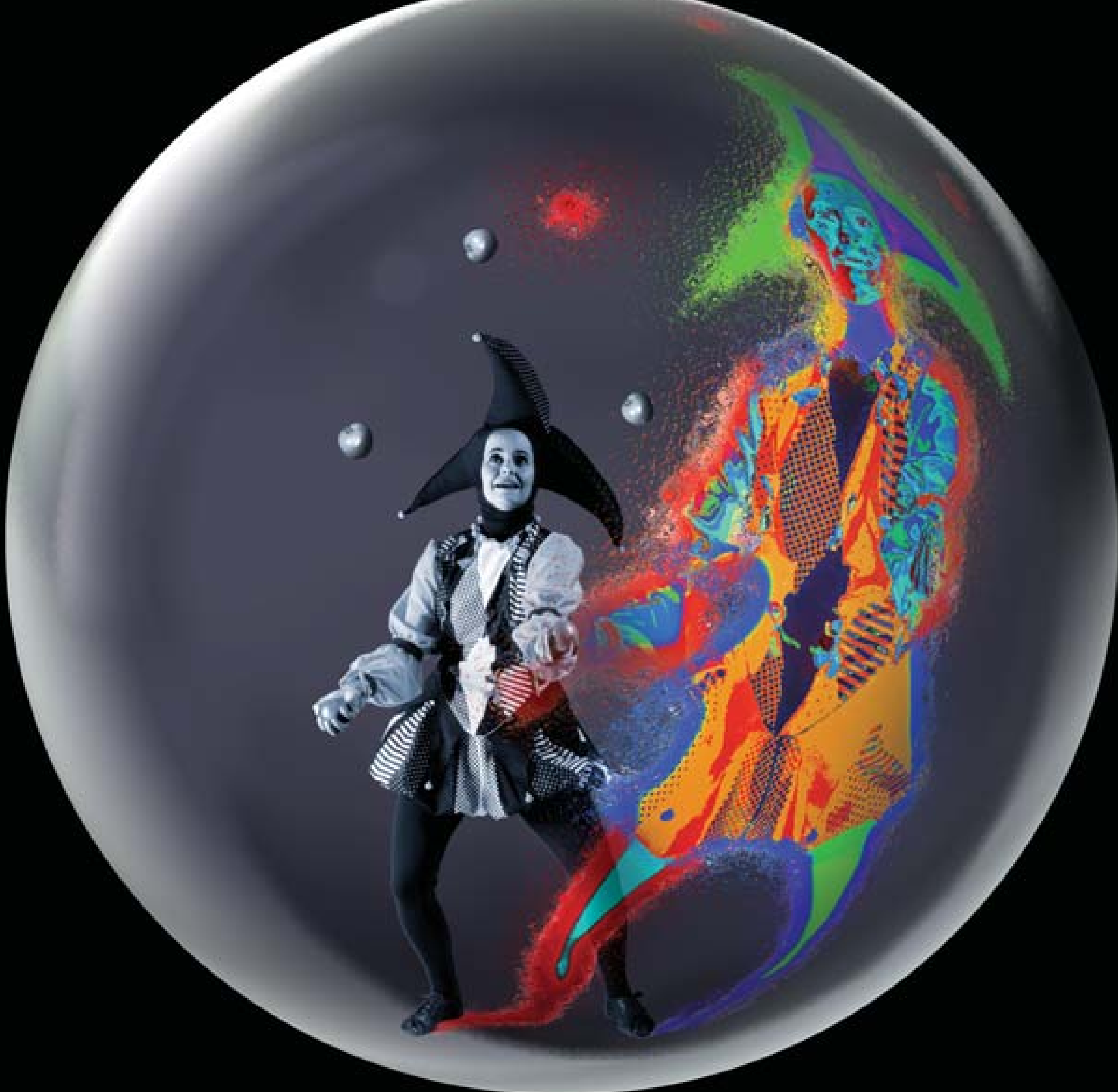
MORE TO EXPLORE

The Great Influenza. Revised edition. John M. Barry. Penguin Books, 2005.

John R. LaMontagne Memorial Symposium on Pandemic Influenza Research: Meeting Proceedings. Institute of Medicine. National Academies Press, 2005.

WHO Global Influenza Preparedness Plan. WHO Department of Communicable Disease Surveillance and Response Global Influenza Program, 2005. **www.who.int/csr/resources/publications/influenza/WHO_CDS_CSR_GIP_2005_5/en/index.html**

Pandemic influenza Web site of the U.S. Department of Health and Human Services, National Vaccine Program Office: **www.hhs.gov/nvpo/pandemics/index.html**



HOLOGRAPHIC THEORY relates one set of physical laws acting in a volume with a different set of physical laws acting on a boundary surface, as represented here by the juggler and her colorful two-dimensional image. The surface laws involve quantum particles that have "color" charges and interact very like the quarks and gluons of standard particle physics. The interior laws are a form of string theory and include the force of gravity (experienced by the juggler), which is hard to describe in terms of quantum mechanics. Nevertheless, the physics on the surface and in the interior are completely equivalent, despite their radically different descriptions.

The Illusion of Gravity

The force of gravity and one of the dimensions of space might be generated out of the peculiar interactions of particles and fields existing in a lower-dimensional realm

By Juan Maldacena

PHOTOGRAPH BY CARY WOLINSKY; PHOTOILLUSTRATION BY BRYAN CHRISTIE DESIGN

Three spatial dimensions are visible all around us—up/down, left/right, forward/backward. Add time to the mix, and the result is a four-dimensional blending of space and time known as spacetime. Thus, we live in a four-dimensional universe. Or do we?

Amazingly, some new theories of physics predict that one of the three dimensions of space could be a kind of an illusion—that in actuality all the particles and fields that make up reality are moving about in a two-dimensional realm like the Flatland of Edwin A. Abbott. Gravity, too, would be part of the illusion: a force that is not present in the two-dimensional world but that materializes along with the emergence of the illusory third dimension.

Or, more precisely, the theories predict that the number of dimensions in reality could be a matter of perspective: physicists could choose to describe reality as obeying one set of laws (including gravity) in three dimensions or, equivalently, as obeying a different set of laws that operates in two dimensions (in the absence of gravity). Despite the radically different descriptions, both theories would describe everything that we see and all the data we could gather about how the universe works. We would have no way to determine which theory was "really" true.

Such a scenario strains the imagination. Yet an analogous phenomenon occurs in everyday life. A hologram is a two-dimensional object, but when viewed under the correct lighting conditions it produces a fully three-dimensional image. All the information describing the three-dimensional image is in essence encoded in the two-dimensional hologram. Similarly, according to the new physics theories, the entire universe could be a kind of a hologram [see "Information in the Holographic Universe," by Jacob D. Bekenstein; Scientific American, August 2003].

The holographic description is more than just an intellectual or philosophical curiosity. A computation that might be very difficult in one realm can turn out to be relatively straightforward in the other, thereby turning some intractable problems of physics into ones that are easily solved. For example, the theory seems useful in analyzing a recent experimental high-energy physics result. Moreover, the holographic theories offer a fresh way to begin constructing a quantum theory of gravity—a theory of gravity that respects the principles of quantum mechanics. A quantum theory of gravity is a key ingredient in any effort to unify all the forces of nature, and it is needed to explain both what goes on in black holes and what happened in the nanoseconds after the big bang. The holographic theories provide potential resolutions of profound mysteries that have dogged attempts to understand how a theory of quantum gravity could work.

A Difficult Marriage

A QUANTUM THEORY of gravity is a holy grail for a certain breed of physicist because all physics except for gravity is well described by quantum laws. The quantum description of physics represents an entire paradigm for physical theories, and it makes no sense for one theory, gravity, to fail to conform to it. Now about 80 years old, quantum mechanics was first developed to describe the behavior of particles and forces in the atomic and subatomic realms. It is at those size scales that quantum effects become significant. In quantum theories, objects do not have definite positions and velocities but instead are described by probabilities and waves that occupy regions of space. In a quantum world, at the most fundamental level everything is in a state of constant flux, even "empty" space, which is in fact filled with virtual particles that perpetually pop in and out of existence.

In contrast, physicists' best theory of gravity, general relativity, is an inherently classical (that is, nonquantum) theory. Einstein's magnum opus, general relativity explains that concentrations of matter or energy cause spacetime to curve and that this curvature deflects the trajectories of particles, just as should happen for particles in a gravitational field. General relativity is a beautiful theory, and many of its predictions have been tested to great accuracy.

In a classical theory such as general relativity, objects have definite locations and velocities, like the planets orbiting the sun. One can plug those locations and velocities (and the masses of the objects) into the equations of general relativity and deduce the curvature of spacetime and from that deduce the effects of gravity on the objects' trajectories. Furthermore, empty spacetime is perfectly smooth no matter how closely one examines it—a seamless arena in which matter and energy can play out their lives.

The problem in devising a quantum version of general relativity is not just that on the scale of atoms and electrons, particles do not have definite locations and velocities. To make matters worse, at the even tinier scale delineated by the Planck length (10^{-33} centimeter), quantum principles imply that spacetime itself should be a seething foam, similar to the sea of virtual particles that fills empty space. When matter and spacetime are so protean, what do the equations of general relativity predict? The answer is that the equations are no longer adequate. If we assume that matter obeys the laws of quantum mechanics and gravity obeys the laws of general relativity, we end up with mathematical contradictions. A quantum theory of gravity (one that fits within the paradigm of quantum theories) is needed.

In most situations, the contradictory requirements of quantum mechanics and general relativity are not a problem, because either the quantum effects or the gravitational effects are so small that they can be neglected or dealt with by approximations. When the curvature of spacetime is very large, however, the quantum aspects of gravity become significant. It takes a very large mass or a great concentration of mass to produce much spacetime curvature. Even the curvature produced near the sun is exceedingly small compared with the amount needed for quantum gravity effects to become apparent.

Though these effects are completely negligible now, they were very important in the beginning of the big bang, which is why a quantum theory of gravity is needed to describe how the big

A quantum theory of gravity will probably provide us with an entirely new perspective on what spacetime is.

Overview/*Equivalent Worlds*

- According to a remarkable theory, a universe that exists in two dimensions and is without gravity may be completely equivalent to a three-dimensional universe with gravity. The three-dimensional universe would emerge from the physics of the two-dimensional universe somewhat like a holographic image arising from a hologram.
- The two-dimensional universe exists on the boundary of the three-dimensional universe. The physics on the boundary looks like strongly interacting quarks and gluons. The physics on the interior includes a quantum theory of gravity—something that string theorists have been developing for decades.
- The equivalence provides a new way to understand properties of black holes, which require a suitable melding of quantum mechanics and gravity. The mathematics of the theory has not yet been rigorously proved, but it seems useful in analyzing a recent experimental high-energy physics result.

bang started. Such a theory is also important for understanding what happens at the center of black holes, because matter there is crushed into a region of extremely high curvature. Because gravity involves spacetime curvature, a quantum gravity theory will also be a theory of quantum spacetime; it should clarify what constitutes the "spacetime foam" mentioned earlier, and it will probably provide us with an entirely new perspective on what spacetime is at the deepest level of reality.

A very promising approach to a quantum theory of gravity is string theory, which some theoretical physicists have been exploring since the 1970s. String theory overcomes some of the obstacles to building a logically consistent quantum theory of gravity. String theory, however, is still under construction and is not yet fully understood. That is, we string theorists have some approximate equations for strings, but we do not know the exact equations. We also do not know the guiding underlying principle that explains the form of the equations, and there are innumerable physical quantities that we do not know how to compute from the equations.

In recent years string theorists have obtained many interesting and surprising results, giving novel ways of understanding what a quantum spacetime is like. I will not describe string theory in much detail here [see "The String Theory Landscape," by Raphael Bousso and Joseph Polchinski; Scientific American, September 2004] but instead will focus on one of the most exciting recent developments emerging from string theory research, which led to a complete, logically consistent, quantum description of gravity in what are called negatively curved spacetimes—the first such description ever developed. For these spacetimes, holographic theories appear to be true.

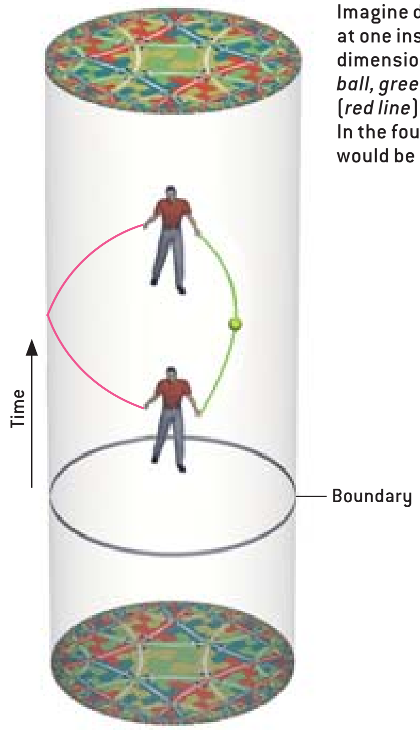
Negatively Curved Spacetimes

All of us are familiar with Euclidean geometry, where space is flat (that is, not curved). It is the geometry of figures drawn on flat sheets of paper. To a very

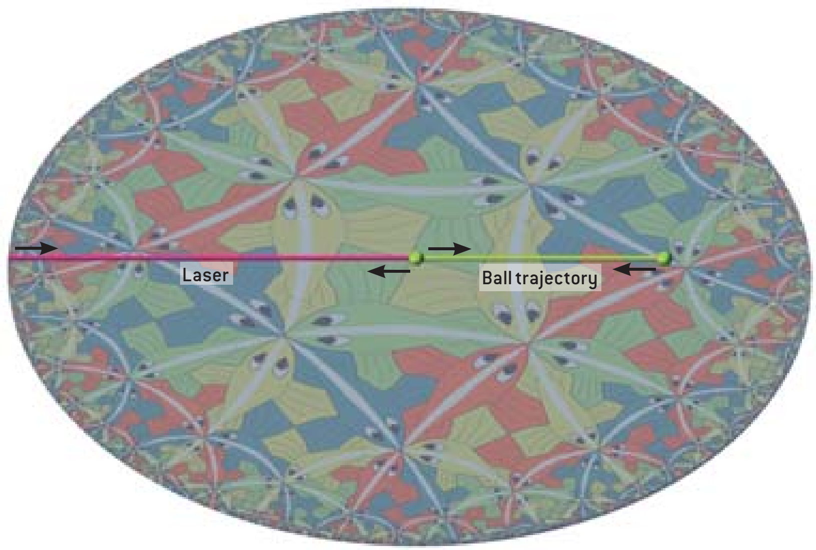
HYPERBOLIC SPACE is depicted in this M. C. Escher drawing (*above*). Each fish is actually the same size, and the circular boundary is infinitely far from the center of the disk. The projection from true hyperbolic space to this representation of it squashes the distant fish to fit the infinite space inside the finite circle. Drawn without that squashing effect, the space is wildly curved, with each small section (*below*) being somewhat like a saddle shape with extra folds.

NEGATIVELY CURVED SPACETIME

The holographic theory involves a negatively curved spacetime known as anti–de Sitter space.



Imagine disks of hyperbolic space stacked atop one another, each representing the state of the universe at one instant. The resulting cylinder is three-dimensional anti–de Sitter space in which the height dimension represents time. Physics operates strangely in such a spacetime: a particle (*such as a tennis ball, green line*) thrown away from the center always falls back in a fixed period of time, and a laser beam (*red line*) can travel to the boundary of the universe and back in that same interval.
In the four-dimensional version, which would be more like our universe, the boundary for each instant would be a sphere instead of a circle.



ALFRED T. KAMAJIAN; SILVIO LEVY AND THE UNIVERSITY OF MINNESOTA (*Escher fish*)

good approximation, it is also the geometry of the world around us: parallel lines never meet, and all the rest of Euclid's axioms hold.

We are also familiar with some curved spaces. Curvature comes in two forms, positive and negative. The simplest space with positive curvature is the surface of a sphere. A sphere has constant positive curvature. That is, it has the same degree of curvature at every location (unlike an egg, say, which has more curvature at the pointy end).

The simplest space with negative curvature is called hyperbolic space, which is defined as space with constant negative curvature. This kind of space has long fascinated scientists and artists alike. Indeed, M. C. Escher produced several beautiful pictures of hyperbolic space, one of which is shown on the preceding page. His picture is like a flat map of the space. The way that the fish become smaller and smaller is just an artifact of how the curved space is squashed to fit on a flat sheet of paper, similar to the way that countries near the poles get stretched on a map of the globe (a sphere).

By including time in the game, physicists can similarly consider space*times* with positive or negative curvature. The simplest spacetime with positive curvature is called de Sitter space, after Willem de Sitter, the Dutch physicist who introduced it. Many cosmologists believe that the very early universe was close to being a de Sitter space. The far future may also be de Sitter–like because of cosmic acceleration. Conversely, the simplest negatively curved spacetime is called anti–de Sitter space. It is similar to hyperbolic space except that it also contains a time direction. Unlike our universe, which is expanding, anti–de Sitter space is neither expanding nor contracting. It looks the same at all times. Despite that difference, anti–de Sitter space turns out to be quite useful in the quest to form quantum theories of spacetime and gravity.

If we picture hyperbolic space as being a disk like Escher's drawing, then anti–de Sitter space is like a stack of those disks, forming a solid cylinder [*see box above*]. Time runs along the cylinder. Hyperbolic space can have more than two spatial dimensions. The anti–de Sitter space most like our spacetime (with three spatial dimensions) would have a three-dimensional "Escher print" as the cross section of its "cylinder."

Physics in anti–de Sitter space has some strange properties. If you were freely floating anywhere in anti–de Sitter space, you would feel as though you were at the bottom of a gravitational

THE AUTHOR

JUAN MALDACENA is a professor in the School of Natural Sciences at the Institute for Advanced Study in Princeton, N.J. Previously he was in the physics department at Harvard University from 1997 to 2001. He is currently studying various aspects of the duality conjecture described in this article. String theorists were so impressed with the conjecture that at the Strings '98 conference they feted him with a song, *The Maldacena*, sung and danced to the tune of *The Macarena*.

well. Any object that you threw out would come back like a boomerang. Surprisingly, the time required for an object to come back would be independent of how hard you threw it. The difference would just be that the harder you threw it, the farther away it would get on its round-trip back to you. If you sent a flash of light, which consists of photons moving at the maximum possible speed (the speed of light), it would actually reach infinity and come back to you, all in a finite amount of time. This can happen because an object experiences a kind of time contraction of ever greater magnitude as it gets farther away from you.

The Hologram

ANTI–DE SITTER SPACE, although it is infinite, has a "boundary," located out at infinity. To draw this boundary, physicists and mathematicians use a distorted length scale similar to Escher's, squeezing an infinite distance into a finite one. This boundary is like the outer circumference of the Escher print or the surface of the solid cylinder I considered earlier. In the cylinder example, the boundary has two dimensions—one is space (looping around the cylinder), and one is time (running along its length). For four-dimensional anti–de Sitter space, the boundary has two space dimensions and one time dimension. Just as the boundary of the Escher print is a circle, the boundary of four-dimensional anti–de Sitter space at any moment in time is a sphere. This boundary is where the hologram of the holographic theory lies.

Stated simply, the idea is as follows: a quantum gravity theory in the interior of an anti–de Sitter spacetime is completely equivalent to an ordinary quantum particle theory living on the boundary. If true, this equivalence means that we can use a quantum particle theory (which is relatively well understood) to define a quantum gravity theory (which is not).

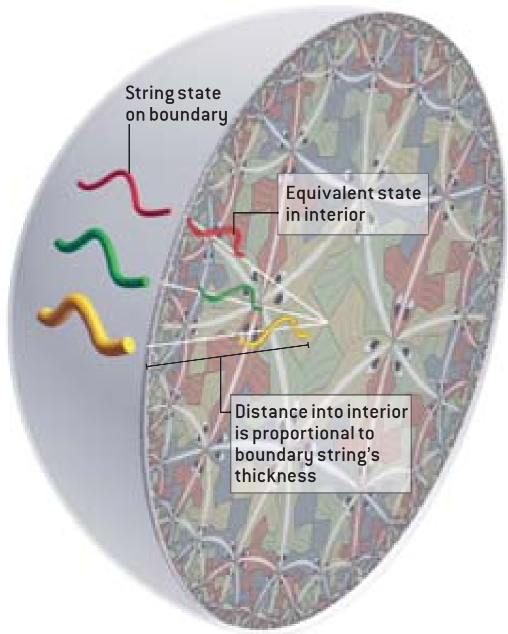
To make an analogy, imagine you have two copies of a movie, one on reels of 70-millimeter film and one on a DVD. The two formats are utterly different, the first a linear ribbon of celluloid with each frame recognizably related to scenes of the movie as we know it, the second a two-dimensional platter with rings of magnetized dots that would form a sequence of 0s and 1s if we could perceive them at all. Yet both "describe" the same movie.

Similarly, the two theories, superficially utterly different in content, describe the same universe. The DVD looks like a metal disk with some glints of rainbowlike patterns. The boundary particle theory "looks like" a theory of particles in the absence of gravity. From the DVD, detailed pictures emerge only when the bits are processed the right way. From the boundary particle theory, quantum gravity and an extra dimension emerge when the equations are analyzed the right way.

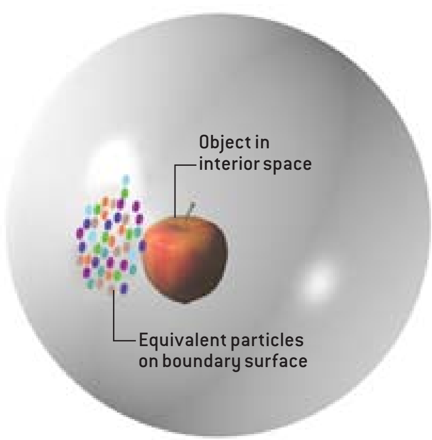
What does it really mean for the two theories to be equivalent? First, for every entity in one theory, the other theory has a counterpart. The entities may be very different in how they are described by the theories: one entity in the interior

CONJURING A DIMENSION

Holographic theory describes how quarks and gluons interacting on the boundary of an anti–de Sitter space could be equivalent to particles in the higher-dimensional interior of the space.



Quarks and gluons on the spherical surface of the anti–de Sitter space interact to form strings of various thicknesses. A holographic interpretation of those strings is that in the interior space they represent elementary particles (which are also strings) whose distance from the boundary corresponds to the string's thickness.



Clouds of quarks and gluons on the boundary surface can thus describe equivalent complex objects (such as this apple) in the interior. The advantage of this holographic theory is that the interior objects experience gravity even though a distinct gravitational interaction does not exist on the surface.

ALFRED T. KAMAJIAN

might be a single particle of some type, corresponding on the boundary to a whole collection of particles of another type, considered as one entity. Second, the predictions for corresponding entities must be identical. Thus, if two particles have a 40 percent chance of colliding in the interior, the two corresponding collections of particles on the boundary should also have a 40 percent chance of colliding.

Here is the equivalence in more detail. The particles that live on the boundary interact in a way that is very similar to how quarks and gluons interact in reality (quarks are the constituents of protons and neutrons; gluons generate the strong nuclear force that binds the quarks together). Quarks have a kind of charge that comes in three varieties, called colors, and the interaction is called chromodynamics. The difference between the boundary particles and ordinary quarks and gluons is that the particles have a large number of colors, not just three.

Gerard 't Hooft of Utrecht University in the Netherlands studied such theories as long ago as 1974 and predicted that the gluons would form chains that behave much like the strings of string theory. The precise nature of these strings remained elusive, but in 1981 Alexander M. Polyakov, now at Princeton University, noticed that the strings effectively live in a higher-dimensional space than the gluons do. As we shall see shortly, in our holographic theories that higher-dimensional space is the interior of anti–de Sitter space.

To understand where the extra dimension comes from, start by considering one of the gluon strings on the boundary. This string has a thickness, related to how much its gluons are smeared out in space. When physicists calculate how these strings on the boundary of anti–de Sitter space interact with one another, they get a very odd result: two strings with different thicknesses do not interact very much with each other. It is as though the strings were separated spatially. One can reinterpret the thickness of the string to be a new spatial coordinate that goes away from the boundary.

Thus, a thin boundary string is like a string close to the boundary, whereas a thick boundary string is like one far away from the boundary [*see box on preceding page*]. The extra coordinate is precisely the coordinate needed to describe motion within the four-dimensional anti–de Sitter spacetime! From the perspective of an observer in the spacetime, boundary strings of different thicknesses appear to be strings (all of them thin) at different radial locations. The number of colors on the boundary determines the size of the interior (the radius of the Escher-like sphere). To have a spacetime as large as the visible universe, the theory must have about 10^{60} colors.

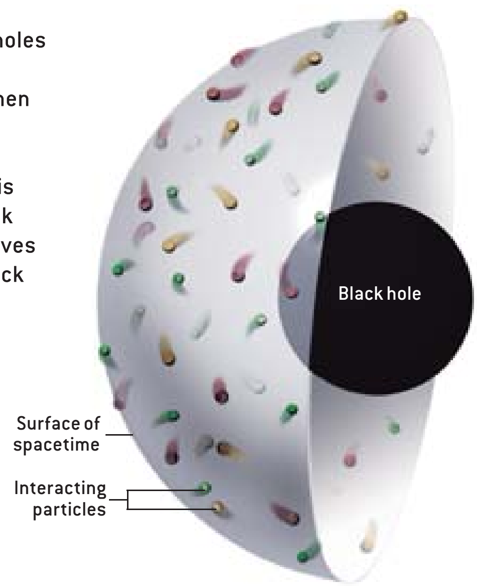
It turns out that one type of gluon chain behaves in the four-dimensional spacetime as the graviton, the fundamental quantum particle of gravity. In this description, gravity in four dimensions is an emergent phenomenon arising from particle interactions in a gravityless, three-dimensional world. The presence of gravitons in the theory should come as no surprise—physicists have known since 1974 that string theories always give rise to quantum gravity. The strings formed by gluons are no exception, but the gravity operates in the higher-dimensional space.

