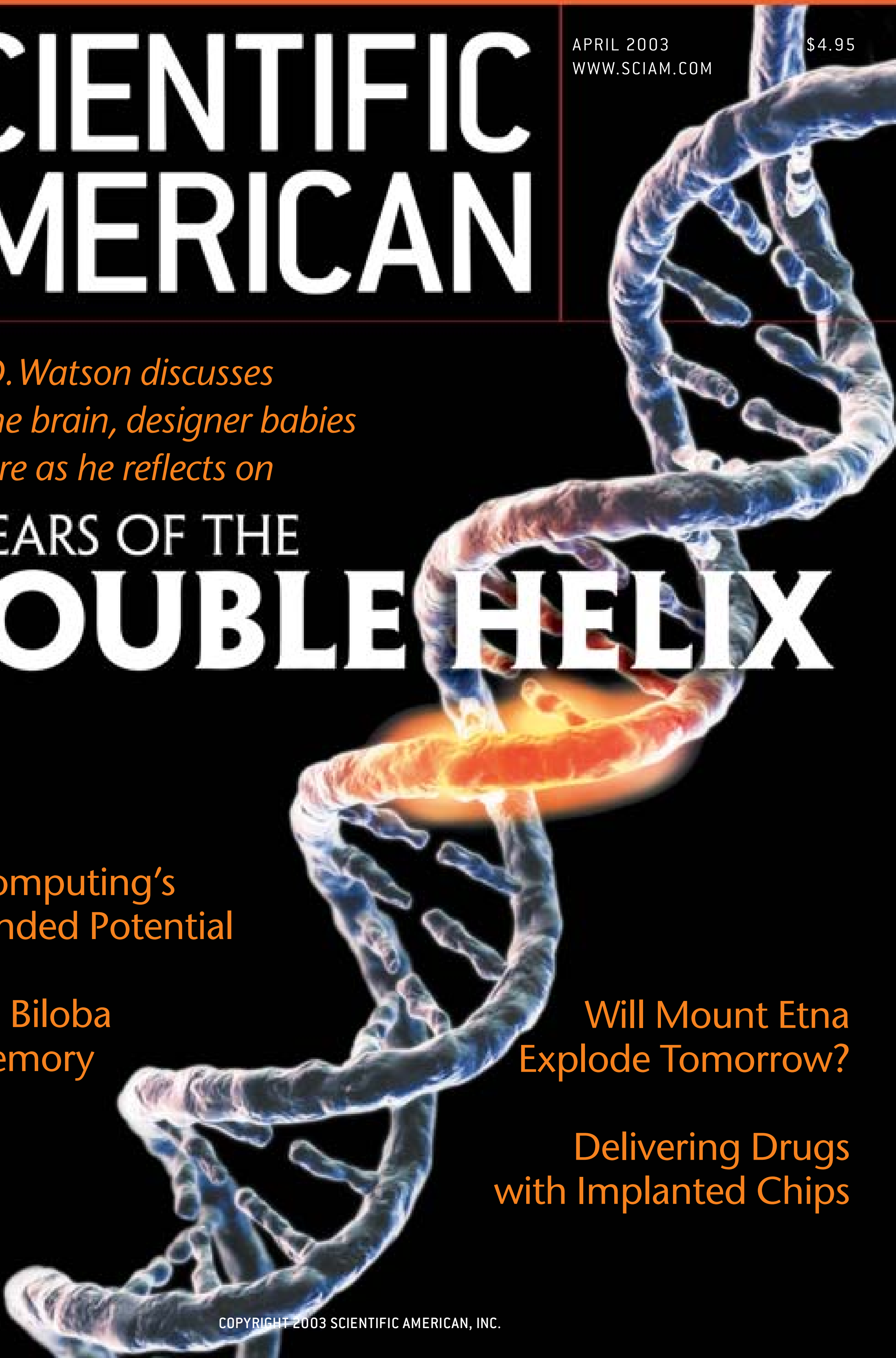
SOLVING THE NEUTRINO MYSTERY • RECOGNIZING ANCIENT LIFE

SCIENTIFIC AMERICAN

APRIL 2003
WWW.SCIAM.COM
$4.95

James D. Watson discusses DNA, the brain, designer babies and more as he reflects on

50 YEARS OF THE DOUBLE HELIX



Grid Computing's Unbounded Potential

Ginkgo Biloba and Memory

Will Mount Etna Explode Tomorrow?

Delivering Drugs with Implanted Chips

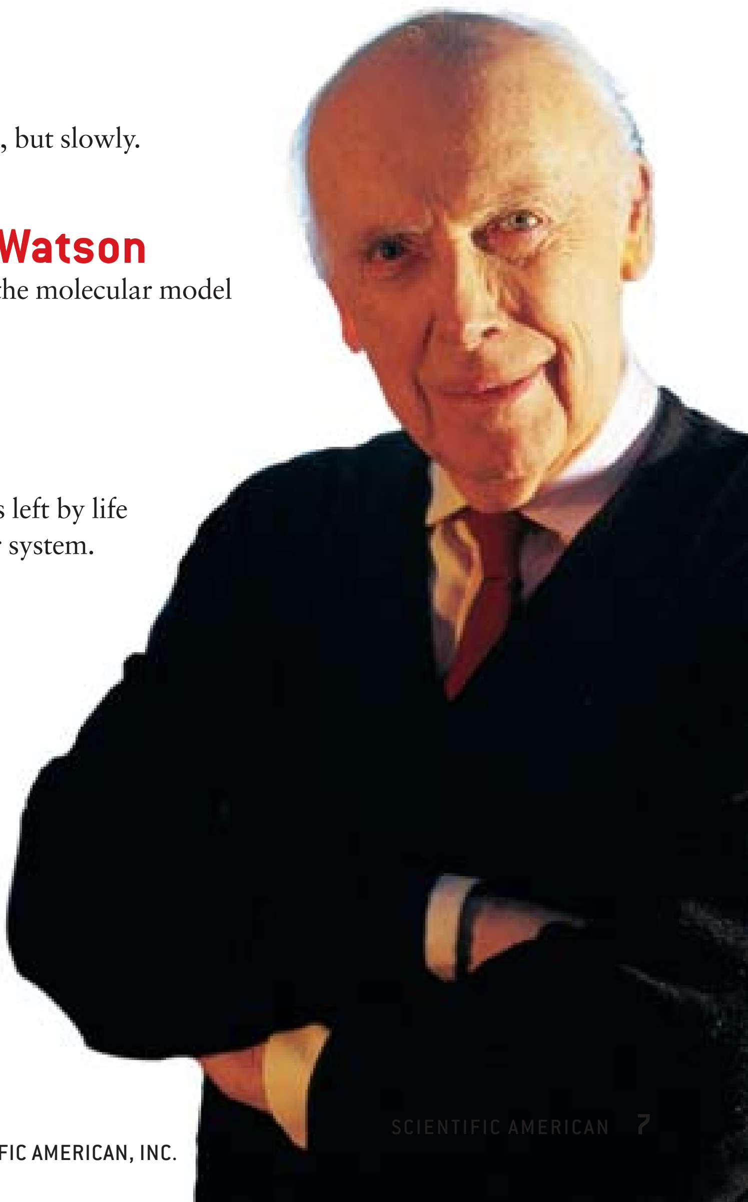
april 2003

contents

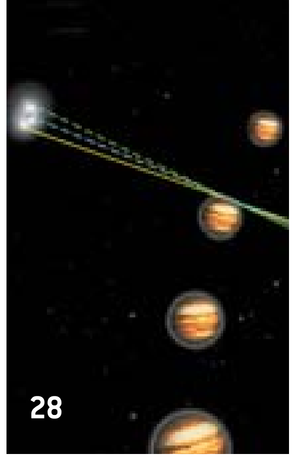
features

SCIENTIFIC AMERICAN Volume 288 Number 4

66 James D. Watson



departments



28

100

94

columns

Cover image by Mike Medicine Horse, Hybrid Medical Animation

Scientific American (ISSN 0036-8733), published monthly by Scientific American, Inc., 415 Madison Avenue, New York, N.Y. 10017-1111. Copyright © 2003 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. 242764. Canadian BN No. 127387652RT; QST No. Q1015332537. 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 or send e-mail to sacust@sciam.com **Subscription inquiries: U.S. and Canada (800) 333-1199; other (515) 247-7631.** Printed in U.S.A.

SA Perspectives

Get Real

When the cloning of a human was announced last December, political and spiritual leaders condemned it as an affront to the "dignity of man." That kind of rhetoric is popping up all over the place. Political scientist Francis Fukuyama warns that genetic engineering and Prozac-like drugs augur "a 'posthuman' stage of history." Bill Joy, co-founder of Sun Microsystems, frets over robotics and nanotechnology: "On this path our humanity may well be lost." Even the *Economist,* a magazine not usually given to apocalyptic predictions, worries that neuroscience could "gut the concept of human nature."

ANXIETY over genetically modified food often reflects abstract worries about science.

SION TOUHIG *Corbis Sygma*

Like their counterparts in earlier ages, these commentators argue that technology is running ahead of our ability to deal with it; although scientific progress is all well and good, we have to rein it in. Such views are often called neo-Luddism, but frankly, that does not do justice to the Luddites. Those machine-smashing textile workers were reacting to immediate threats, such as losing their jobs. Today's concerns tend to be abstract, and that is their problem.

A science magazine is all in favor of abstract thinking, but at some point abstraction needs to make contact with reality. And the reality of research bears little resemblance to the technocynics' horror stories. Will cloning, for example, open the door to "designer babies"? Maybe one day. For now, though, researchers are struggling to develop cloning just to grow tissues that a patient's immune system won't reject. Even would-be baby cloners don't purport to fiddle with the genome.

Are people supposed to give up the prospect of lifesaving therapies to avoid a distant, hypothetical threat? The answer from technocynics is yes. In his book last year Fukuyama drew a line between medical therapy (OK) and genetic enhancement (not OK) but went on to advocate a ban on all cloning, even the therapeutic kind. Similarly, Joy has called for a "relinquishment" of all—*all*—research into robotics, nanotechnology and genetic engineering. Where does this absolutist stand leave the rest of us? We have watched our parents and grandparents waste away from cancer and Alzheimer's disease. We have seen children die of diabetes and friends fall to depression, malaria and HIV. If it comes down to a choice between the vague unease that emerging technologies conjure up or the very unvague suffering they could cure, we know how we would decide.

The technocynics basically want us to grin and bear it, lest our attempts at self-improvement do more harm than good. Yet if history is any guide, fears about the impact of new technologies generally wind up sounding pretty silly. Thoreau regarded trains, telegraphs, newspapers and even mail delivery as dehumanizing. Late Victorians predicted that industrialization and urbanization would cause our species to degenerate to a prehuman state. In the 1970s critics of in vitro fertilization said it would create monstrous or deranged babies. In all these cases, abstract worries gave way to mundane ones. New technologies did bring new problems, but people worked around them. Few would, in retrospect, ditch the technologies altogether.

The biggest danger, then, is not that science will run ahead of ethics, but the opposite: that ethical hair triggers will paralyze worthy research. Striking a balance is not easy. Bioethicist Gregory Stock offers a sound prescription: "We should deal with actual rather than imagined problems." To stop research is to give up trying to make the world a better place. It denies human nature in order to save it.

THE EDITORS editors@sciam.com

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
212-754-0550
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
515-247-7631
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

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

Southeast/Southwest
MancheeMedia
972-662-2503
fax: 972-662-2577

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

Canada
Fenn Company, Inc.
905-833-6200
fax: 905-833-2116

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

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

Sweden
Publicitas Nordic AB
+46-8-442-7050
fax: +49-8-442-7059

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

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

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

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

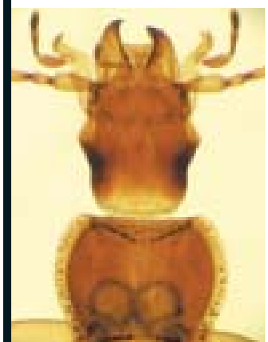
Hong Kong
Hutton Media Limited
+852-2528-9135
fax: +852-2528-9281

On the Web

WWW.SCIENTIFICAMERICAN.COM

FEATURED THIS MONTH

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



MARK W. WESTNEAT *Courtesy of the Field Museum*

Ultrapowerful X-rays Reveal How Beetles Really Breathe

Even the most up-to-date biology textbooks, if they address insect respiration, now need revision. With the help of a high-energy particle accelerator, researchers have documented bugs breathing in a manner never before thought possible: like mammals. The x-ray video technology used to conduct the examinations could have applications in robotics and medicine.

Parasite's Plant Genes Could Be Achilles' Heel

The sleeping-sickness parasite kills nearly 66,000 people annually—and silently infects almost 450,000 more. New work suggests that the parasite carries algae genes and thus could succumb to drugs based on herbicides.

Dairy Farming, Old and New

Got milk? Although the drink is ubiquitous today, it is unclear when ancient farmers started using their cattle for dairy products (as opposed to meat). But recent archaeological analyses show that Britons were harvesting milk as early as 6,000 years ago. Meanwhile the results of another study indicate that current cheese-making practices could be enhanced by genetic engineering.

Our Galaxy's Next Supernova?

Astronomers have identified the best candidate yet for our galaxy's next supernova explosion. It seems that the star Rho Cassiopeiae, located 10,000 light-years from Earth, is the one most likely to run out of fuel and meet a violent fate in the near future.

Ask the Experts

Why do hangovers occur?

Sant Singh of the Chicago Medical School explains.
www.sciam.com/askexpert_directory.cfm

SCIENTIFIC AMERICAN DIGITAL

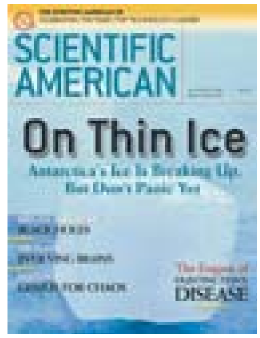
Access every issue of *Scientific American* magazine from 1993 to the present.

SUBSCRIBE TODAY! www.sciamdigital.com

SCIENTIFIC AMERICAN®
Established 1845

Letters

EDITORS@SCIAM.COM



IF APRIL IS THE CRUELEST MONTH, December may make us feel the most reflective, as we recall the past year's events. So it was with the December 2002 Perspectives, "In Science We Trust." The column reviewed some of the achievements—and regrettable setbacks—of science, which the editors nonetheless praised for "its incremental progress toward a more complete understanding of the observable world." The commentary resonated with many, including James Edgar of Melville, Saskatchewan, who responded: "I think I'll photocopy your editorial and add it to my collection of wise words—a collection that helps me to explain my beliefs about science, astronomy, evolution and life." Other writers express their beliefs concerning the December issue on the following pages.

POLLOCK'S FRACTALS

I don't understand some aspects of "Order in Pollock's Chaos," by Richard P. Taylor. If the computer measures squares in Jackson Pollock's works that have paint in them and those that don't, I don't think that the numerical ratio between the haves and the have-nots would change, no matter the scale.

The second thing that bothers me is the straight-line graph, when squares from 10 to zero millimeters are analyzed. Paintings are not like computer fractals, in which the locations of edges can be determined at every scale. At such sizes, I cannot imagine how one would know where the edge of the line is, given that paint bleeds, runs, is absorbed by the surface and mixes with other colors. It would seem, too, that the hills and valleys of the canvas would become the dominant features. Also, if a photograph of the painting was scanned into the computer, doesn't the analysis exceed the resolution capabilities of the photograph and the scanner?

Michael Burke
New York City

TAYLOR REPLIES: The fractal character of a pattern does in fact reveal itself through the way the number of filled squares changes with magnification. For something to be fractal, the number of filled squares, N, must scale with the square length, L, according to the power law relation $N \sim L^{-D}$. D is the fractal dimension—it quantifies the scaling relation among the patterns observed at different magnifications. This power law relation is true also for smooth Euclidean shapes. The distinguishing property is that for a smooth Euclidean line $D = 1$, whereas for a fractal line $1 < D < 2$.

Regarding the second point: as noted in the text, we examine the fractal behavior over a range from about a meter down to a millimeter. For the fakes, the biggest distortion away from fractal behavior occurs at the small scales. After we established the fractal character of Pollock's paintings, we then went back to the film to determine the physical processes that created them. For large scales, the key was in the way that Pollock moved around the canvas (he actually followed motions called Levy flights). At smaller scales (10 centimeters and below), the fluid dynamics become important: how the paint was launched from the brush, how it fell and how it seeped into the canvas.

Fractals in the real world are different from mathematical fractals: they can't go on forever. In fact, most fractals in nature continue over a magnification range of only about 20 times. Pollock is extraordinary in this regard, because his fractals are charted over a magnification of 1,000 times! His patterns are fractal down to the finest speck of paint, about one millimeter in size.

Finally, we use high-resolution images in which distortion doesn't occur until 0.8 millimeter. Also, before sending images through the computer, we visually inspect them for any distortions caused by bumpiness.

LAGGING PHOTONS?

I read with interest the Profile of Fotini Markopoulou Kalamara ["Throwing Einstein for a Loop"], by Amanda Gefter, and the article "The Brightest Explosions in the Universe," by Neil Gehrels, Luigi Piro and Peter J. T. Leonard. I was particularly intrigued by the following quotations. From the Profile: "One experiment could be to track gamma-ray photons from billions of light-years away. If spacetime is in fact discrete, then individual photons should travel at slightly different speeds, depending on their wavelength." From the article: "Roughly 90 of the [gamma-ray] bursts seen by BATSE [the Burst and Transient Source Experiment onboard the Compton Gamma Ray Observatory] form a distinct class of their own, defined by ultralow luminosities and long spectral lags, meaning that the high- and low-energy gamma-ray pulses arrive several seconds apart. No one knows why the pulses are out of sync."

This may just be coincidence. I have no idea what Markopoulou Kalamara's theories suggest the arrival-time difference should be for various wavelengths of photons, and there must be myriad possible explanations for the BATSE results. But it struck me.

Jonathan Leete
Arlington, Va.

GEHRELS AND LEONARD REPLY: The spectral lags observed in gamma-ray bursts by BATSE are quite different from what is predicted by quantum gravity. The BATSE lags observed between energies of 100 and 300 kilo-electron-volts (keV) ranged up to several seconds in length, with higher-energy photons arriving before lower-energy ones.

But quantum gravity predicts an effect on the order of about three milliseconds per giga-electron-volt (GeV) per billion light-years distance. This amounts to a lag of less than 0.001 millisecond for a burst source at one billion light-years observed between 100 and 300 keV; such small lags were undetectable by BATSE. Also, quantum gravity predicts that higher-energy photons lag behind lower-energy ones—contrary to the effect seen by BATSE.

The quantum gravity lags would be easier to observe at GeV energies. We are excitedly awaiting the 2006 launch of the Gamma-ray Large Area Space Telescope (GLAST), with the hope that it will detect such lags in gamma-ray bursts.

GAMMA-RAY BURST produces intriguing photons.

YOU WIN SOME . . .

I really enjoyed the "Scientific American 50." A once-a-year summary of major developments is a great way to get the big picture. Don't change it (much) next year!

Mike Steiner
via e-mail

As a longtime subscriber, I have never seen so much space wasted as in the "Scientific American 50." Surely you can find better articles. I hope this won't be an annual waste.

Peter Tiley
Dundas, Ontario

SECRETS OF SPECIES SUCCESS

In "Food for Thought," William Leonard states that "the goal of all organisms is the same: to devote sufficient funds to reproduction to ensure the long-term success of the species." This implies that individuals act for "the good of the species"—a notion that has long been shown to be false. If one can speak of a "goal" for individual organisms, it would be to maximize their genetic contribution to future generations.

Don Luce
Bell Museum of Natural History
Minneapolis

LEONARD REPLIES: I did not mean to imply that organisms act for the good of the species. It's true that an individual's motivation is to maximize its own reproductive success. That said, from the long-term lens of evolution (and the perspective of the population), the act of individuals allocating energy to the next generation is what enables species to persist and succeed.

3-D MEMORIES

Memory is plastic, as Mark Alpert demonstrates in Technicalities ["Getting Real"] when he recalls viewing the 1983 film *Jaws 3-D* through cardboard goggles with red and blue filters. He's describing the anaglyph process, which used one red filter and one green (or blue) one for 3-D viewing of projected monochrome images.

The process used in *Jaws 3-D,* however, was different; it permitted stereoscopic projection in full color. It employs polarizing filters at the projector and goggles with polarizing filters. I also remember red and blue goggles from a series of 3-D Batman comic books in the early 1960s, however. Maybe the lenses Alpert remembers were not from a movie at all.

Tom Flynn
Buffalo, N.Y.

ALPERT REPLIES: You're absolutely right. That's what happens when you read too many comic books.

ERRATUM In "On Thin Ice," by Robert A. Bindschadler and Charles R. Bentley, a statement about global warming should have read: "Around the world, temperatures have risen gradually since the end of the last ice age, but the trend has accelerated markedly since the mid-1900s"; we mistakenly printed "since the mid-1990s."

MARK A. GARLICK

50, 100 & 150 Years Ago

FROM SCIENTIFIC AMERICAN

The Wily Flu ▪ Frozen Continent ▪ Fishy Aviation

APRIL 1953

INFLUENZA VS. IMMUNITY—"The serological character of the A virus has changed seven or eight times since 1933, and each change in character has within a year been evident all over the earth. Soon after influenza A2 was found in the U.S., it appeared in Australia and England as well. After it had taken hold, no A1 strains were found anywhere. And so for each successive change. It is a parasite whose only natural host is man. To survive, it must pass continually from one human being to another—it is inhaled and lodges in the respiratory tract. But it soon finds itself in the position epidemiologists call 'exhaustion of susceptible hosts.' In other words, almost the entire population becomes immune. This highly transmissible virus meets the situation by a transformation of character—a mutation that enables it to overcome its host's immunity. —Sir Macfarlane Burnet" [*Editors' note: Burnet was a co-winner of the 1960 Nobel Prize for Physiology or Medicine "for discovery of acquired immunological tolerance."*]

A MASSIVE SEARCH—"The elementary particle corresponding to the gravitational field has been named the graviton. There can be little doubt that in a formal mathematical sense the graviton exists. However, nobody has ever observed an individual graviton. Because of the extreme weakness of the gravitation interaction, in practice only large masses produce observable gravitational effects. In the case of large masses, the number of gravitons involved in the interaction is very large, and the field behaves like a classical field. Consequently, many physicists believe that the individual graviton never will be observed. Whether the graviton has a real existence is one of the most important open questions in physics. —Freeman J. Dyson"

THE INDIAN motor bicycle, 1903

APRIL 1903

BRAVING ANTARCTICA—"Reports say the 'Discovery' entered the ice pack December 23, 1901, in latitude 67. On March 24 the ship was frozen in, but the expedition passed a comfortable winter near Mounts Erebus and Terror. On September 2 two sledge parties were sent out. The best record made was that of Capt. Robert F. Scott, Dr. Edward Wilson, and Lieut. Ernest Shackleton. These intrepid explorers traveled 94 miles to the south, reaching land in latitude 80 deg. 7 min. This is the most southerly point yet attained. The expedition proved a most severe test of the endurance of both men and animals. All the dogs perished, so that several men had to drag the sledges back. Lieut. Shackleton almost died from exposure."

EASIER RIDER—"The increasing interest in motor bicycles manifested of late among cyclists is directly attributable to the numerous improvements which have brought various makes of these machines up to a high standard of excellence. The 'Indian' motocycle is one type of machine which has become quite popular in the cycling world. Great care has been exercised in the construction of the motor used in this machine, and by thorough testing under all conditions, it has been brought up to a high state of efficiency." [*Editors' note: Before World War I, the Indian Motocycle Company was the largest manufacturer of motorcycles in the world.*]

APRIL 1853

FLIES LIKE A FISH—"Theodore Poesche has presented a plan for navigating the atmosphere with a car propelled by a steam engine without employing a balloon. His plan is to build a long, narrow, and light wooden vessel, with wings of canvas, and propel it by a screw propeller driven by steam power. 'My ship,' he says, 'most nearly resembles the flying fish, which progresses by the spiral action of the tail, while its extended fins support it in the air.' The screw propeller was proposed long ago to drive aerial ships with balloons, but could not do it then, and to do so now without a balloon is an impossibility."

SPACEFLIGHT

Rethinking the Shuttle

IN FUTURE MANNED FLIGHTS, SMALLER WILL BE SAFER **BY MARK ALPERT**

As NASA investigates why *Columbia* broke up during its reentry into the atmosphere on February 1, killing all seven astronauts onboard, the space agency faces some difficult choices. For more than a decade, aerospace experts had warned about the vulnerability of the aging, 100-ton space shuttles to the superheated gases that envelop the craft as they descend to Earth. If investigators determine that a breach in *Columbia*'s heat shield or aluminum skin doomed the mission, NASA might require shuttle crews to inspect the craft's exterior before reentry and perhaps devise a strategy for repairing damage while in orbit. But if the accident's cause cannot be pinpointed or if a major redesign of the three remaining shuttles is required, NASA may have to accelerate its development of a smaller, more reliable spacecraft.

Previous efforts to replace the shuttle fleet have been expensive failures [see "Has the Space Age Stalled?" by Mark Alpert; SCIENTIFIC AMERICAN, April 2002]. Last November the agency committed $2.4 billion to producing a design for an orbital space plane (OSP) that could ferry a crew of at least four astronauts to the International Space Station. (With the shuttles grounded, NASA lost access to the station; only the Russian *Soyuz* and *Progress* spacecraft can ferry crews and supplies to the orbital outpost.) NASA's plans, however, are still vague; the agency has not yet decided whether the OSP will be a winged vehicle like the shuttle, a lifting body (a squat craft shaped to maximize aerodynamic lift), or a capsule like *Soyuz*. And even if Congress approved an additional $10 billion to build the space plane, it would not be ready to carry crews into orbit until 2012. Dennis E. Smith, manager of the OSP program, is looking for ways to speed up the schedule, but he cautions, "I don't think we can save a lot of time."

The orbital space plane promises to be much safer than the shuttle. The OSP would hold only astronauts, not heavy cargo, so it

REENTRY TRAGEDY: Fallen debris from the space shuttle *Columbia* leans against a fence near Douglass, Tex.

DONNA McWILLIAM *AP Photo*

NASA ARTIST CONCEPT

FOUR DESIGNS for the orbital space plane (*clockwise from top left*): lifting body; winged vehicle with sharp leading edges; shuttlelike vehicle; and capsule, which may be the safest for reentry.

would be compact and light enough to be launched by a single-use commercial rocket such as the Delta 4. The shuttle, in contrast, requires three rocket engines built into the vehicle, an external tank of liquid hydrogen and oxygen to feed those engines, and a pair of solid-fuel boosters.

The immense strain of a shuttle launch invites hazards: a leak in a solid-fuel booster caused the loss of the shuttle *Challenger* in 1986, and a piece of foam insulation falling from the external tank may have damaged *Columbia*'s left wing shortly after its launch on January 16. The smaller size of the OSP would also reduce the chance of a collision with micrometeoroids and man-made debris while the craft is in orbit. (Such debris could have struck *Columbia* during its final mission.) And the OSP's heat shield could be fashioned from newly developed metallic panels, making it more resilient than the shuttle's patchwork of ceramic tiles.

The main disadvantage of the space plane is that it could not perform all the shuttle's tasks. NASA would have to develop an unmanned launch and docking system to send heavy payloads to the space station. And the OSP would have its own risks, of course. The safety record of even the most successful rockets is not perfect—the Delta 2, for example, has carried payloads into orbit 104 times since 1989 but did explode once, in 1997. NASA would need to extensively test and upgrade the boosters chosen to launch the OSP.

To minimize the dangers of atmospheric reentry, the best design choice for the OSP may be a capsule shape. According to Theodore A. Postol, a space systems expert at the Massachusetts Institute of Technology, a blunt capsule falling through the atmosphere heats up much less than a winged vehicle does. And by eliminating wings, wheels and control surfaces, engineers could devote more of the craft's mass to the all-important thermal shield. After descending to the lower atmosphere, the capsule could float on parachutes to an ocean landing, just as the *Apollo* modules did in the 1960s. "Given all those benefits, is it really worth landing on a runway?" Postol asks.

NASA officials, though, do not seem enthusiastic about the capsule design. Smith, the OSP manager, expresses concern about the reliability of parachute mechanisms and the cost of retrieving the spacecraft from the ocean. Postol thinks a different factor may explain the agency's reluctance: "I expect that NASA will resist the capsule for emotional reasons. The astronauts want to fly the vehicle." Even if that makes for a riskier reentry.

A NEW VEHICLE FOR ASTRONAUTS

Boeing, Lockheed Martin, and a team consisting of Orbital Sciences and Northrop Grumman have already begun work on preliminary designs for the orbital space plane. The design proposed by Orbital Sciences is based on the HL-20 Lifting Body, a vehicle concept that NASA studied in the early 1990s. The space plane would be about 37 feet long (compared with 122 feet for the shuttle) and have a wingspan of about 35 feet (compared with 78 feet for the shuttle). It could carry a crew of five astronauts to and from the International Space Station (the shuttle typically carries seven). Also, one of the space planes could be continuously docked to the station in case it is needed for emergency evacuation.

An editor's commentary about the odds against *Columbia* appears at www.sciam.com, under the "Explore" link.

ENVIRONMENT

Oiling Up Spain

A SUNKEN TANKER COULD TARNISH SPAIN FOR DECADES BY LUIS MIGUEL ARIZA

Thousands of tons of heavy fuel remain in the bow and stern sections of the *Prestige,* the oil tanker that split in half off the northwestern coast of Spain on November 19, 2002. It sank to the seabed, more than 3,500 meters deep in the Atlantic Ocean some 200 kilometers from Galicia. Tons of toxic fuel have oozed from 20 cracks in the hulls as semisolid black strings, like toothpaste being squeezed from a tube, and have drifted toward the sea surface. It has become Spain's worst ecological disaster ever, halting coastal fishing and polluting beaches. The ship has already spilled at least 30,000 tons of oil, and researchers aren't sure when the seepage will stop.

At first, scientists thought that the fuel would freeze: oil remains fluid to six degrees Celsius, and the deep water around the wreck is at 2.75 degrees C. Experts guess that the *Prestige* left port from Ventspils, Latvia, in the Baltic Sea with 77,000 tons at 50 degrees C. But no one knows what the oil temperature is now. According to Malcolm L. Spaulding, professor of ocean engineering at the University of Rhode Island, the calculations of the cooling time for the oil in the bow and the stern, which lie 2.5 kilometers apart on the seafloor, have proved extraordinarily difficult. In principle, the fuel in contact with the external walls of the hulls should cool faster than that in the center of the tank. The time it takes for all the oil to cool down depends on the amount and rate of mixing, a critical factor "that complicates heat transfer from the oil to the tank walls and finally to the seawater," Spaulding explains. If the oil mixed together well, the entire cargo should have cooled off and frozen in 40 days. That did not happen, Spaulding remarks, probably because the mixing is substantially reduced. He now anticipates "cooling times of many months to several years."

Michel Girin, director of the French Center of Documentation, Research and Experimentation on Accidental Water Pollution, notes that because the viscous fuel that escapes from the cracks probably does not come from the ship's middle, the *Prestige* will effectively be a "permanent source of pollution." Simulations in a pressure chamber that mimics the conditions at the tanker (about 100 atmospheres) have revealed that the fuel will never solidify. "We have seen that even at –10 degrees Celsius, it continues to flow," remarks Jean Croquette of the French Research Institute for Exploitation of the Sea. "The density of the fuel is lower than that of the seawater, so it retains the capacity to flow." That means, he says, that the leakage won't stop.

Many sunken ships have leaked significant amounts of oil decades after their wreckage, according to the National Oceanic and Atmospheric Administration. The SS *Jacob Luckenbach,* a freighter that sank in 1953 off the coast of San Francisco, caused periodic "mystery spills" in the Bay Area until it was identified and much of its oil removed just last year. The oil tanker *Nakhodka,* which sank in 1997 to a depth of 2,500 meters near the Japanese island of Honshu, continues to seep small quantities of fuel, Girin says.

To prevent any more oil from the *Prestige* from reaching Spain's coast, a commission of Spanish scientists used the French submersible *Nautile* to patch the cracks. This past January the *Nautile* blocked 17 of the 20 ruptures with steel plates or bell-shaped caps, slowing down the leakage rate of 120 tons a day soon after the sinking to a couple tons by mid-February. Emilio Lora Tamayo, head of the commission, admits that the patches will not last forever and will need continual maintenance. Even a complete sealing would be a temporary solution: researchers calculate that the hulls of the tanker will erode and break apart within 23 to 40 years. The neutralization of the *Prestige,* Tamayo observes, "is a technological challenge never attempted before."

Luis Miguel Ariza is a science writer based in Barcelona, Spain.

IGNOMINIOUS END to the *Prestige*, carrying 77,000 tons of fuel oil, as it broke in half off the coast of Spain in November 2002 and sank to a depth of 3,500 meters. Its cargo could contaminate the environment for decades.

REUTERS NEWMEDIA, INC. *Corbis*

TOO SLICK TO HANDLE?

Researchers have proposed several ways to deal with the remaining oil in the *Prestige,* none of which is ideal. Entombing the wreckage with concrete, sand or plastic would be difficult to do underwater. Exploding the hulls and successfully retrieving the freed oil at the surface with ship-mounted skimmers depend strongly on weather conditions; rough seas could lead to uncontrolled spreading.

The best bet may be direct pumping of the fuel. Smit, the Dutch salvage company that raised the Russian submarine *Kursk,* engineered a system that retrieved the oil from two tankers lying 80 meters underwater near Bussan, Korea. "Basically, we go down with little robots that drill holes in the tank," explains Lars Walder, a Smit spokesperson. "By mixing the oil with hot water or a special type of seed oil, we can make the fuel more fluid in order to be able to pump it out." The company admits, however, that the technology to work at depths in excess of 2,000 meters has yet to be developed.

GEOCHEMISTRY

Foiling a Faint Sun

HUNGRY BACTERIA MAY HAVE WARMED AN ANCIENT EARTH BY SARAH SIMPSON

With its often telltale stench of rotten eggs, sulfur would not top most people's list of favorite culinary seasonings, but plenty of bacteria adore it. What's more, their predictable preferences for certain flavors of sulfur can be preserved inside rocks for billions of years, revealing detailed stories about the earth's distant past—most recently, a scenario for why the planet didn't freeze when the sun was considerably cooler.

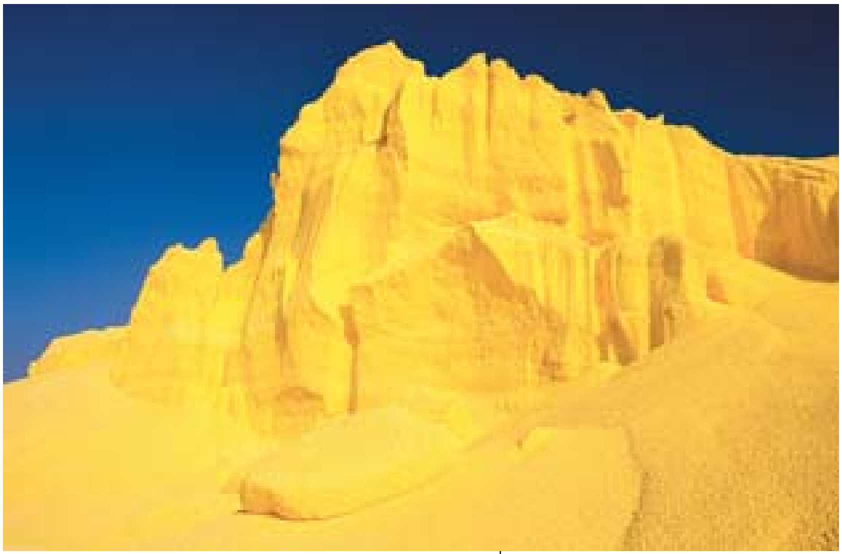
It all boils down to who's eating what. All organisms are finicky, and the sulfur-loving bacteria are no exception. They prefer sulfur 32, the element's lightest isotope, to heavier sulfur 34 when given the choice. In the modern world, they leave a waste product (usually recorded in the mineral pyrite) with a surplus of sulfur 32 that is as much as 7 percent greater than that of their food source, a sulfur compound called sulfate (today a common salt in seawater). But the surplus of sulfur 32 in pyrite that formed before 2.5 billion years ago rarely exceeds 1 percent. This discrepancy suggests to geochemists that back then, during a geologic eon known as the Archean, sulfate was scarce: with less food available, sulfur lovers couldn't afford to be as picky about the isotope they favored.