Thus, the holographic correspondence is not just a wild new possibility for a quantum theory of gravity. Rather, in a fundamental way, it connects string theory, the most studied approach to quantum gravity, with theories of quarks and gluons, which are the cornerstone of particle physics. What is more, the holographic theory seems to provide some insight into the elusive exact equations of string theory. String theory was actually invented in the late 1960s for the purpose of describing strong interactions, but it was later abandoned (for that purpose) when the theory of chromodynamics entered the scene. The correspondence between string theory and chromodynamics implies that these early efforts were not misguided; the two descriptions are different faces of the same coin.

Varying the boundary chromodynamics theory by changing the details of how the boundary particles interact gives rise to an assortment of interior theories. The resulting interior theory can have only gravitational forces, or gravity plus some extra force such as the electromagnetic force, and so on. Unfortunately, we do not yet know of a boundary theory that gives rise to an interior theory that includes exactly

UNDERSTANDING BLACK HOLES

Physicist Stephen W. Hawking showed in the 1970s that black holes have a temperature and give off radiation, but physicists since then have been deeply puzzled. Temperature is a property of a collection of particles, but what is the collection that defines a black hole? The holographic theory solves this puzzle by showing that a black hole is equivalent to a swarm of interacting particles on the boundary surface of spacetime.



ALFRED T. KAMAJIAN

the four forces we have in our universe.

I first conjectured that this holographic correspondence might hold for a specific theory (a simplified chromodynamics in a four-dimensional boundary spacetime) in 1997. This immediately excited great interest from the string theory community. The conjecture was made more precise by Polyakov, Stephen S. Gubser and Igor R. Klebanov of Princeton and Edward Witten of the Institute for Advanced Study in Princeton, N.J. Since then, many researchers have contributed to exploring the conjecture and generalizing it to other dimensions and other chromodynamics theories, providing mounting evidence that it is correct. So far, however, no example has been rigorously proved—the mathematics is too difficult.

So far no example of the holographic correspondence has been rigorously proved—the mathematics is too difficult.

Mysteries of Black Holes

HOW DOES THE holographic description of gravity help to explain aspects of black holes? Black holes are predicted to emit Hawking radiation, named after Stephen W. Hawking of the University of Cambridge, who discovered this result. This radiation comes out of the black hole at a specific temperature. For all ordinary physical systems, a theory called statistical mechanics explains temperature in terms of the motion of the microscopic constituents. This theory explains the temperature of a glass of water or the temperature of the sun. What about the temperature of a black hole? To understand it, we would need to know what the microscopic constituents of the black hole are and how they behave. Only a theory of quantum gravity can tell us that.

Some aspects of the thermodynamics of black holes have raised doubts as to whether a quantum-mechanical theory of gravity could be developed at all. It seemed as if quantum mechanics itself might break down in the face of effects taking place in black holes. For a black hole in an anti–de Sitter spacetime, we now know that quantum mechanics remains intact, thanks to the boundary theory. Such a black hole corresponds to a configuration of particles on the boundary. The number of particles is very large, and they are all zipping around, so that theorists can apply the usual rules of statistical mechanics to compute the temperature. The result is the same as the temperature that Hawking computed by very different means, indicating that the results can be trusted. Most important, the boundary theory obeys the ordinary rules of quantum mechanics; no inconsistency arises.

Physicists have also used the holographic correspondence in the opposite direction—employing known properties of black holes in the interior spacetime to deduce the behavior of quarks and gluons at very high temperatures on the boundary. Dam Son of the University of Washington and his collaborators studied a quantity called the shear viscosity, which is small for a fluid that flows very easily and large for a substance more like molasses. They found that black holes have an extremely low shear viscosity—smaller than any known fluid. Because of the holographic equivalence, strongly interacting quarks and gluons at high temperatures should also have very low viscosity.

A test of this prediction comes from the Relativistic Heavy Ion Collider (RHIC) at Brookhaven National Laboratory, which has been colliding gold nuclei at very high energies. A preliminary analysis of these experiments indicates the collisions are creating a fluid with very low viscosity. Even though Son and his co-workers studied a simplified version of chromodynamics, they seem to have come up with a property that is shared by the real world. Does this mean that RHIC is creating small five-dimensional black holes? It is really too early to tell, both experimentally and theoretically. (Even if so, there is nothing to fear from these tiny black holes—they evaporate almost as fast as they are formed, and they "live" in five dimensions, not in our own four-dimensional world.)

Many questions about the holographic theories remain to be answered. In particular, does anything similar hold for a universe like ours in place of the anti–de Sitter space? A crucial aspect of anti–de Sitter space is that it has a boundary where time is well defined. The boundary has existed and will exist forever. An expanding universe, like ours, that comes from a big bang does not have such a well-behaved boundary. Consequently, it is not clear how to define a holographic theory for our universe; there is no convenient place to put the hologram.

An important lesson that one can draw from the holographic conjecture, however, is that quantum gravity, which has perplexed some of the best minds on the planet for decades, can be very simple when viewed in terms of the right variables. Let's hope we will soon find a simple description for the big bang! SA

MORE TO EXPLORE

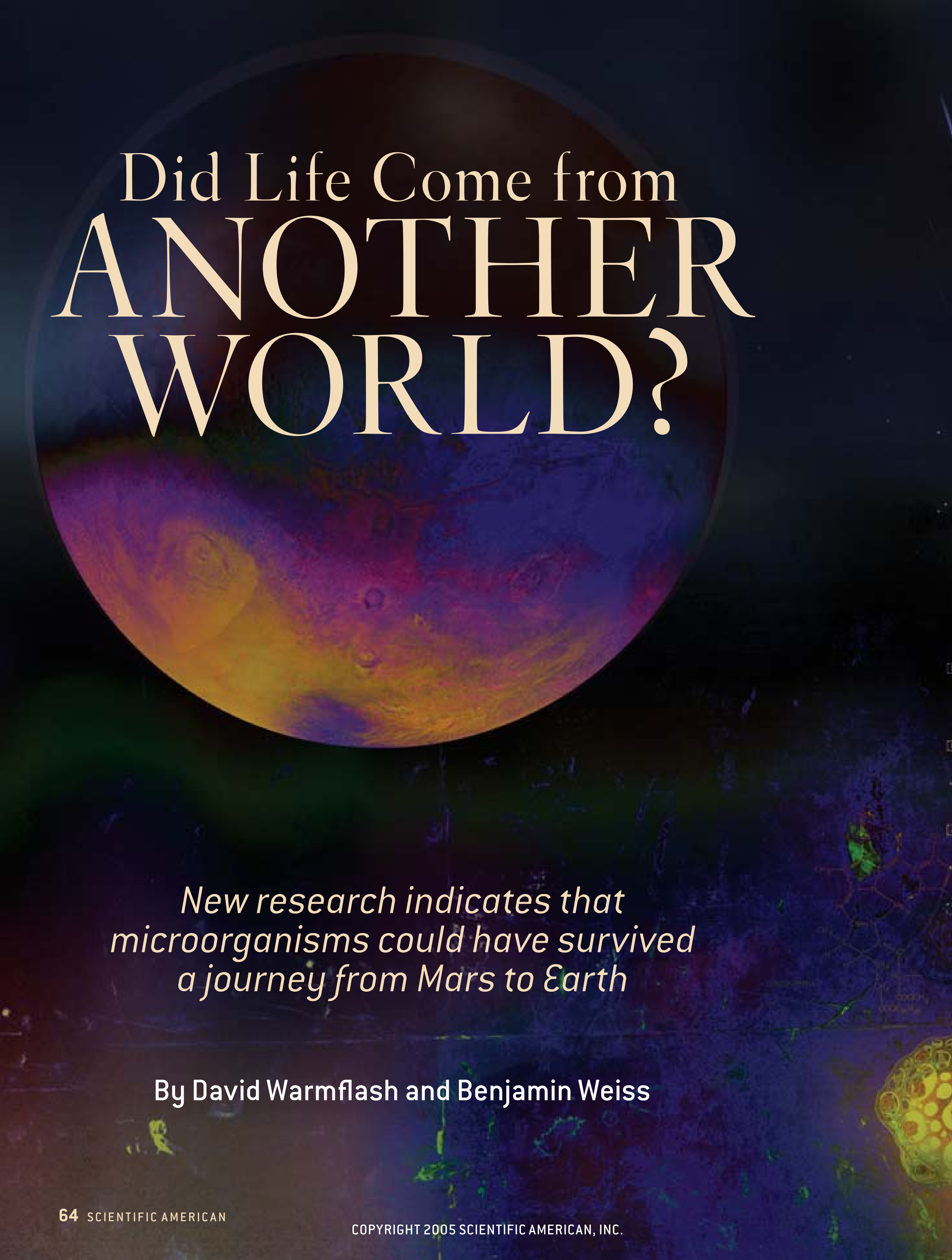
Anti–de Sitter Space and Holography. Edward Witten in *Advances in Theoretical and Mathematical Physics,* Vol. 2, pages 253–291; 1998. Available online at **http://arxiv.org/abs/hep-th/9802150**

Gauge Theory Correlators from Non-Critical String Theory. S. Gubser, I. R. Klebanov and A. M. Polyakov in *Applied Physics Letters B,* Vol. 428, pages 105–114; 1998. **http://arxiv.org/abs/hep-th/9802109**

The Theory Formerly Known as Strings. Michael J. Duff in *Scientific American,* Vol. 278, No. 2, pages 64–69; February 1998.

The Elegant Universe. Brian Greene. Reissue edition. W. W. Norton and Company, 2003.

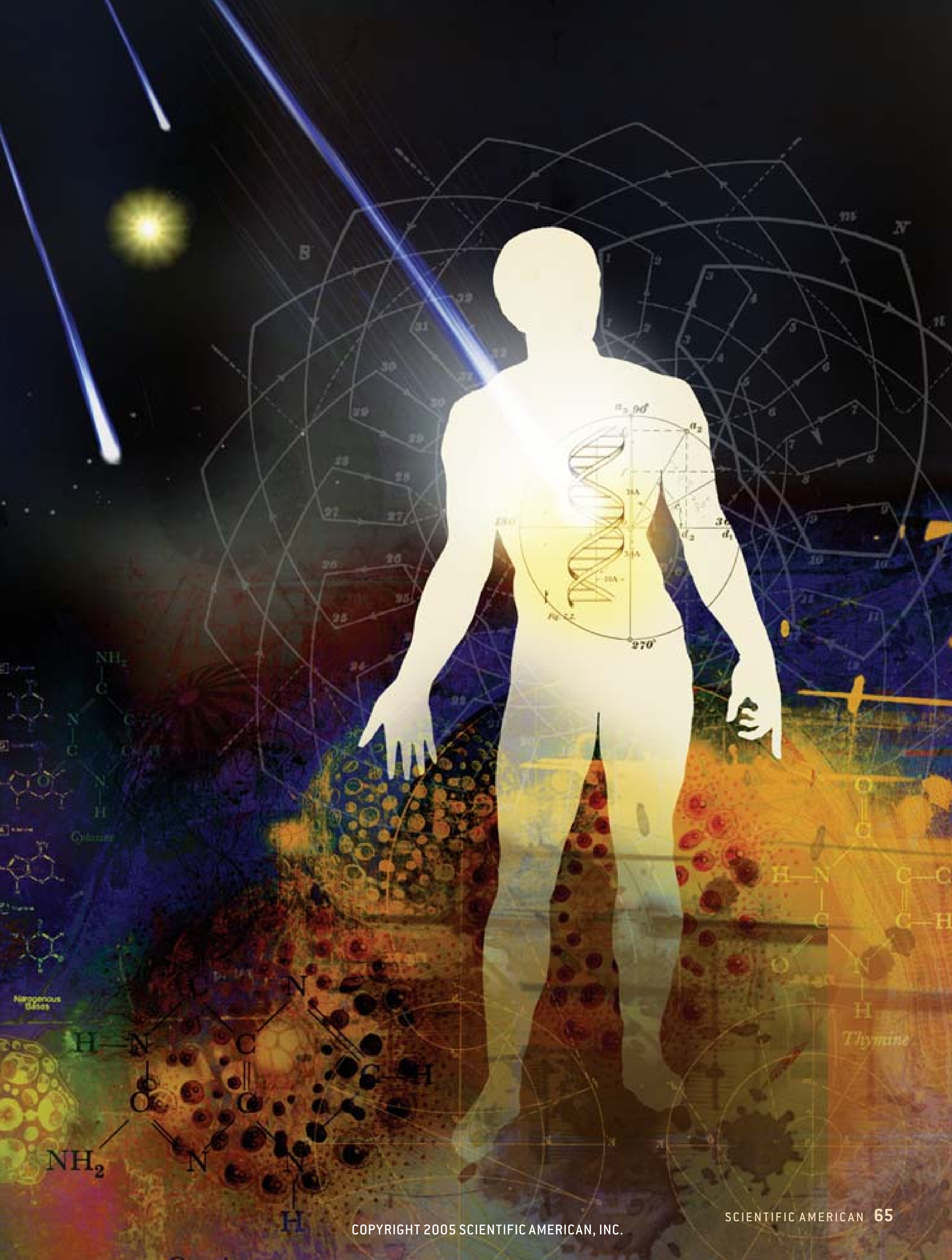
A string theory Web site is at **superstringtheory.com**



Did Life Come from ANOTHER WORLD?

New research indicates that microorganisms could have survived a journey from Mars to Earth

By David Warmflash and Benjamin Weiss



Most scientists have long assumed that life on Earth is a homegrown phenomenon. According to the conventional hypothesis, the earliest living cells emerged as a result of chemical evolution on our planet billions of years ago in a process called abiogenesis. The alternative possibility—that living cells or their precursors arrived from space—strikes many people as science fiction. Developments over the past decade, however, have given new credibility to the idea that Earth's biosphere could have arisen from an extraterrestrial seed.

Planetary scientists have learned that early in its history our solar system could have included many worlds with liquid water, the essential ingredient for life as we know it. Recent data from NASA's Mars Exploration Rovers corroborate previous suspicions that water has at least intermittently flowed on the Red Planet in the past. It is not unreasonable to hypothesize that life existed on Mars long ago and perhaps continues there. Life may have also evolved on Europa, Jupiter's fourth-largest moon, which appears to possess liquid water under its icy surface. Saturn's biggest satellite, Titan, is rich in organic compounds; given the moon's frigid temperatures, it would be highly surprising to find living forms there, but they cannot be ruled out. Life may have even gained a toehold on torrid Venus. The Venusian surface is probably too hot and under too much atmospheric pressure to be habitable, but the planet could conceivably support microbial life high in its atmosphere. And, most likely, the surface conditions on Venus were not always so harsh. Venus may have once been similar to early Earth.

Moreover, the expanses of interplanetary space are not the forbidding barrier they once seemed. Over the past 20 years scientists have determined that more than 30 meteorites found on Earth originally came from the Martian crust, based on the composition of gases trapped within some of the rocks. Meanwhile biologists have discovered organisms durable enough to survive at least a short journey inside such meteorites. Although no one is suggesting that these particular organisms actually made the trip, they serve as a proof of principle. It is not implausible that life could have arisen on Mars and then come to Earth, or the reverse. Researchers are now intently studying the transport of biological materials between planets to get a better sense of whether it ever occurred. This effort may shed light on some of modern science's most compelling questions: Where and how did life originate? Are radically different forms of life possible? And how common is life in the universe?

From Philosophy to the Laboratory

TO THE ANCIENT philosophers, the creation of life from nonliving matter seemed so magical, so much the realm of the gods, that some actually preferred the idea that ready-made living forms had come to Earth from elsewhere. Anaxagoras, a Greek philosopher who lived 2,500 years ago, proposed a hypothesis called "panspermia" (Greek for "all seeds"), which posited that all life, and indeed all things, originated from the combination of tiny seeds pervading the cosmos. In modern times, several leading scientists—including British physicist Lord Kelvin, Swedish chemist Svante Arrhenius and Francis Crick, co-discoverer of the structure of DNA—have advocated various conceptions of panspermia. To be sure, the idea has also had less reputable proponents, but they should not detract from the fact that panspermia is a serious hypothesis, a potential phenomenon that we should not ignore when considering the distribution and evolution of life in the universe and how life came to exist specifically on Earth.

In its modern form, the panspermia hypothesis addresses how biological material might have arrived on our planet but not how life originated in the first place. No matter where it started, life had to arise from nonliving matter. Abiogenesis moved from the realm of philosophy to that of experimentation in the 1950s, when chemists Stanley L. Miller and Harold C. Urey of the University of Chicago demonstrated that amino acids and other molecules important to life could be generated from simple compounds believed to exist on early Earth. It is now thought that molecules of ribonucleic acid (RNA) could have also assembled from smaller compounds and played a vital role in the development of life.

In present-day cells, specialized RNA molecules help to build proteins. Some RNAs act as messengers between the genes, which are made of deoxyribonucleic acid (DNA), and the ribosomes, the protein factories of the cell. Other RNAs bring amino acids—the building blocks of proteins—to the ribosomes, which in turn contain yet another type of RNA. The RNAs work in concert with protein enzymes that aid in linking the amino

Overview/*Life from Space*

- The panspermia hypothesis posits that living cells or their precursors could have emerged on another planet or moon billions of years ago and hitched a ride to Earth on a meteorite.
- A small fraction of the rocks blasted off Mars by asteroid or comet impacts could have reached Earth in just a few years.
- Researchers plan to evaluate the likelihood of panspermia by studying whether microorganisms can survive an interplanetary journey.

DUNG HUANG (*preceding pages*)

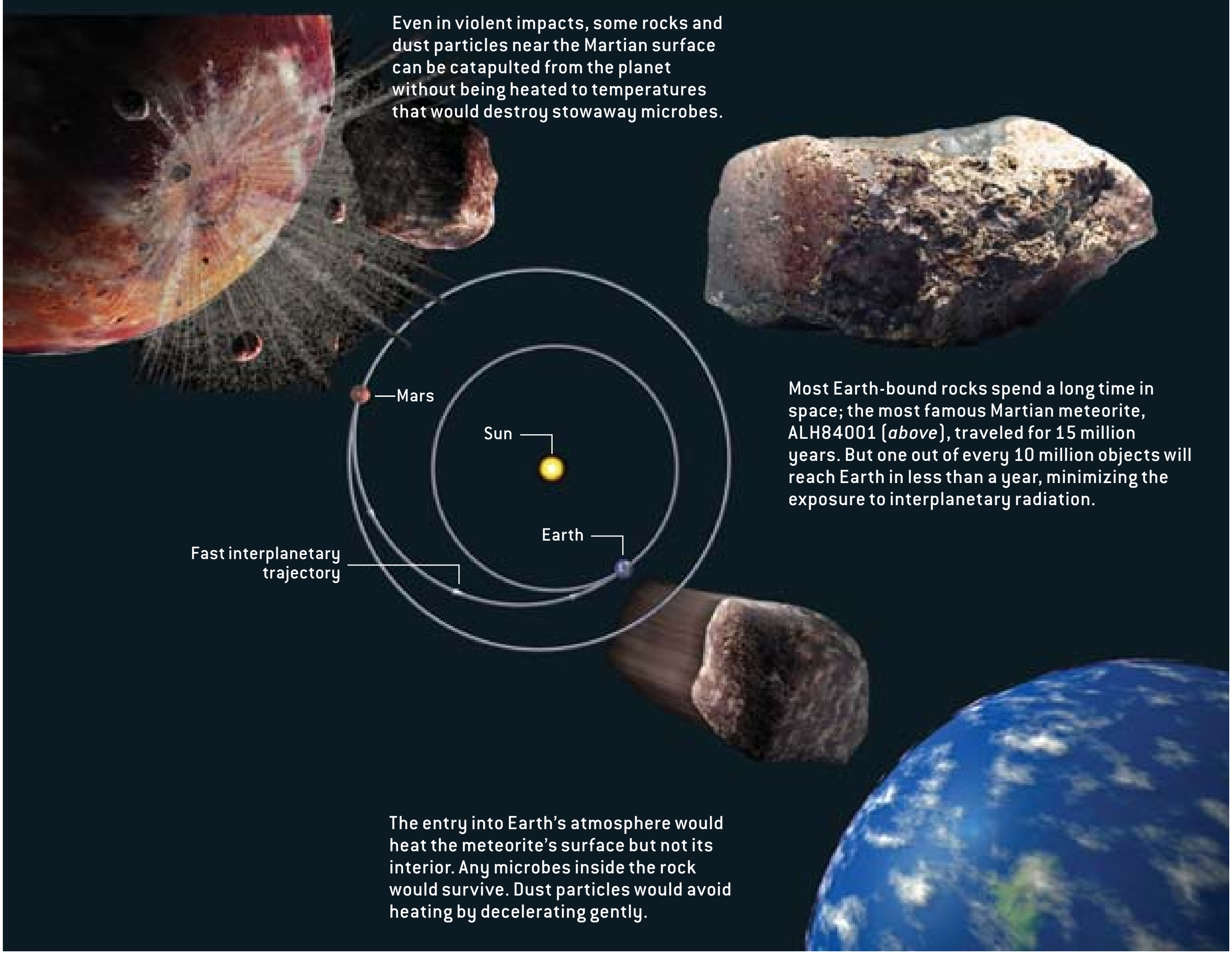
acids together, but researchers have found that the RNAs in the ribosome can perform the crucial step of protein synthesis alone. In the early stages of life's evolution, all the enzymes may have been RNAs, not proteins. Because RNA enzymes could have manufactured the first proteins without the need for preexisting protein enzymes to initiate the process, abiogenesis is not the chicken-and-egg problem that it was once thought to be. A prebiotic system of RNAs and proteins could have gradually developed the ability to replicate its molecular parts, crudely at first but then ever more efficiently.

This new understanding of life's origins has transformed the scientific debate over panspermia. It is no longer an either-or question of whether the first microbes arose on Earth or arrived from space. In the chaotic early history of the solar system, our planet was subject to intense bombardment by meteorites containing simple organic compounds. The young Earth could have also received more complex molecules with enzymatic functions, molecules that were prebiotic but part of a system that was already well on its way to biology. After landing in a suitable habitat on our planet, these molecules could have continued their evolution to living cells. In other words, an intermediate scenario is possible: life could have roots both on Earth and in space. But which steps in the development of life occurred where? And once life took hold, how far did it spread?

Scientists who study panspermia used to concentrate only on assessing the basic plausibility of the idea, but they have recently sought to estimate the probability that biological materials made the journey to Earth from other planets or moons. To begin their interplanetary trip, the materials would have to be ejected from their planet of origin into space by the impact of a comet or asteroid [*see box below*]. While traveling through space, the ejected rocks or dust particles would need to be captured by the gravity of another planet or moon,

THE INTERPLANETARY EXPRESS

Every few million years an asteroid or comet hits Mars with enough energy to eject rocks that can escape the Red Planet's gravity and eventually reach Earth. If life evolved on Mars billions of years ago, it is conceivable that rocks containing biological materials could have made the voyage quickly enough to transplant the extraterrestrial seeds.



Even in violent impacts, some rocks and dust particles near the Martian surface can be catapulted from the planet without being heated to temperatures that would destroy stowaway microbes.

Most Earth-bound rocks spend a long time in space; the most famous Martian meteorite, ALH84001 (*above*), traveled for 15 million years. But one out of every 10 million objects will reach Earth in less than a year, minimizing the exposure to interplanetary radiation.

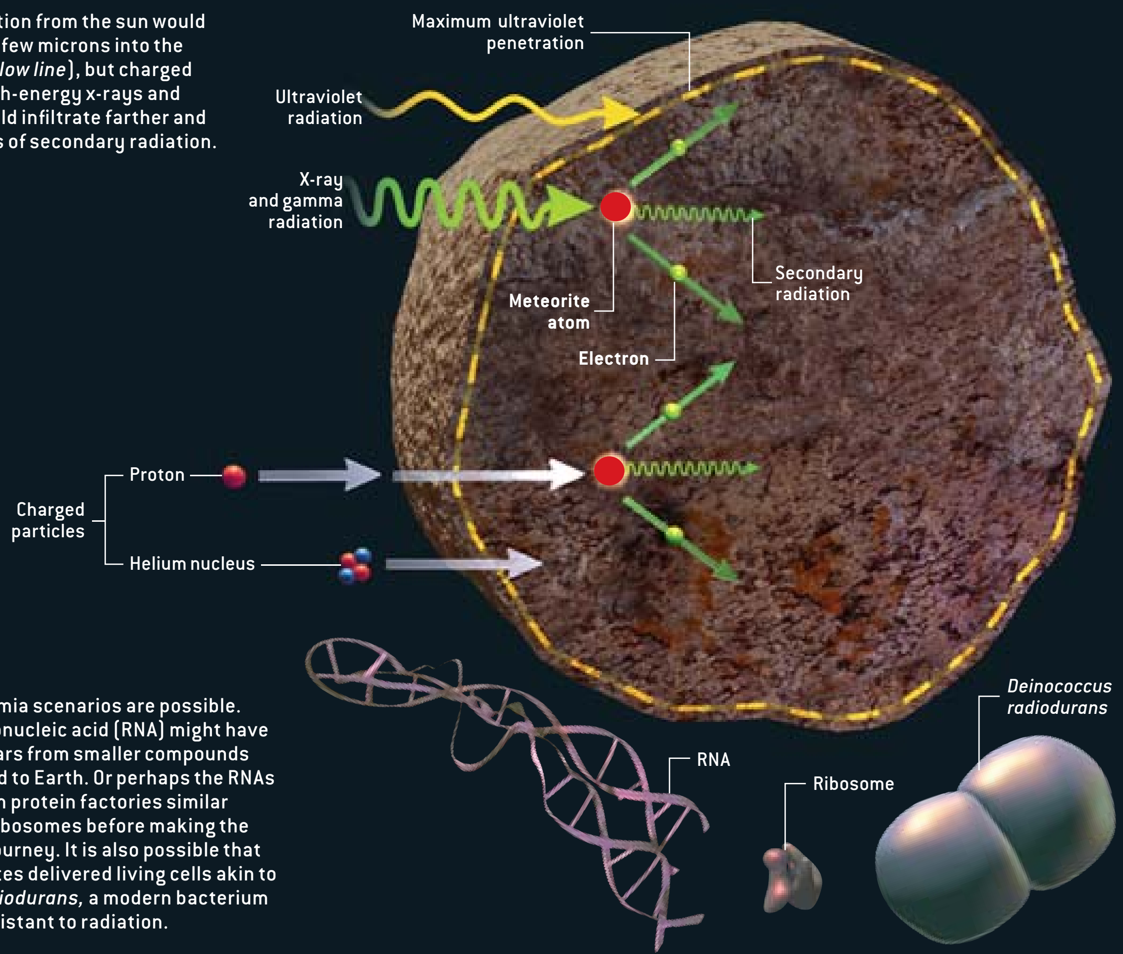
The entry into Earth's atmosphere would heat the meteorite's surface but not its interior. Any microbes inside the rock would survive. Dust particles would avoid heating by decelerating gently.