Indeed, when Kirsten S. Habicht of the University of Southern Denmark in Odense and her colleagues grew cultures of three kinds of modern sulfur-craving bacteria, they found that the sulfur 32 surplus dropped to the low value recorded in Archean rocks only when the researchers reduced sulfate concentrations below 200 micromolar—not even 1/100 the amount present in oceans now.

Such low levels of sulfate indicate that the Archean atmosphere was virtually devoid of oxygen. Without oxygen to help wear down the continents, sulfur minerals stay locked in the rocks and unavailable to hungry bacteria. Extremely low sulfate levels could also explain why the oceans did not freeze, considering that the sun was about 25 percent dimmer. Scientists have long believed that greenhouse gases must have been insulating the earth much more efficiently than they do today. They have argued that methane—with many times the heat-trapping capacity of carbon dioxide—may have been more abundant, in part because it could stick around a whopping 10,000 years in the oxygen-poor atmosphere of the past. But what generated the continuous supply of the gas was still uncertain.

That's where the sulfur lovers reenter the story. Habicht's team calculated that their sulfate-starved bacteria would perform their daily chores, such as decomposing the remains of other organisms, much more slowly—at rates 30 to 90 percent lower than those fed a modern serving of sulfate. In the ancient world, that would have left countless open seats at the dinner table for another set of hungry microorganisms, the methanogens. A proliferation of these methane producers could have sustained a strong greenhouse effect.

Uwe H. Wiechert of ETH Zentrum in Zürich says that the calculations of Habicht's team are the first clear suggestion that methanogens provided that needed greenhouse-gas supply. But to know for sure will require independent verification of methane's behavior in the Archean atmosphere, which might have resembled present-day conditions on Saturn's methane-shrouded moon, Titan—the target for NASA's Huygens probe in 2005. In the meantime, stinky sulfur will undoubtedly reveal more sweet secrets.



CORBIS

WEALTH OF SULFUR feeds a prolific suite of bacteria on earth today, but its scarcity in the ancient ocean kept the microbes at bay. The mound shown, in Wilmington, Calif., is a stockpile of refined sulfur.

GEMSTONE GOSSIP

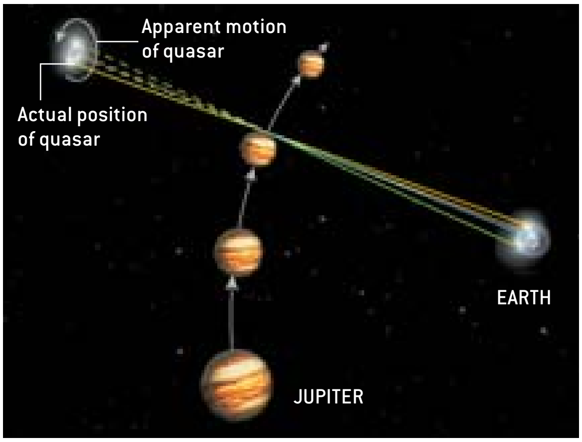
Who would have thought that a girl's best friend could whisper tales of an ancient earth choking from a lack of oxygen? James Farquhar of the University of Maryland and his colleagues recently found that tiny specks of iron sulfide inside African diamonds bear a bizarre ratio of sulfur isotopes that results only when ultraviolet radiation from the sun breaks apart sulfur-rich gases in the atmosphere. Farquhar had shown previously that this atmospheric breakup diminishes dramatically once oxygen builds up in the atmosphere and the ozone layer begins blocking the radiation. The diamond discovery further implies that this unique sulfur signature can survive a geologic roller coaster: riding from the atmosphere to deep within the earth's mantle, where diamonds form, and then back to the planet's surface.

PHYSICS

A Tale of Two C's

GRAVITY SPEED TEST RAISES SOME RELATIVISTIC EYEBROWS **BY GEORGE MUSSER**

The hardest thing to prove is something you think you already know. How can you be sure that you're proving it, rather than merely reasserting your belief? So it is with the latest test of Einstein's general theory of relativity—a measurement of the speed at which changes in a gravitational field propagate. If the sun suddenly shattered into a million pieces, this speed would determine how many minutes of blissful ignorance the denizens of Earth would have until our orbit went haywire. In Einstein's theory, the speed of gravity (abbreviated c_g) exactly equals the speed of light in a vacuum (c).



LIKE A LENS, Jupiter bends the light rays from a distant quasar. The yellow ray is unaffected by Jupiter and takes a direct path to Earth; the dotted lines show the illusory paths of the ray. Jupiter's motion causes the quasar image to trace out a circle. Relativistic effects skew the circle (*not shown*).

Lo and behold, that is what a physicist-astronomer duo announced at the American Astronomical Society meeting in January. Einstein, they concluded, was right once again. Yet most relativity researchers are skeptical. "It's a beautiful experiment that gives a very nice new confirmation of general relativity, but it's still unclear whether it's testing the speed of gravity," says Steven Carlip of the University of California at Davis.

No one questions the basic experimental setup, devised by Sergei Kopeikin of the University of Missouri and Edward Fomalont of the National Radio Astronomy Observatory. The idea was to look for the effect that a nearby celestial body has on the light rays from a more distant object. The nearby body should bend the light rays, temporarily shifting the image of the distant object. In a famous (if controversial) expedition in 1919, English astronomer Arthur Eddington detected the deflection of starlight by the sun. Just over a decade ago high-precision radio astronomy—in particular, very long baseline interferometry, which links together far-flung radio dishes into a single globe-spanning telescope—saw the minute bending caused by Jupiter.

Since then, radio interferometry has gotten 10 times more precise. So Kopeikin and Fomalont went one step further: to look not only for the bending caused by a static body but also for relativistic effects caused by the motion of that body. Such effects depend on the ratio of the body's velocity to c. For Jupiter, which orbits the sun at 13 kilometers a second, the ratio is about one part in 20,000. That seems awfully small, but the researchers calculated that geometric factors would magnify any effects to detectable levels.

Last September they put their plan into action when Jupiter passed close to the line of sight between Earth and a quasar. The quasar image scooted 1,300 microarcseconds across the sky—with a 50-microarcsecond skew, just as expected from relativistic effects.

So far, so uncontroversial. The fun begins when you ask which relativistic effect was operating. There are oodles of possibilities, and Einstein's notoriously subtle equations do not specify which mathematical term corresponds to which physical effect. Kopeikin and Fomalont contend that the dominant effect was the propagation of gravity. As Jupiter travels, its gravitational force on the ray varies, and the variation takes a little while to travel through space to the ray. To isolate this effect, the scientists constructed an alternative version of relativity, in which c_g could differ from c. They were then able to infer a value for c_g from the data, without presuming it. The two c's turned out to have the same numerical value, with a precision of 20 percent.

But others, notably Clifford M. Will of Washington University, take a different approach to extending relativity and attribute the observed skew to the better-known relativistic effects of time dilation and length contraction. From the vantage point of Earth, Jupiter's moving gravitational field looks slightly flattened, which alters the amount of light deflection that we perceive. This flattening depends on c but not on c_g. The propagation of Jupiter's gravity does play a role, but Will argues that it corresponds to a different (and much smaller) term in the equations. If so, Kopeikin and Fomalont cannot infer a value for c_g.

The disagreement will not be easy to

EINSTEIN VERSUS NEWTON

In Newton's theory of gravity, the speed of gravity (c_g) is infinite; if the sun blew up, Earth's orbit would change instantaneously. But Einstein's special theory of relativity wouldn't look too kindly on that. To preserve the distinction between cause and effect, the speed of light (c) must be the ultimate speed limit. Special relativity also suggests that c_g cannot be less than c: if it were, gravity would behave differently for different observers. Unfortunately, Newton's theory of gravity cannot accommodate a finite c_g without making orbits unstable. The conflict between Newton's theory and special relativity led Einstein to devise an entirely new theory of gravity, general relativity.

SLIM FILMS

resolve. Most researchers lean toward Will's approach, which builds in consistency with other experimental tests. Some go so far as to say that the entire debate is pointless, because there are tests that have higher precision, but others think Kopeikin and Fomalont could be probing something unique. Sorting things out will take more theoretical work as well as direct measurement of gravitational radiation. No mainstream physicist doubts that c_g equals c. But in science, it is not enough to be right. You have to be right for the right reasons.

WILDLIFE

Out of the Woods

MOVING THE GRAY WOLF OFF THE ENDANGERED LIST **BY EMILY HARRISON**

Only six domestic animal species have ever earned their way off the U.S. Endangered Species List. The gray wolf is closing in on becoming the seventh. Although many wolf biologists back the decision, not all wildlife advocates are cheering the pending status change.

In 1974, after a century of aggressive extermination efforts had nearly extinguished gray wolves in the lower 48 states, the Endangered Species Act took effect, and the dwindling species was whisked onto the list. The U.S. Fish and Wildlife Service (FWS) subsequently initiated recovery programs for the gray wolf in three regions, setting population goals for the West, East and Southwest.

With federal protection and reintroduction programs to seed the West with ecologically appropriate wolves from Canada, gray wolf populations burgeoned. Today there are 44 breeding pairs (a total of 664 wolves) in the Western zone, exceeding the target of 30. In the Eastern area, over 3,800 wolves live in the Great Lakes region—almost triple the target number—and a new population of over 600 wolves teems in the states around Minnesota. (In the Southwest, recovery of the Mexican gray wolf is still in its infancy, and the animal's endangered status will remain intact.)

Because the wolf populations have now met their goals for the West and East, the FWS wants to reclassify the wolf from endangered to threatened and delist the species in all states outside its historical range. The FWS fully expects the reclassification proposal to pass this spring and hopes to delist the populations in the Northwest and Midwest in the next year.

Several wildlife groups, however, protest that the proposed status change is premature. The wolf has not returned to the Northeast, where it was formerly an important predator in that ecosystem. They also argue that out West the population is too thin for the wolves to set out from the recovery zones and into their former ranges in the southern Rockies and the Pacific Northwest.

W. PERRY CONWAY *Corbis*

TWO SIDES: Gray wolf supporters disagree about the next steps in the species' recovery.

FWS biologists respond that their job is to ensure that the wolf is no longer in danger of extinction, not to restore the species to every place it could live. "The Endangered Species List is not a tool for other agendas. The act mandates that if a species doesn't need protection anymore, you must remove it," insists Ed Bangs, wolf recovery coordinator for the West. L. David Mech, wolf expert and senior research scientist with the U.S. Geological

WATCHING THE WOLF

Delisting the gray wolf does not mean that it will be left to its own devices. The U.S. Fish and Wildlife Service must monitor the wolf populations in the delisted areas for at least five years and can "emergency relist" the species if necessary. The FWS also requires that all key states within a recovery zone submit biologically acceptable wolf management plans before the species is removed from the list. The plans deal with issues such as population control, compensation for loss of livestock, recreational hunting, and permission to defend property. "There's no requirement for state management plans in the Endangered Species Act," comments Midwest recovery coordinator Ron Refsnider. "We're imposing that because of the wolf's unique situation."

U.S. species removed from the Endangered Species List:

Brown pelican, 1985

American alligator, 1987

Arctic peregrine falcon, 1994

Gray whale, 1994

American peregrine falcon, 1999

Aleutian Canada goose, 2001

Survey, agrees that the gray wolf is no longer at risk of extinction in the lower 48 states. "When recovery goals were planned, certain numbers were set that would signify recovery. I see no evidence that those numbers were too low." Mech believes the genetic diversity and population growth rate in these numbers are sufficient for maintaining viable populations.

Even so, opponents of the change contend that the wolf is unique among recovered species: it is the only one that was deliberately exterminated. Because deep-seated animosities against the wolf still exist, wolves face fiercer threats than other recovered species.

Despite some persisting hostility toward the gray wolf, experts assert that the situation is not the same as it was before 1974. Attitudes have grown more tempered with public education, which Mech expects will continue to serve the animal: "We started off 20 years ago saying, 'Save the wolf.' We've done that. Now the thing to say is, 'Manage the wolf.'"

CLONING

Ma's Eyes, Not Her Ways

CLONES CAN VARY IN BEHAVIORAL—AND PHYSICAL—TRAITS BY CAROL EZZELL

One pig savors a ripe banana, whereas its cloned sister turns up its snout. Another always thrashes its trotters to get away when it is picked up, whereas the others nuzzle into a human embrace. Although clones have been described as identical twins, studies of the behavioral—and even physical—traits of cloned animals are showing that that is not necessarily the case.

Ted Friend and Greg Archer of Texas A&M University created the cloned piglets. They observed as much physical and behavioral variation among the members of two litters of cloned pigs (of four and five individuals, respectively) as among those of two litters of eight pigs bred naturally. Not only did the cloned siblings show distinct food preferences and temperaments, but they also varied in physical characteristics: some had more bristly coats or fewer teeth than others did.

The clones are "just like normal pigs," Friend concludes. "They're not at all like identical twins." Conditions in the uterus could play a role, he speculates. The two cloned litters were borne by different surrogate sows, and the dissimilarities are even more pronounced between the litters.

The poorly understood phenomena called epigenetics also has an effect. Each individual carries two copies of a given gene. In certain instances, one copy is turned off if it is inherited from the mother, whereas other genes have the paternal copy silenced. Researchers are now trying to figure out what determines such selective gene expression and how the pattern is established in the developing embryo.

The pigs are just the latest evidence that clones aren't mere duplicates. The world's first cloned cat, Cc (for "carbon copy"), is no copycat, according to her creators, Duane C. Kraemer and his co-workers, who are also at Texas A&M. Cc is more curious and playful than Rainbow, the cat from which she was cloned. Their coats are also different (although that may have more to do with the way calico coats are inherited—a calico's spots result from the random migration of pigment-carrying cells during development).

Robert P. Lanza of Advanced Cell Technology in Worcester, Mass., says it should come as no surprise that clones are not exact replicas, because among humans, identical twins often have strikingly varying personalities. He and his colleagues have observed the same phenomenon among their herds of cloned cattle. Far from behaving similarly, herds cloned from the same individual develop the usual social hierarchy, with some cows being more skittish and others more bold.

Lanza suggests, though, that cloning might enable scientists to study the importance of genetics in behavior. His group collaborated with a team at Wake Forest University two years ago to attempt to clone an alcoholic monkey named Buttercup. Their objective was to determine whether the cloned offspring of an alcoholic animal would also become addicted to alcohol and to examine the degree to which alcoholism has a genetic basis. The cloning attempt failed, but such studies might one day allow researchers to probe more precisely the influence of genes on behavior.

CLONED CAT Cc gets a nuzzle in the ear from her genetic twin.

PAT SULLIVAN *AP Photo*

IN MEMORIAM: GOOD-BYE, DOLLY

The world's first cloned sheep, Dolly, is destined for permanent display at the National Museum of Scotland—after her date with a taxidermist. Dolly was euthanized in February at the age of six, roughly half the average life span of a normal sheep. She was suffering from a lung infection that usually afflicts newborn cloned animals as well as from chronic arthritis, another seemingly common affliction of clones. Her poor health—and now death—has fueled debates over the safety of cloning and the ethics of using the technology to create human babies.

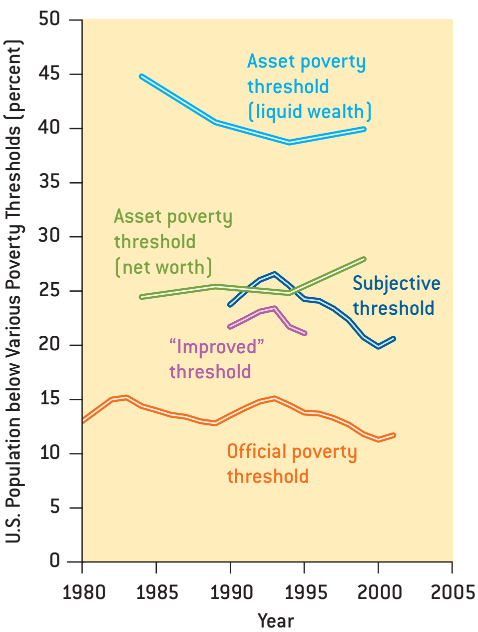
Defining Poverty

OFFICIAL POVERTY STATISTICS MAY BE MISLEADING **BY RODGER DOYLE**

The poverty threshold—the level of income that separates the poor from the not poor—was the brainchild of economist Mollie Orshanksy of the Social Security Administration, who developed it in the early 1960s. Orshanksy did not have the extensive data on the income and consumption habits of Americans needed to fashion a completely satisfactory formula; as a result, it had certain built-in inequities, such as an underestimation of the cost of nonfood items.

After a time other problems became apparent. The formula did not allow for changing demographics, including the substantial rise in the number of working mothers, whose costs for child care were not factored into the formula. Nor did it take into account higher Social Security payroll and other taxes levied on the working poor. Nor did it adjust for geographic variations in the cost of living, such as the two-to-one differential between San Francisco and Houston. The only significant adjustment made was for cost-of-living increases.