ALFRED T. KAMAJIAN; NASA JOHNSON SPACE CENTER (*ALH84001 meteorite*)

A COSMIC NOAH'S ARK

Biological materials might best survive the perils of interplanetary space if carried inside meteorites. Radiation would be the primary threat.

Ultraviolet radiation from the sun would penetrate only a few microns into the rock (*dashed yellow line*), but charged particles and high-energy x-rays and gamma rays would infiltrate farther and produce showers of secondary radiation.



Several panspermia scenarios are possible. Molecules of ribonucleic acid (RNA) might have assembled on Mars from smaller compounds and then traveled to Earth. Or perhaps the RNAs combined to form protein factories similar to present-day ribosomes before making the interplanetary journey. It is also possible that Martian meteorites delivered living cells akin to *Deinococcus radiodurans,* a modern bacterium that is highly resistant to radiation.

ALFRED T. KAMAJIAN

then decelerated enough to fall to the surface, passing through the atmosphere if one were present. Such transfers happen frequently throughout the solar system, although it is easier for ejected material to travel from bodies more distant from the sun to those closer in and easier for materials to end up on a more massive body. Indeed, dynamic simulations by University of British Columbia astrophysicist Brett Gladman suggest that the mass transferred from Earth to Mars is only a few percent of that delivered from Mars to Earth. For this reason, the most commonly discussed panspermia scenario involves the transport of microbes or their precursors from Mars to Earth.

Simulations of asteroid or comet impacts on Mars indicate that materials can be launched into a wide variety of orbits. Gladman and his colleagues have estimated that every few million years Mars undergoes an impact powerful enough to eject rocks that could eventually reach Earth. The interplanetary journey is usually a long one: most of the approximately one ton of Martian ejecta that lands on Earth every year has spent several million years in space. But a tiny percentage of the Martian rocks arriving on Earth's surface—about one out of every 10 million—will have spent less than a year in space. Within three years of the impact event, about 10 fist-size rocks weighing more than 100 grams complete the voyage from Mars to Earth. Smaller debris, such as pebble-size rocks and dust particles, are even more likely to make a quick trip between planets; very large rocks do so much less frequently.

THE AUTHORS

DAVID WARMFLASH and *BENJAMIN WEISS* have taken different yet complementary approaches to studying whether life could have come to Earth from another world. Warmflash is an astrobiologist at the University of Houston and the NASA Johnson Space Center. He is currently helping to develop molecular tests to search for microorganisms on Mars and the Jovian moon Europa. Weiss is assistant professor of planetary sciences at the Massachusetts Institute of Technology. His recent studies of various Martian meteorites suggest that they were not heat-sterilized during their transfer from Mars to Earth.

Could biological entities survive this journey? First, let us consider whether microorganisms could live through the ejection process from the meteorite's parent body. Recent laboratory impact experiments have found that certain strains of bacteria can survive the accelerations and jerks (rates of changes of acceleration) that would be encountered during a typical high-pressure ejection from Mars. It is crucial, however, that the impact and ejection do not heat the meteorites enough to destroy the biological materials within them.

Planetary geologists formerly believed that any impact ejecta with speeds exceeding the Martian escape velocity would almost certainly be vaporized or at least completely melted. This idea was later discounted, though, following the discovery of unmelted, largely intact meteorites from the moon and Mars. These findings led H. Jay Melosh of the University of Arizona to calculate that a small percentage of ejected rocks could indeed be catapulted from Mars via impact without any heating at all. In short, Melosh proposed that when the upward-propagating pressure wave resulting from an impact reaches the planetary surface, it undergoes a 180-degree phase change that nearly cancels the pressure within a thin layer of rock just below the surface. Because this "spall zone" experiences very little compression while the layers below are put under enormous pressure, rocks near the surface can be ejected relatively undeformed at high speeds.

Next, let us consider survivability during the entry into Earth's atmosphere. Edward Anders, formerly of the Enrico Fermi Institute at the the University of Chicago, has shown that interplanetary dust particles decelerate gently in Earth's upper atmosphere, thus avoiding heating. Meteorites, in contrast, experience significant friction, so their surfaces typically melt during atmospheric passage. The heat pulse, however, has time to travel a few millimeters at most into the meteorite's interior, so organisms buried deep in the rock would certainly survive.

Over the past five years a series of papers by one of us (Weiss) and his colleagues analyzed two types of Martian meteorites: the nakhlites, a set of rocks blasted off Mars by an asteroid or comet impact 11 million years ago, and ALH84001, which left the Red Planet four million years earlier. (ALH84001 became famous in 1996 when a group of scientists led by David McKay of the NASA Johnson Space Center claimed that the rock showed traces of fossilized microorganisms akin to Earth's bacteria; a decade later researchers are still debating whether the meteorite contains evidence of Martian life.) By studying the magnetic properties of the meteorites and the composition of the gases trapped within them, Weiss and his collaborators found that ALH84001 and at least two of the seven nakhlites discovered so far were not heated more than a few hundred degrees Celsius since they were part of the Martian surface. Furthermore, the fact that the nakhlites are nearly pristine rocks, untouched by high-pressure shock waves, implies that the Martian impact did not heat them above 100 degrees C.

Many, though not all, terrestrial prokaryotes (simple one-celled organisms such as bacteria that lack a membrane-bound nucleus) and eukaryotes (organisms with well-defined nuclei) could survive this temperature range. This result was the first direct experimental evidence that material could travel from planet to planet without being thermally sterilized at any point from ejection to landing.

The Problem of Radiation

FOR PANSPERMIA to occur, however, microorganisms need to survive not only ejection from the first planet and atmospheric entry to the second but the interplanetary voyage itself. Life-bearing meteoroids and dust particles would be exposed to the vacuum of space, extremes in temperature and several different kinds of radiation. Of particular concern is the sun's high-energy ultraviolet (UV) light, which breaks the bonds that hold together the carbon atoms of organic molecules. It is very easy to shield against UV, though; just a few millionths of a meter of opaque material is enough to protect bacteria.

Indeed, a European study using NASA's Long Duration Exposure Facility (LDEF), a satellite deployed by the space shuttle in 1984 and retrieved from orbit by the shuttle six years later, showed that a thin aluminum cover afforded adequate UV shielding to spores of the bacterial species *Bacillus subtilis*. Of the spores protected by the aluminum but exposed to the vacuum and temperature extremes of space, 80 percent remained viable—researchers reanimated them into active bacterial cells at the end of the mission. As for the spores not covered by aluminum and therefore directly exposed to solar UV radiation, most were destroyed, but not all. About one in 10,000 unshielded spores stayed viable, and the presence of substances such as glucose and salts increased their survival rates. Even within an object as small as a dust particle, solar UV would not necessarily render an entire microbial colony sterile. And if the colony were inside something as large as a pebble, UV protection would be sharply increased.

Informative as it was, the LDEF study was conducted in low Earth orbit, well within our planet's protective magnetic field. Thus, this research could not say much about the effects of interplanetary charged particles, which cannot penetrate Earth's magnetosphere. From time to time, the sun produces bursts of energetic ions and electrons; furthermore, charged particles are a major com-

Earth's biosphere could have arisen from an extraterrestrial seed.

ponent of the galactic cosmic radiation that constantly bombards our solar system. Protecting living things from charged particles, as well as from high-energy radiation such as gamma rays, is trickier than shielding against UV. A layer of rock just a few microns thick blocks UV, but adding more shielding actually *increases* the dose of other types of radiation. The reason is that charged particles and high-energy photons interact with the rocky shielding material, producing showers of secondary radiation within the meteorite.

These showers could reach any microbes inside the rock unless it was very big, about two meters or more in diameter. As we have noted above, though, large rocks make fast interplanetary voyages very infrequently. Consequently, in addition to UV protection, what really matters is how resistant a microbe is to all components of space radiation and how quickly the life-bearing meteorite moves from planet to planet. The shorter the journey, the lower the total radiation dose and hence the greater the chance of survival.

In fact, *B. subtilis* is fairly robust in terms of its radiation resistance. Even more hardy is *Deinococcus radiodurans,* a bacterial species that was discovered during the 1950s by agricultural scientist Arthur W. Anderson. This organism survives radiation doses given to sterilize food products and even thrives inside nuclear reactors. The same cellular mechanisms that help *D. radiodurans* repair its DNA, build extra-thick cell walls and otherwise protect itself from radiation also mitigate damage from dehydration. Theoretically, if organisms with such capabilities were embedded within material catapulted from Mars the way that the nakhlites and ALH84001 apparently were (that is, without excessive heating), some fraction of the organisms would still be viable after many years, perhaps several decades, in interplanetary space.

Yet the actual long-term survival of active organisms, spores or complex organic molecules beyond Earth's magnetosphere has never been tested. Such experiments, which would put the biological materials within simulated meteoritic materials and expose them to the environment of interplanetary space, could be conducted on the surface of the moon. In fact, biological samples were carried onboard the Apollo lunar missions as part of an early incarnation of the European radiation study. The longest Apollo mission, though, lasted no more than 12 days, and samples were kept within the Apollo spacecraft and thus not exposed to the full space-radiation environment. In the future, scientists could place experimental packages on the lunar surface or on interplanetary trajectories for several years before returning them to Earth for laboratory analysis. Researchers are currently considering these approaches.

Meanwhile a long-term study known as the Martian Radiation Environment Experiment (MARIE) is under way. Launched by NASA in 2001 as part of the Mars Odyssey Orbiter, MARIE's instruments are measuring doses of galactic cosmic rays and energetic solar particles as the spacecraft circles the Red Planet. Although MARIE includes no biological material, its sensors are designed to focus on the range of space radiation that is most harmful to DNA.

Future Studies

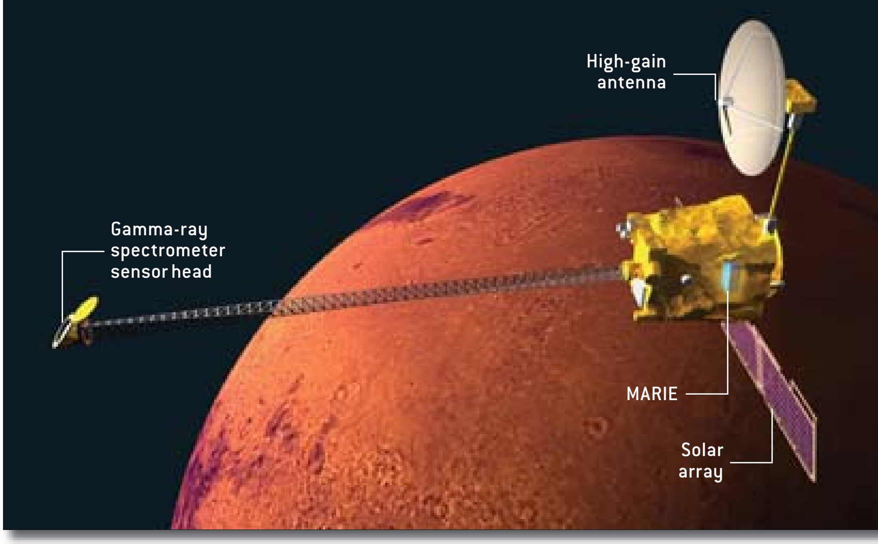
AS WE HAVE SHOWN, panspermia is plausible theoretically. But in addition, important aspects of the hypothesis have made the transition from plausibility to quantitative science. Meteorite evidence shows that material has been transferred between planets throughout the history of the solar system and that this process still occurs at a well-established rate. Furthermore, laboratory studies have demonstrated that a sizable fraction of microorganisms within a piece of planetary material ejected from a Mars-size planet could survive ejection into space and entry through Earth's atmosphere. But other parts of the panspermia hypothesis are harder to pin down. Investigators need more data to determine whether radiation-resistant organisms

NASA's Long Duration Exposure Facility carried spores of the bacterial species *Bacillus subtilis* (*inset*) in orbit for six years. Researchers found that a thin aluminum cover was enough to shield the spores from damaging ultraviolet radiation, enabling 80 percent of them to survive.

NASA LANGLEY RESEARCH CENTER; TONY BRAIN *Photo Researchers, Inc.* (*inset*)

MONITORING MARS

The Mars Odyssey Orbiter gauges the dangers of the interplanetary environment with an instrument called MARIE, which measures galactic cosmic rays and high-energy solar particles as the spacecraft circles the Red Planet.



ALFRED T. KAMAJIAN

such as *B. subtilis* or *D. radiodurans* could live through an interplanetary journey. And even this research would not reveal the likelihood that it actually happened in the case of Earth's biosphere, because the studies involve present-day terrestrial life-forms; the organisms living billions of years ago could have fared much worse or much better.

Moreover, scientists cannot quantify the likelihood that life exists or once existed on planets other than Earth. Researchers simply do not know enough about the origin of any system of life, including that of Earth, to draw solid conclusions about the probability of abiogenesis occurring on any particular world. Given suitable ingredients and conditions, perhaps life needs hundreds of millions of years to get started. Or perhaps five minutes is enough. All we can say with any certainty is that by 2.7 billion years ago, or perhaps several hundred million years earlier, life-forms were thriving on Earth.

Because it is not possible at this time to quantify all the steps of the panspermia scenario, investigators cannot estimate how much biological material or how many living cells most likely arrived at Earth's surface in a given period. Moreover, the transfer of viable organisms does not automatically imply the successful seeding of the planet that receives them, particularly if the planet already has life. If, for example, Martian microbes arrived on Earth after life independently arose on our planet, the extraterrestrial organisms may not have been able to replace or coexist with the homegrown species. It is also conceivable that Martian life did find a suitable niche on Earth but that scientists have simply not identified it yet. Researchers have inventoried no more than a few percent of the total number of bacterial species on this planet. Groups of organisms that are genetically unrelated to the known life on Earth might exist unrecognized right under our noses.

Ultimately, scientists may not be able to know whether and to what extent panspermia has occurred until they discover life on another planet or moon. For example, if future space missions find life on the Red Planet and report that Martian biochemistry is very different from our own, researchers would know immediately that life on Earth did not come from Mars. If the biochemistries were similar, however, scientists might begin to wonder if perhaps the two biospheres had a common origin. Assuming that Martian life-forms used DNA to store genetic information, investigators could study the nucleotide sequences to settle the question. If the Martian DNA sequences did not follow the same genetic code used by living cells on Earth to make proteins, researchers would conclude that Mars-Earth panspermia is doubtful. But many other scenarios are possible. Investigators might find that Martian life uses RNA or something else entirely to guide its replication. Indeed, yet-to-be-discovered organisms on Earth may fall into this category as well, and the exotic terrestrial creatures might turn out to be related to the Martian life-forms.

Whether terrestrial life emerged on Earth or through biological seeding from space or as the result of some intermediate scenario, the answer would be meaningful. The confirmation of Mars-Earth panspermia would suggest that life, once started, could readily spread within a star system. If, on the other hand, researchers find evidence of Martian organisms that emerged independently of terrestrial life, it would suggest that abiogenesis can occur with ease throughout the cosmos. What is more, biologists would be able to compare Earth organisms with alien forms and develop a more general definition of life. We would finally begin to understand the laws of biology the way we understand the laws of chemistry and physics—as fundamental properties of nature. SA

MORE TO EXPLORE

Worlds in the Making: The Evolution of the Universe. Svante Arrhenius. Harper, 1908.

The Structural Basis of Ribosome Activity in Peptide Bond Synthesis. P. Nissen, J. Hansen, N. Ban, P. B. Moore and T. A. Steitz in *Science*, Vol. 289, pages 878–879; August 11, 2000.

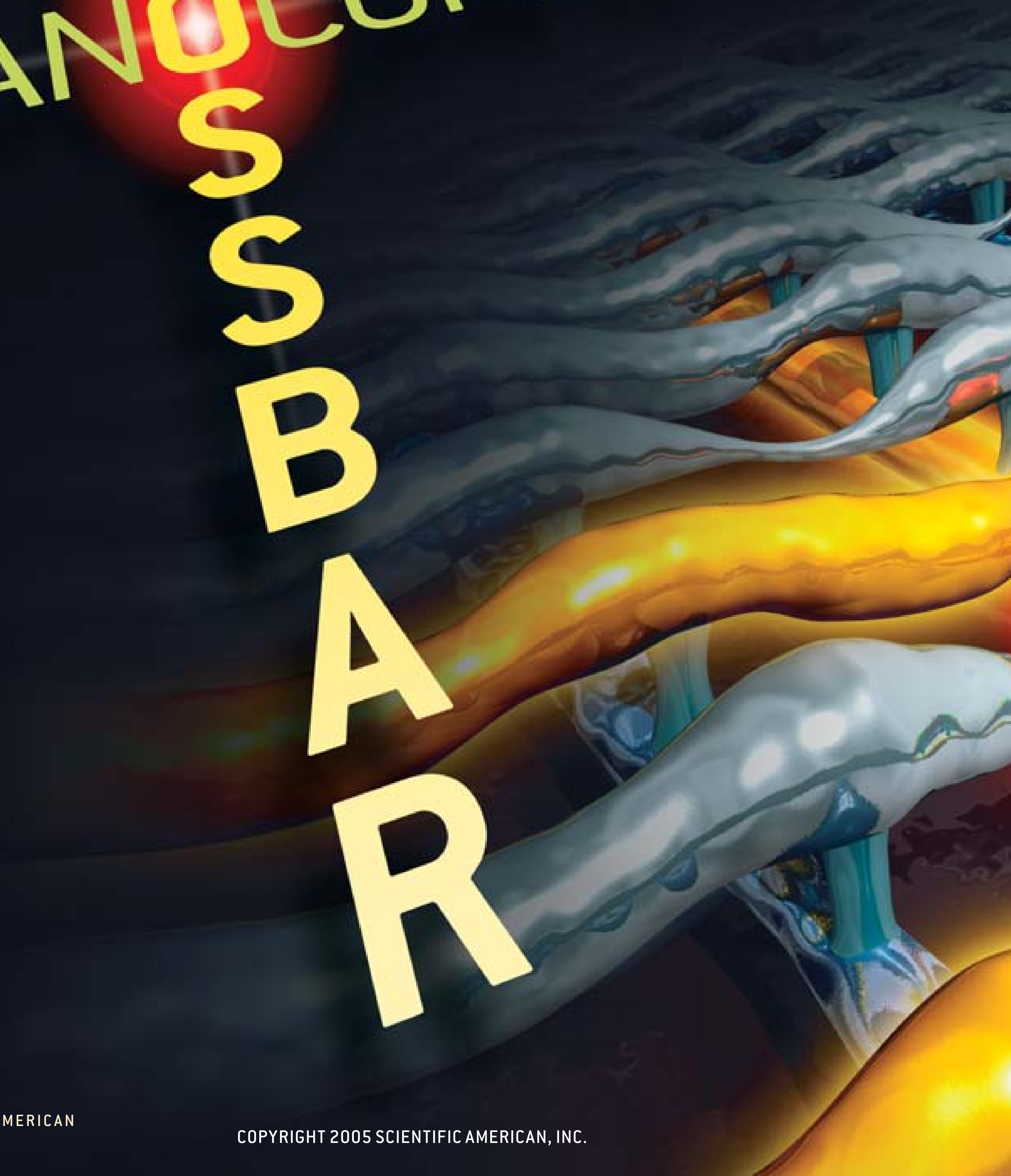
Risks Threatening Viable Transfer of Microbes between Bodies in Our Solar System. C. Mileikowsky, F. A. Cucinotta, J. W. Wilson, B. Gladman, G. Horneck, L. Lindegren, H. J. Melosh, H. Rickman, M. Valtonen and J. Q. Zheng in *Planetary and Space Science*, Vol. 48, Issue 11, pages 1107–1115; September 2000.

Martian Surface Paleotemperatures from Thermochronology of Meteorites. D. L. Shuster and B. P. Weiss in *Science*, Vol. 309, pages 594–600; July 22, 2005.

Origins of the Genetic Code: The Escaped Triplet Theory. M. Yarus, J. G. Caporaso and R. Knight in *Annual Review of Biochemistry*, Vol. 74, pages 179–198; July 2005.

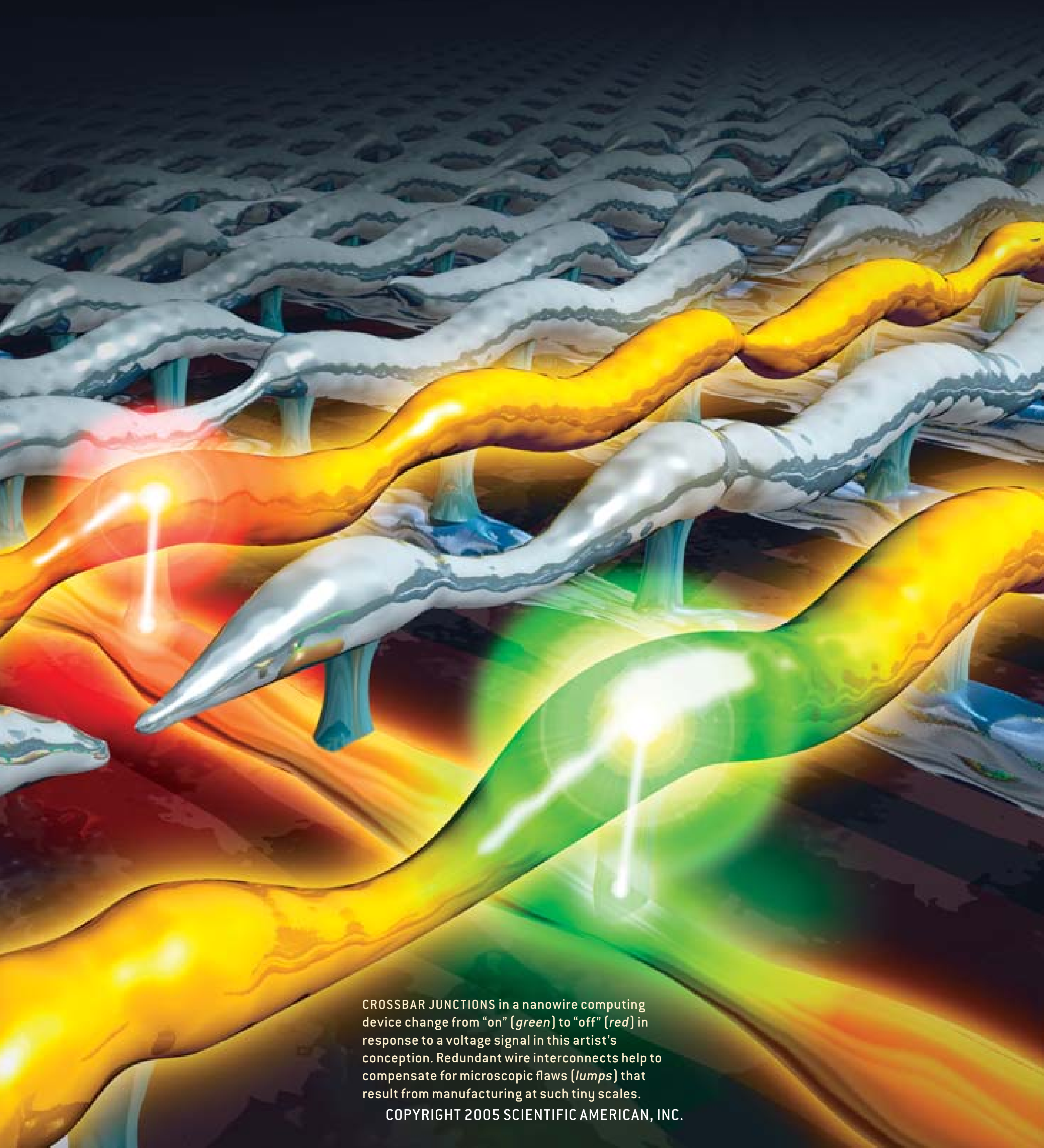
CROSSBAR NANOCOMPUTERS

Crisscrossing assemblies of defect-prone nanowires



could succeed today's silicon-based circuits

By Philip J. Kuekes, Gregory S. Snider and R. Stanley Williams



CROSSBAR JUNCTIONS in a nanowire computing device change from "on" (*green*) to "off" (*red*) in response to a voltage signal in this artist's conception. Redundant wire interconnects help to compensate for microscopic flaws (*lumps*) that result from manufacturing at such tiny scales.

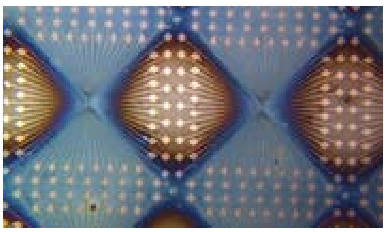
IN A LITTLE OVER HALF A CENTURY, THE NUMBER OF TRANSISTORS ON A SILICON CHIP has grown from just one to nearly a billion—an accomplishment celebrated as Moore's Law. By greatly enhancing digital machines' ability to crunch numbers, execute logical operations and store data, this unprecedented manufacturing success has enabled revolutionary changes in our day-to-day lives while spawning one of the planet's largest and most influential industries.

As more and more transistors are packed onto silicon integrated circuits (ICs) during the next decade and a half, the lengths of the smallest chip features will shrink to nearly the molecular scale. Even the most optimistic proponents of ICs believe that major innovations will be required to reach the ultimate operating limit of the silicon transistor: a length for functional features around 10 nanometers (nm), or about 30 atoms long. Finding alternative technologies that can further shrink computing devices is crucial to maintaining technological progress. But because of the silicon IC's amazingly successful track record, the performance bar for any successor is so high it will take at least a decade to develop candidates that will be available when they are needed.

Researchers worldwide are exploring several exciting alternatives. Quantum computing, for instance, is a novel technique that takes advantage of "spooky" quantum-mechanical properties to process information. Quantum computing is, however, decades away from realization, and even then it remains unclear how useful it would be for most applications. Many research groups are therefore searching for a midterm alternative that could be ready for commercialization in about 10 years. To be economically viable, such a technology must share a great deal with the existing IC processor infrastructure, including critical items like fabrication foundries and software platforms.

Our research team at Hewlett-Packard (HP) Laboratories views the crossbar architecture as the most likely path forward. A crossbar consists of one set of parallel nanowires (less than 100 atoms wide) that cross over a second set. A material that can be stimulated electrically to conduct either more electricity or less is sandwiched between the two sets of wires. The resulting interwire junctions form a switch at each intersection between crossing wires that can hold its "on" or "off" status over time.

Crossbars offer several benefits: The regular pattern of crisscrossing nanowires makes manufacturing relatively straightforward, especially compared with the complex structures of microprocessors. Its arraylike composition provides clear ways to instill defect tolerance in circuits. The structure can be built using a wide range of substances and processes, which provides tremendous flexibility in adapting existing designs to new materials. Finally, this single geometry can provide memory, logic and interconnection, making it very adaptable.



CROSSBAR memories with their test pads.

Random fluctuations make it impossible to build a perfect machine from nanocomponents.

Overview/*Nanoelectronics*

- Moving beyond today's silicon integrated chip technology will require shrinking logic and memory circuits to the scale of a few nanometers. Large arrays of intersecting nanowires called crossbars provide the basis for one of the best candidate technologies for nanocomputing success.
- The nanowires that comprise crossbars are so small that atomic defects and flaws in their manufacture are unavoidable and serious. Building redundancy into the circuitry and using coding theory techniques compensate for the many imperfections.

(Criss)crossing Over

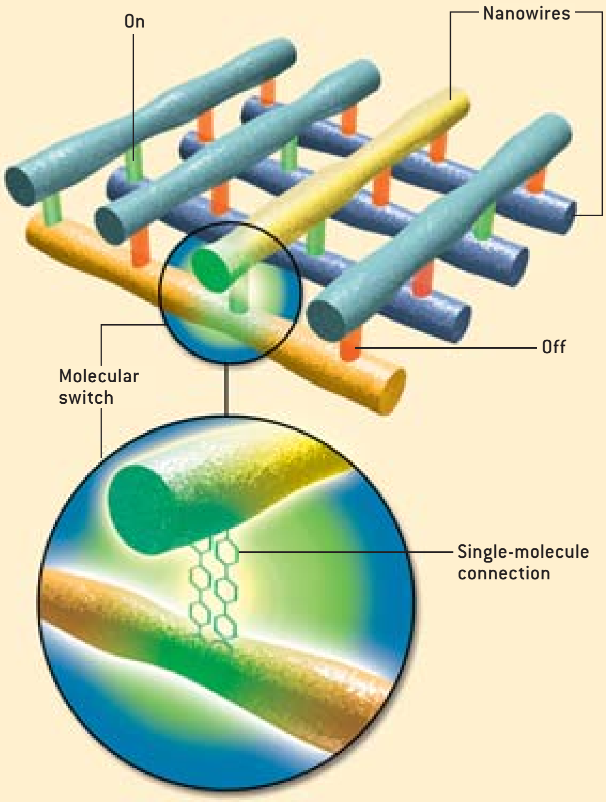
OUR TEAM'S JOURNEY toward this avenue of research began in 1995, when one of us (Williams) moved to HP from the chemistry department at the University of California, Los Angeles. Though not a computer expert, he did know a few things about electronics: one, that a computer's circuits had to be perfect to operate correctly and, two, that random atomic fluctuations at room temperature and above (caused by entropy) would make it impossible to build a perfect machine from billions of components, each composed of only a few atoms. Even

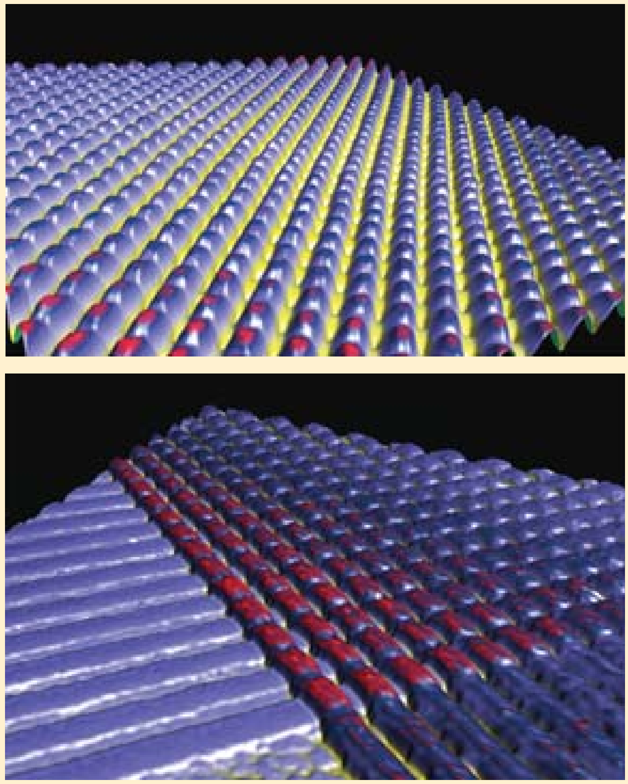
SLIM FILMS (*preceding pages*); G.-Y. JUNG AND B. BAUSK *HP Labs* (*this page*)

ON AND OFF AT THE CROSSROADS

The key component of the crossbar architecture is a nanoscale switch that can be turned "on" or "off" by applying an appropriate voltage to the wires it connects.

In Hewlett-Packard (HP) Laboratories's version, the switch is formed at the junction between two crossing nanowires that are separated by a single monolayer of molecules. The switch starts out in a high-resistance state, blocking the flow of electrons between its two nanowires ("off," *red highlights, below*). But when a large enough voltage of the appropriate polarity is placed across it (*indicated by the yellow and orange nanowires*), the switch changes abruptly to a much lower resistance state, allowing electrons to flow more easily ("on," *green highlights*). The switch stays in this low-resistance state until a large enough negative voltage makes it revert to its original state. As long as the voltage is maintained between these positive and negative thresholds, the switch remains in the state in which it was last set. Some switches the authors have examined have retained their set states for more than three years so far. If the switches can be toggled back and forth many times, they are reconfigurable and can be used in a random-access memory or a reprogrammable logic circuit.





PROTOTYPE ARRAY for a crossbar computing device, depicted in an atomic-force micrograph (*top*), has 34 nanowires (each 30 nanometers wide) intersecting with 34 others. The detail (*bottom*) shows how one set of nanowires crosses over the other set. A junction of two nanowires is smaller than a typical virus.

JANA BRENNING (*illustrations*); COURTESY OF W. WU AND G. MEDEIROS-RIBEIRO *HP Labs* (*micrographs*)

atomic-scale irregularities impose significant variations on the size of nanodevices, which can destroy their electrical properties. Consequently, some sizable fraction of the tiny devices will not work. It was natural for Williams to conclude that nanoelectronics were therefore impossible and that his research at HP should focus on other technologies.

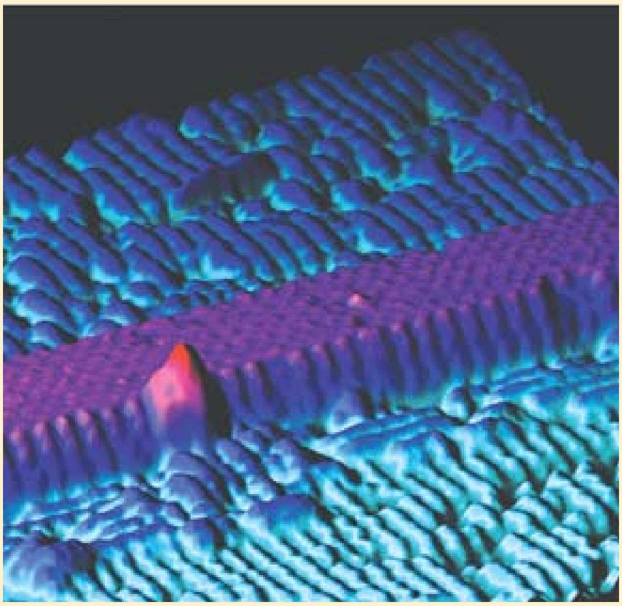
A chance meeting the following year with an HP computer architect (Kuekes) changed that perception dramatically and set the pair on an unexpected path. Kuekes told Williams about a supercomputer called Teramac that he and others (including Snider) had built. Teramac operated perfectly, even though about 220,000 of its components (approximately 3 percent of the total) were defective. The trick, Kuekes said, was that the supercomputer's design had significant redundancy in its interconnect circuitry. After all the flaws were located and catalogued, programs run on the computer were compiled to avoid the broken parts, essentially by routing around the defects via the extra connections.

Williams saw at once that Teramac's tolerance of defects provided a way to construct computers that operate perfectly despite huge numbers of "broken" nanoscale parts. That summer Williams and visiting U.C.L.A. chemist James R. Heath worked on applying the concepts of nanoparticle assembly (assembling complex structures out of tiny building blocks) to computers. After much discussion with Kuekes and Snider about the defect tolerance of chemically assembled computing

BUILD TOP-DOWN OR BOTTOM-UP?

The field of nanoscale fabrication is extremely active today, with many competing techniques under study. These approaches can be classified into two categories: top-down and bottom-up (*below*). The former examples resemble conventional IC manufacturing methods that use photolithography followed by chemical etching or deposition of materials to create the desired features. The latter approaches are based on extensions of chemical or biochemical processes by which atoms or molecules self-assemble into a desired configuration because of their planned, inherent properties. Most investigators in this field agree that some combination of the two approaches will be required to build future nanoscale circuits.

At HP, our team uses imprint lithography to create the crossbars. We and our collaborators employ electron beam lithography to construct molds for the circuits. Although this process is slow and costly, we can make duplicates of the final product, which then are used to stamp out large quantities of circuits, much as vinyl LP records were made. A thin layer of a polymer or polymer precursor coats a substrate, the mold is pressed into this soft layer, and the impressed pattern hardens under exposure to heat or ultraviolet light. The advantage of this approach is that electron-beam lithography can fabricate arbitrary wire geometries on the mold. The drawback is that the present resolution of the features in a set of parallel wires is limited to roughly 30-nanometer half-pitch (half the distance between the centers of two wires, a standard industry measure), although we are working on a number of techniques to improve on this performance. *—P.J.K., G.S.S. and R.S.W.*



ATOMIC FLAWS AND DEFECTS appear in this scanning tunneling microscope image of a nanowire of erbium silicide that was grown on a silicon surface using a chemical (bottom-up) method. Bumps on the surface of the wire, which is approximately three nanometers (or 10 atoms) wide, are individual atoms. The bulge on the side of the nanowire is a defect where the width changes from 10 atoms wide to only nine.

D.A.A. OHLBERG AND G. MEDEIROS-RIBEIRO *HP Labs*

systems, Williams and Heath wrote a paper about the topic as an educational exercise. To the surprise of all involved, it was taken seriously and eventually published in *Science* in 1998.

Rapid Results Required

THAT SAME YEAR Bruce E. Gnade and William L. Warren, then program directors at the Defense Advanced Research Projects Agency (DARPA), recognized that effective architecture was critical for developing the new nanoscale device technologies the agency was supporting. At the time, interest in molecular electronics was enjoying a resurgence, years after it had first been proposed in 1974 by Avi Aviram of IBM and Mark A. Ratner of Northwestern University. It was not, however, until the early 1990s that Mark A. Reed of Yale University and James M. Tour of Rice University actually started measuring the electrical properties and synthesizing new molecules for electronics. Gnade and Warren understood that electronic devices without an architecture to link them into a useful circuit were mere intellectual curiosities. Their challenge to the research community to define a workable architecture for molecular devices kick-started the research efforts of many groups and encouraged the formation of several significant collaborations.

Our HP/U.C.L.A. team immediately embraced that challenge, but we faced a dilemma. The Teramac-inspired architecture that we had proposed would have required five years to develop, but DARPA wanted tangible results (as a 16-bit memory device) in only two. Heath, Kuekes and Williams brainstormed during the next several weeks to come up with a concept that could meet the deadline. Kuekes and Williams were aware of HP's magnetic random-access memory project and understood that the simple crossbar structure on which it was based was the ultimate abstraction of the Teramac configuration.

Heath pointed out that a crossbar "looked like a crystal" and that it should therefore be possible to build such a system chemically. What was needed was some way to connect each pair of intersecting wires in the crossbar with a switch that could be turned on and off at will. Williams suggested that an electrochemically active material sandwiched between the wires should make it possible to change the electrical resistance of the contacts substantially and reversibly, by applying the appropriate voltages across the two nanowires. That is, the switch would be closed by electrochemically shrinking the quantum-mechanical "tunneling" gap that the electrons have to jump across to get from one electrode to the other. Applying the opposite voltage bias to widen the tunneling gap and raise the electrical resistance would reopen the switch.

Heath provided the material we needed. He introduced our collaboration to molecules that had been designed by J. Fraser Stoddart, then a new U.C.L.A. faculty member, to operate as electrochemically actuated mechanical switches. The concept was that anything that will change shape between two wires should also affect the ability of electrons to tunnel from one wire to another. A key step was persuading a very busy Stoddart to chemically modify his molecules, which he had christened rotaxanes, to make them oily. This alteration enabled Heath to place a small drop of rotaxanes on a water surface so they would spread out to form a film one molecule thick, which was transferred onto a substrate (a process called the Langmuir-Blodgett technique) on which the bottom set of wires had been formed. After that, we deposited the top set of wires by evaporating metal through a mask, which completed the circuit. These early experiments led to several U.S. patent applications, a proposal to DARPA and another paper in *Science*.

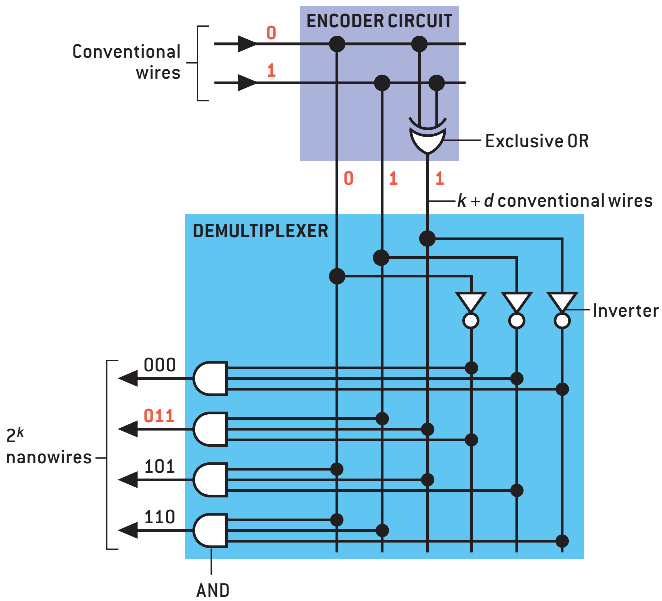
Making the Cut

DESPITE CONSIDERABLE SKEPTICISM on the part of the research community, our crossbar and electrochemical switch concept was accepted by DARPA for its two-year trial, along with several others. Early in this effort the Heath and Stoddart research groups demonstrated that rotaxane molecules sandwiched between electrodes could indeed toggle between high- and low-resistance states. We and others, including Charles Lieber's group at Harvard University as well as the Reed and Tour groups, have since seen a range of different nanoscale switching mechanisms. The diverse observations and approaches generated some confusion within the broader research community, and the various switching phenomena have yet to be sorted out, but the existence of electrical switching is today widely recognized. Dozens of research teams across the globe are now working to develop robust nanoscale electrical switches based on atoms or molecules [*see box on page 80*].

Using the crossbar structure, our U.C.L.A. partners became the first group to demonstrate a working 16-bit memory for the DARPA program in 2000. Their success encouraged the agency to fund a successor program with a far more ambitious goal: the fabrication of a 16-kilobit memory with a density of 100 billion bits per square centimeter. This objective sets the bar extremely high, because it requires fabrication capabilities that are not expected to become available to the semiconductor industry until around 2018.

Our group at HP continues to invent new types of circuits based on the crossbar—notably, defect-tolerant memories and different families of logic circuits. Interesting modifications of the original architectural concept also have been developed by André DeHon of the California Institute of Technology, who collaborates with the Lieber group, and Konstantin K. Likharev of Stony Brook University. Although the crossbar and switch architecture started off as the dark-horse candidate in the DARPA challenge, it has now been adopted and adapted by many research groups worldwide, including those of Masakazu Aono of the National Institute for Materials Science in Japan and Rainer Waser of the Research Center Jülich in Germany.

To understand the crossbar approach, we must discuss the nature of the switch and crossbar structure, the fabrication of crossbars from nanoscale wires [*see box on opposite page*] and the possibility of building reliable circuits from unreliable components.



DEMULTIPLEXER enables conventional wires on silicon chips to control a far greater number of nanowires. If k is the number of conventional wires, the multiplexer can control 2^k nanowires. An additional d conventional wires provide sufficient redundancy for the control to work despite broken connections between nanowires and conventional wires. In this simplified diagram, $k = 2$ and $d = 1$; two micron-scale wires control four nanowires with one bit of redundancy . In this example, the conventional wires input the switch address 01 (*red*), to which the encoder circuit adds a redundant bit, yielding the coded address 011. The coded address then activates the nanowire in the demultiplexer designated 011 [*see next page for further explanation*].

THE AUTHORS

PHILIP J. KUEKES, GREGORY S. SNIDER and *R. STANLEY WILLIAMS* develop next-generation computing technologies at the Quantum Science Research (QSR) program at Hewlett-Packard Laboratories in Palo Alto, Calif. Kuekes devises novel ideas in the areas of computation, electronic circuits and devices, and quantum information research. The chief architect of the QSR program has designed and built leading-edge computers for more than 30 years. Snider, currently a consultant with HP, is exploring ways to improve the architectural design of nanoelectronics. He has worked previously on logic circuit design, compilers, operating systems, logic synthesis, digital signal processing, computer security and networking systems. Williams, as HP Senior Fellow and director of HP's QSR program, guides the multidisciplinary team that designs, builds and tests new nanocircuits. In the past Williams focused on solid-state chemistry and physics, but his primary interest now is the study of the intersection of nanoscience and information technology.

SLIM FILMS; SOURCE: WARREN ROBINETT

From Micro to Nano and Back

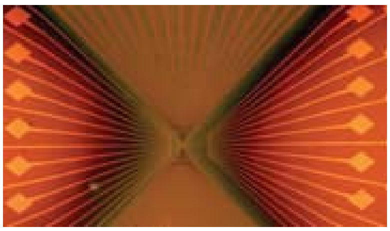
THE PHILOSOPHY BEHIND the nanoscale crossbar is that we must learn to live with its unavoidable imperfections and work around them. The "find and avoid" strategy of Teramac will work as long as it is possible to communicate with the nanowires. This, however, poses another question: How does one bridge the gaps in size and number of wires between nanoelectronics and the conventional-scale silicon ICs that will be required to control the crossbars? If only one-to-one connections can be made, nanoscale crossbars would afford no net advantage. We can solve that problem by making the electrical connections through a demultiplexer, a subcircuit that takes as an input a binary number (such as 1010) and selects a single nanowire that has that binary string as a unique identifier [*see illustration on preceding page*]. In our case, the demultiplexer is a special type of crossbar in which many nanowires connect to a small number of conventional wires.

The number of wires required to input a binary address is the same as the length of the digital names, but the quantity of nanowires that can be addressed equals the number of unique addresses. For example, a string four bits long (0000, 0001, 0010 and so on) can specify 16 unique addresses. Therefore, four micron-scale wires can control 16 nanowires. This fact is important because, to make building the nanoscale circuits worthwhile, one needs to be able to control a lot of nanocircuitry with little conventional electronics. In general, if k is the number of conventional wires feeding into the demultiplexer, it can individually control 2^k nanowires, which is very favorable exponential scaling.

A big problem occurs, however, if one of the connections between a nanowire in the demultiplexer and a conventional wire is broken. Then it is no longer possible to distinguish among the k different nanowires that share that defective bit in its address. (For instance, if the last bit in the string is broken, then 0000 and 0001 seem identical, as do 1110 and 1111 and other pairs.) Hence, one bad connection in the demultiplexer leads to the loss of all the nanocircuitry that is downstream from k nanowires—a catastrophic failure. This result appears to require that demultiplexers, which are half nanocircuitry, be perfect—a violation of our guiding principle that nanoelectronics will be defective.

We find the solution to this quandary in the field of coding theory, which engineers apply when transmitting digital information through noisy environments (as in orbital satellite communications). The general idea is: first, break up a message into blocks of binary data—strings of 0s and 1s. Each block is then extended by adding some extra bits to make a larger block, the code. The extra bits are calculated with an algebraic expression utilizing the bits in the original message block as inputs. When this larger message is sent through the air or some other noisy environment, a few of the bits in the coded message may be garbled (some 1s are turned into 0s and vice versa). By running the code backward on the receiving end, however, the original message can be recovered exactly (as long as the number of flipped bits does not overwhelm the code).

With the guidance of Gadiel Seroussi, Ronnie Roth and Warren Robinett of HP, our team has applied this concept to protect our nanowires from broken connections in the demultiplexer. Rather than numbering the nanowires consecutively, we use an extended address space in which the number of wires entering the demultiplexer is larger than the minimum number needed to specify each nanowire uniquely (by an additional d wires). In this case, it turns out that each nanowire can have several broken connections to the conventional wires, yet the demultiplexer can still successfully address all the nanowires. The degree of redundancy required depends on the probability of connection defects; a relatively small amount of redundancy (about 40 percent) can improve the effective manufacturing yield of a demultiplexer from 0.0001 to 0.9999 if the defect rate of the connections in the demultiplexer is 0.01.



FAN-OUT WIRES from crossbar to test pads.

We managed to create logic functions without the use of transistors in a crossbar circuit.

Making Memories

SINCE THAT FIRST 16-bit memory device, both the Heath group and our team at HP have demonstrated 64-bit memories at 62-nm half-pitch (which is half the distance between the centers of two adjacent wires, a standard semiconductor industry measure) in 2002 and last year a one-kilobit crossbar at 30-nm half-pitch, using different approaches for the wires and switches. (In comparison, the half-pitch of the most advanced semiconductor IC in 2005 is 90 nm.) Each nanowire in these demonstration memories was connected to an individual contact. We wrote a bit as a 1 (low-resistance) or a 0 (high-resistance) simply by applying a bias voltage that exceeded the threshold for directly toggling the desired switch across its two wires. As long as the voltage threshold for recording a 1 or a 0 is relatively sharp and the variation in "write" voltages among the junctions in the array is less than one half the switching voltage, this procedure ensures that only the desired bit in the array is written (and no others are accidentally written or erased). We read the bit stored in the switch by applying a much lower volt-

Y. CHEN AND B. BAUSK *HP Labs*

Groups Researching Crossbar Architectures

GROUP(S)	INSTITUTION(S)	SWITCH
J. R. Heath/ J. F. Stoddart	Caltech/U.C.L.A.	Rotaxane monolayer between silicon and titanium nanowires
C. Lieber/ A. DeHon	Harvard University/Caltech	Silicon nanowire field-effect transistors
M. Aono	National Institute for Materials Science, Japan	Silver sulfide ionic conductor (silver-based atomic switch)
R. Waser	Research Center Jülich, Germany	Defect motion in ferroelectric thin films
K. K. Likharev	Stony Brook University	Molecular single-electron transistor
Quantum Science Research	Hewlett-Packard Laboratories	Metal nanowire oxidation/reduction

age across the selected crossing wires and measuring the resistance at their junction. These initial results proved promising—in HP's 64-bit memory, the resistance ratio between 1 ("on") and 0 ("off") exceeded 100, making the bits very easy to read.

With the goal of nanoscale memory within reach (the DARPA challenge requires a half-pitch of 16 nm), our next big hurdle is to perform universal computation with nanoscale logic circuits. With Duncan R. Stewart at HP, we have configured crossbars to perform simple logic (Boolean AND and OR operations) by setting the resistance values of switches in a crossbar. The range of logic that can be performed, however, is limited without the NOT operation, or signal inversion, which changes a 1 to a 0 and a 0 to a 1. The wired logic functions also necessarily cause the voltage levels to trail off; if one tries to use too many in a series circuit, 1s and 0s would no longer be distinguishable and computation would not be possible.

In silicon ICs, transistors perform both signal restoration and inversion. This fact has motivated the Heath and Lieber groups to fabricate transistors made from silicon nanowires. We and DeHon have described logic circuits with a "tile and mosaic" topology that can be built using transistors and other elements that are fabricated into a crossbar. Because this approach employs current IC technology, however, eventually it suffers from the aforementioned limitations, so it does not offer an extension beyond Moore's Law. As an alternative, we are investigating signal inversion and restoration without transistors.

Our team is building an unusual form of crossbar logic circuit with arrays of switches and wired ANDs and ORs. In this case, the switches perform a latching operation, which we recently demonstrated with Stewart. We define the voltage level needed to turn a switch on as a 1 and that to turn it off as a 0. Any wire connected to the input of a switch will perforce set that switch to the wire's present logical state, thus transferring one bit of information from "logic" to "memory."