To remedy these and other shortcomings, economists tried to devise a better formula. One, constructed by Thesia I. Garner of the Bureau of Labor Statistics and her colleagues, builds on the more extensive statistics now available. The results, depicted by the purple line in the chart, yield an improved threshold, 55 percent higher than the official level.



SOURCES: U.S. Census Bureau; Thesia I. Garner et al.; Asena Caner and Edward N. Wolff; Gallup Organization

Another approach is to base the threshold on purely subjective perceptions. In 1992 the Gallup Organization asked Americans, "What is the smallest amount of money a family ... needs each week to get along in this community?" The average of their answers is the basis of the subjective threshold calculation, indicated by the dark blue line. On average the respondents named a figure 76 percent higher than the official poverty level.

Still another method focuses on assets rather than income. Economists Asena Caner of the Jerome Levy Economics Institute at Bard College and Edward N. Wolff of New York University have calculated several measures of asset poverty. In one, indicated by the green line, they define asset poverty as a net worth insufficient to cover minimal living expenses for three months. In a similar measure, indicated by the light blue line, they define the term as insufficient liquid wealth to cover these expenses. (Liquid wealth is cash and other easily monetized assets.) Under the latter definition, a significant proportion of middle-class people would be considered at risk for poverty. Both the asset poverty lines show a different trend than the income poverty lines, possibly because of changes in savings rates over time.

The official threshold data were highly useful in the 1960s, but now they are outdated and, according to some, greatly understate the problem. Since at least 1995, when a panel of experts under the aegis of the National Research Council recommended new guidelines, a growing consensus has emerged that the official measure is inadequate. Most economists argue that it should be discontinued and replaced by a revised measure or perhaps even several measures, including at least one indicator of asset poverty.

Rodger Doyle can be reached at rdoyle2@adelphia.net

MORE TO EXPLORE

Measuring Poverty: A New Approach. Edited by Constance F. Citro and Robert T. Michael. National Academy Press, 1995.

Experimental Poverty Measurement for the 1990s. Thesia I. Garner et al. in *Monthly Labor Review*, Vol. 121, No. 3; March 1998. (http://stats.bls.gov/opub/ml/mlrhome.htm)

Asset Poverty in the United States, 1984–1999: Evidence from the Panel Study of Income Dynamics. Asena Caner and Edward N. Wolff. Levy Economics Institute, Working Paper No. 356; 2002. (www.levy.org)

United States Poverty Studies and Poverty Measurement: The Past Twenty-Five Years. Howard Glennerster in *Social Services Review*, Vol. 76, No. 1; March 2002.

RODGER DOYLE

DATA POINTS: INVADED NATION

Life in a new country can be a lot easier. Two studies, which examined 473 European plant species and 26 animal species that have invaded the U.S., confirm long-standing thinking that such species tend to have fewer enemies and infections in their new digs. They are therefore better able to survive and to crowd out indigenous flora and fauna. Invasive species are considered the second biggest threat to biodiversity, after habitat destruction.

Percent drop in fungal infections among European plants after invading the U.S.: 84

Percent drop in viral infections: 24

Percent drop in all diseases: 77

Average number of parasites on a species in its indigenous range: 16

Number of parasites that accompany an invader to its new range: 3

Number the invader subsequently picks up: 4

Number of nonindigenous species in the U.S.: at least 30,000

Annual cost of damage done by these species: $123 billion

SOURCES: Nature, *February 6, 2003; economic costs from a January 1999 report by David Pimentel of Cornell University and his colleagues.*

ELECTRONICS

Bits through Ballistics

The hard part of constructing superdense computer chips isn't crowding bits onto a silicon wafer; it's reading each individual magnetic state, 0 or 1. The smaller the bit, the weaker its magnetic field. Now Susan Hua and Harsh Deep Chopra of the State University of New York at Buffalo describe a device whose electrical resistance changes by 100,000 times in a magnetic field—in principle, sensitive enough by far to read terabit-dense bits (1,000 times denser than bits on current chips). The device, a nanoscale nickel "whisker," relies on an effect called ballistic magnetoresistance (BMR). In BMR, an applied magnetic field causes electrons to shoot through a wire with little ricocheting off atoms. Key to the effect were pinch points that result from the whisker's manufacture and seem to make electrons go ballistic. The team reports in the February *Physical Review B* that it can reliably craft such structures. Chopra adds that BMR devices could also make useful biological sensors. *—JR Minkel*

PHYSICS

Skipping for Smarties

SKIP THIS: The physics of a stone's throw.

An inquisitive seven-year-old son has led a physicist to lay out the science behind stone skipping, a.k.a. smutting, skiffing, ricochet, or "ducks and drakes." Lydéric Bocquet of the University of Claude Bernard Lyon in France reduced the problem to its essentials: a thin, flat stone rebounding off a uniform surface of water at a shallow angle, like a water ski skimming over a lake. By formulating equations of motion for the stone, Bocquet found that its initial speed and especially its rate of spin are key to achieving that satisfying string of splashes. Giving the stone an initial spin generates a gyroscopic effect that minimizes tilting after each impact. If the stone goes too vertical, it sinks. According to his equations, the world record for skips, 38, corresponds to a stone flying 12 meters and whirling 14 times a second. The work appears in the February *American Journal of Physics*. *—JR Minkel*

MOLECULAR BIOLOGY

Motoring with RNA

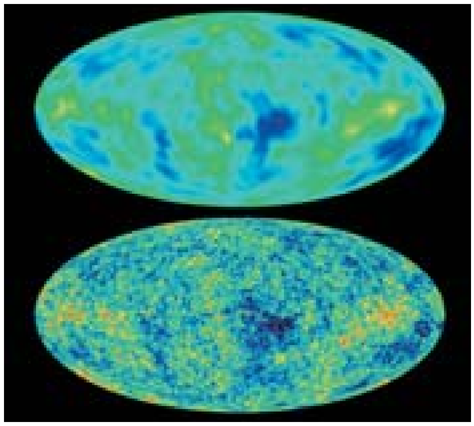
Put RNA and ATP together, and you have more than a spoonful of alphabet soup—you have one of the strongest nanometer-size motors and what may have been a crucial step in the origin of life. For decades, scientists knew that RNA ferried information in cells and could help catalyze reactions. Such a central role suggested that RNA existed as the underpinning for DNA and proteins. Now molecular virologists at Purdue University have found that in the bacteria-killing virus Phi 29, ATP can bind to and fuel a remarkably powerful motor made of six RNA molecules. The virus subsequently relies on this ATP-powered RNA to shuffle its genes. The ability to bind to ATP was once thought to be limited to proteins. The researchers suggest that RNA not only seeded life but also could have directed how ATP was used. They hope to devise machines that fuel themselves with organic molecules. The findings appear in the February *Journal of Biological Chemistry*. *—Charles Choi*

CHRIS VINCENT *Corbis*; ILLUSTRATION BY MATT COLLINS

COSMOLOGY

Unfolding the MAP

The much anticipated Wilkinson Microwave Anisotropy Probe (WMAP), recently named for the late Princeton University astrophysicist David Wilkinson, has observed the cosmic microwave background radiation with about 30 times the resolution of NASA's earlier COBE mission. Besides bearing out prevailing paradigms and nailing down key parameters, WMAP has delved into a whole new level of physics. Polarization of the radiation, never before measured with such precision, provides novel evidence for the theory of cosmological inflation and reveals when intergalactic gas became ionized (presumably by the very first stars). Combined with other data, WMAP finds a relative dearth of small spots in the radiation, a phenomenon known as tilt. Not only does that winnow out certain models of inflation, it could solve a nagging mystery: why the universe has fewer small galaxies than models predict. Finally, WMAP confirms a puzzling discrepancy noted by COBE: a lack of features on the largest angular scales, which might mean the universe is finite in size with a strange topology. The main WMAP paper is available at arXiv.org/abs/astro-ph/0302207. —*George Musser*



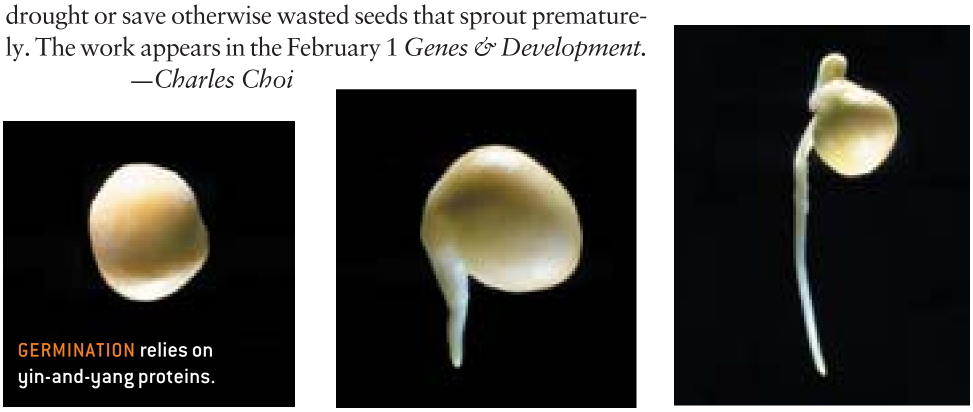
UNIVERSE AT AGE 380,000 YEARS was seen by the COBE satellite a decade ago (*top*) and recently by the sharper-eyed WMAP satellite (*bottom*). Yellows and reds are warmer (hence, denser) regions; blues are cooler (sparser).

PLANT GENETICS

Another Chance at Life

Reliving the teenage years may not be possible, but plants actually can get a second chance at growing up. Newborn sprouts just out from their seeds can suspend growth for up to 30 days if their environs turn risky, such as during Indian summers, in winter or in untimely droughts. Plant molecular biologists at the Rockefeller University investigating the weed *Arabidopsis thaliana* have found a protein that is involved in granting plants this new lease on life. This compound, named AFP, restores arrested development by binding to another protein called ABI5, which the researchers previously discovered stalled the growth. "ABI5 and AFP are kind of like yin and yang," says Luis Lopez-Molina, one of the protein's discoverers. Further studies to locate molecules that promote one protein or the other could help crops resist drought or save otherwise wasted seeds that sprout prematurely. The work appears in the February 1 *Genes & Development*.

—*Charles Choi*



GERMINATION relies on yin-and-yang proteins.

NASA/WMAP SCIENCE TEAM (*top*); J. BURGESS *Photo Researchers, Inc.* (*bottom*)

BRIEF POINTS

- **Food for thought: recent increases in American crop yields may have stemmed from short-term climate changes rather than farm management practices such as crop diversification, pesticide spraying and fertilizer use.**

  Science, February 14, 2003

- **After dropping out in 1998, the U.S. will rejoin the $5-billion International Thermonuclear Experimental Reactor (ITER), an attempt to create self-sustaining fusion by 2014.**

  U.S. Department of Energy press release, January 30, 2003

- **A new kind of solar cell, in which a dye on a metal surface absorbs light and produces electricity, has a surprisingly high efficiency. Its constituent materials are cheaper and more durable than conventional photovoltaic sources, such as silicon.**

  Nature, February 6, 2003

- **Pure oxygen helps to repair wounds that resist healing with sutures, creams and other standard treatments. The oxygen was kept in contact with the skin via an adhesive-edged plastic wrap and applied for 90 minutes a day.**

  Pathophysiology, January 2003

Working Weeds

A German company develops a way to peek into plant metabolism By KATHRYN BROWN

Herbicide chemistry may not sound particularly inspiring. But in 1995, when Richard N. Trethewey heard a talk on how herbicides affect the metabolism of barley, he was struck by an idea. At the time, the chemical company BASF was using gas chromatography to characterize the metabolic effects of potential herbicides. Could this technique, Trethewey wondered, be expanded to profile everyday metabolism in plants?

Like many, Trethewey, then a biochemist at the Max Planck Institute of Molecular Plant Physiology in Golm, Germany, had hopes of engineering plants to grow in different environments, grow new drugs or simply grow better. The notion of metabolism was well understood: like people, plants take in nutrients (sun, water and air) and then metabolize, or transform, those nutrients into lots of other things (metabolites), from vitamins to defense toxins. But how does this chemical choreography play out?

MODEL PLANT: The weed *Arabidopsis thaliana* has helped Metanomics reveal key pathways that could let plants tolerate cold or generate extra oil.

METANOMICS

That was largely a mystery. And although the emerging field of functional genomics had just begun to link genes to their protein products, most techniques stopped short of defining metabolic pathways. In contrast, metabolic profiling—or tallying a plant's metabolites, after altering its genes—seemed capable of capturing information about those pathways more quickly. Done systematically, profiling might identify how genes regulate a metabolic process and screen for those with desirable traits.

After some preliminary experiments with transgenic potatoes, Trethewey was convinced that the method could work. He and five colleagues drafted a business plan, sent it to venture capitalists and biotech companies, and finally struck a deal with BASF, which promised them $26 million over five years to develop the idea into a viable product. In 1998 the team founded the Berlin-based firm Metanomics ("metabolism" combined with "genomics"), which aims to identify key plant metabolism genes—for instance, those that allow plants to tolerate cold or generate extra oil—and then license its novel plant collection and genomics database to breeders and biotech firms. To do so, the company has built a broad technology platform, from genetically modifying plants to growing, testing and cataloguing them in its MetaMap database. The endeavor involves 90 staff members, 50 mass spectrometers, 4,500 square meters of labs, offices, greenhouses and growth chambers, and more than 140,000 plant lines.

At Metanomics, the hunt for major metabolites is straightforward, if exhausting. The company has focused early efforts on the weed *Arabidopsis thaliana*, the first plant to have its genome fully sequenced. By packaging a gene silencer into an agrobacterium, which then infects *Arabidopsis* seeds, scientists are engineering a random, "knockout" group of plant lines, each missing the activity of just one gene or DNA segment usually found in *Arabidopsis*. Using a similar technique, the staff is introducing novel genes from yeast and other organisms to create an "overexpression" group of plant lines, each bearing one additional gene.

After being bred for two generations, modified plants go head to head in tests with unmodified plants, which serve as controls. Researchers feed ground-up plant tissue to mass spectrometers, which use gas or liquid chromatography to dissolve and separate the metabolites inside. Those metabolites then get chopped into ion fragments and sorted by size. After all this is done, scientists know which metabolites each plant had and in what amounts. If a plant leaf brims with a food oil, for instance, Metanomics can detect that change—and backtrack to the responsible gene. So far the company has bred more than 40,000 knockout and 100,000 overexpression plant lines.

Typical of functional genomics, metabolic profiling involves genetically altering organisms and then rapidly screening them for chemical or physical changes. What is unique is the focus: these profilers seek a snapshot of a plant's complex metabolome, or network of metabolic pathways. Like factory inspectors, they want to know precisely how plants churn out chemical products. Metanomics, for instance, hopes to learn which stress-tolerant pathways help plants to survive the cruel outdoors. So the company exposes both genetically modified and unmodified plants to extreme conditions, depriving them of water or light, for example. Scientists note individual plant responses in the MetaMap database—and then use bioinformatics software to look for biochemical patterns in those responses. How do unmodified plants usually respond to chilly temperatures? Which modified plants do better—and how do they differ metabolically? Or, as Trethewey puts it: "Which of *Arabidopsis*'s 27,000 genes are really important for influencing these metabolic networks, these stress functions? That's a clear goal, where we will be able to make substantial progress."

The numbers are substantial. During an average week, Metanomics generates about 150 cloned and transformed plant genes, 10,000 chromatographic and mass-spectrometric measurements, and 150 gigabytes of data. Alone, each *Arabidopsis* cell houses an estimated 15,000 to 30,000 chemicals. So perhaps it's not surprising that Metanomics CEO Arno J. Krotzky says that

During an average week, Metanomics generates about 150 cloned and transformed plant genes, 10,000 chromatographic and mass-spectrometric measurements, and 150 gigabytes of data.

the company's biggest hurdle has been organization. "It took us a while to adjust this really large technology platform," says Krotzky, who compares the data juggling with that done by Celera, which sequenced the human genome. "We're working with thousands of living organisms. To control plant variability at a level that shows minor changes in metabolic networks is a significant scientific challenge."

Mindful of the stumbles of other biotech companies, Krotzky adds, Metanomics has worked quietly for four years, completing internal proof-of-concept tests before making any public scientific claims. He says that the firm has identified the functions of roughly 300 novel plant genes and intends to publish early results soon. The MetaMap database will also become available to customers later this year. Metanomics is not disclosing any precise finds, although the discoveries include some stress-tolerance and amino acid synthesis genes.

Richard A. Dixon, director of plant biology at the Samuel Roberts Noble Foundation, a nonprofit organization in Ardmore, Okla., that works to improve plant productivity, believes that Metanomics is on the right track. "They have a critical mass of well-trained scientists, state-of-the-art instruments, and collaborations with important academic groups," Dixon says. Although other bioinformatics companies exist, Metanomics joins North Carolina–based Paradigm Genetics in an elite group of businesses that do both extensive plant metabolism experiments and software development. "In metabolic profiling, I think we are currently the reference stick other companies must measure up to," Krotzky remarks.

Metanomics plans to branch into commercial crops and plant products with the potential to produce drugs. It may also begin testing the safety of genetically modified plants for biotech clients. After all, Krotzky notes, Metanomics already has the technology to compare conventional and modified plants at the metabolic level. Along the way, he'd like the company's scientists to uncover the functions behind all the estimated 27,000 genes in *Arabidopsis*. But even for Metanomics, he concedes, that is a milestone that won't be reached today or tomorrow. SA

Razing the Tollbooths

A call for restricting patents on basic biomedical research By GARY STIX

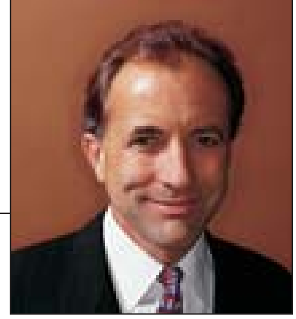
The Bayh-Dole Act, a 1980 law intended to prod the commercialization of government-supported research, gave universities a major role in ushering in the new era of biotechnology. The law fulfilled legislators' most ambitious expectations by encouraging the patenting of academic research—and the exclusive licensing of those patents to industry. In 1979 universities received a mere 264 patents—a number that in 2000 rose to 3,764, about half of which went to biomedical discoveries. The 14-fold increase far outpaced the overall growth in patents during that period. A few voices in the intellectual-property community have now charged that Bayh-Dole has gone too far. Patents, they claim, have been granted on the fruits of biomedical research that should remain in the public domain. In recent co-authored articles, Arti K. Rai of the University of Pennsylvania and Rebecca S. Eisenberg of the University of Michigan at Ann Arbor have proposed reform of the law, contending that development of new biopharmaceuticals and related technologies has been hindered by extending patent coverage beyond actual products to basic research findings. DNA sequences, protein structures and disease pathways should, in many cases, serve as a general knowledge base that can be used freely by everyone.

JENNIFER KANE

Rai and Eisenberg cite the case of a patent obtained by teams at Harvard University, the Massachusetts Institute of Technology and the Whitehead Institute for Biomedical Research in Cambridge, Mass. It covers methods of treating disease by regulating cell-signaling activity involving nuclear factor kappa B (NF-κB), which controls genes for processes ranging from cell proliferation to inflammation in various maladies. Those institutions and Ariad Pharmaceuticals (also in Cambridge), the exclusive licensee of the patent, are now suing Eli Lilly, claiming that two of its drugs—one for osteoporosis, one for sepsis—infringe the patent. Ariad has contacted more than 50 other companies that are researching or commercializing drugs that work through this pathway, asking them for licensing fees and royalties. The broad-based patent does not protect specific drugs. Instead it has become a tollbooth for commercial drug research and development on the NF-κB pathway. "In this case, as in many others, upstream [precommercial] patents issued to academic institutions serve as a tax on innovation, diluting rather than fortifying incentives for product development," the authors wrote in the winter-spring issue of *Law and Contemporary Problems*. (Their other article on the Bayh-Dole Act appeared in the January-February issue of *American Scientist*.)

Rai and Eisenberg suggest that the law should be altered to make it easier for the government—in particular, the National Institutes of Health—to specify that such upstream research remain public and not be subject to patents. They also recommend facilitating the government's ability to mandate the nonexclusive licensing of a patent at reasonable rates. Both actions are permitted under the current law but have almost never been exercised; the law makes it cumbersome to do so.

Fiddling with Bayh-Dole does bear risks. For instance, an executive-branch agency such as the NIH could be subject to political pressure in barring patents: an administration opposed to using embryos in scientific investigations might order an agency to withhold patents on such research. But university technology-transfer offices, Rai and Eisenberg contend, cannot be entrusted to make decisions about when to forgo patenting, given that a big part of their mission is to bring in licensing revenues. So more leverage is needed to ensure that basic biomedical research remains open to all. SA



I, Clone

The Three Laws of Cloning will protect clones and advance science By MICHAEL SHERMER

In his 1950 science-fiction novel *I, Robot,* Isaac Asimov presented the Three Laws of Robotics: "1. A robot may not injure a human being, or, through inaction, allow a human being to come to harm. 2. A robot must obey the orders given it by human beings except where such orders would conflict with the First Law. 3. A robot must protect its own existence as long as such protection does not conflict with the First or Second Law."

The irrational fears people express today about cloning parallel those surrounding robotics half a century ago. So I would like to propose Three Laws of Cloning that also clarify three misunderstandings: 1. A human clone is a human being no less unique in his or her personhood than an identical twin. 2. A human clone has all the rights and privileges that accompany this legal and moral status. 3. A human clone is to be accorded the dignity and respect due any member of our species.

Clone all you like—you'll never produce another you.

Although such simplifications risk erasing the rich nuances found in ethical debates over pioneering research, they do aid in attenuating risible fears often associated with such advances. It appears that the Raelians have not succeeded in Xeroxing themselves, but it is clear that someone, somewhere, sometime soon is going to generate a human clone. And once one team has succeeded, it will be Katy bar the door for others to bring on the clones.

If cloning produces genetic monstrosities that render it impractical as another form of fertility enhancement, then it will not be necessary to ban it, because no one will use it. If cloning does work, however, there is no reason to forbid it, because the three common reasons given for implementing restrictions are myths. I call them the Identical Personhood Myth, the Playing God Myth, and the Human Rights and Dignity Myth.

The Identical Personhood Myth is well represented by activist Jeremy Rifkin: "It's a horrendous crime to make a Xerox of someone. You're putting a human into a genetic straitjacket." *Baloney*. He and fellow cloning critics have the argument bass ackward. As environmental determinists, they should be arguing: "Clone all you like—you'll never produce another you, because environment matters as much as heredity." The best scientific evidence to date indicates that roughly half the variance among us is accounted for by genetics and the rest by environment. It is impossible to duplicate the near-infinite number of permutations that come into play during the development of each individual, so cloning is no threat to unique personhood.

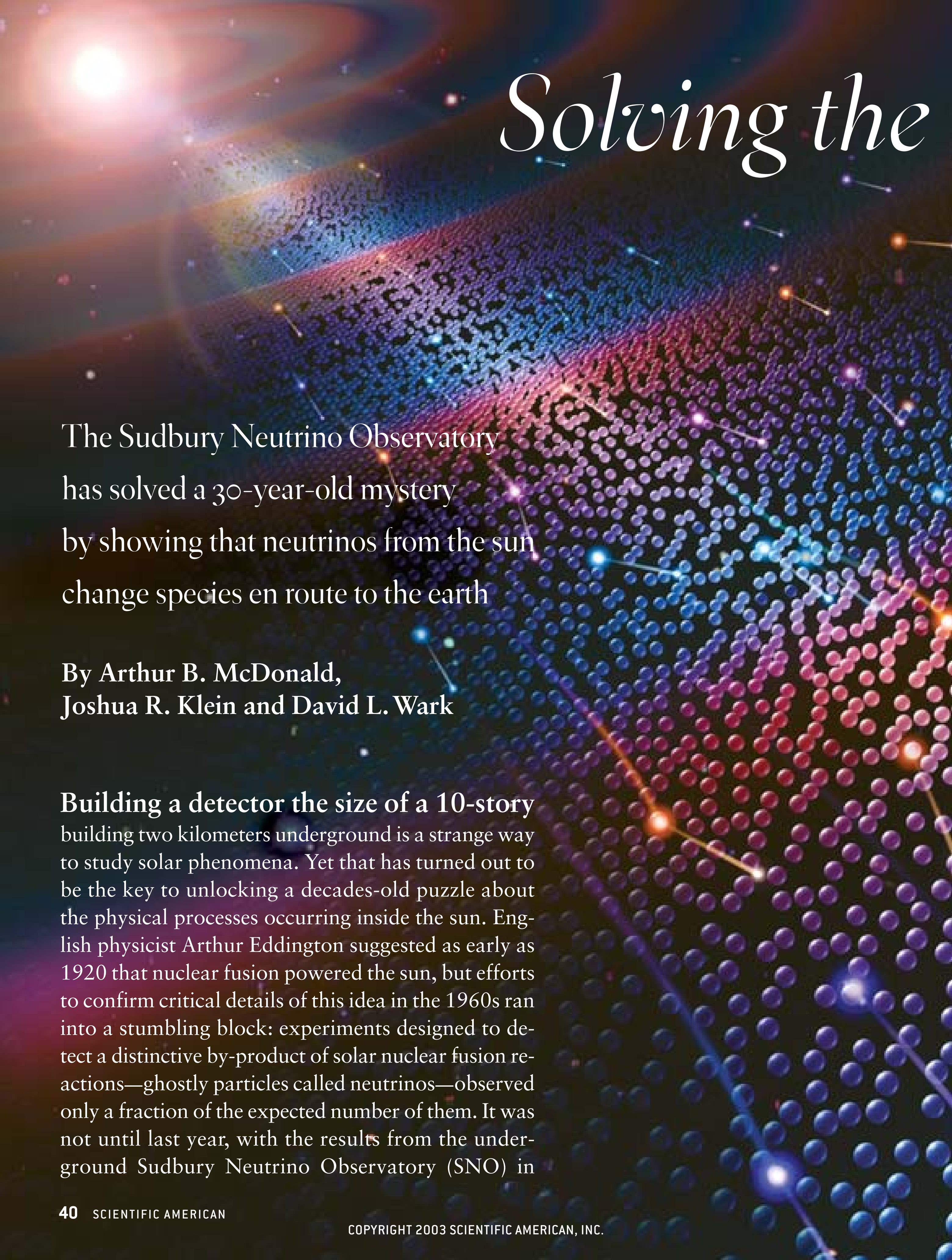
The Playing God Myth has numerous promoters, among the latest being Stanley M. Hauerwas, a professor of theological ethics at Duke University: "The very attempt to clone a human being is evil. The assumption that we must do what we can do is fueled by the Promethean desire to be our own creators." In support of this myth, he is not alone. A 1997 *Time*/CNN poll revealed that 74 percent of 1,005 Americans answered "yes" to the question "Is it against God's will to clone human beings?" *Balderdash*. Cloning may seem to be "playing God" only because it is unfamiliar. Consider earlier examples of once "God-like" fertility technologies that are now cheerfully embraced because we have become accustomed to them, such as in vitro fertilization and embryo transfer.

The Human Rights and Dignity Myth is embodied in the Roman Catholic Church's official statement against cloning, based on the belief that it denies "the dignity of human procreation and of the conjugal union," as well as in a Sunni Muslim cleric's demand that "science must be regulated by firm laws to preserve humanity and its dignity." *Bunkum*. Clones will be no more alike than twins raised in separate environments, and no one is suggesting that twins do not have rights or dignity or that they should be banned.

Instead of restricting or preventing the technology, I propose that we adopt the Three Laws of Cloning, the principles of which are already incorporated in the laws and language of the U.S. Constitution, and allow science to run its course. The soul of science is found in courageous thought and creative experiment, not in restrictive fear and prohibitions. For science to progress, it must be given the opportunity to succeed or fail. Let's run the cloning experiment and see what happens. SA

Michael Shermer is publisher of Skeptic *magazine (www.skeptic.com) and general editor of* The Skeptic Encyclopedia of Pseudoscience.

BRAD HINES



Solving the

The Sudbury Neutrino Observatory has solved a 30-year-old mystery by showing that neutrinos from the sun change species en route to the earth

By Arthur B. McDonald, Joshua R. Klein and David L. Wark

Building a detector the size of a 10-story building two kilometers underground is a strange way to study solar phenomena. Yet that has turned out to be the key to unlocking a decades-old puzzle about the physical processes occurring inside the sun. English physicist Arthur Eddington suggested as early as 1920 that nuclear fusion powered the sun, but efforts to confirm critical details of this idea in the 1960s ran into a stumbling block: experiments designed to detect a distinctive by-product of solar nuclear fusion reactions—ghostly particles called neutrinos—observed only a fraction of the expected number of them. It was not until last year, with the results from the underground Sudbury Neutrino Observatory (SNO) in

Solar Neutrino Problem

Ontario, that physicists resolved this conundrum and thereby fully confirmed Eddington's proposal.

Like all underground experiments designed to study the sun, SNO's primary goal is to detect neutrinos, which are produced in great numbers in the solar core. But unlike most of the other experiments built over the previous three decades, SNO detects solar neutrinos using heavy water, in which a neutron has been added to each of the water molecules' hydrogen atoms (making deuterium). The additional neutrons allow SNO to observe solar neutrinos in a way never done before, by counting all three types, or "flavors," of neutrino equally. Using this ability, SNO has demonstrated that the deficit of solar neutrinos seen by earlier experiments resulted not from poor measurements or a misunderstanding of the sun but from a newly discovered property of the neutrinos themselves.

Ironically, the confirmation of our best theory of the sun exposes the first flaw in the Standard Model of particle physics—our best theory of how the most fundamental constituents of matter behave. We now understand the workings of the sun better than we do the workings of the microscopic universe.

The Problem

THE FIRST SOLAR NEUTRINO EXPERIMENT, conducted in the mid-1960s by Raymond Davis, Jr., of the University of Pennsylvania and his co-workers, was intended to be a triumphant confirmation of the fusion theory of solar power generation and the start of a new field in which neutrinos could be used to learn more about the sun. Davis's experiment, located in the Homestake gold mine near Lead, S.D., detected neutrinos by a radiochemical technique. The detector contained 615 metric tons of liquid tetrachloroethylene, or dry-cleaning fluid, and neutrinos transformed atoms of chlorine in this fluid into atoms of argon. But rather than seeing one atom of argon created each day, as theory predicted, Davis observed just one every 2.5 days. (In 2002 Davis shared the Nobel Prize with Masatoshi Koshiba of the University of Tokyo for pioneering work in neutrino physics.) Thirty years of experiments following Davis's all found similar results despite using a variety of different techniques. The number of neutrinos arriving from the sun was always significantly less than the predicted total, in some cases as low as one third, in others as high as three fifths, depending on the energies of the neutrinos studied. With no understanding of why the predictions and the measurements were so different, physicists had to put on hold the original goal of studying the solar core by observing neutrinos.

While experimenters continued to run their neutrino experiments, theorists improved the models used to predict the rate of solar neutrino production. Those theoretical models are complex, but they make only a few assumptions—that the sun is powered by nuclear reactions that change the element abundances, that this power creates an outward pressure that is balanced by the inward pull of gravity, and that energy is transported by photons and convection. The solar models continued to predict neutrino fluxes that exceeded the measurements, but other projections they made, such as the spectrum of helioseismologic vibrations seen on the solar surface, agreed very well with observations.

The mysterious difference between the predictions and the measurements became known as the solar neutrino problem. Although many physicists still believed that inherent difficulties in detecting neutrinos and calculating their production rate in the sun were somehow the cause of the discrepancy, a third alternative became widely favored despite its somewhat revolutionary implications. The Standard Model of particle physics holds that there are three completely distinct, massless flavors of neutrinos: the electron-neutrino, muon-neutrino and tau-neutrino. The fusion reactions in the center of the sun can produce only electron-neutrinos, and experiments like Davis's were designed to look exclusively for this one flavor—at solar neutrino energies, only electron-neutrinos can convert chlorine atoms to argon. But if the Standard Model were incomplete, and the neutrino flavors were not distinct but instead mixed in some way, then an electron-neutrino from the sun might transform into one of the other flavors and thus escape detection.

The most favored mechanism for a change in neutrino flavor is neutrino oscillation [*see illustration on page 44*], which requires that the neutrino flavors (electron-, muon- and tau-neutrinos) are made up of mixtures of neutrino states (denoted as 1, 2 and 3) that have different masses. An electron-neutrino could then be a mixture of states 1 and 2, and a muon-neutrino could be a different mixture of the same two states. Theory predicts that as they travel from the sun to the earth, such mixed neutrinos will oscillate between one flavor and another.

Particularly strong evidence of neutrino oscillation was provided by the Super-Kamiokande collaboration in 1998, which found that muon-neutrinos produced in the upper atmosphere by cosmic rays were disappearing with a probability that depended on the distance they traveled. This disappearance is explained extremely well by neutrino oscillations, in this case muon-neutrinos that are probably turning into tau-neutrinos. The former are easily detected by Super-Kamiokande at cosmic-ray energies, but the latter mostly evade detection [see "De-

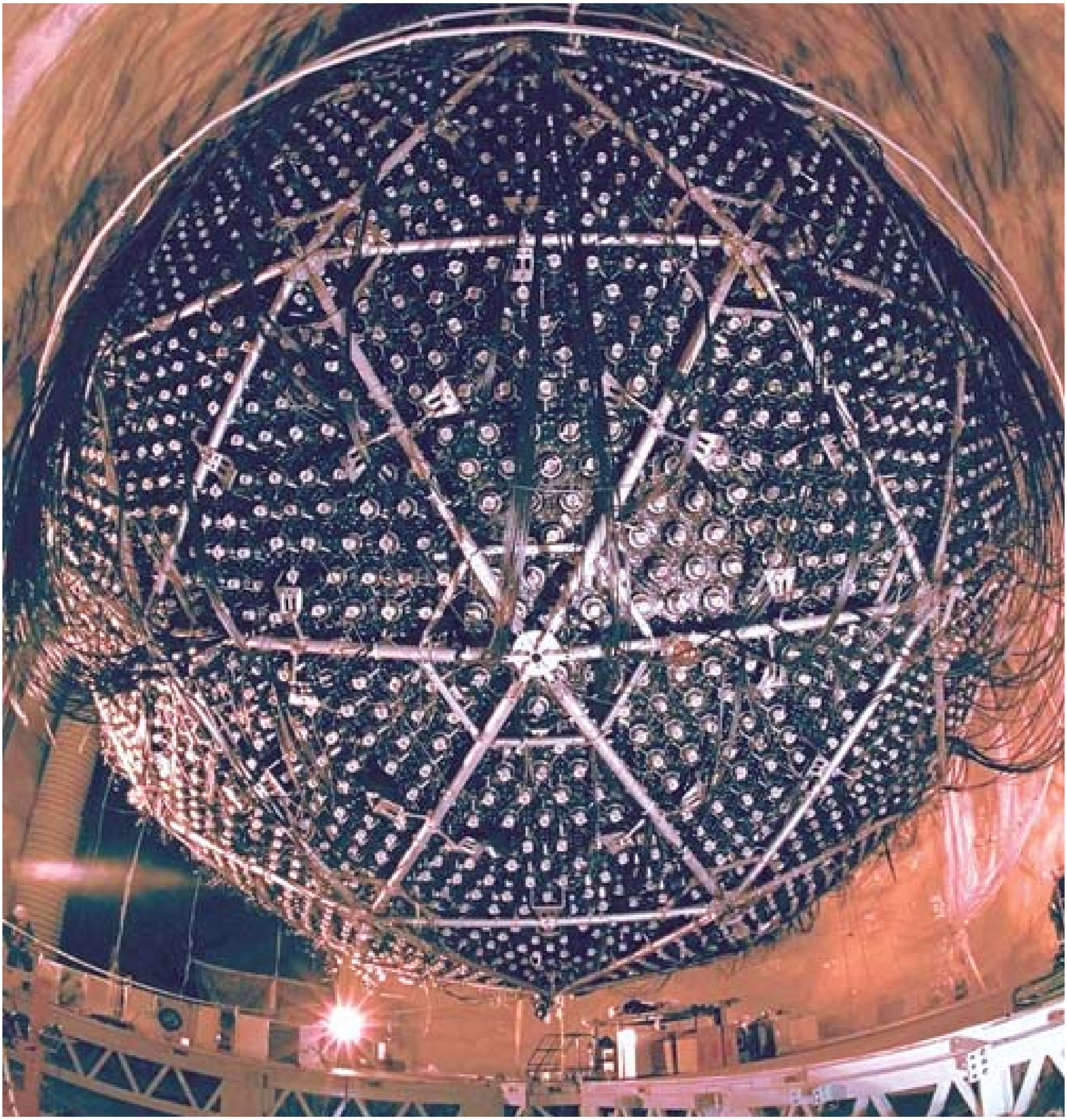
Overview/*Oscillating Neutrinos*

- Since the 1960s, underground experiments have been detecting far fewer electron-neutrinos from the sun than theory predicts. The mystery came to be known as the solar neutrino problem.
- In 2002 the Sudbury Neutrino Observatory (SNO) resolved the solar neutrino problem by determining that many of the electron-neutrinos produced inside the sun change to other flavors of neutrinos before reaching the earth, causing them to go undetected by past experiments.
- SNO's result confirms that we understand how the sun is powered and implies that neutrinos, long thought to be massless, have masses. The Standard Model of particle physics, which is otherwise extraordinarily successful, must be modified to accommodate this change.