Once stored as a memory state, that bit can be employed in further logic operations by connecting the output wire from the switch to a voltage supply (in our case, a wire from the clock that controls the timing of the operations). This new connection can then be used to restore the voltage of the logic state to its desired value when it has degraded. Another trick is to switch the voltages representing a 1 and a 0 on the output wires, which inverts the logic signal. This change supplies the logical NOT operation and, combined with either ANDs or ORs, is sufficient to perform any computation. Hence, we managed to create the signal restoration and inversion functions without the use of transistors or their semiconductor properties in a crossbar logic circuit.

Beyond Silicon ICs

THE PATH TO universal computing beyond transistor integrated circuits is still highly uncertain, but the crossbar architecture has emerged during the past several years as a principal contender for a new computing paradigm. Much remains to be done. Three different areas of research must advance rapidly and together: architecture, device physics and nanomanufacturing. Ensuring good communications across disciplinary boundaries will be as challenging as solving the technical issues. Success will require multiple teams of researchers who are simultaneously competing against and cooperating with one another, such as the participants in the DARPA challenge.

Defect tolerance will be a necessary element of any future strategy for nanoelectronics. The crossbar architecture is ideal for implementing strategies based on finding and avoiding bad components and on coding theory to compensate for mistakes. Future circuits may actually be more robust than current electronics, even though they will start out with a high fraction of defective components. The built-in redundancy will make them resistant to forces (such as radiation exposure) that cause catastrophic failures in conventional circuits and instead enable their performance to degrade gracefully.

The quantum-mechanical nature of tunneling switches is suited for nanoscale circuits. As the feature sizes of devices shrink, the electrons in them behave more like quantum-mechanical bodies. Such switches should be able to scale down to nearly single-atom dimensions—which suggests just how far the future miniaturization of electronic circuitry might someday go. SA

MORE TO EXPLORE

Configurable Computing. John Villasenor and William H. Mangione-Smith in *Scientific American,* Vol. 276, No. 6, pages 66–71; June 1997.

A Defect-Tolerant Computer Architecture: Opportunities for Nanotechnology. J. R. Heath, P. J. Kuekes, G. S. Snider and R. S. Williams in *Science,* Vol. 280, pages 1716–1721; June 1998.

Computing with Molecules. Mark A. Reed and James M. Tour in *Scientific American,* Vol. 282, No. 6, pages 86–93; June 2000.

Feynman Lectures in Computation. Paperbound edition. Richard P. Feynman. Edited by Tony Hey and Robin W. Allen. Perseus Books Group, 2000.

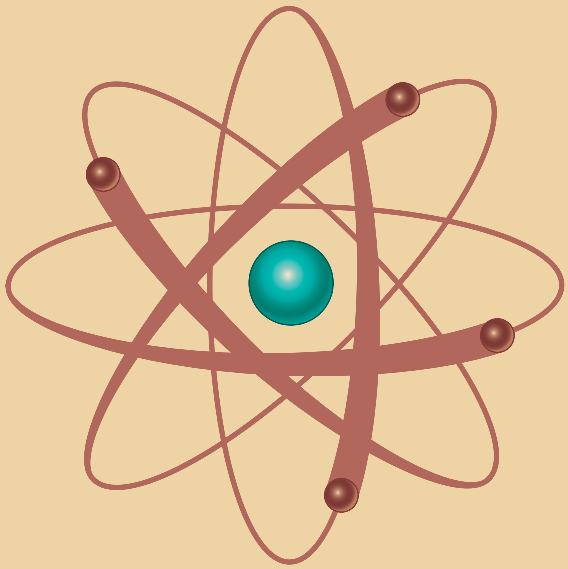
The International Technology Roadmap for Semiconductors (ITRS) Web site is at **http://public.itrs.net/**

OPEN-PIT MINING of the Oklo uranium deposit in Gabon revealed more than a dozen zones where nuclear fission had once taken place.

THE WORKINGS OF AN

ANCIENT NUCLEAR REACTOR

BY ALEX P. MESHIK



Two billion years ago parts of an African uranium deposit spontaneously underwent nuclear fission. The details of this remarkable phenomenon are just now becoming clear

FRANÇOIS GAUTHIER-LAFAYE

In May 1972 a worker at a nuclear fuel–processing plant in France noticed something suspicious.

He had been conducting a routine analysis of uranium derived from a seemingly ordinary source of ore. As is the case with all natural uranium, the material under study contained three isotopes—that is to say, three forms with differing atomic masses: uranium 238, the most abundant variety; uranium 234, the rarest; and uranium 235, the isotope that is coveted because it can sustain a nuclear chain reaction. Elsewhere in the earth's crust, on the moon and even in meteorites, uranium 235 atoms make up 0.720 percent of the total. But in these samples, which came from the Oklo deposit in Gabon (a former French colony in west equatorial Africa), uranium 235 constituted just 0.717 percent. That tiny discrepancy was enough to alert French scientists that something strange had happened. Further analyses showed that ore from at least one part of the mine was far short on uranium 235: some 200 kilograms appeared to be missing—enough to make half a dozen or so nuclear bombs.

For weeks, specialists at the French Atomic Energy Commission (CEA) remained perplexed. The answer came only when someone recalled a prediction published 19 years earlier. In 1953 George W. Wetherill of the University of California at Los Angeles and Mark G. Inghram of the University of Chicago pointed out that some uranium deposits might have once operated as natural versions of the nuclear fission reactors that were then becoming popular. Shortly thereafter, Paul K. Kuroda, a chemist from the University of Arkansas, calculated what it would take for a uranium-ore body spontaneously to undergo self-sustained fission. In this process, a stray neutron causes a uranium 235 nucleus to split, which gives off more neutrons, causing others of these atoms to break apart in a nuclear chain reaction.

NATURAL FISSION REACTORS have been found only in the African nation of Gabon, at the Oklo and adjacent Okelobondo uranium mines, and at a site named Bangombe, located some 35 kilometers away.

Kuroda's first condition was that the size of the uranium deposit should exceed the average length that fission-inducing neutrons travel, about two thirds of a meter. This requirement helps to ensure that the neutrons given off by one fissioning nucleus are absorbed by another before escaping from the uranium vein.

A second prerequisite is that uranium 235 must be present in sufficient abundance. Today even the most massive and concentrated uranium deposit cannot become a nuclear reactor, because the uranium 235 concentration, at less than 1 percent, is just too low. But this isotope is radioactive and decays about six times faster than does uranium 238, which indicates that the fissile fraction was much higher in the distant past. For example, two billion years ago (about when the Oklo deposit formed) uranium 235 must have constituted approximately 3 percent, which is roughly the level provided artificially in the enriched uranium used to fuel most nuclear power stations.

The third important ingredient is a neutron "moderator," a substance that can slow the neutrons given off when a uranium nucleus splits so that they are more apt to induce other uranium nuclei to break apart. Finally, there should be no significant amounts of boron, lithium or other so-called poisons, which absorb neutrons and would thus bring any nuclear reaction to a swift halt.

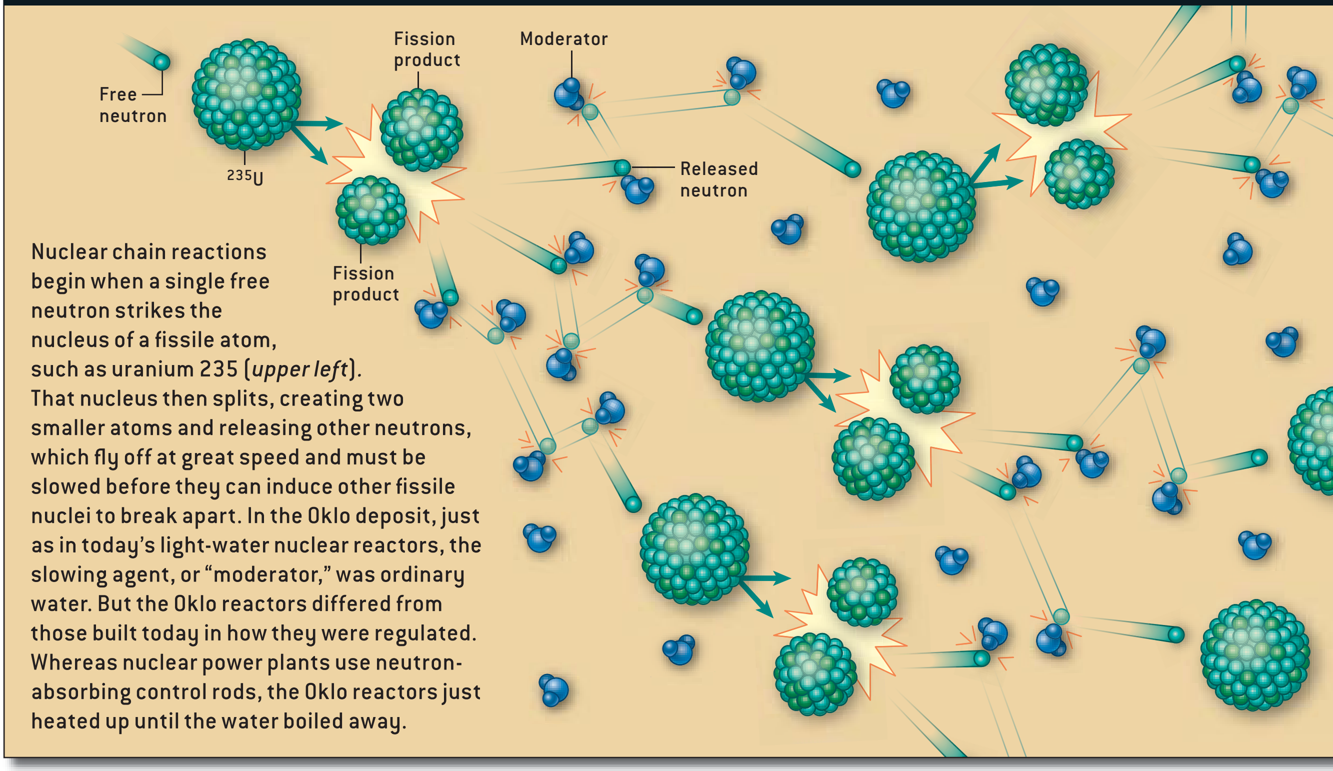
Amazingly, the actual conditions that prevailed two billion years ago in what researchers eventually determined to be 16 separate areas within the Oklo and adjacent Okelobondo uranium

Overview/*Fossil Reactors*

- Three decades ago French scientists discovered that parts of a uranium deposit then being mined in Gabon had long ago functioned as natural fission reactors.
- The author and two colleagues recently used measurements of xenon gas (a product of uranium fission) to deduce that one of these ancient reactors must have operated with a duty cycle of about half an hour on and at least two and a half hours off.
- Perhaps further studies of xenon retained within minerals will reveal natural nuclear reactors elsewhere. But for the moment, the examples discovered in Gabon remain unique windows on possible changes in fundamental physical constants and on how buried nuclear waste migrates over time.

JEN CHRISTIANSEN

FISSION UP CLOSE



Nuclear chain reactions begin when a single free neutron strikes the nucleus of a fissile atom, such as uranium 235 (*upper left*). That nucleus then splits, creating two smaller atoms and releasing other neutrons, which fly off at great speed and must be slowed before they can induce other fissile nuclei to break apart. In the Oklo deposit, just as in today's light-water nuclear reactors, the slowing agent, or "moderator," was ordinary water. But the Oklo reactors differed from those built today in how they were regulated. Whereas nuclear power plants use neutron-absorbing control rods, the Oklo reactors just heated up until the water boiled away.

JEN CHRISTIANSEN

mines were very close to what Kuroda outlined. These zones were all identified decades ago. But only recently did my colleagues and I finally clarify major details of what exactly went on inside one of those ancient reactors.

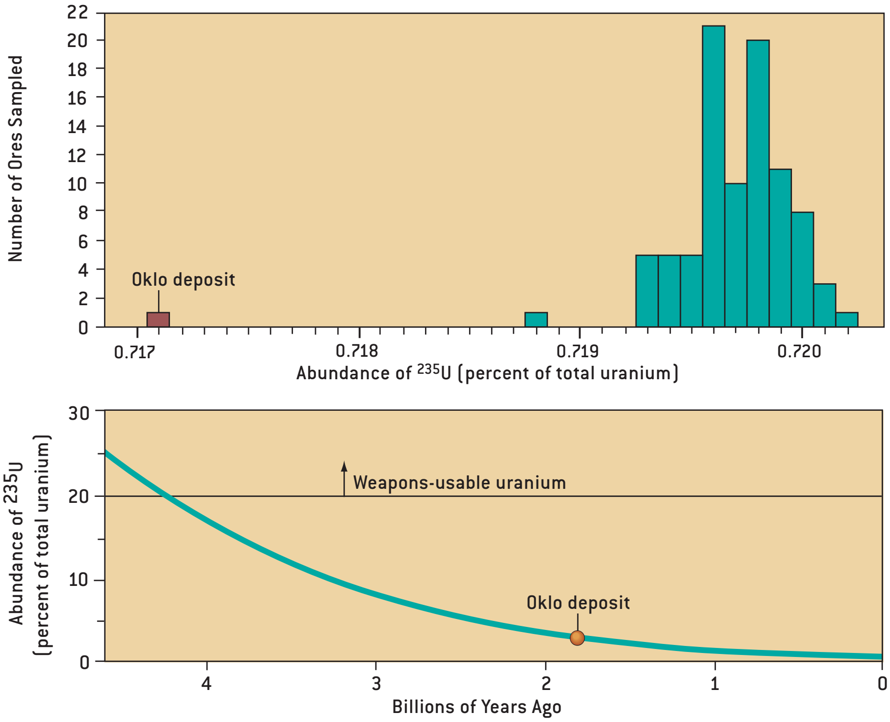
Proof in the Light Elements

PHYSICISTS CONFIRMED the basic idea that natural fission reactions were responsible for the depletion in uranium 235 at Oklo quite soon after the anomalous uranium was discovered. Indisputable proof came from an examination of the new, lighter elements created when a heavy nucleus is broken in two. The abundance of these fission products proved so high that no other conclusion could be drawn. A nuclear chain reaction very much like the one that Enrico Fermi and his colleagues famously demonstrated in 1942 had certainly taken place, all on its own and some two billion years before.

Shortly after this astonishing discovery, physicists from around the world studied the evidence for these natural nuclear reactors and came together to share their work on "the Oklo phenomenon" at a special 1975 conference held in Libreville, the capital of Gabon. The next year George A. Cowan, who represented the U.S. at that meeting (and who, incidentally, is one of the founders of the renowned Santa Fe Institute, where he is still affiliated), wrote an article for *Scientific American* [see "A Natural Fission Reactor," by George A. Cowan, July 1976] in which he explained what scientists had surmised about the operation of these ancient reactors.

Cowan described, for example, how some of the neutrons released during the fission of uranium 235 were captured by the more abundant uranium 238, which became uranium 239 and, after emitting two electrons, turned into plutonium 239. More than two tons of this plutonium isotope were generated within the Oklo deposit. Although almost all this material, which has a 24,000-year half-life, has since disappeared (primarily through natural radioactive decay), some of the plutonium itself underwent fission, as attested by the presence of its characteristic fission products. The abundance of those lighter elements allowed scientists to deduce that fission reactions must have gone on for hundreds of thousands of years. From the amount of uranium 235 consumed, they calculated the total energy released, 15,000 megawatt-years, and from this and other evidence were able to work out the average power output, which was

The Oklo reactors managed to maintain a modest power output for perhaps A FEW HUNDRED MILLENNIA.



URANIUM 235 ATOMS constitute about 0.720 percent of natural uranium—almost universally. So when workers discovered that uranium taken from the Oklo mine contained little more than 0.717 percent, they were justifiably startled. This result indeed falls far outside the range found for other uranium ores (*top*). The explanation is that in the past the ratio of uranium 235 to uranium 238 was much higher, as can be deduced from uranium 235's much shorter half-life. That higher ratio made fission possible, and fission consumed much of the uranium 235. When the Oklo deposit formed 1.8 billion years ago, the natural level of uranium 235 was about 3 percent—similar to what is now provided in many nuclear-reactor fuels. When the earth first formed, some 4.6 billion years ago, the figure exceeded 20 percent, a level at which uranium would today be considered "weapons-usable."

probably less than 100 kilowatts—say, enough to run a few dozen toasters.

It is truly amazing that more than a dozen natural reactors spontaneously sprang into existence and that they managed to maintain a modest power output for perhaps a few hundred millennia. Why is it that these parts of the deposit did not explode and destroy themselves right after nuclear chain reactions began? What mechanism provided the necessary self-regulation? Did these reactors run steadily or in fits and starts? The solutions to these puzzles emerged slowly after initial discovery of the Oklo phenomenon. Indeed, the last question lingered for more than three decades before my colleagues and I at Washington University in St. Louis began to address it by examining a piece of this enigmatic African ore.

Noble-Gas Epiphanies

OUR RECENT WORK on one of the Oklo reactors centered on an analysis of xenon, a heavy inert gas, which can remain imprisoned within minerals for billions of years. Xenon possesses nine stable isotopes, produced in various proportions by different nuclear processes. Being a noble gas, it resists chemical bonding with other elements and is thus easy to purify for isotopic analysis. Xenon is extremely rare, which allows scientists to use it to detect and trace nuclear reactions, even those that occurred in primitive meteorites before the solar system came into existence.

To analyze the isotopic composition of xenon requires a mass spectrometer, an instrument that can separate atoms according to their atomic weight. I was fortunate to have access to an extremely accurate xenon mass spectrometer, one built by my Washington colleague Charles M. Hohenberg. But before using his apparatus, we had to extract the xenon from our sample. Scientists usually just heat the host material, often above the melting point, so that the rock loses its crystalline structure and cannot hold on to its hidden cache of xenon. To glean greater information about the genesis and retention of this gas, we adopted a more delicate approach called laser extraction, which releases xenon selectively from a single mineral grain, leaving adjacent areas intact.

We applied this technique to many tiny spots on our lone available fragment of Oklo rock, only one millimeter thick and four millimeters across. Of course, we first needed to decide where exactly to aim the laser beam. Here Hohenberg and I relied on our colleague Olga Pravdivtseva, who had constructed a detailed x-ray map of our sample and identified the constituent minerals. After each extraction, we purified the resulting gas and passed the xenon into Hohenberg's mass spectrometer, which indicated the number of atoms of each isotope present.

Our first surprise was the location of the xenon. It was not, as we had expected, found to a significant extent in the uranium-rich mineral grains. Rather the lion's share was trapped in aluminum phosphate minerals, which contain no uranium at all. Remarkably, these grains showed the highest concentration of xenon ever found in any natural material. The second epiphany was that the extracted gas had a significantly different isotopic makeup from what is usually produced in nuclear reactors. It had

THE AUTHOR

ALEX P. MESHIK began his study of physics at St. Petersburg State University in Russia. He obtained his Ph.D. at the Vernadsky Institute of the Russian Academy of Sciences in 1988. His doctoral thesis was devoted to the geochemistry, geochronology and nuclear chemistry of the noble gases xenon and krypton. In 1996 Meshik joined the Laboratory for Space Sciences at Washington University in St. Louis, where he is currently studying, among other things, noble gases from the solar wind that were collected and returned to the earth by the Genesis spacecraft.

JEN CHRISTIANSEN; SOURCE FOR HISTOGRAM: "THE VARIABILITY OF THE NATURAL ABUNDANCE OF ^{235}U," BY GEORGE A. COWAN AND HANS H. ADLER IN *GEOCHIMICA AND COSMOCHIMICA ACTA*, VOL. 40; 1976 seemingly lost a large portion of the xenon 136 and 134 that would certainly have been created from fission, whereas the lighter varieties of the element were modified to a lesser extent.

How could such a change in isotopic composition have come about? Chemical reactions would not do the trick, because all isotopes are chemically identical. Perhaps nuclear reactions, such as neutron capture? Careful analysis allowed my colleagues and me to reject this possibility as well. We also considered the physical sorting of different isotopes that sometimes takes place: heavier atoms move a bit more slowly than their lighter counterparts and can thus sometimes separate from them. Uranium-enrichment plants—industrial facilities that require considerable skill to construct—take advantage of this property to produce reactor fuel. But even if nature could miraculously create a similar process on a microscopic scale, the mix of xenon isotopes in the aluminum phosphate grains we studied would have been different from what we found. For example, measured with respect to the amount of xenon 132 present, the depletion of xenon 136 (being four atomic-mass units heavier) would have been twice that of xenon 134 (two atomic mass units heavier) if physical sorting had operated. We did not see that pattern.

Our understanding of the anomalous composition of the xenon came only after we thought harder about how this gas was born. None of the xenon isotopes we measured were the direct result of uranium fission. Rather they were the products of the decay of radioactive isotopes of iodine, which in turn were formed from radioactive tellurium and so forth, according to a well-known sequence of nuclear reactions that gives rise to stable xenon.

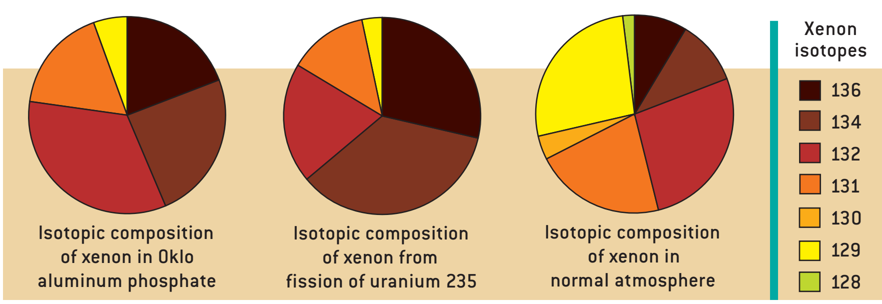
Our key insight was the realization that different xenon isotopes in our Oklo sample were created at different times—following a schedule that depended on the half-lives of their iodine parents and tellurium grandparents. The longer a particular radioactive precursor lives, the longer xenon formation from it is held off. For example, production of xenon 136 began at Oklo only about a minute after the onset of self-sustained fission. An hour later the next lighter stable isotope, xenon 134, appeared. Then, some days after the start of fission, xenon 132 and 131 came on the scene. Finally, after millions of years, and well after the nuclear chain reactions terminated, xenon 129 formed.

The Oklo reactors very likely PULSED ON AND OFF in some fashion.

Had the Oklo deposit remained a closed system, the xenon accumulated during operation of its natural reactors would have preserved the normal isotopic composition produced by fission. But scientists have no reason to think that the system was closed. Indeed, there is good cause to suspect the opposite. The evidence comes from a consideration of the simple fact that the Oklo reactors somehow regulated themselves. The most likely mechanism involves the action of groundwater, which presumably boiled away after the temperature reached some critical level. Without water present to act as a neutron moderator, nuclear chain reactions would have temporarily ceased. Only after things cooled off and sufficient groundwater once again permeated the zone of reaction could fission resume.

This picture of how the Oklo reactors probably worked highlights two important points: very likely they pulsed on and off in some fashion, and large quantities of water must have been moving through these rocks—enough to wash away some of the xenon precursors, tellurium and iodine, which are water-soluble. The presence of water also helps to explain why most of the xenon now resides in grains of aluminum phosphate rather than in the uranium-rich minerals where fission first created these radioactive precursors. The xenon did not simply migrate from one set of preexisting minerals to another—it is unlikely that aluminum phosphate minerals were present before the Oklo reactors began operating. Instead those grains of aluminum phosphate probably formed in place through the action of the nuclear-heated water, once it had cooled to about 300 degrees Celsius.

During each active period of operation of an Oklo reactor and for some time afterward, while the temperature remained high, much of the xenon gas (including xenon 136 and 134, which were generated relatively quickly) was driven off. When the reactor cooled down, the longer-lived xenon precursors (those that would later spawn xenon 132, 131 and 129, which we found in relative abundance) were preferentially



XENON GAS extracted from aluminum phosphate minerals in an Oklo sample showed a curious isotopic composition (*left*). The pattern does not match what is expected from the fission of uranium 235 (*center*), nor does it resemble the isotopic composition of atmospheric xenon (*right*). Notably, the amounts of xenon 131 and 132 are higher and the amounts of 134 and 136 are lower than would be expected from the fission of uranium 235. Although these observations were at first quite puzzling to the author, he later realized that they held the key to understanding the workings of this ancient nuclear reactor.

JEN CHRISTIANSEN

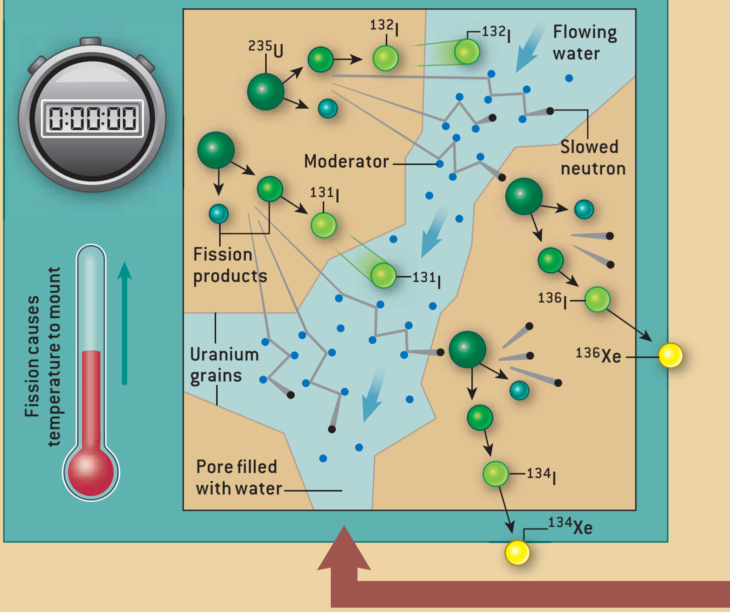
incorporated into growing grains of aluminum phosphate. Then, as more water returned to the reaction zone, neutrons became properly moderated and fission once again resumed, allowing the cycle of heating and cooling to repeat. The result was the peculiar segregation of xenon isotopes we uncovered.