SLIM FILMS (*preceding pages*)

tecting Massive Neutrinos," by Edward Kearns, Takaaki Kajita and Yoji Totsuka; Scientific American, August 1999].

A similar process could explain the solar neutrino deficit. In one scenario, the neutrinos would oscillate during their eight-minute journey through the vacuum of space from the sun to the earth. In another model, the oscillation is enhanced during the first two seconds of travel through the sun itself, an effect caused by the different ways in which each neutrino flavor interacts with matter. Each scenario requires its own specific range of neutrino parameters—mass differences and the amount of intrinsic mixing of the flavors. Despite the evidence from Super-Kamiokande and other experiments, however, it remained possible



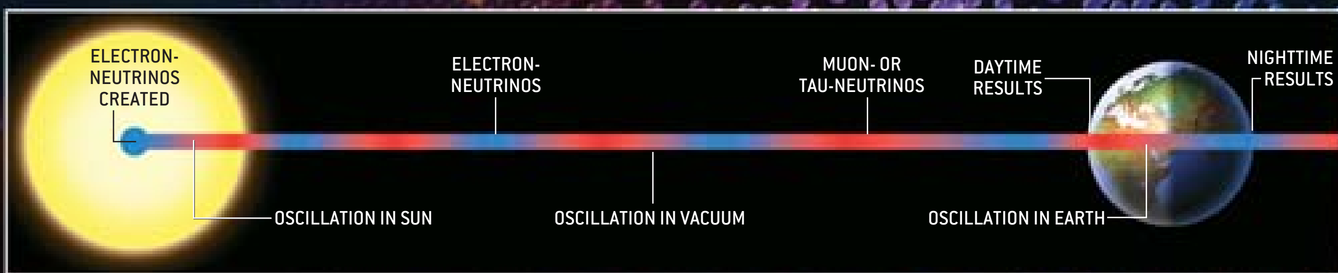
PHOTOMULTIPLIER TUBES—more than 9,500 of them—on a geodesic sphere 18 meters in diameter act as the eyes of the Sudbury Neutrino Observatory. The tubes surround and monitor a 12-meter-diameter acrylic sphere that contains 1,000 tons of heavy water. Each tube can detect a single photon of light. The entire assembly is suspended in ordinary water. All the materials that make up the detector must be extraordinarily free of natural traces of radioactive elements to avoid overwhelming the tubes with false solar neutrino counts.

COURTESY OF SUDBURY NEUTRINO OBSERVATORY (*photograph*); SLIM FILMS (*next two pages*)

DETECTING FICKLE NEUTRINOS

HOW NEUTRINOS OSCILLATE

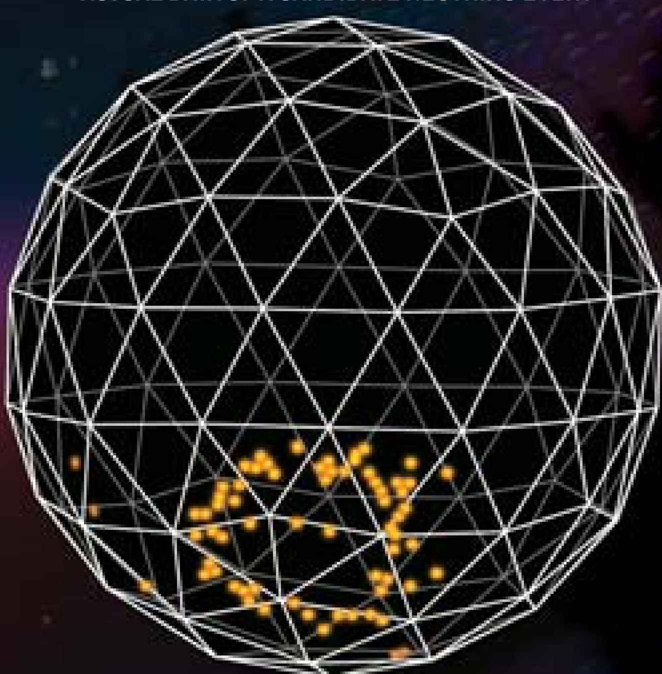
An electron-neutrino (*left*) is actually a superposition of a type 1 and a type 2 neutrino with their quantum waves in phase. Because the type 1 and type 2 waves have different wavelengths, after traveling a distance they go out of phase, making a muon- or a tau-neutrino (*middle*). With further travel the neutrino oscillates back to being an electron-neutrino (*right*).



WHERE NEUTRINOS OSCILLATE

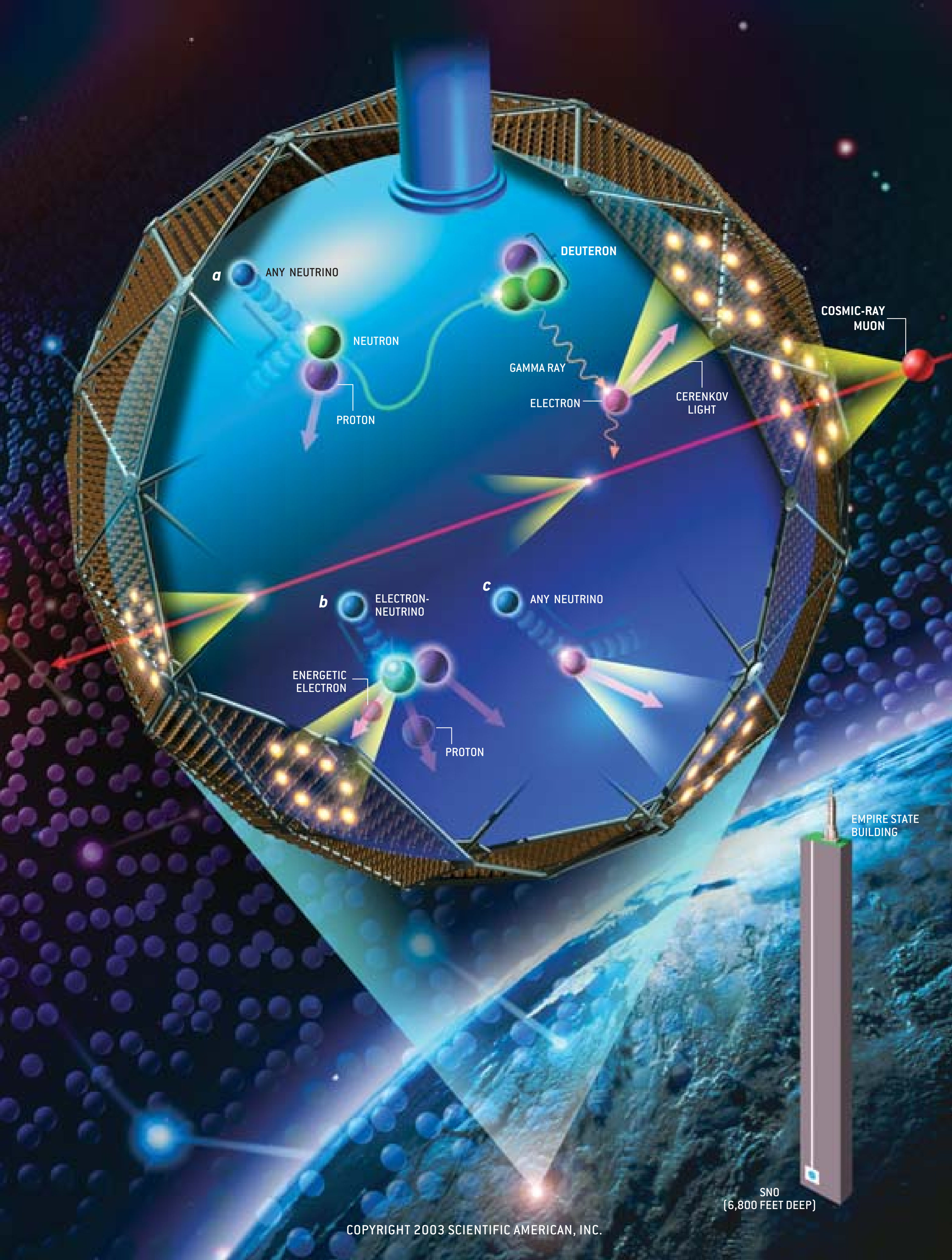
The electron-neutrinos produced at the center of the sun may oscillate while they are still inside the sun or after they emerge on their eight-minute journey to the earth. Which oscillation occurs depends on details such as the mass differences and the intrinsic degree of mixing of type 1 and 2 neutrinos. Extra oscillation may also occur inside the earth, which manifests as a difference between daytime and nighttime results.

ACTUAL DATA OF A CANDIDATE NEUTRINO EVENT



HOW SNO DETECTS NEUTRINOS

The Sudbury Neutrino Observatory, or SNO (*opposite page*), detects a neutrino by seeing a characteristic ring of Cerenkov light emitted by a high-speed electron. The neutrino produces the energetic electron in SNO's heavy water (*large blue sphere*) in one of three ways. In deuteron breakup (*a*), the neutrino (*blue*) splits a deuterium nucleus into its component proton (*purple*) and neutron (*green*). The neutron eventually combines with another deuteron, releasing a gamma ray (*wavy line*), which in turn knocks free an electron (*pink*) whose Cerenkov light (*yellow*) is detected. In neutrino absorption (*b*) a neutron absorbs the neutrino and is thereby turned into a proton and an energetic electron. Only electron-neutrinos can be absorbed in this way. Less often the neutrino may collide directly with an electron (*c*). Cosmic-ray muons (*red*) are distinguished from neutrinos by the amount of Cerenkov light they produce and where they produce it—outside the detector as well as inside. The number of muons is reduced to manageable levels by positioning the detector two kilometers underground.


a
ANY NEUTRINO
NEUTRON
PROTON
DEUTERON
GAMMA RAY
ELECTRON
CERENKOV LIGHT
COSMIC-RAY MUON
b
ELECTRON-NEUTRINO
ENERGETIC ELECTRON
PROTON
c
ANY NEUTRINO
EMPIRE STATE BUILDING
SNO (6,800 FEET DEEP)

that neutrinos were disappearing by some process other than oscillation. Until 2001 scientists had no *direct* evidence of solar neutrino oscillation, in which the transformed solar neutrinos themselves were detected.

The Observatory

THE SUDBURY NEUTRINO OBSERVATORY was designed to search for this direct evidence, by detecting neutrinos using several different interactions with its 1,000 tons of heavy water. One of these reactions exclusively counts electron-neutrinos; the others count all flavors without distinguishing among them. If the solar neutrinos arriving at the earth consisted only of electron-neutrinos—and therefore no flavor transformation was occurring—then the count of neutrinos of all flavors would be the same as the count of electron-neutrinos alone. On the other hand, if the count of all flavors was far in excess of the count of the electron-neutrinos, that would prove that neutrinos from the sun were changing flavor.

The key to SNO's ability to count both electron-neutrinos alone and all flavors is the heavy water's deuterium nuclei, also called deuterons. The neutron in a deuteron produces two separate neutrino reactions: neutrino absorption, in which an electron-neutrino is absorbed by a neutron and an electron is created, and deuteron breakup, in which a deuterium nucleus is broken apart and the neutron liberated. Only electron-neutrinos can undergo neutrino absorption, but neutrinos of any flavor can break up deuterons. A third reaction detected by SNO, the scattering of electrons by neutrinos, can also be used to count neutrinos other than electron-neutrinos but is much less sensitive to muon- and tau-neutrinos than the deuteron breakup reaction [*see illustration on preceding page*].

SNO was not the first experiment to use heavy water. In the 1960s T. J. Jenkins and F. W. Dix of Case Western Reserve University used heavy water in a very early attempt to observe neutrinos from the sun. They used about 2,000 liters (two tons) of heavy water aboveground, but the signs of solar neutrinos were swamped by the effects of cosmic rays. In 1984 Herb Chen of the University of California at Irvine proposed bringing 1,000 tons of heavy water from Canada's CANDU nuclear reactor to the bottom of INCO Ltd.'s Creighton nickel mine in Sudbury—a location that was deep enough to enable a clear measurement of both neutrino absorption and deuteron breakup for solar neutrinos.

Chen's proposal led to the establishment of the SNO scientific collaboration and ultimately to the creation of the SNO detector. The 1,000 tons of heavy water are held in a 12-meter-diameter transparent acrylic vessel. The heavy water is viewed by more than 9,500 photomultiplier tubes held on an 18-meter-diameter geodesic sphere [*see illustration on page 43*]. Each tube is capable of detecting a single photon of light. The entire structure is submerged in ultrapure ordinary water filling a cavity carved out of the rock two kilometers below the surface of the earth.

SNO's Measurement

SOLAR NEUTRINOS CAN BE OBSERVED deep underground because of the extreme weakness of their interaction with matter. During the day, neutrinos easily travel down to SNO through two kilometers of rock, and at night they are almost equally unaffected by the thousands of kilometers that they travel up through the earth. Such feeble coupling makes them interesting from the perspective of solar astrophysics. Most of the energy created in the center of the sun takes millions of years to reach the solar surface and leave as sunlight. Neutrinos, in contrast, emerge after two seconds, coming to us directly from the point where solar power is created.

With neither the whole sun nor the entire earth able to impede the passage of neutrinos, capturing them with a detector weighing just 1,000 tons poses something of a challenge. But although the vast majority of neutrinos that enter SNO pass through it, on very rare occasions, one will—by chance alone—collide with an electron or an atomic nucleus and deposit enough energy to be observed. With enough neutrinos, even the rarity of these interactions can be overcome. Luckily, the sun's neutrino output is enormous—five million high-energy solar neutrinos pass through every square centimeter of the earth every second—which leads to about 10 neutrino events, or interactions, in SNO's 1,000 tons of heavy water every day. The three types of neutrino reaction that occur in SNO all generate energetic electrons, which are detectable through their production of Cerenkov light—a cone of light emitted like a shock wave by the fast-moving particle.

This small number of neutrino events, however, has to be

EIGHT DECADES OF THE SUN AND NEUTRINOS

IT HAS TAKEN MOST OF A CENTURY to verify fully that we understand how the sun generates its power. Along the way, neutrinos have gone from speculative hypothesis to key experimental tool. Their oscillations point to fundamental new physics to be discovered in the decades to come.

1920 — 1940

1920 **Arthur Eddington** proposes that the sun is powered by nuclear fusion converting hydrogen atoms into helium

1930 **Wolfgang Pauli** rescues conservation of energy by hypothesizing an unseen particle, the neutrino, that carries away energy from some radioactive decays

1938 **Hans Bethe** analyzes the basic nuclear processes that could power the sun and accurately estimates the sun's central temperature

1956 **Frederick Reines** and **Clyde Cowan** first detect the neutrino using the Savannah River nuclear reactor distinguished from flashes of Cerenkov light caused by other particles. In particular, cosmic-ray muons are created continually in the upper atmosphere, and when they enter the detector they can produce enough Cerenkov light to illuminate every photomultiplier tube. The intervening kilometers of rock between the earth's surface and SNO reduce the deluge of cosmic-ray muons to a mere trickle of just three an hour. And although three muons an hour is a far greater rate than the 10 neutrino interactions a day, these muons are easy to distinguish from neutrino events by the Cerenkov light they produce in the ordinary water outside the detector.

A far more sinister source of false neutrino counts is the intrinsic radioactivity in the detector materials themselves. Everything inside the detector—from the heavy water itself to the acrylic vessel that holds it to the glass and steel of the photomultiplier tubes and support structure—has trace amounts of naturally occurring radioactive elements. Similarly, the air in the mine contains radioactive radon gas. Every time a nucleus in these radioactive elements decays inside the SNO detector, it can release an energetic electron or gamma ray and ultimately produce Cerenkov light that mimics the signal of a neutrino. The water and the other materials used in SNO are purified to remove the bulk of the radioactive contaminants (or were chosen to be naturally pure), but even parts per billion are enough to overwhelm the true neutrino signal with false counts.

The task before SNO is therefore very complex—it must count neutrino events, determine how many are caused by each of the three reactions, and estimate how many of the apparent neutrinos are caused by something else, such as radioactive contamination. Errors as small as a few percent in any of the steps of analysis would render meaningless SNO's comparison of the electron-neutrino flux to the total neutrino flux. Over the 306 days of running, from November 1999 to May 2001, SNO recorded nearly half a billion events. By the time the data reduction was complete, only 2,928 of these remained as candidate neutrino events.