It is not entirely obvious what forces kept this xenon inside the aluminum phosphate minerals for almost half the planet's lifetime. In particular, why was the xenon generated during a given operational pulse not driven off during the next one? Presumably it became imprisoned in the cagelike structure of the aluminum phosphate minerals, which were able to hold on to the xenon gas created within them, even at high temperatures. The details remain fuzzy, but whatever the final answers are, one thing is clear: the capacity of aluminum phosphate for capturing xenon is truly amazing.

XENON REVEALS CYCLIC OPERATION

Efforts to account for the isotopic composition of xenon at Oklo required a consideration of other elements as well. Iodine in particular drew attention because xenon arises from its radioactive decay. Modeling the creation of fission products and their radioactive decay revealed that the peculiar isotopic composition of xenon resulted from the cyclic operation of the reactor. That cycle is depicted in the three panels at the right.

1 Groundwater permeating the deposit acted as a moderator, allowing the fission of uranium 235 to begin. The decay of certain radioactive fission products quickly gave rise to xenon 134 and 136, but these gas atoms had a tendency to be driven off by the rising heat of the reactor. The longer-lived xenon precursors iodine 131 and 132 went into solution and were washed away about as swiftly as they were created.



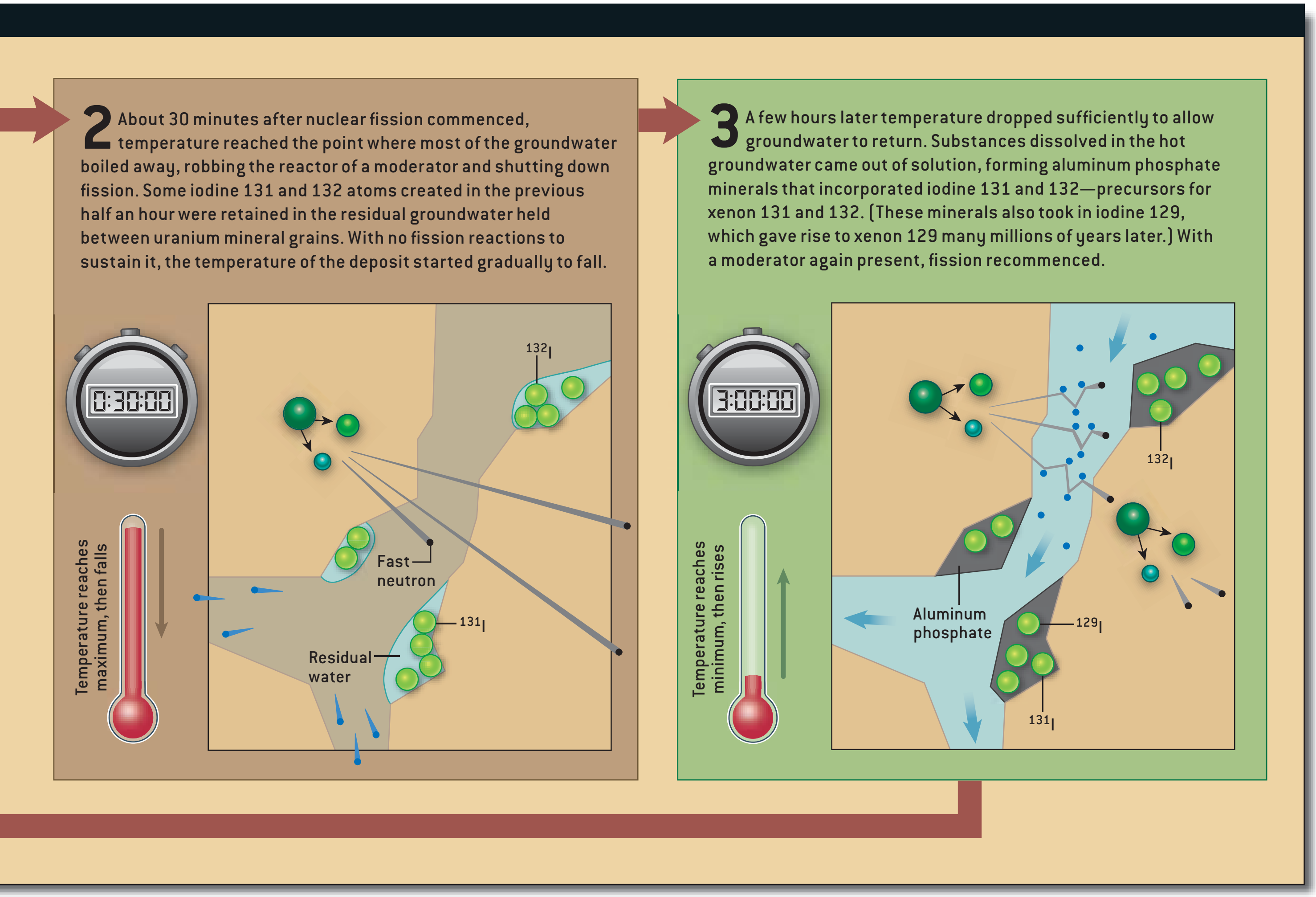
Nature's Operating Schedule

AFTER MY COLLEAGUES and I had worked out in a general way how the observed set of xenon isotopes was created inside the aluminum phosphate grains, we attempted to model the process mathematically. This exercise revealed much about the timing of reactor operation, with all xenon isotopes providing pretty much the same answer. The Oklo reactor we studied had switched "on" for 30 minutes and "off" for at least 2.5 hours. The pattern is not unlike what one sees in some geysers, which slowly heat up, boil off their supply of groundwater in a spectacular display, refill, and repeat the cycle, day in and day out, year after year. This similarity supports the notion not only that groundwater passing through the Oklo deposit was a neutron moderator but also that its boiling away at times accounted for the self-regulation that protected these natural reactors from destruction. In this regard, it was extremely effective, allowing not a single meltdown or explosion during hundreds of thousands of years.

One would imagine that engineers working in the nuclear power industry could learn a thing or two from Oklo. And they certainly can, though not necessarily about reactor design. The more important lessons may be about how to handle nuclear waste. Oklo, after all, serves as a good analogue for a long-term geologic repository, which is why scientists have examined in great detail how the various products of fission have migrated away from these natural reactors over time. They have also scrutinized a similar zone of ancient nuclear fission found in exploratory boreholes drilled at a site called Bangombe, located some 35 kilometers away. The Bangombe reactor is of special interest because it was more shallowly buried than those unearthed at the Oklo and Okelobondo mines and thus has had more water moving through it in recent times. In all, the observations boost confidence that many kinds of dangerous nuclear waste can be

The self-regulation was extremely effective, allowing NOT A SINGLE MELTDOWN or explosion during hundreds of thousands of years.

JEN CHRISTIANSEN



successfully sequestered underground.

Oklo also demonstrates a way to store some forms of nuclear waste that were once thought to be almost impossible to prevent from contaminating the environment. Since the advent of nuclear power generation, huge amounts of radioactive xenon 135, krypton 85 and other inert gases that nuclear plants generate have been released into the atmosphere. Nature's fission reactors suggest the possibility of locking those waste products away in aluminum phosphate minerals, which have a unique ability to capture and retain such gases for billions of years.

The Oklo reactors may also teach scientists about possible shifts in what was formerly thought to be a fundamental physical constant, one called α (alpha), which controls such universal quantities as the speed of light [see "Inconstant Constants," by John D. Barrow and John K. Webb; Scientific American, June]. For three decades, the two-billion-year-old Oklo phenomenon has been used to argue against α having changed. But last year Steven K. Lamoreaux and Justin R. Torgerson of Los Alamos National Laboratory drew on Oklo to posit that this "constant" has, in fact, varied significantly (and, strangely enough, in the opposite sense from what others have recently proposed). Lamoreaux and Torgerson's calculations hinge on certain details about how Oklo operated, and in that respect the work my colleagues and I have done might help elucidate this perplexing issue.

Were these ancient reactors in Gabon the only ones ever to have formed on the earth? Two billion years ago the conditions necessary for self-sustained fission must not have been too rare, so perhaps other natural reactors will one day be discovered. I expect that a few telltale wisps of xenon could aid immensely in this search. SA

MORE TO EXPLORE

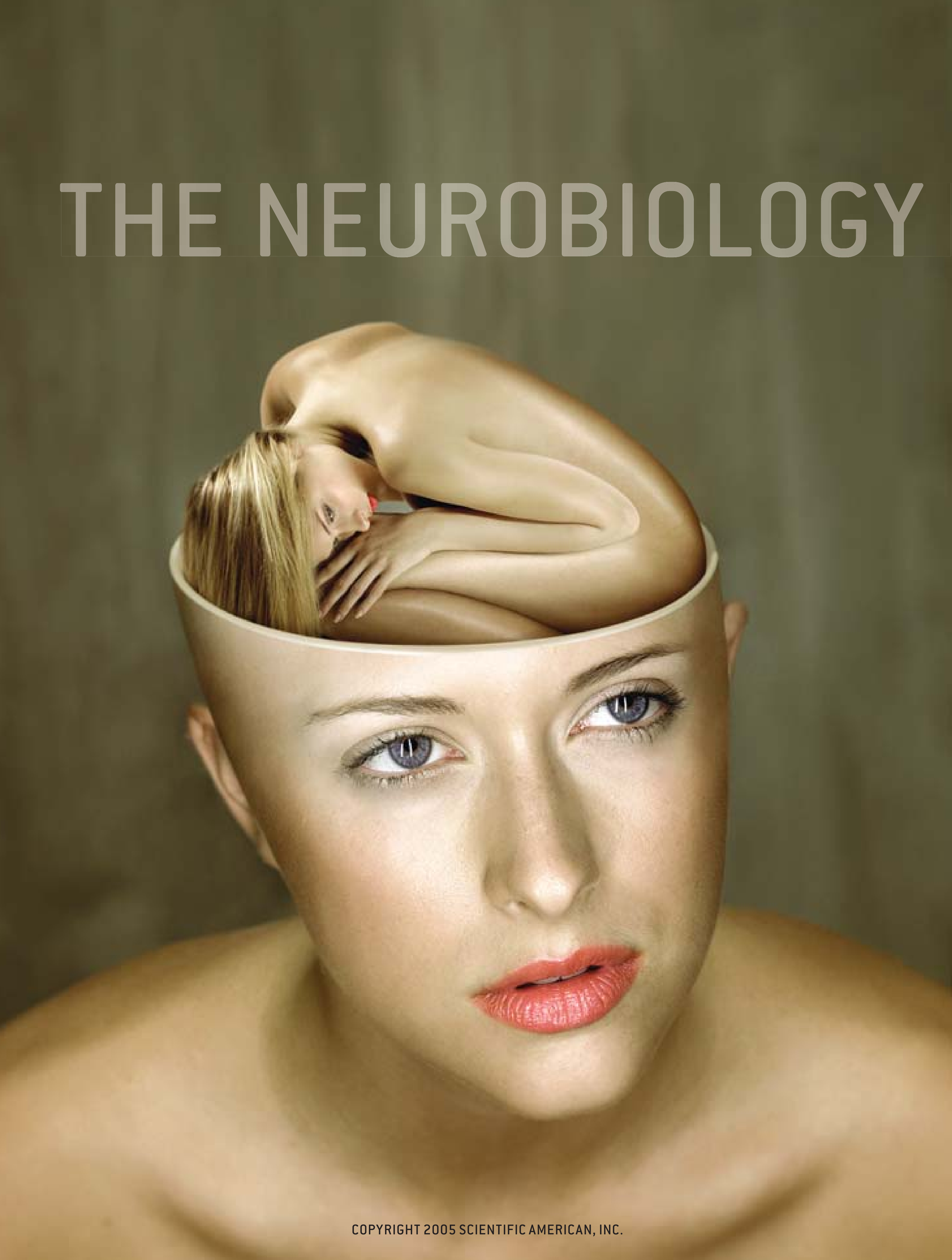
On the Nuclear Physical Stability of the Uranium Minerals. Paul Kazuo Kuroda in *Journal of Chemical Physics,* Vol. 25, No. 4, pages 781–782; 1956.

A Natural Fission Reactor. George A. Cowan in *Scientific American,* Vol. 235, No. 1, pages 36–47; July 1976.

Neutron Moderation in the Oklo Natural Reactor and the Time Variation of Alpha. S. K. Lamoreaux and J. R. Torgerson in *Physical Review D,* Vol. 69, No. 12, Paper 121701(R); June 2004. Preprint available at **arxiv.org/abs/nucl-th/0309048**

Record of Cycling Operation of the Natural Nuclear Reactor in the Oklo/Okelobondo Area in Gabon. A. P. Meshik, C. M. Hohenberg and O. V. Pravdivtseva in *Physical Review Letters,* Vol. 93, No. 18, Paper 182302; October 29, 2004.

THE NEUROBIOLOGY



OF THE *Self*

Biologists are beginning to tease out how the brain gives rise to a constant sense of being oneself

The most obvious thing about yourself is your self. "You look down at your body and know it's yours," says Todd Heatherton, a psychologist at Dartmouth University. "You know it's your hand you're controlling when you reach out. When you have memories, you know that they are yours and not someone else's. When you wake up in the morning, you don't have to interrogate yourself for a long time about who you are."

The self may be obvious, but it is also an enigma. Heatherton himself shied away from direct study of it for years, even though he had been exploring self-control, self-esteem and other related issues since graduate school. "My interests were all around the self but not around the philosophical issue of what is the self," he explains. "I avoided speculations about what it means. Or I tried to, anyway."

Things have changed. Today Heatherton, along with a growing number of other scientists, is tackling this question head-on, seeking to figure out how the self emerges from the brain. In the past few years, they have begun to identify certain brain activities that may be essential for producing different aspects of self-awareness. They are now trying to determine how these activities give rise to the unified feeling we each have of being a single entity. This research is yielding clues to how the self may have evolved in our hominid ancestors. It may even help scientists treat Alzheimer's disease and other disorders that erode the knowledge of self and, in some cases, destroy it altogether.

By Carl Zimmer

The Self Is Special

AMERICAN PSYCHOLOGIST William James launched the modern study of this area in 1890 with his landmark book, *The Principles of Psychology.* "Let us begin with the Self in its widest acceptation, and follow it up to its most delicate and subtle form," he proposed. James argued that although the self might feel like a unitary thing, it has many facets—from awareness of one's own body to memories of oneself to the sense of where one fits into society. But James confessed to being baffled as to how the brain produced these self-related thoughts and wove them into a single ego.

Since then, scientists have found some telling clues through psychological experiments. Researchers interested in memories of the self, for instance, have asked volunteers questions about themselves, as well as about other people. Later the researchers gave the volunteers a pop quiz to see how well they remembered the questions. People consistently did a better job of remembering questions about themselves than about others. "When we tag things as relevant to the self, we remember them better," Heatherton says.

Some psychologists argued that these results simply meant that we are more familiar to ourselves than other people are to us. Some concluded instead that the self is special; the brain uses a different, more efficient system to process information about it. But psychological tests could not pick a winner from these competing explanations, because in many cases the hypotheses made the same predictions about experimental outcomes.

Further clues have emerged from injuries that affect some of the brain regions involved in the self. Perhaps the most famous case was that of Phineas Gage, a 19th-century railroad construction foreman who was standing in the wrong place at the wrong time when a dynamite blast sent a tamping iron through the air. It passed right through Gage's head, and yet, astonishingly, Gage survived.

Gage's friends, though, noticed something had changed. Before the accident, he had been considered an efficient worker and a shrewd businessman. Afterward he became profane, showed little respect for others and had a hard time settling on plans for the future. His friends said he was "no longer Gage."

Cases such as Gage's showed that the self is not the same as consciousness. People can have an impaired sense of themselves without being unconscious. Brain injuries have also revealed that the self is constructed in a complicated way. In 2002, for example, Stan B. Klein of the University of California at Santa Barbara and his colleagues reported on an amnesiac known as D.B. The man was 75 years old when he suffered brain damage from a heart attack and lost the ability to recall anything he had done or experienced before it. Klein tested D.B.'s self-knowledge by giving him a list of 60 traits and asking him whether they applied to him somewhat, quite a bit, definitely, or not at all. Then Klein gave the same questionnaire to D.B.'s daughter and asked her to use it to describe her father. D.B.'s choices significantly correlated with his daughter's. Somehow D.B. had retained an awareness of himself without any access to memories of who he was.

> **The sight of someone being touched made her feel as if someone were touching her in the same place on her own body. She thought everyone had that experience.**

Clues from Healthy Brains

IN RECENT YEARS, scientists have moved beyond injured brains to healthy ones, thanks to advances in brain imaging. At University College London, researchers have been using brain scans to decipher how we become aware of our own bodies. "This is the very basic, low-level first point of the self," UCL's Sarah-Jayne Blakemore says.

When our brains issue a command to move a part of our bodies, two signals are sent. One goes to the brain regions that control the particular parts of the body that need to move, and another goes to regions that monitor the movements. "I like to think of it as a 'cc' of an e-mail," Blakemore observes. "It's all the same information sent to a different place."

Our brains then use this copy to predict what kind of sensation the action will produce. A flick of an eye will make objects appear to move across our field of vision. Speaking will make us hear our own voice. Reaching for a doorknob will make us feel the cold touch of brass. If the actual sensation we receive does not closely match our prediction, our brains become aware of the difference. The mismatch may make us pay more attention to what we are doing or prompt us to adjust our actions to get the results we want.

Overview/*My Brain and Me*

- Increasing numbers of neurobiologists are exploring how the brain manages to form and maintain a sense of self.
- Several brain regions have been found to respond differently to information relating to the self than they do to information relating to others, even to very familiar others. For instance, such regions may be more active when people think about their own attributes than when they think about the characteristics of other individuals. These regions could be part of a self-network.
- For some, the goal of this research is to better understand, and to find new therapies for, dementia.

AARON GOODMAN (*page 92*)

AARON GOODMAN

But if the sensation does not match our predictions at all, our brains interpret them as being caused by something other than ourselves. Blakemore and her colleagues documented this shift by scanning the brains of subjects they had hypnotized. When the researchers told the subjects their arms were being lifted by a rope and pulley, the subjects lifted their arms. But their brains responded as if someone else were lifting their arms, not themselves.

A similar lack of self-awareness may underlie certain symptoms of schizophrenia. Some schizophrenics become convinced that they cannot control their own bodies. "They reach over to grab a glass, and their movement is totally normal. But they say, 'It wasn't me. That machine over there controlled me and made me do it,'" Blakemore explains.

Studies on schizophrenics suggest that bad predictions of their own actions may be the source of these delusions. Because their sensations do not match their predictions, it feels as if something else is responsible. Bad predictions may also create the auditory hallucinations that some schizophrenics experience. Unable to predict the sound of their inner voice, they think it belongs to someone else.

One reason the sense of self can be so fragile may be that the human mind is continually trying to get inside the minds of other people. Scientists have discovered that so-called mirror neurons mimic the experiences of others. The sight of someone being painfully poked, for example, stimulates neurons in the pain region of our own brains. Blakemore and her colleagues have found that even seeing someone touched can activate mirror neurons.

They recently showed a group of volunteers videos of other people being touched on the left or right side of the face or neck. The videos elicited the same responses in some areas of the volunteers' brains as occurred when the volunteers were touched on the corresponding parts of their own bodies. Blakemore was inspired to carry out the study when she met a 41-year-old woman, known as C., who took this empathy to a surprising extreme. The sight of someone being touched made C. feel as if someone were touching her in the same place on her own body. "She thought everyone had that experience," Blakemore remarks.

Blakemore scanned the woman's brain and compared its responses to those of normal volunteers. C.'s touch-sensitive regions reacted more strongly to the sight of someone else being touched than those regions did in the normal subjects. In addition, a site called the anterior insula (located on the brain's surface not far from the ear) became active in C. but not in the normal volunteers. Blakemore finds it telling that the anterior insula has also displayed activity in brain scans of people who are shown pictures of their own faces or who are identifying their own memories. It is possible that the anterior insula helps to designate some information as relating to ourselves instead of to other people. In the case of C., it simply assigns information incorrectly.

Brain scans have also shed light on other aspects of the self. Heatherton and his colleagues at Dartmouth have been using the technology to probe the mystery of why people remember information about themselves better than details about other people. They imaged the brains of volunteers who viewed a series of adjectives. In some cases, the researchers asked the subjects whether a word applied to the subjects themselves.

THE AUTHOR

CARL ZIMMER is a journalist based in Connecticut. His latest book, *Soul Made Flesh: The Discovery of the Brain—and How It Changed the World,* was recently published in paperback. He also writes The Loom, a blog about biology (www.corante.com/loom/).

In others, they asked if a word applied to George W. Bush. In still other cases, they asked simply whether the word was shown in uppercase letters.

The researchers then compared the patterns of brain activity triggered by each kind of question. They found that questions about the self activated some regions of the brain that questions about someone else did not. Their results bolstered the "self is special" hypothesis over the "self is familiar" view.

A Common Denominator

ONE REGION that Heatherton's team found to be important to thinking about oneself was the medial prefrontal cortex, a patch of neurons located in the cleft between the hemispheres of the brain, directly behind the eyes. The same region has also drawn attention in studies on the self carried out by other laboratories. Heatherton is now trying to figure out what role it serves.

"It's ludicrous to think that there's any spot in the brain that's 'the self,' " he says. Instead he suspects that the area may bind together all the perceptions and memories that help to produce a sense of self, creating a unitary feeling of who we are. "Maybe it's something that brings information together in a meaningful way," Heatherton notes.

If he is right, the medial prefrontal cortex may play the same role for the self as the hippocampus plays in memory. The hippocampus is essential for forming new memories, but people can still retain old memories even after it is injured. Rather than storing information on its own, the hippocampus is believed to create memories by linking together far-flung parts of the brain.

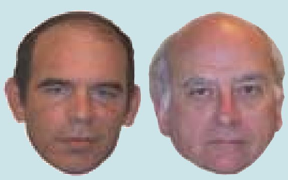
The medial prefrontal cortex could be continuously stitching together a sense of who we are. Debra A. Gusnard of Washington University and her coworkers have investigated what occurs in the brain when it is at rest—that is, not engaged in any particular task. It turns out that the medial prefrontal cortex becomes more active at rest than during many kinds of thinking.

"Most of the time we daydream—we think about something that happened to us or what we think about other people. All this involves self-reflection," Heatherton says.

Other scientists are investigating the brain networks that may be organized by the medial prefrontal cortex. Matthew Lieberman of the University of California at Los Angeles has been using brain scans to solve the mystery of D.B., the man who knew himself even though he had amnesia. Lieberman and his colleagues scanned the brains of two sets of volunteers: soccer players and improvisational actors. The researchers then wrote up two lists of words, each of which was relevant to one of the groups. (Soccer players: athletic, strong, swift; actors: performer, dramatic, and so on.) They also composed a third list of words that did not apply specifically to either (messy and reliable, for example). Then they showed their subjects the words and asked them to decide whether

Just Another Pretty Face?

As Carl Zimmer notes in the accompanying article, investigators disagree over whether the brain treats the self as special—processing information about the self differently from information about other aspects of life. Some argue that parts of our brain that change their activity when we think about ourselves do so simply because we are familiar with ourselves, not specifically because the self is involved; anything else that was familiar would evoke the same response.



J.W. Gazzaniga

In one study addressing this question, researchers photographed a man referred to as J.W., whose right and left cerebral hemispheres operated independently as a result of surgery that had severed the connections (to treat intractable epilepsy). They also photographed someone very familiar to the man—Michael Gazzaniga, a well-known brain researcher who had spent a lot of time with J.W. Next they created a series of images in which J.W.'s face morphed into Gazzaniga's (*below*) and displayed them in random order. For each image, they had J.W. answer the question "Is it me?" Then they repeated the process, having him answer, "Is it Mike?" They also performed the test with the faces of others well known to J.W.

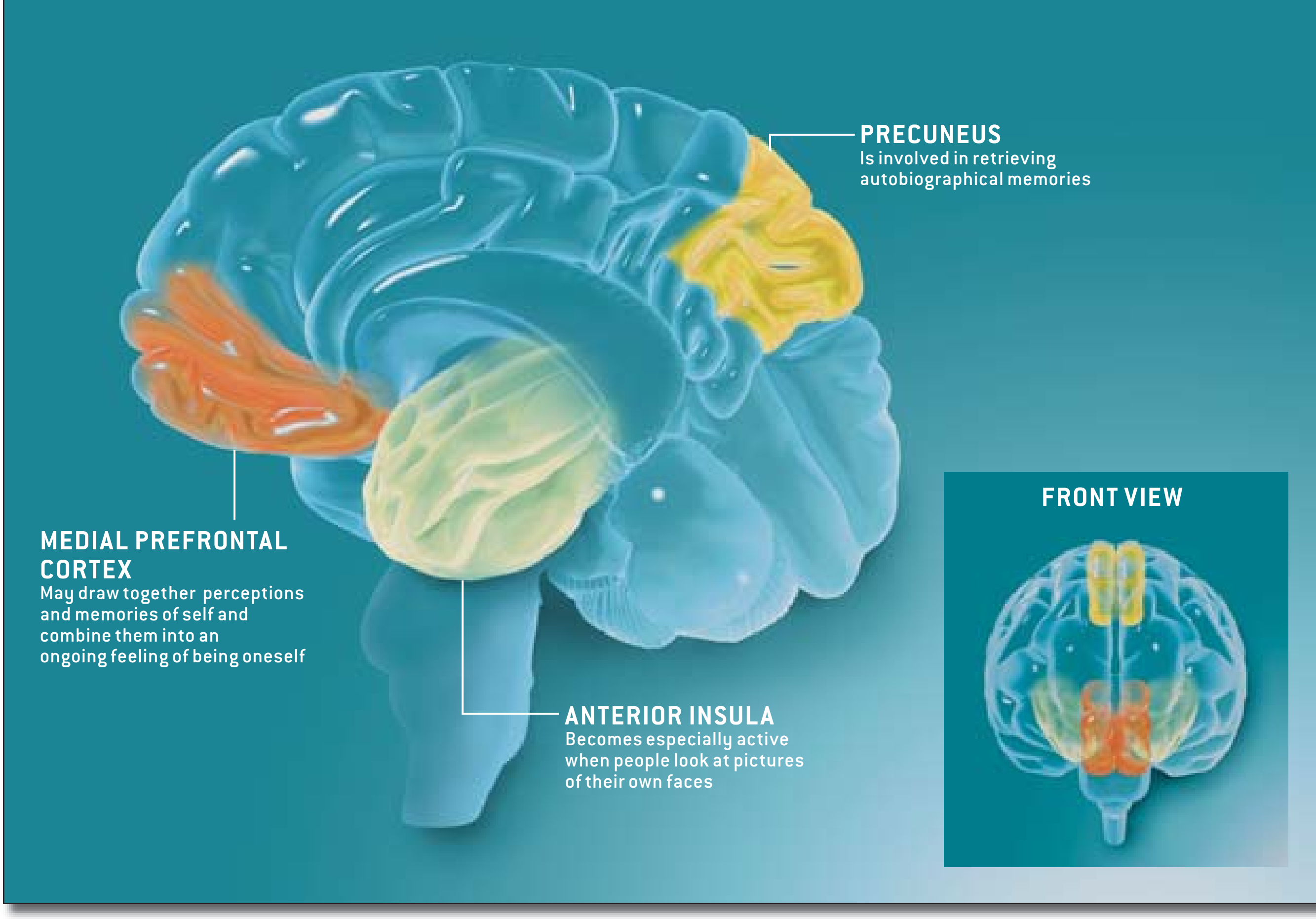
They found that J.W.'s right hemisphere was more active when he recognized familiar others, but his left hemisphere was most active when he saw himself in the photographs. These findings lend support to the self-is-special hypothesis. The issue, though, is far from solved: both camps have evidence in their favor. *—Ricki Rusting, managing editor*

90 percent J.W. —— Morphing ——→ 10 percent J.W.

"MIKE OR ME? SELF-RECOGNITION IN A SPLIT-BRAIN PATIENT," BY DAVID J. TURK ET AL., IN *NATURE NEUROSCIENCE*, VOL. 5, NO. 9; SEPTEMBER 2002

COMPONENTS OF A SELF-NETWORK

The brain regions highlighted below are among those that have been implicated, at least by some studies, as participating in processing or retrieving information specifically related to the self or in maintaining a cohesive sense of self across all situations. For clarity, the view below omits the left hemisphere, except for its anterior insula region.



each one applied to themselves or not.

The volunteers' brains varied in their responses to the different words. Soccer-related words tended to increase activity in a distinctive network in the brains of soccer players, the same one that became more active in response to actor-related words in actors. When they were shown words related to the other group, a different network became more active. Lieberman refers to these two networks as the reflective system (or C system) and the reflexive system (or X system).

The C system taps into the hippocampus and other parts of the brain already known to retrieve memories. It also includes regions that can consciously hold pieces of information in mind. When we are in new circumstances, our sense of our self depends on thinking explicitly about our experiences.

But Lieberman argues that over time, the X system takes over. Instead of memories, the X system encodes intuitions, tapping into regions that produce quick emotional responses based not on explicit reasoning but on statistical associations. The X system is slow to form its self-knowledge, because it needs many experiences to form these associations. But once it takes shape, it becomes very powerful. Soccer players know whether they are athletic, strong or swift without having to consult their memories. Those qualities are intimately wrapped up with who they are. On the other hand, they do not have the same gut instinct about whether they are dramatic, and in these cases they must think explicitly about their experiences. Lieberman's results may solve the mystery of D.B.'s paradoxical self-knowledge. It is conceivable that his brain damage wiped out his reflective system but not his reflexive system.

Although the neuroscience of the self is now something of a boom industry, it has its critics. "A lot of these studies aren't constrained, so they don't say anything," says Martha Farah, a cognitive neuroscientist at the University of Pennsylvania. The experiments, she argues, have not been designed carefully enough to eliminate other explanations—for example, that we use certain brain regions to think about any person, including ourselves. "I don't think there's any 'there' there," she says.

Heatherton and other scientists involved in this research think that Farah

3FX, INC.

is being too tough on a young field. Still, they agree that they have yet to figure out much about the self-network and how it functions.

The Evolving Self

UNCOVERING THIS NETWORK may allow scientists to understand how our sense of self evolved. The primate ancestors of humans probably had the basic bodily self-awareness that is studied by Blakemore and her associates. (Studies on monkeys suggest that they make predictions about their own actions.) But humans have evolved a sense of self that is unparalleled in its complexity. It may be significant that the medial prefrontal cortex is "one of the most distinctly human brain regions," according to Lieberman. Not only is it larger in humans than in nonhuman primates, but it also has a greater concentration of uniquely shaped neurons called spindle cells. Scientists do not yet know what these neurons do but suspect that they play an important role in processing information. "It does seem like there's something special there," he comments.

Heatherton thinks that the human self-network may have evolved in response to the complex social life of our ancestors. For millions of years hominids lived in small bands, cooperating to find food and sharing what they found. "The only way that works is through self-control," he says. "You have to have cooperation, and you have to have trust." And these kinds of behaviors, he argues, require a sophisticated awareness of oneself.

If the full-fledged human self were a product of hominid society, that link would explain why there are so many tantalizing overlaps between how we think about ourselves and how we think about others. This overlap is not limited to the physical empathy that Blakemore studies. Humans are also uniquely skilled at inferring the intentions and thoughts of other members of their species. Scientists have scanned people engaged in using this so-called theory of mind, and some of the regions of the brain that become active are part of the network used in thinking about oneself (including the medial prefrontal cortex). "Understanding ourselves and having a theory of mind are closely related," Heatherton says. "You need both to be a functioning human being."

AARON GOODMAN

The self requires time to develop fully. Psychologists have long recognized that it takes a while for children to acquire a stable sense of who they are. "They have conflicts in their self-concepts that don't bother them at all," Lieberman comments. "Little kids don't try to tell themselves, 'I'm still the same person.' They just don't seem to connect up the little pieces of the self-concept."

Lieberman and his colleagues wondered if they could track children's changing self-concept with brain imaging. They have begun studying a group of children and plan to scan them every 18 months from ages nine to 15. "We asked kids to think about themselves and to think about Harry Potter," he says. He and his team have compared the brain activity in each task and compared the results with those in adults.

"When you look at 10-year-olds, they show this same medial prefrontal cortex activation as adults do," Lieberman notes. But another region that becomes active in adults, known as the precuneus, is a different story. "When they think about themselves, they activate this region *less* than they do when they think about Harry Potter."

Lieberman suspects that in children, the self-network is still coming online. "They've got the stuff, but they're not applying it like adults do."

Insights into Alzheimer's

ONCE THE SELF-NETWORK does come online, however, it works very hard. "Even with the visual system, I can close my eyes and give it something of a rest," comments William Seeley, a neurologist at the University of California, San Francisco. "But I can never get away from living in my body or representing the fact that I'm the same person I was 10 seconds or 10 years ago. I can never escape that, so that network must be busy."

One patient, described by Seeley and others in the journal *Neurology* in 2001, had collected jewelry and fine crystal for much of her life before abruptly starting to gather stuffed animals at age 62. A lifelong conservative, she began to berate people in stores who were buying conservative books and declared that "Republicans should be taken off the earth." Other patients have suddenly converted to new religions or become obsessed with painting or photography. Yet they have little insight into why they are no longer their old selves. "They say pretty shallow things, like 'This is just the way I am now,'" Seeley says. Within a few years, frontotemporal dementia can lead to death.

Michael Gazzaniga, director of Dartmouth's Center for Cognitive Neuroscience and a member of the President's Council on Bioethics, believes that deciphering the self may pose a new kind of ethical challenge. "I think there's going to be the working out of the circuits of self—self-referential memory, self-description, personality, self-awareness," Gazzaniga says. "There's going to be a sense of what has to be in place for the self to be active."

It is even possible, Gazzaniga suggests, that someday a brain scan might determine whether Alzheimer's or some other dementia has destroyed a person's self. "Someone's going to say, 'Where's Gramps?'" he predicts. "And they're going to be able to take a picture of Gramps under certain conditions and say, 'Those circuits are not working.'"

Gazzaniga wonders whether people will begin to consider the loss of the self when they write out their living wills. "Advanced directives will come into play," he predicts. "The issue will be whether you deliver health care. If people catch pneumonia, do you give them antibiotics or let them go?"

Seeley offers a more conservative forecast, arguing that a brain scan on its own probably will not change people's minds about life-and-death decisions. Seeley thinks the real value of the science of the self will come in treatments for Alzheimer's and other dementias. "Once we know which brain regions are involved in self-representation, I think we can take an even closer look at which cells in that brain region are important and then look deeper and say which molecules within cells and which genes that govern them lead to this vulnerability," he says. "And if we've done that, we've gotten closer to disease mechanisms and cures. That's the best reason to study all this. It's not just to inform philosophers." SA

Someday a brain scan might determine whether dementia has destroyed a person's self. "Someone's going to say, 'Where's Gramps?' and they're going to be able to ... say, 'Those circuits are not working.'"

The more energy that a cell consumes, the greater its risk of damaging itself with toxic by-products. Seeley suspects that the hardworking neurons in the self-network are particularly vulnerable to this damage over the life span. Their vulnerability, he argues, may help neurologists make sense of some brain disorders that erode the self. "It is curious that we can't find certain pathological changes of Alzheimer's or other dementias in nonhuman species," Seeley says.

According to Seeley, the results of recent brain-imaging studies of the self agree with findings by him and others on people with Alzheimer's and other dementias. People with Alzheimer's develop tangled proteins in their neurons. Some of the first regions to be damaged are the hippocampus and precuneus, which are among the areas involved in autobiographical memories. "They help you bring images of your past and future into mind and play with them," Seeley notes. "People with Alzheimer's are just less able to move smoothly back and forth through time."

As agonizing as it may be for family members to watch a loved one succumb to Alzheimer's, other kinds of dementia can have even more drastic effects on the self. In a condition known as frontotemporal dementia, swaths of the frontal and temporal lobes degenerate. In many cases, the medial prefrontal cortex is damaged. As the disease begins to ravage the self-network, people may undergo strange changes in personality.

MORE TO EXPLORE

A Self Less Ordinary: The Medial Prefrontal Cortex and You. C. Neil Macrae, Todd F. Heatherton and William M. Kelley in *Cognitive Neurosciences III.* Edited by Michael S. Gazzaniga. MIT Press, 2004.

Is Self Special? A Critical Review of Evidence from Experimental Psychology and Cognitive Neuroscience. Seth J. Gillihan and Martha J. Farah in *Psychological Bulletin,* Vol. 131, No. 1, pages 76–97; January 2005.

The Lost Self: Pathologies of the Brain and Identity. Edited by Todd E. Feinberg and Julian Paul Keenan. Oxford University Press, 2005.

Conflict and Habit: A Social Cognitive Neuroscience Approach to the Self. Matthew D. Lieberman and Naomi I. Eisenberger in *Psychological Perspectives on Self and Identity,* Vol. 4. Edited by A. Tesser, J. V. Wood and D. A. Stapel. American Psychological Association (in press). Available online at **www.scn.ucla.edu/pdf/rt4053_c004Lieberman.pdf**

The Land of MILK & MONEY

By Gary Stix

The first drug from a transgenic animal may be nearing approval

Proteins are biotechnology's raw crude. For much of its 30-year history, the industry has struggled to come up with a steady source of supply, squeezing the maximum out of these large-molecule commodities from cell lines isolated from hamster ovaries and the like. In the late 1990s—with the advent of a new class of protein-based drugs, monoclonal antibodies—demand sometimes outstripped supply. For decades, the scientists who created recombinant erythropoietin to rejuvenate red blood cells and monoclonal antibodies to combat cancer have sought out alternative forms of manufacture.

A new bioreactor—an animal genetically engineered to produce a therapeutic protein in its milk—may finally be ready to fulfill its long-awaited promise. The European Medicines Evaluation Agency (EMEA) may decide early next year on approval of an anticoagulant protein, human antithrombin, that is produced in goat's milk to treat a hereditary disorder. If the drug, ATryn, finally gets a nod from regulators, its approval will mark the culmination of a meandering 15-year journey for GTC Biotherapeutics, a Framingham, Mass., spin-off of the biotech giant Genzyme.

The idea of making transgenic drugs occurred to a number of scientists during the mid-1980s, when the new industry began to wrestle with the challenge of making complex proteins: ensuring that these big molecules were folded into the proper shape and that they had all their sugars in the right places on the surface of the proteins' amino acids. Chinese hamster ovary cells do the job, but getting enough product has been a constant frustration and one reason why biotech drugs today cost so much. In addition, mammalian cell cultures are not always an ideal medium: at times, it is simply too hard to produce proteins in this manner.

In their quest for greater efficiencies, researchers noticed that the mammary glands of cows, rabbits and goats, among others, are capable of becoming ideal protein manufacturing plants because of their ability to make high volumes of complex proteins. Milk glands, moreover, do not need the constant coddling required for cell cultures.

Genzyme got involved after its purchase in 1989 of Integrated Genetics, which had a portfolio of drugs and diagnostics products. To head up its program, Genzyme recruited one of the pioneers in this technology from another company, Biogen. Harry Meade, along with Nils Lonberg, had patented a method of extracting therapeutic proteins from mice.

In the early 1990s Genzyme's program was targeted at producing drugs in goat's milk. Genzyme, though, was not focusing on transgenics and decided to spin off its operation into a separate entity, Genzyme Transgenics (later re-

BRAD DeCECCO

TRANSGENIC GOAT gets milked at a farm owned by GTC Biotherapeutics, headquartered in Framingham, Mass. The animal secretes a valuable pharmaceutical protein in its milk.

named GTC Biotherapeutics), in which the parent still holds an equity interest. The new company could thus produce its drugs for other firms without the inevitable conflicts of interest that would have arisen had it remained within the bosom of a large drugmaker.

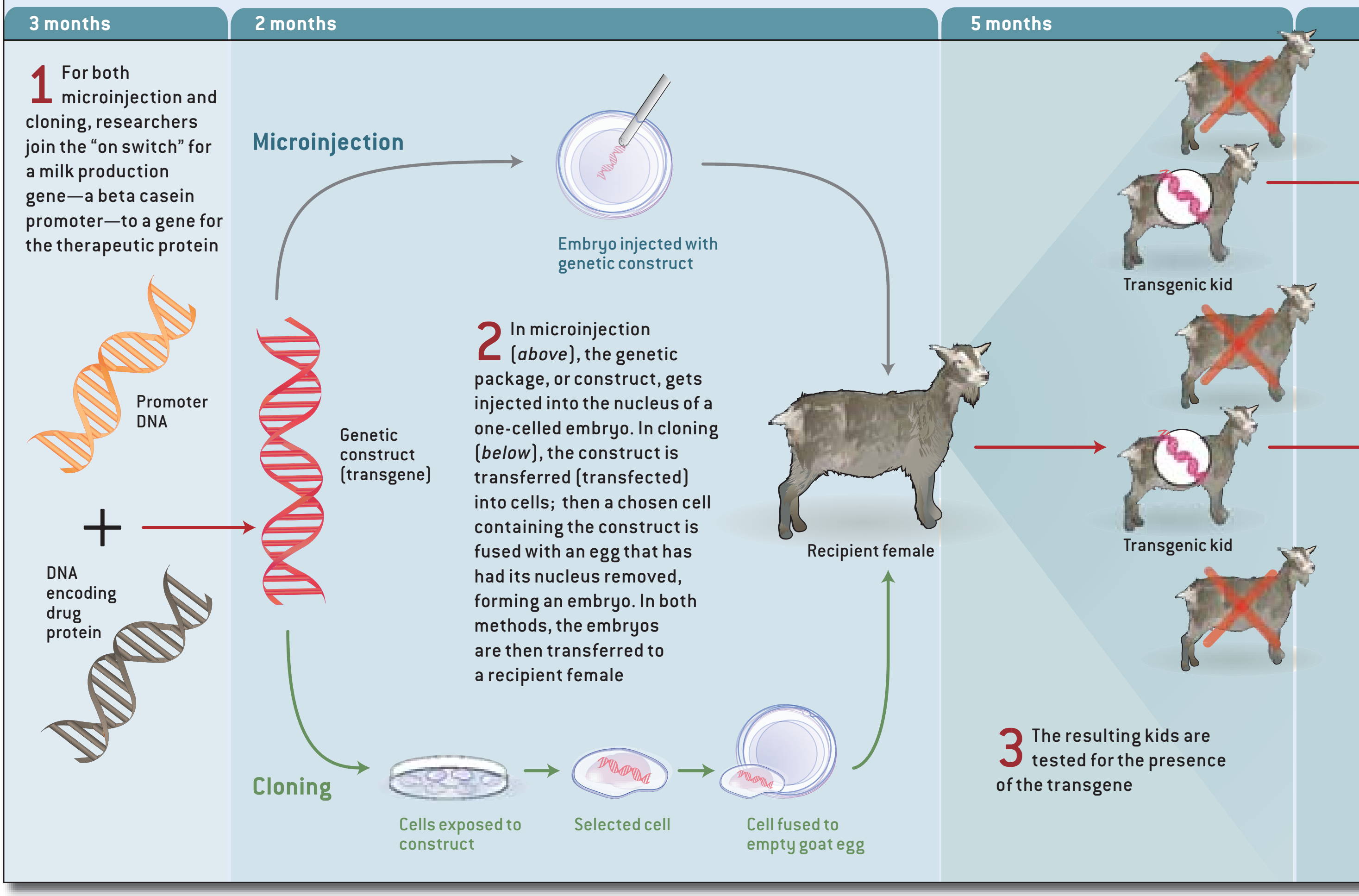
Goats as Drug Factories

INITIALLY, GTC generated transgenic goats by microinjecting into the developing nucleus of a one-cell embryo a gene encoding the desired human protein (along with DNA that promotes activation of that gene in milk). Such embryos were transferred into female goats, which produced offspring that were then tested for the presence of the newly integrated gene. The milk of these "founder" animals contains the therapeutic protein, which must then undergo a purification process. The mature transgenic animals were bred usually with nontransgenic goats as a first step toward producing a herd [*see box on next two pages*]. Microinjection, however, is an inefficient process. Only 1 to 5 percent of the embryos result in transgenic animals. For newer drugs in its portfolio, GTC has adopted somatic cell nuclear transfer, a.k.a. cloning, which ensures that an animal will carry the desired transgene. Dolly the sheep was cloned, in fact, with the intention of eventually using this procedure to create transgenic animals having useful properties, not as a means to make carbon copies of baseball legend Ted Williams or a favorite dead pet.

GTC stuck with goats because they reproduce more rapidly than cows and can yield more protein than mice or rabbits. Other efforts, including a more nascent GTC endeavor, have opted for cows. Pharming, a Netherlands-based company, aims to milk both cows and rabbits for drugs. Yet others have pursued distinctive forms of bioreactors: making drugs in chicken eggs, for instance. After undertaking basic development of the technology during the 1990s,

MILKING GOATS FOR DRUGS

GTC Biotherapeutics, which is counting on European approval of an anticlotting drug produced in goats, has used two major approaches to create a transgenic animal. The older technique, microinjection, employed for the drug undergoing regulatory review called ATryn,



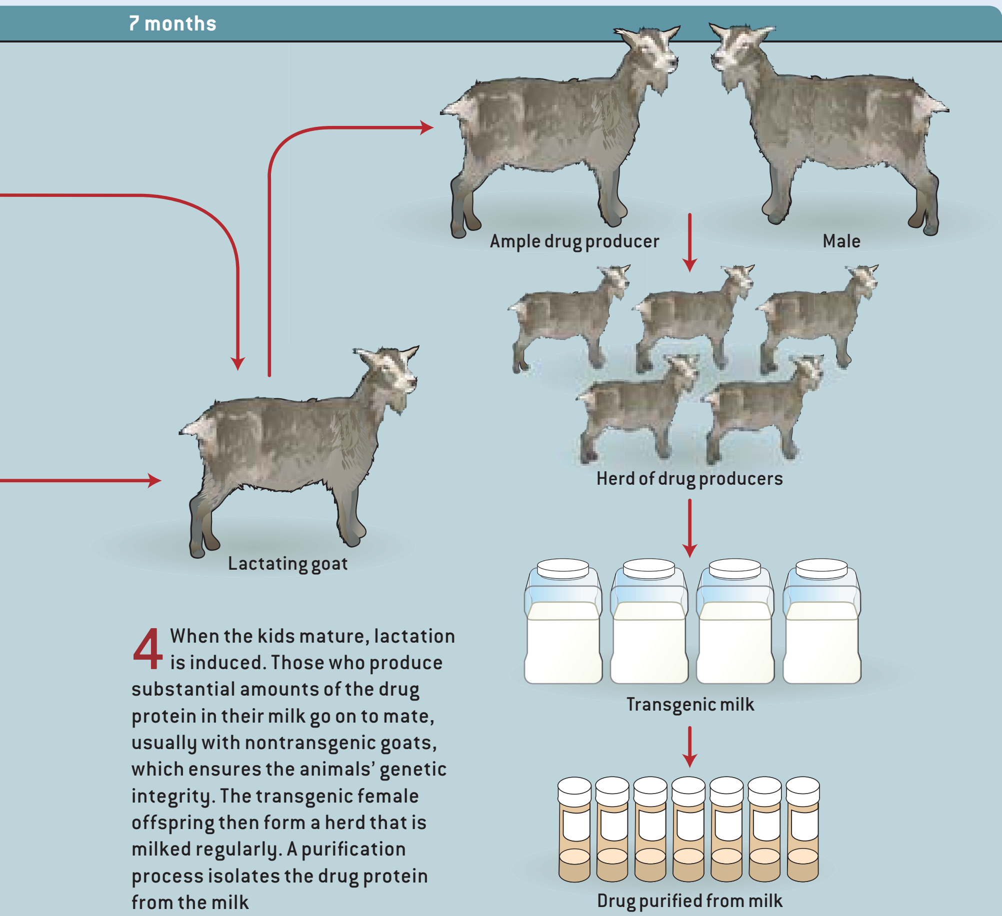
GTC hung out a shingle, marketing itself as a technology platform for companies that either wanted to produce difficult-to-make pharmaceutical proteins or needed large quantities at low cost. The one catch was that regulators had never approved a transgenically produced drug, and the more than a dozen partners that GTC took on tended to view the technology as a backup in case other protein-drug development strategies did not work out. They were unwilling to accept the expense and risk of an arduous regulatory process for a pioneering form of drug manufacture.

GTC recognized the need to demonstrate on its own the potential for the technology and, in the late 1990s, began a clinical trial of human antithrombin for patients undergoing bypass surgery who develop resistance to the anticoagulant drug heparin. Transgenic antithrombin was intended to improve supply and address concerns about pathogens in the form of the drug isolated from human blood. The company completed the required clinical trials. But when the Food and Drug Administration asked for more data in late 2000—which would have necessitated additional testing—then chief executive Sandra Nusinoff Lehrman scrapped the effort. In mid-2001 Nusinoff Lehrman left, and her replacement, Geoffrey Cox, decided to proceed with development of transgenic antithrombin—this time in European clinical trials for patients with inherited antithrombin deficiency. Regulators there had recently issued guidelines that set out the requirements for getting approval for antithrombin.

The company still has a few partnerships. It also has a preliminary program to make other blood proteins, such as alpha-1 antitrypsin, and a clinical trial in the U.S. for ATryn. But its future hinges on the European approval. The company, which went public in 1993, has flirted with penny-stock status (less than $1 a share), and its cash levels are much depleted from what they were at the start of the decade. It has also experienced "restructurings," layoffs that occurred in 2003 and 2004. "This is an important moment," says Cox of the upcoming EMEA decision. "This isn't a business for the faint of heart."

LUCY READING-IKKANDA

involves introducing a gene directly into an embryo. The company has also been a pioneer in producing transgenic drugs through cloning.



4 When the kids mature, lactation is induced. Those who produce substantial amounts of the drug protein in their milk go on to mate, usually with nontransgenic goats, which ensures the animals' genetic integrity. The transgenic female offspring then form a herd that is milked regularly. A purification process isolates the drug protein from the milk

Bioreactor Blues

OTHER TRANSGENIC companies have also had a rough haul. The Scottish company PPL Therapeutics, which helped to clone Dolly, encountered difficulties and sold its remaining intellectual property to Pharming in 2004. The latter has staged a comeback since filing for protection from creditors in 2001. It hopes to get approval soon for a treatment for hereditary angioedema, a genetic disease that causes swelling from the absence of the C1 inhibitor protein.

If GTC survives, it could become the leader in transgenics. The impetus for starting the company still appears justified. The capital costs for a drug production facility using hamster cells can amount to $400 million to $500 million, Cox says, whereas a herd of goats can produce the comparable amount of drug for $50 million. "There's still a need for alternative production methodologies," says Philip Nadeau, who tracks GTC as an analyst with S. G. Cowen. "There are still proteins that are difficult to produce using traditional methods, and therefore a company like GTC should certainly have a niche." ATryn's uses could be broadened to encompass an array of treatments—for coronary bypass, burn or sepsis patients—that might, in total, bring in as much as $700 million annually, Cox estimates.

The drug appears to have surmounted an important technical hurdle: so far it has not created any adverse immune response in patients. But such events will always remain a worry. Researchers administering inhaled transgenic alpha-1 antitrypsin from sheep bred by PPL discovered that some patients suffered pulmonary symptoms that caused them to leave the trial—a possible immune reaction to residual proteins from the animal that remained after purification of the drug. The PPL drug, given on a longer-term basis than ATryn is, needed to be better purified, notes Meade, GTC's chief scientific officer.

Producing drugs in goats has so far elicited less criticism than the debate over genetically modified plants. Goats cannot drift with the wind like corn pollen, spreading their transgenes to unexpected places. "If it's able to make drugs available that are not otherwise available by other methods and if it would make drugs cheaper, it would be certainly advantageous to consumers," notes Jane Rissler of the Union of Concerned Scientists. "Frankly, consumers have not benefited very much [so far] from biotechnology in the agricultural sector."