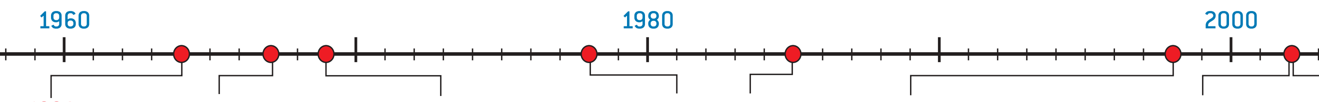
SNO cannot uniquely determine whether a given candidate neutrino event was the result of a particular reaction. Typically an event like the one shown on page 44 could equally well be the result of deuteron breakup as neutrino absorption. Fortunately, differences between the reactions show up when we examine many events. For example, deuteron breakup, the splitting of a deuterium nucleus in the heavy water, always leads to a gamma ray of the same energy, whereas the electrons produced by neutrino absorption and electron scattering have a broad spectrum of energies. Similarly, electron scattering produces electrons that travel away from the sun, whereas the Cerenkov light from deuteron breakup can point in any direction. Finally, the locations where the reactions occur differ as well—electron scattering, for instance, occurs as easily in the outer layer of light water as in the heavy water; the other reactions do not. With an understanding of those details, SNO researchers can statistically determine how many of the observed events to assign to each reaction.

Five million high-energy **solar neutrinos** pass through every square centimeter of your body every second.

Such an understanding is the result of measurements that were complete nuclear physics experiments in their own right: to determine how to measure energy using Cerenkov light, sources of radioactivity with known energies were inserted inside the detector. To measure how the Cerenkov light travels through and reflects off the various media in the detector (the water, the acrylic, the photomultiplier tubes), a variable wavelength laser light source was used. The effects of radioactive contamination were assessed by similar experiments, including radioassays of the water using new techniques designed specifically for SNO.

For the final SNO data set, after statistical analysis, 576 events were assigned to deuteron breakup, 1,967 events to neutrino absorption and 263 to electron scattering. Radioactivity and other backgrounds caused the remaining 122. From these numbers of events, we must calculate how many actual neutrinos must be passing through SNO, based on the tiny probabilities that any particular neutrino will break up a deuteron, be absorbed or scatter an electron. The upshot of all the calcu-

BRYAN CHRISTIE DESIGN



1964 **John Bahcall** predicts the neutrino flux expected to be observed from the sun

1967 **Raymond Davis, Jr.**, first measures neutrinos from the sun, using 600 tons of dry-cleaning fluid in a mine in Lead, S.D.

1969 **Vladimir Gribov** and **Bruno Pontecorvo** propose that neutrino oscillations explain the anomalously small number of neutrinos detected

1978 and 1985 **Stanislav Mikheyev, Alexei Smirnov** and **Lincoln Wolfenstein** posit that matter can enhance neutrino oscillations

1998 **Super-Kamiokande** assembles evidence of neutrino oscillation in cosmic-ray neutrinos

2002 **SNO** confirms that electron-neutrinos from the sun are oscillating to another flavor, fully resolving the solar neutrino problem

2002 **KamLAND** experiment detects oscillation of antineutrinos emitted from nuclear reactors

Some Other Neutrino Experiments

HOMESTAKE: Solar neutrino detector located in the Homestake gold mine in Lead, S.D. The original chlorine experiment, started in 1966, used 600 tons of dry-cleaning fluid. Supplemented in 1996 by a radiochemical sodium iodide experiment using 100 tons of iodine.

KAMIOKA: Houses Super-Kamiokande, a 50,000-ton light-water detector studying cosmic-ray and solar neutrinos, as well as muon-neutrinos beamed from the KEK facility 250 kilometers away ("K2K" experiment). Also houses KamLAND, a smaller detector (1,000 tons of liquid scintillator, which emits light when a charged particle passes through) that counts anti-electron-neutrinos emitted by all the nuclear reactors nearby in South Korea and Japan.

SAGE (Russian-American Gallium Solar Neutrino Experiment): Located at Baksan in the Caucasus Mountains in Russia. Uses 50 tons of gallium, which is capable of detecting the low-energy neutrinos produced by proton-proton fusion in the sun.

GRAN SASSO: The world's largest underground laboratory, accessed via a highway tunnel, located under the Gran Sasso Mountains about 150 kilometers east of Rome. Solar neutrino experiments include Gallex/GNO, which began in 1991 and uses 30 tons of gallium (as aqueous gallium trichloride), and Borexino, a sphere of 300 tons of scintillator viewed by 2,200 photomultipliers, scheduled for completion this year.

MINIBOONE (Booster Neutrino Experiment): Located at Fermilab in Illinois. Beams of muon-neutrinos and anti-muon-neutrinos travel through 500 meters of earth to be detected in an 800-ton tank of mineral oil. Endeavoring to test a controversial result reported by the LSND experiment at Los Alamos National Lab in 1995. Began collecting data in September 2002.

MINOS: Will beam neutrinos from Fermilab to the Soudan detector, 735 kilometers away in Minnesota. Detector is 5,400 tons of iron laced with plastic particle detectors. Projected to begin taking data in 2005.

lations is that the observed 1,967 neutrino absorption events represent 1.75 million electron-neutrinos passing through each square centimeter of the SNO detector every second. That is only 35 percent of the neutrino flux predicted by solar models. SNO thus first confirms what other solar neutrino experiments have seen—that the number of electron-neutrinos arriving from the sun is far smaller than solar models predict.

The critical question, however, is whether the number of electron-neutrinos arriving from the sun is significantly smaller than the number of neutrinos of all flavors. Indeed, the 576 events assigned to deuteron breakup represent a total neutrino flux of 5.09 million per square centimeter per second—far larger than the 1.75 million electron-neutrinos measured by neutrino absorption. These numbers are determined with high accuracy. The difference between them is more than five times the experimental uncertainty.

The excess of neutrinos measured by deuteron breakup means that nearly two thirds of the total 5.09 million neutrinos arriving from the sun are either muon- or tau-neutrinos. The sun's fusion reactions can produce only electron-neutrinos, so some of them must be transformed on their way to the earth. SNO has therefore demonstrated directly that neutrinos do not behave according to the simple scheme of three distinct massless flavors described by the Standard Model. In 20 years of trying, only experiments such as Super-Kamiokande and SNO have shown that the fundamental particles have properties not contained in the Standard Model. The observations of neutrino flavor transformation provide direct experimental evidence that there is yet more to be discovered about the microscopic universe.

But what of the solar neutrino problem itself—does the discovery that electron-neutrinos transform into another flavor completely explain the deficit observed for the past 30 years? It does: the deduced 5.09 million neutrinos agrees remarkably well with the predictions of solar models. We can now claim that we really do understand the way the sun generates its power. Having taken a detour lasting three decades, in which we found that the sun could tell us something new about neutrinos, we can finally return to Davis's original goal and begin to use neutrinos to understand the sun. For example, neutrino studies could determine how much of the sun's energy is produced by direct nuclear fusion of hydrogen atoms and how much is catalyzed by carbon atoms.

The Future

THE IMPLICATIONS OF SNO'S DISCOVERY go even further. If neutrinos change flavor through oscillation, then they cannot be massless. After photons, neutrinos are the second most numerous known particles in the universe, so even a tiny mass could have a significant cosmological significance. Neutrino oscillation experiments such as SNO and Super-Kami-

THE AUTHORS

ARTHUR B. McDONALD, JOSHUA R. KLEIN and *DAVID L. WARK* are members of the 130-strong Sudbury Neutrino Observatory (SNO) collaboration. McDonald, a native of Sydney, Nova Scotia, has been the director of the SNO Institute since its inception in 1989. He is also professor of physics at Queen's University in Kingston, Ontario. Klein received his Ph.D. from Princeton University in 1994 and began his work on SNO at the University of Pennsylvania. He is now assistant professor of physics at the University of Texas at Austin. Wark has spent the past 13 years in the U.K., at the University of Oxford, the University of Sussex and the Rutherford Appleton Laboratory, trying to explain the infield fly rule to cricket fans. He has worked on a number of neutrino experiments in addition to SNO.

okande measure only mass differences, not masses themselves. Showing that mass differences are not zero, however, proves that at least some of the masses are not zero. Combining the oscillation results for mass differences with upper limits for the electron-neutrino mass from other experiments shows that neutrinos make up something between 0.3 and 21 percent of the critical density for a flat universe. (Other cosmological data strongly indicate that the universe is flat.) This amount is not negligible (it is roughly comparable to the 4 percent density that arises from gas, dust and stars), but it is not quite enough to explain all the matter that seems to be present in the universe. Because neutrinos were the last known particles that could have made up the missing dark matter, some particle or particles not currently known to physics must exist—and with a density far in excess of everything we *do* know.

SNO has also been searching for direct evidence of the effects of matter on neutrino oscillations. As mentioned earlier, travel through the sun can enhance the probability of oscillations. If this occurs, the passage of neutrinos through thousands of kilometers of the earth could lead to a small reversal in the process—the sun might shine more brightly in electron-neutrinos at night than during the day. SNO's data show a small excess of electron-neutrinos arriving at night compared with during the day, but as of now the measurement is not significant enough to decide whether the effect is real.

The results reported by SNO so far are just the beginning. For the observations cited here, we detected the neutrons from the critical deuteron breakup events by observing their capture by other deuterium atoms—an inefficient process that produces little light. In May 2001 two tons of highly purified sodium chloride (table salt) were added to the heavy water. Chlorine nuclei capture neutrons with much higher efficiency than deuterium nuclei do, producing events that have more light and are easier to distinguish from background. Thus, SNO will make a separate and more sensitive measurement of the deuteron breakup rate to check the first results. The SNO collaboration has also built an array of ultraclean detectors called proportional counters, which will be deployed throughout the heavy water in mid-2003 to look for the neutrons directly. Making these detectors was a technical challenge of the first order because they must have a spectacularly low level of intrinsic radioactive background—corresponding to about one count per meter of detector per year. Those devices will essentially check SNO's earlier results by an independent experiment.

SNO has unique capabilities, but it is not the only game in town. In December 2002 the first results from a new Japanese-American experiment called KamLAND were reported. The KamLAND detector is at the Super-Kamiokande site and studies electron-antineutrinos produced by all the nuclear reactors nearby in Japan and Korea. If matter-enhanced neutrino oscillations explain the flavor change seen by SNO, theory predicts that these antineutrinos should also change flavor over distances of tens or hundreds of kilometers. Indeed, KamLAND has seen too few electron-antineutrinos, implying that they are oscillating en route from the nuclear reactors to the detector.

Future neutrino experiments might help explain why the universe is made of matter rather than antimatter.

The KamLAND results imply neutrino mass differences and mixing parameters similar to those seen by SNO.

Future neutrino experiments might probe one of the biggest mysteries in the cosmos: Why is the universe made of matter rather than antimatter? Russian physicist Andrei Sakharov first pointed out that to get from a big bang of pure energy to the current matter-dominated universe requires the laws of physics to be different for particles and antiparticles. This is called CP (charge-parity) violation, and sensitive measurements of particle decays have verified that the laws of physics violate CP. The problem is that the CP violation seen so far is not enough to explain the amount of matter around us, so phenomena we have not yet observed must be hiding more CP violation. One possible hiding place is neutrino oscillations.

To observe CP-violating neutrino oscillations will be a multistage process. First physicists must see electron-neutrinos appear in intense beams of muon-neutrinos. Second, higher-intensity accelerators must be built to produce beams of neutrinos so intense and pure that their oscillations can be observed in detectors located across continents or on the other side of the earth. Studies of a rare radioactive process called neutrinoless double beta decay will provide further information about neutrino masses and CP violation.

It will probably be more than a decade before these experiments become a reality. A decade may seem a long way off, but the past 30 years, and the sagas of experiments such as SNO, have shown that neutrino physicists are patient and very persistent—one has to be to pry out the secrets of these elusive particles. These secrets are intimately tied up with our next level of understanding of particle physics, astrophysics and cosmology, and thus persist we must. SA

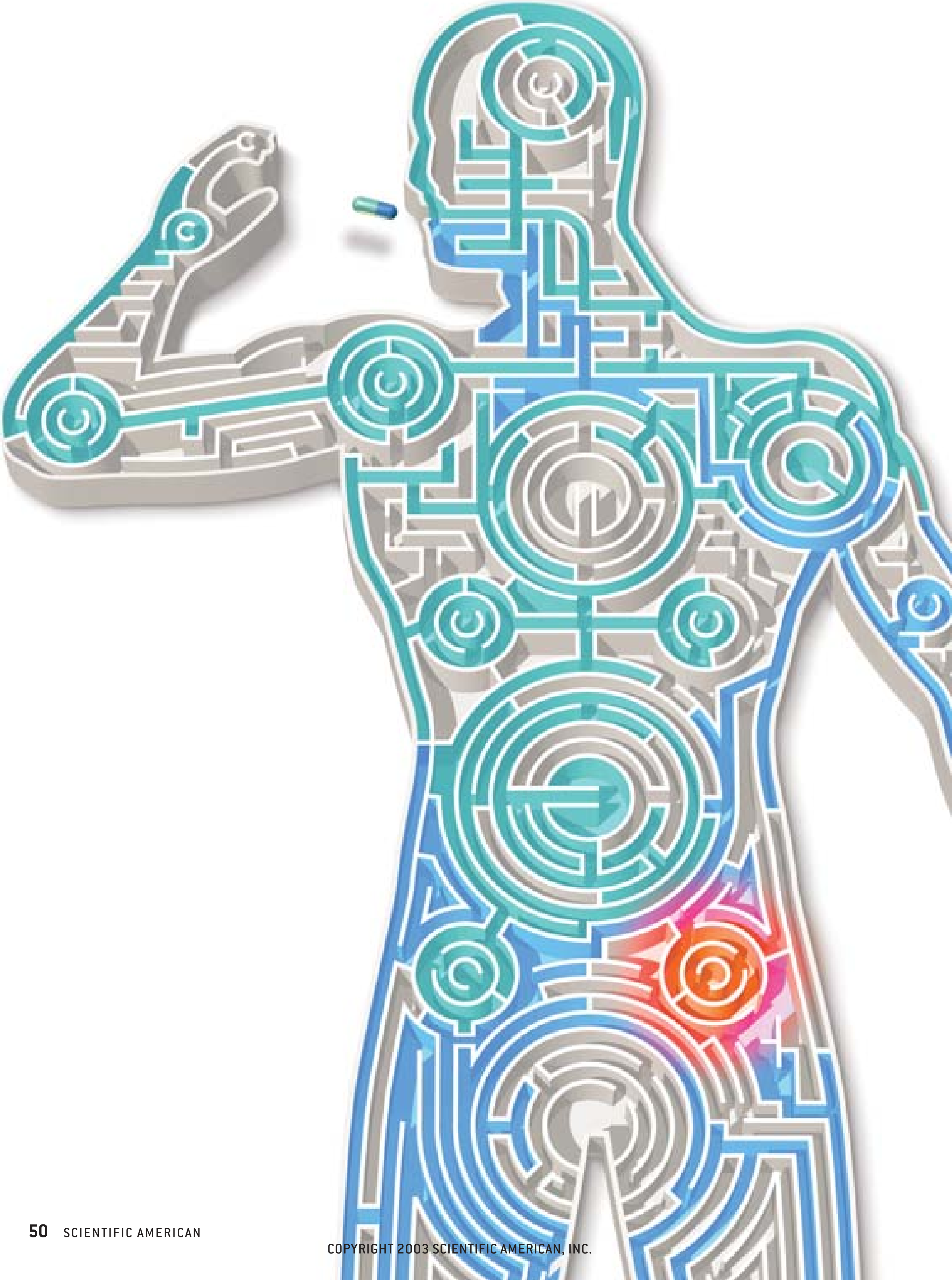
MORE TO EXPLORE

The Origin of Neutrino Mass. Hitoshi Murayama in *Physics World*, Vol. 15, No. 5, pages 35–39; May 2002.

The Asymmetry between Matter and Antimatter. Helen R. Quinn in *Physics Today*, Vol. 56, No. 2, pages 30–35; February 2003.

The SNO Web site is at **www.sno.phy.queensu.ca**

The Neutrino Oscillation Industry Web site, maintained by Argonne National Laboratory, is at **www.neutrinooscillation.org**



By Robert Langer

ILLUSTRATIONS BY SLIM FILMS

How to get drugs where they need to go

Where a Pill Won't Reach

From the point of view of a drug, it's a long trip from the pill bottle to an ache or the site of an infection.

After someone swallows a medicine, the chemical must traverse a veritable maze. It has to survive a journey through the stomach and reach the intestines intact before crossing the intestinal wall into circulation. Once in the blood, it gets filtered through the liver before it can travel to the rest of the body. At each "way station," the compound must resist the acids of digestive juices, jump membrane barriers or fend off enzymes designed to chop it into useless bits.

Pharmaceutical makers have come up with various solutions to help certain medicines on the market today surpass these obstacles; however, these approaches do not work for many other drugs. One strategy, for instance, coats pills with a shell that is insoluble in stomach secretions but that dissolves readily once it hits the more alkaline environment of the small intestine. But if a drug is made of protein—as most agents produced using biotechnology are—it also has to evade the activity of protein-destroying enzymes called proteases. Packaging pills with their own bodyguards (in this case, molecules called protease inhibitors) could enable protein-based drugs to survive, but it would not aid them in crossing the gut lining; they are too big to slip into the blood as easily as more typical drugs, which generally consist of small molecules. Coatings also have a limited ability to control a drug's pharmacokinetics—the rate at which it enters circulation and the time it persists in the body's tissues and organs. A drug can be toxic if it gets into the bloodstream too quickly at high concentrations or if it sticks around too long; conversely, it can be ineffective if a protracted delay occurs before it begins circulating.