At GTC, the scrapie-free goats brought in from New Zealand are penned within a 190-acre enclosure on a 300-acre plot in Charlton, Mass. The animals are fed—and not permitted to graze—to diminish the possibility of contracting disease from contaminants in other animals. Thirty goats are devoted to making ATryn among a transgenic herd of more than 300, and an additional 1,200 nontransgenic animals are kept for breeding. "We have more veterinarians than M.D.s," Cox says. If ATryn finally receives approval, traditional dairy farmers flirting with insolvency may gaze in astonishment at a product made in milk that commands thousands of dollars per gallon. SA

MORE TO EXPLORE

Transgenic Animals: Generation and Use. Edited by Louis Marie Houdebine. CRC Press, 1997.

Production of Goats by Somatic Cell Nuclear Transfer. Alexander Baguisi et al. in *Nature Biotechnology*, Vol. 17, No. 5, pages 456–461; May 1999.

GTC Biotherapeutics: **www.transgenics.com**

NUTS

Case Cracked

"How are nuts shelled commercially?" asks *Scientific American* reader Bill Lush. "Every time I struggle with a Brazil nut or a pecan I wonder. It must be tough, since we're dealing with a natural product that varies in size and shape. And don't tell me it's done by hand—I would be soooo disappointed."

Lush's note expresses a frustration many people share during Thanksgiving and the December holidays—the stressful attempt to crack a nut just enough to open it without crushing the prize inside. Production plants can process 30 to 60 tons of nuts a day, smashing fewer than one tenth of 1 percent.

Most nuts are harvested by shaking their trees with a bulldozerlike contraption. The jewels are scooped up from the ground along with dirt, grass, leaves, sticks and stones. Screens and shakers sieve this mess to leave a reasonably junk-free bin of nuts, which are then mechanically sorted by size. To handle the size variations, a processing plant operates a number of cracking machines in parallel, each one accepting nuts of a specified size—for almonds, say, 8/16 inch wide, 9/16, 10/16 and so on. Pecans are typically sifted into five size ranges, peanuts into six. "The more precise the match, the less damage to the kernel," says Bill Hoskins, director of quality assurance at Blue Diamond Growers in Sacramento, Calif.

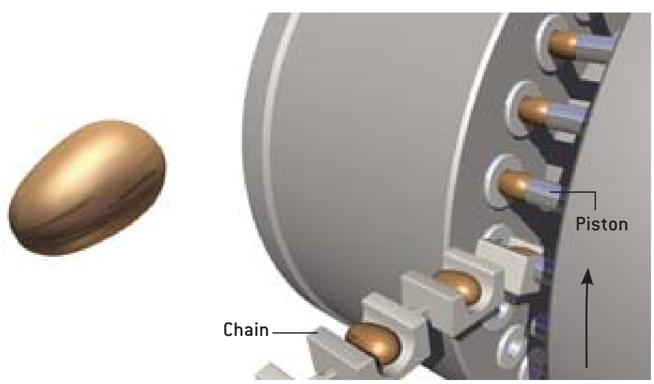
Although machines are tailored to each nut type, a few basic techniques—screen impurities, tear off shells and aspirate both away—underlie the processes [*see illustrations*]. "The most important objective is to keep cleaning the product flow," says Lewis Carter, Jr., chairman of Lewis M. Carter Manufacturing in Donalsonville, Ga., which makes a large share of American machines.

The technology has evolved slowly for decades. For example, "the almond industry has adapted much of its machinery from the peanut industry," Hoskins says. Similar equipment and procedures are used for grains and beans, too.

A few nuts pose unique challenges. Black walnut shells are so hard they require special crackers. Brazil nuts are actually seeds that grow in groups of eight to 24 inside a small coconutlike pod that must itself be cracked. —*Mark Fischetti*

SOFT-SHELL NUTS, such as almonds (*shown*), are separated from harvest debris and sorted by size through a series of perforated pans. Nuts of similar diameter fall between a pair of rollers spaced slightly closer than that diameter. One 10-inch-diameter roller rotates faster than the other, creating a shearing action that tears the shell away from the kernel. In contrast, peanuts are pushed through sharp gates that slice off the shell.

HARD-SHELL NUTS, such as pecans (*shown*) and hazelnuts, are submerged in water at 190 degrees Fahrenheit for three to 12 minutes, which softens (and pasteurizes) them. Once they are extracted, the shell hardens within minutes, but the meat remains pliable. The nuts settle into a chain that pulls each one past a piston, which strikes the shell at 35 to 45 pounds per square inch, cracking it; the soft core remains intact.



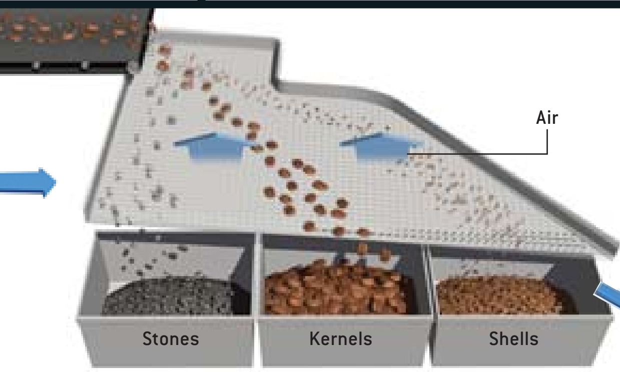
GEORGE RETSECK

DID YOU KNOW. . .

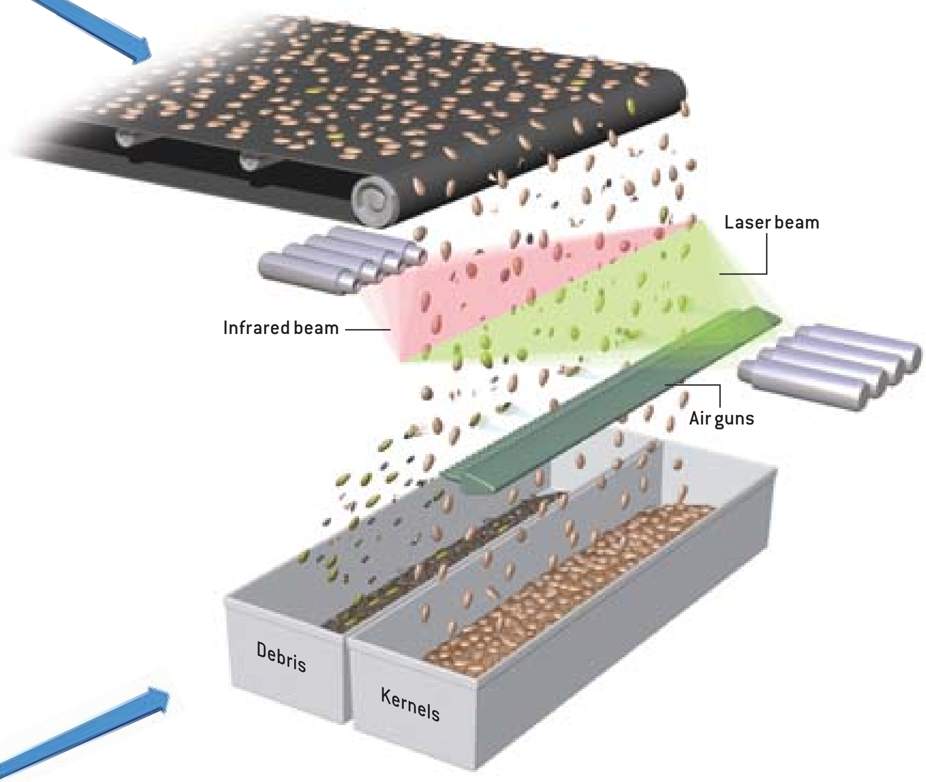
FRESH COLD ONE: Peak U.S. harvests are in the autumn, but processing continues year-round. Almonds can be stored for six to eight months at 32 to 60 degrees Fahrenheit before shelling. Pecans can keep in freezers for two to three years; they are 80 percent oil, which remains very stable when frozen, says Brenda Lara, plant manager at Green Valley Pecan Company in Sahuarita, Ariz.

SQUIRREL MANIA: Ninety-nine percent of U.S. hazelnuts are grown in the Willamette Valley of Oregon. In the past, the shells were burned as fuel; the hot market today is for mulch. Because the shells absorb very little moisture, they drain quickly yet block sun from drying the soil underneath, and they decompose slowly. They are also dense, preventing weed seeds from germinating. But the scent drives squirrels nuts—no kernels to be found.

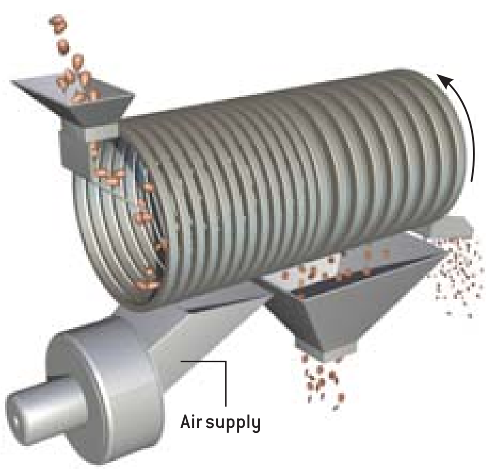
TOUGH NUT TO CRACK: Many of the world's cashews are cultivated in India, Malaysia, Indonesia and eastern Africa. Each nut grows inside a small, pear-shaped peduncle—a false fruit. In developing countries workers often tear the fruit away by hand. Oil on the nutshell surface is toxic and corrosive to human skin, so the nut is dried and roasted to extract the oil and to make the shell brittle. Most shells are cracked by hand, too; workers dust their fingers in wood ash or linseed or castor oil, for protection. Gradually, machinery is being deployed.



INCLINED GRAVITY TABLE oscillates and vibrates as air is blown up through the screen tray. The material stratifies within a fluidized bed: shell fragments (low specific gravity) float down to one end and are drawn off. Heavier, whole kernels migrate to the central part of the screen. Dense objects (stones) climb to the high end and are combed away.



SHELLER DRUM rotates as cracked nuts are conveyed inside, drawn through the cylinder by negative air pressure. Rings strike the nuts, knocking the cracked shells free.



FOR MOST NUTS, kernels and remaining impurities fall through a scanner. Infrared or laser beams reflect off each object to check for poor shape (partial kernel) or bad color (rotten kernel). An air gun blows a suspect nugget out of the four-foot-wide stream. Production workers spot-check whole kernels before packaging.

Have a topic idea? Send it to workingknowledge@sciam.com

Making Tracks on Mars

AT LAST, SOME FORM OF INTELLIGENCE IS ALTERING THE DUSTY FACE OF OUR NEIGHBORING WORLD

BY DAVID GRINSPOON

ROVING MARS: SPIRIT, OPPORTUNITY, AND THE EXPLORATION OF THE RED PLANET
by Steve Squyres. Hyperion, 2005 ($25.95)

DYING PLANET: MARS IN SCIENCE AND THE IMAGINATION
by Robert Markley. Duke University Press, 2005, paperbound ($24.95)

The planet Mars—crimson and bright, filling our telescopes with vague intimations of almost-familiar landforms—has long formed a celestial tabula rasa on which we have inscribed our planetological theories, utopian fantasies, and fears of alien invasion or ecological ruin. In the past few years we have been begun inscribing something new in the sands of Mars: tire tracks.

Two recent books come at the planet from very different perspectives. In *Roving Mars,* Steve Squyres gives us a vivid, intimate travelogue of the spectacularly successful (and as of this writing, ongoing) mission of the Mars Exploration Rovers, for which he serves as principal investigator. Robert Markley's *Dying Planet* takes a more distant view of the human relationship with Mars. An English professor at the University of Illinois, Markley writes about Mars science as a knowledgeable outsider, weaving in cultural history and science fiction.

Roving Mars is a page-turner. Squyres's writing is clear, with folksy touches, including flashes of dry humor and brief but revealing vignettes illuminating the personalities of a number of the key players. Many books wax eloquent about planetary exploration. Squyres shows how it is actually done. Spirit and Opportunity may not be the best spacecraft names ever, but they do provide the perfect subtitle for this book. We learn just how much spirit, in the form of human ingenuity, perseverance, joy and tears, is distilled into those craft, and how for years opportunity knocked at Squyres's door and then quickly ran off and left him, as proposals were rejected, missions were canceled, and NASA's Mars plans went through countless revisions. His perseverance seems quixotic at times, yet he never stopped answering the door, and when opportunity finally came knocking for real, Squyres was given a chance, with a ridiculously compressed schedule, to get two rovers built, tested and ready for launch in 34 months.

He and his team understandably got attached to the rovers. You, too, will come to identify with the little mechanical puppies as you learn about their troubled gestation, hurried birth and solitary departures. I was initially surprised to find that 65 percent of the book takes place before the first picture is returned from Mars. I came to appreciate this focus on the lesser-known early story.

I already thought of Spirit and Opportunity as the "miracle rovers" for their great longevity. But I had no idea. Even though you know that ultimately nothing serious is going to go wrong before they get to Mars, the tale of tribulations is gripping. Working under intense time and budget pressure, the team wrestles with faulty parachutes, defective instruments, flubbed tests, mysterious short circuits, and design flaws discovered as the pages fly off the calendar toward a launch window that could not be changed, because of the laws of celestial mechanics. I found myself wiping away tears as Squyres described the strange sadness finally evoked by the launch, as he realizes that however it all ends, they ain't never coming back home.

MARS ROVER Opportunity, June 2005.

NASA/JPL

Roving Mars is often refreshingly candid. Describing a design error that got his team's camera rejected from an earlier mission, he confesses, "With one terrible, boneheaded mistake we had thrown away five years of work." He shares some of the less than exemplary behavior needed to win the game of big science: the artful formation of science teams put together, in part, to neutralize the competition, the changing of mission designs to win over potential reviewers, and the sometimes brutal necessity of competing ruthlessly against one's long-time friends and colleagues. He also writes revealingly about the tense battles pitting "the idealistic, impractical scientists against the stubborn, practical engineers." This disarming honesty is one of the aspects of the book I appreciate the most. What are they going to do now, take his mission away? Squyres can tell it like it is because he has succeeded.

A very different perspective on the culture of science can be found in *Dying Planet*. This book is not a page-turner. The style is hyperacademic and wordy. As a scientist, I found Markley's approach, which examines science as a cultural enterprise and critically dissects our language, methods and beliefs, to be both enlightening and occasionally annoying—the latter mostly when the language became opaque and the analysis long-winded. The amount of serious scrutiny devoted to mediocre (but fun) films like *Total Recall* overestimates their importance as barometers of societal views and attitudes about the human future on Mars. The book is also sprinkled with numerous small but irritating errors that would have been caught in a careful once-over by a planetary scientist. But if you can deal with occasional passages about bringing Mars "within the semiotics of historical and experiential time" and "the ways in which Mars exists as a complex multidisciplinary object," there are many historical, literary, political and cultural nuggets—in particular, a reanalysis of the lingering impact on planetary science of the Martian canal theory of Percival Lowell, a planetologist who early in the 20th century suggested that the lines visible on Mars were the engineering works of an advanced race struggling to irrigate their drying, dying world.

Whereas previous writers have suggested that Lowell's influence quickly waned after 1920, Markley argues that this standard account "underestimates the ways in which Lowell's paradigm of a dying planet influenced scientific speculation about the composition of the

Martian atmosphere, the character of its surface, and the nature of its putative life-forms." The idea of the canals, and of a relatively wet and clement Mars, persisted in scientific discourse up to, and even into, the Space Age. Markley shows that throughout the 20th century, reputable scientists held out hope for water and plant life on Mars even while spectroscopic evidence of temperature and atmospheric composition suggested otherwise. This history should give us pause today when we rejoice in the latest evidence for a possible bio-friendly Mars.

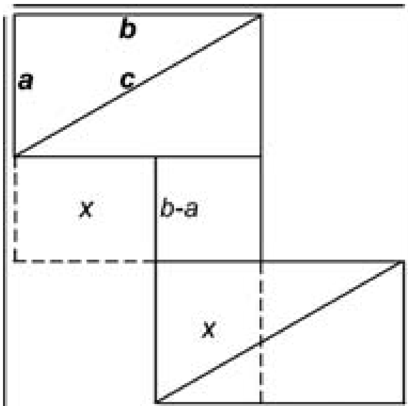
More than a century after Lowell's controversial observations, Mars remains the planet that lost its water and, probably, its capacity for life. It is now obvious that Mars cannot be flagrantly alive in the sense that Earth is. Today the question is, Can Mars be at least barely alive, with colonies holding out in underground lakes and hot springs? Possible signs of life continue to be the subject of debate and spur for further observations: 100 years ago it was the canals; 50 years ago, ephemeral hints of chlorophyll were seen and later discredited. Today we see traces of atmospheric methane that just might be the breath of underground survivors. Meanwhile the rocks, mute witnesses to the real Martian story, are beginning to talk under the patient scrutiny of the Mars Exploration Rovers.

Markley might have a field day with some of the language in *Roving Mars*. Squyres seems obsessed with finding a "water story," even to the point of repeatedly implying that the success of each rover's mission hinges on finding concrete evidence of rocks formed in or altered by liquid water. Now of course, "following the water" is important for figuring out the natural history of Mars and whether it might have been more life-friendly in the past. But isn't a robot geologist, especially the first one on a new world, successful if it simply learns any story the rocks have to tell? It doesn't seem quite right to launch an investigation committed to finding a certain answer. Even less so when that answer, "Yes, there was water here, at some uncertain time and for some uncertain duration," basically confirms the paradigm that has been developing since 1971, as Mariner 9 and subsequent orbiters have provided images and global maps showing abundant evidence for flooding and channel formation. Ultimately the tremendous success of this mission goes much deeper than a simple verification of this widespread belief. Mars has myriad stories to tell. If water is what we seek, then that's what we'll find.

It is obvious that Squyres knows all this, so why the dumbed-down rhetoric, where success is equated with proving a certain hypothesis? This is, in part, playing along with a certain line that NASA is using to package our current Mars exploration program. "Follow the Water" makes a fine slogan, and it does capture the essence of the questions of comparative planetology and habitability, both past and future, that form the core goals of our Mars program. Yet the actual motivations and strategies are never that simple. Given Squyres's admirable candor in addressing other aspects of NASA politics, I was surprised to see him uncritically adopt what seems like a simplified motivational mantra when describing the success criteria for his complex, multifaceted mission. Maybe he's just haunted by Lowell's ghost.

Squyres's book ends in September 2004, with Opportunity rooting among the sedimentary rock walls of Endurance crater, and Spirit steadfastly climbing up West Spur in the Columbia hills, where it has finally found some honest-to-goodness bedrock. I am hoping that next year there will be a paperback edition with the further exploits of the miracle rovers and their dogged, intelligent designers. SA

David Grinspoon is curator of astrobiology at the Denver Museum of Nature and Science and author of Lonely Planets: The Natural Philosophy of Alien Life.



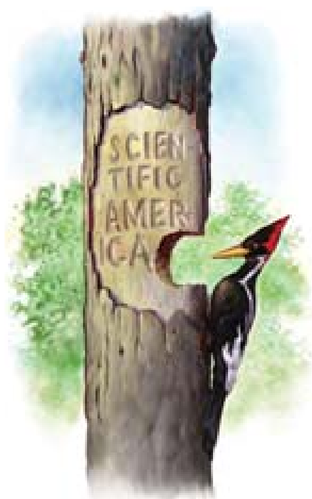
A multiple pictorial proof of the Pythagorean theorem, $a^2 + b^2 = c^2$. On the z shape (solid lines, the duckling), with the required area, the two outer triangles rotate to give c^2, the two top triangles rotate to give $a^2 + b^2$ at bottom, and (dashed lines) the lower rectangle x rotates to give a diagonal $a^2 + b^2$.

More work by the author, Paul Vjecsner, is on his website, http://vjecsner.net.

Naming Names

MUCH ADO ABOUT *N* BY STEVEN MIRSKY

There had actually been some good news for a change, for a while. The ivory-billed woodpecker, long thought to be extinct, turned out not to be. As a bird-watcher who thrills to any fleeting glimpse of a plain old pileated woodpecker, not to mention the redbellies, hairies and downies that commonly slam their heads into trees in my own Bronx backyard, I imagined how great it would be to see an ivory bill. And then I imagined that it's probably in the best long-term survival interests of any species to keep the word "ivory" out of its name. I therefore mused that the ivory-billed woodpecker should henceforth be known as the cheap-Formica-billed woodpecker. You know, for its own good.



Those happy thoughts disappeared along with New Orleans. Like most of the civilized world, I went from sad to stunned to seething as the city was swamped, just days ago as I write. Then I heard a highly respected journalist make a small mistake on national television. And, in direct contrast to the agencies that deal with disasters, I probably overreacted. Because I went berserk: full-out John-McEnroe-bad-line-call insane. Temporarily, I hope. Here's what happened.

The October 2001 issue of *Scientific American* featured a depressingly prescient article by contributing editor Mark Fischetti entitled "Drowning New Orleans." The piece reads as if Fischetti wrote it in the days following Hurricane Katrina's assault on the city, describing in detail almost exactly the scenario that engulfed the gulf—and what could have been done to prevent it. Fischetti thus became a frequent guest on television news programs in the early days of September.

One such TV appearance came on Sunday, September 4, 2005 (a date which will live inflaming me), on *Meet the Press*. NBC News Washington bureau chief and *Meet the Press* moderator Tim Russert introduced Fischetti and noted that his article had appeared in—buckle up—*Scientific America*. (The written transcript of the program corrects the error, but it's clear in the audio available through the NBC Web site.)

This mistake is common. I've lost count of the number of people I've met who say, "Oh, I read *Scientific America* all the time." To which I say to myself, "No, you don't."

So why did I get so mad this time? There were certainly much more pressing issues that day than the correct name of this publication. Russert's mistake might have been a mere tongue slip. But I'm betting that the more likely explanation for the name truncation is that most muckety-mucks, whether in politics or journalism, are unfamiliar with *Scientific American* because they don't pay all that much attention to science in general. Which is ironic, because Fischetti's piece isn't the only one in these pages that can tell the reader what will be happening years down the road, whereas most news outlets only tell you what happened yesterday.

Actually, *Moot the Press* and I have had issues for some time. Back in the dark days before the turn of the millennium, Y2K was an alleged impending crisis. The renowned computer scientist, systems analyst and cultural anthropologist Pat Robertson appeared with Russert to explain that the coming computer meltdown was going to cripple us. (Yes, this is the same Pat Robertson who recently pulpitized in favor of assassinating the duly elected president of a sovereign nation, then denied it and then apologized for it.) I wrote a letter to the producer suggesting that a serious discussion of the Y2K issue would require guests with greater expertise in the field than a televangelist has.

So it is definitely a positive development that people like Mark Fischetti, people who actually know stuff, get invited to appear on national news programs, along with spinning politicians. Having some scientists on would also be good. And it would be nice if you got our name right. Because it's inconceivable that a national news host could refer to *The Wall Street Journey*, *Newswake* or *Tim*, Tim. SA

ILLUSTRATION BY MATT COLLINS; PHOTOGRAPH BY FRANK VERONSKY

How does the slingshot effect work to change the orbit of a spacecraft?

Jeremy B. Jones, Cassini Navigation Team chief at the Jet Propulsion Laboratory in Pasadena, Calif., explains:

Spacecraft taking advantage of a gravity assist use the same principles that underlie orbital changes occurring regularly among moons and smaller bodies in the solar system. Comets from outlying regions, for instance, are often thrown into the inner solar system by the major planets, frequently Jupiter.

Absent any other influence, a moon or a spacecraft traces an elliptical path around a larger body, called the primary body, with constant orbital energy and angular momentum. But when a spacecraft comes close to a moon that is also circling the same primary body, the two smaller objects exchange orbital energy and angular momentum. Because the total orbital energy remains constant, if the spacecraft gains orbital energy, that of the moon decreases. Orbital period, the time required to complete one revolution, is proportional to orbital energy. Therefore, as the spacecraft's orbital period lengthens (the slingshot effect), that of the moon grows shorter.

Because a spacecraft is much, much smaller than a moon, the effect on its orbit is far greater than on that of a moon. For example, the Cassini spacecraft to Saturn is about 3,000 kilograms, whereas Titan, the largest of the ringed planet's satellites, weighs some 10^{23} kilograms. The effect of a slingshot maneuver on Cassini is thus about 20 orders of magnitude greater than that on Titan.

A spacecraft that passes "behind" the moon gets an increase in its velocity (and orbital energy) relative to the primary body, which gives the appearance of a slingshot throwing it into a larger orbit. We can also fly a spacecraft "in front" of a moon, to decrease its velocity (and orbital energy). Moreover, traveling "above" or "below" a moon can alter the direction of the spacecraft's velocity, modifying only its orbital orientation (and angular momentum magnitude). Intermediate flyby orientations change both energy and angular momentum. Of course, all such adjustments precipitate an inverse change in the energy and angular momentum of the moon, but its larger mass results in changes so small that they are undetectable among all the other forces that affect a moon's orbit.

Where does wind come from?

Chris Weiss, assistant professor of atmospheric science at Texas Tech University, provides this answer:

Simply put, wind is the motion of air molecules. Two concepts are central to understanding what causes wind: air and air pressure.

Air contains molecules of nitrogen (about 78 percent by volume), oxygen (about 21 percent), water vapor and other trace elements. All these air molecules move about very quickly, colliding readily with one another and with any objects at ground level.

Air pressure is defined as the amount of force that these molecules impart on a given area. In general, the more air molecules present, the greater the air pressure. Wind, in turn, is driven by what is called the pressure gradient force.

Changes in air pressure, such as those caused by the dynamics of storm systems and uneven solar heating, in a given horizontal area force air molecules from the region of relatively high air pressure to rush toward the area of low pressure.

The areas of high and low pressure displayed on a weather map in large part drive the gentle breezes we usually experience. The pressure differences behind this wind are only about 1 percent of the total atmospheric pressure, and these changes occur over the range of multiple states. The winds in severe storms, in contrast, result from much larger and more concentrated areas of pressure change. SA

For a complete text of these and other answers from scientists in diverse fields, visit www.sciam.com/askexpert

MATT COLLINS