Injecting a drug avoids the obstacles posed by the stomach and the intestinal tract, but many patients are understandably reluctant to give themselves shots repeatedly or to visit a doctor every day. Accordingly, scientists have sought a better way. Over the past two decades a spate of alternative drug delivery systems have been designed: sales of drugs administered by patch, implant, long-acting injection, topical gel, controlled-release pill, or nasal or lung spray now exceed $20 billion a year in the U.S. alone. Two notable examples approved recently by the Food and Drug Administration are Nutropin Depot—injectable, degradable polymer microspheres that secrete human growth hormone for up to four weeks between shots—and Gliadel, a wafer that can be implanted into the brain to administer chemotherapy directly to brain tumors. Already available in Europe and coming soon in the U.S. will be polymer-coated, drug-releasing stents, which so far have shown remarkable results in keeping blood vessels open after a clot-clearing procedure called angioplasty.

Indeed, scientists have been exploring nearly every part of the body—the skin, the nose, the lungs, as well as the intestine—as portals for introducing drug payloads. In the process, they have devised noninvasive ways to deliver complex molecules, such as using ultrasound to blast drugs through the skin painlessly. They have also combined advances in nanotechnology and microfabrication to make implantable microchips that can deliver drugs precisely and on schedule.

Overview/*Drug Delivery*

- Many drugs—especially the protein-based pharmaceuticals made using biotechnology—are broken down quickly if taken orally.
- Accordingly, researchers are developing new means of administering medicines, including wearable devices that use pulses of electricity or ultrasound to drive drugs through the skin painlessly.
- The future promises implantable microchips that deliver drugs in preprogrammed doses and that communicate with computers in physicians' offices.

Breaching the Wall

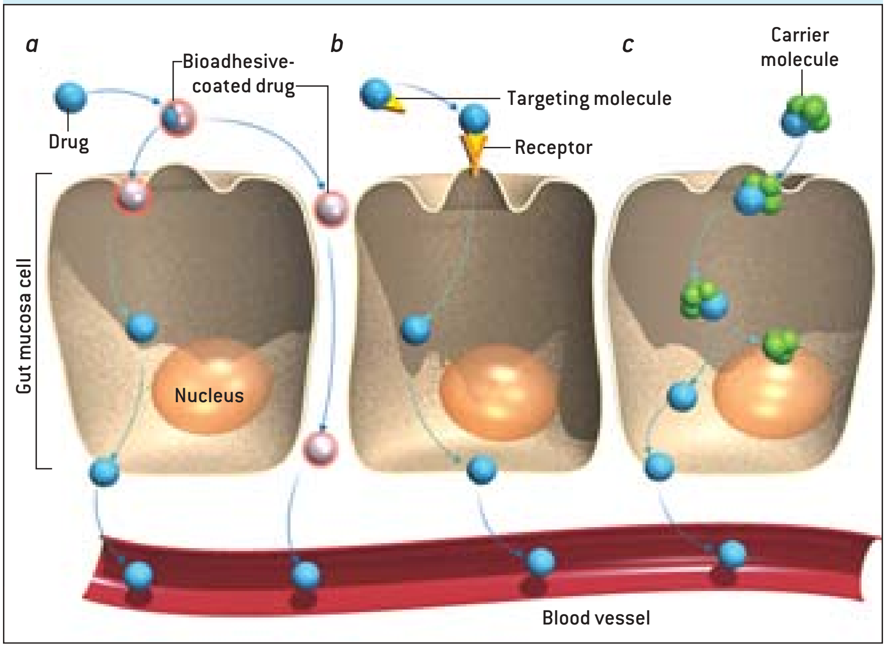
A VARIETY OF GROUPS have been using the new technologies to solve the problem of penetrating the intestinal wall. Edith Mathiowitz of Brown University and her colleagues, for example, have developed a way to entrap proteins in extremely tiny blobs of a gluey substance called a bioadhesive, which can penetrate through and between intestinal cells. The concept of using bioadhesion to enable orally administered drugs to attach to mucous membranes had its origins in work conducted in the 1970s and 1980s in the laboratories of Tsuneji Nagai of Hoshi University in Tokyo, Joseph R. Robinson of the University of Wisconsin–Madison and Nicholas A. Peppas of Purdue University (he is now at the University of Texas at Austin). Until 10 years ago the

GETTING PAST THE GUT

HURDLES: Digestive acids and enzymes degrade drugs before they can reach their targets; drugs have difficulty piercing the intestinal wall

SOLUTIONS: (*a*) Coat drugs with bioadhesive polymers that bind to the gut lining and squeeze between cells; (*b*) link drugs to targeting molecules that home in on receptors on intestinal cells that allow the drugs to be taken up; (*c*) attach drugs to carrier molecules that escort them into cells

SOME COMPANIES WORKING ON THE PROBLEM:
Biotech Australia, Roseville, New South Wales
Emisphere Technologies, Tarrytown, N.Y.



Scientists have been exploring nearly every part of the body—the skin, the nose, the lungs, as well as the intestine—as portals for introducing drug payloads.

most promising bioadhesive polymers seemed to be water-loving plastics called hydrophilic polymers and hydrogels. From these early studies, researchers inferred that the most "wettable" polymers—those with the highest concentrations of carboxyl groups—were the materials of choice for bioadhesion. Although they would stick to the intestinal mucosa, however, they did not penetrate it very well, and they released the protein too quickly.

In 1997 Mathiowitz found that hydrophobic (water-repelling) bioadhesive polymers called polyanhydrides, which expose carboxyl groups on their exteriors as their surfaces erode, can bind to the gut lining just as well as hydrophilic polymers but can cross the intestinal mucosa and enter the bloodstream more readily. One polyanhydride in particular—poly(fumaric-co-sebacic anhydride)—showed greater adhesive forces than any other material tested. The technology is now being explored to make a form of insulin that can be given orally. (Many new drug delivery approaches focus on insulin because it is a protein that must be injected regularly by people with type 1 diabetes.) In animal tests, the polyanhydride shows promise with both hydrophilic and hydrophobic proteins.

Peppas and his co-workers have also developed polymers that are not only bioadhesive but swell in response to a pH change. They are able to protect a protein drug such as insulin from the acidic pH of the stomach and then release it in the more alkaline pH of the intestine. The polymers are also able to protect the protein from proteases in the upper small intestine and can temporarily open the connections between intestinal cells, allowing the protein to pass through.

Another strategy to deliver protein-based drugs orally is to encase them in carrier molecules that can ferry them across the gut lining. Emisphere Technologies in Tarrytown, N.Y., has developed a series of molecular carriers that appear to squeeze proteins to make them smaller so they can cross cell membranes more readily. Once the carrier does its job of getting the drug inside a cell, it breaks away and allows the protein to spring back into its native—and therefore active—form. Emisphere is testing the strategy to administer insulin to diabetics and to deliver the blood-thinning protein heparin to people undergoing hip-replacement surgery, who sometimes experience blood clots.

Other scientists have been working to couple drug proteins to molecules that target specific receptors in the gastrointestinal tract. One of the early examples is the work of Gregory J. Russell-Jones of Biotech Australia in Roseville, New South Wales, who exploited the fact that cells in the intestine use receptors to grab vitamin B12 and transport it through the gut wall. He found that by linking a protein to vitamin B12, he could trick the vitamin receptors into taking up the protein as well as the vitamin. But there are only so many vitamin B12 receptors in the gut, and they might not be abundant enough to drag a given protein drug into the blood in the amounts needed for a therapeutic effect. Other scientists are now looking to harness lectins—plentiful, sticky molecules that constitute part of the connective tissue between intestinal cells—or other substances to do the job.

Patching It Up

THE INTESTINE is a fairly direct route to the bloodstream, but the skin is much more accessible. Although skin can be a relatively impermeable barrier, a few drugs have just the right physical and chemical characteristics to cross it at reasonable rates. Transdermal patches that last up to seven days are now on the market: nicotine to help people stop smoking and estradiol (estrogen) to counter the symptoms of menopause or to act as part of a contraceptive [see "Potent Patches," Working Knowledge, page 92].

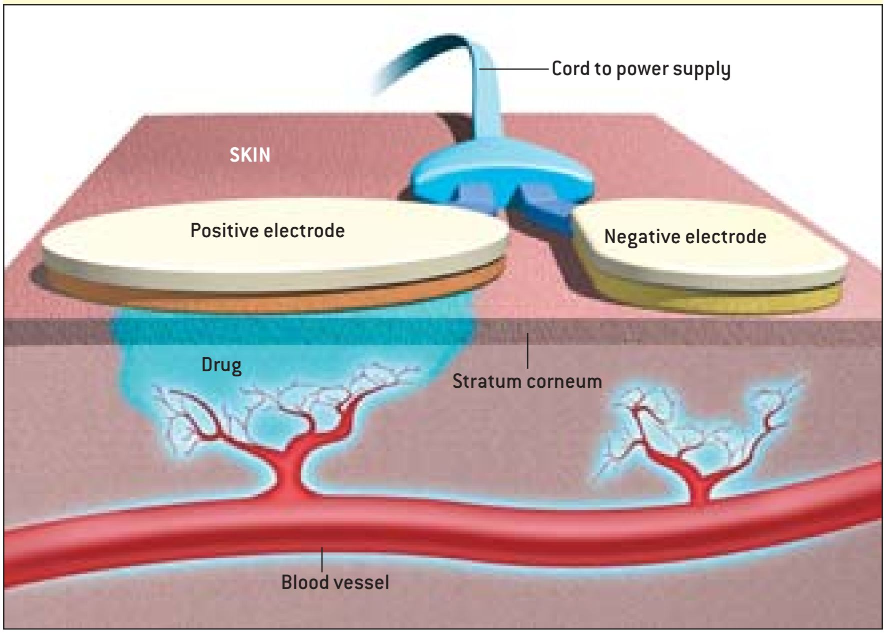
Passing a small, direct electric current through the skin can make the epidermis permeable to many other drugs—including proteins. Scientists at ALZA in Mountain View, Calif., and at Vyteris, a spin-off of Becton Dickinson based in Fair Lawn, N.J., are independently conducting advanced clinical trials using the technique, known as iontophoresis. In general, iontophoresis employs two patches—one negatively charged and one positively charged—that are linked to a reservoir of a given drug. A painless pulse of electricity can drive drugs, which tend to be charged, through the impermeable outer layer of the epidermis and into the blood vessels of the dermis. Vyteris, for example, has applied for FDA approval to market its iontophoresis system to deliver the painkiller lidocaine. The system's battery source is small enough to wear underneath clothing. The company is planning clinical tests of its device's ability to administer daily doses of parathyroid hormone to patients with osteoporosis or pulses of gonadotropin-releasing hormone every 90 minutes to women preparing for in vitro fertilization procedures.

PENETRATING THE SKIN

HURDLES: Tough stratum corneum (skin's outer layer) blocks drug entry; large molecules have difficulty crossing the epidermis to the blood vessels in the dermis

SOLUTIONS: Iontophoresis (*below*) uses painless pulses of electricity to make skin permeable; ultrasound employs sound waves that temporarily make tiny channels in the stratum corneum

SOME COMPANIES WORKING ON THE PROBLEM:
ALZA, Mountain View, Calif.
Sontra Medical, Cambridge, Mass.
Vyteris, Fair Lawn, N.J.



Ultrasound has also been harnessed to enhance skin permeability. Joseph Kost—a former postdoctoral fellow and visiting scientist in my laboratory and now a professor at Ben-Gurion University in Israel—discovered that ultrasound can temporarily disorder the skin's outermost layer, the stratum corneum, the principal barrier to drug diffusion. My M.I.T. colleague Daniel Blankschtein, Samir Mitragotri, a former graduate student in my lab, and I have used ultrasound to increase up to 5,000 times the ability of proteins the size of insulin to diffuse through the skin. Sontra Medical in Cambridge, Mass., which I co-founded and where Kost is chief scientific officer, is testing the technique for administering insulin and pain medications. The ultrasound device uses a short (15-second) burst of energy

THE AUTHOR

ROBERT LANGER is Kenneth J. Germeshausen Professor of Chemical and Biomedical Engineering at the Massachusetts Institute of Technology. He is also a director of a number of companies, including Sontra Medical and MicroCHIPS. In 2002 he received the Charles Stark Draper Prize, which is considered the equivalent of the Nobel Prize for engineering. In 1998 he was granted the Lemelson-M.I.T. Prize for being "one of history's most prolific inventors in medicine." Langer is one of the few individuals to have ever been elected to all three of the U.S. national academies: the Institute of Medicine, the National Academy of Sciences, and the National Academy of Engineering.

much weaker than that employed for diagnostic imaging to make the skin more permeable in a particular spot for up to 24 hours. The ultrasonic horn of the handheld device vibrates at 55,000 cycles a second (55 kilohertz) in a liquid medium coupled to the skin. The low-frequency ultrasonic energy creates tiny bubbles that expand and contract in the coupling medium and in the cell membranes of the stratum corneum, in effect drilling temporary miniature channels through which drugs can enter.

Inhaling the Future

LUNG DELIVERY represents another important opportunity and challenge—whether for treating lung conditions or for administering a drug to the bloodstream quickly to treat diseases elsewhere in the body. The lungs consist of microscopic sacs called alveoli that are linked directly to blood vessels. During breathing, oxygen enters the blood through the alveoli, and the waste product carbon dioxide exits. A similar process can also admit aerosols of larger molecules, such as protein-based drugs. It has been difficult, however, to design inhaler devices that can produce a sufficient number of aerosol particles small enough to penetrate deeply into the lung, without wasting the drug. (Most conventional inhalers, such as those used to treat asthma, deliver less than 10 percent of their contents.) Immune cells in the lung called macrophages can also clear most drugs rapidly.

A variety of researchers and companies are now working on improved inhaler designs that administer extremely fine mists in an efficient manner. One such inhaler, developed by Aradigm in Hayward, Calif., for liquid formulations, pushes a given drug through small nozzles that can be preprogrammed to deliver particular doses. Another approach, developed by Nektar Therapeutics (previously called Inhale Therapeutic Systems) in San Carlos, Calif., can generate an aerosol cloud from a dry powder by compressing air into it and breaking it into tiny particles capable of reaching the deepest areas of the lung. Both companies are now testing their devices in diabetics to deliver insulin without the need for injections.

Until the mid-1990s, however, scientists paid little attention to the aerosol particles themselves. At that time, David A. Edwards, who was then a postdoctoral fellow in my lab (he is now a professor at Harvard University), began to think of ways to design better aerosol mists. He reasoned that by dramatically lowering the density of aerosol particles while increasing their size and porosity, he might reduce the tendency of the particles to aggregate, thus enabling them to enter the lungs through an air stream produced by an extremely small, simple inhaler. Consider the difference between wet basketballs, for example, and wet grains of sand. The former have essentially no propensity for sticking together,

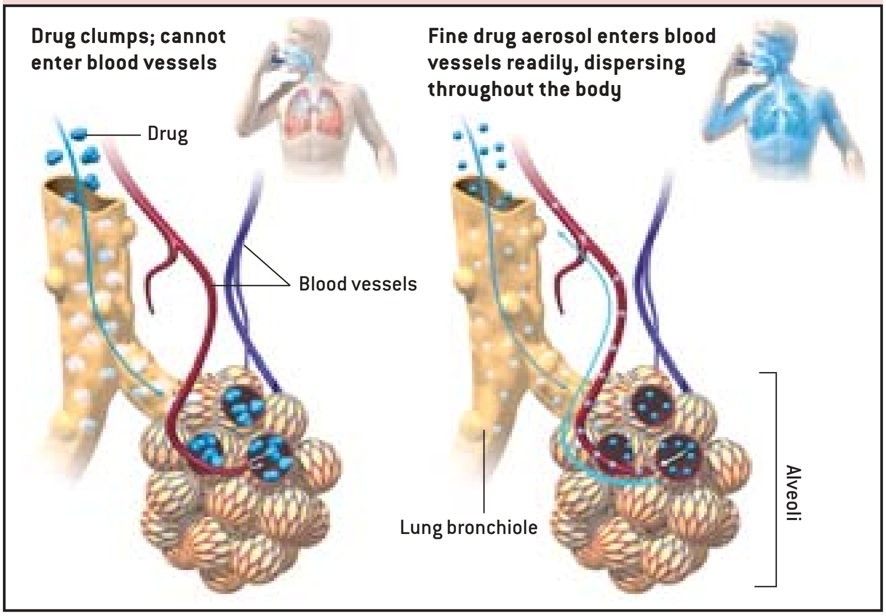
A painless pulse of electricity can drive drugs, which tend to be charged, through the outer layer of the epidermis and into the blood vessels of the dermis.

ENTERING THE LUNGS

HURDLES: Penetrating the air sacs, or alveoli; avoiding destruction by immune cells called macrophages

SOLUTIONS: Optimize the size of aerosol particles (liquid or powder) so they can reach deep into the lungs; prevent the aerosol particles from sticking together so they can form a fine mist

SOME COMPANIES WORKING ON THE PROBLEM:
Alkermes, Cambridge, Mass.
Aradigm, Hayward, Calif.
Nektar Therapeutics (formerly Inhale Therapeutic Systems), San Carlos, Calif.



Delivering Genes

Gene therapy depends on ferrying new genetic material into body cells

As the noted biologist Inder M. Verma of the Salk Institute for Biological Studies in San Diego has stated, there are three challenges in gene therapy: delivery, delivery, delivery. To introduce a new gene into the body, scientists must condense the corresponding DNA into small packages that can be taken up by a cell. But that is not all: they must also protect the gene from the cell's destructive enzymes, deliver it to the nucleus and release it in active form. For years, scientists have harnessed viruses as vectors, Trojan horses that sneak foreign genes into cells. But even viruses that have been disabled carry risks, as the tragic death in 1999 of gene-therapy trial volunteer Jesse Gelsinger made all too clear.

As researchers work to understand and reduce the risks of viral vectors for gene therapy, they are also devising alternative means for delivering genes that are based on polymers or on fatty molecules called lipids. One interesting approach, developed by Mark E. Davis of the California Institute of Technology, involves polymers named cationic, or positively charged, B-cyclodextrins (CDs).

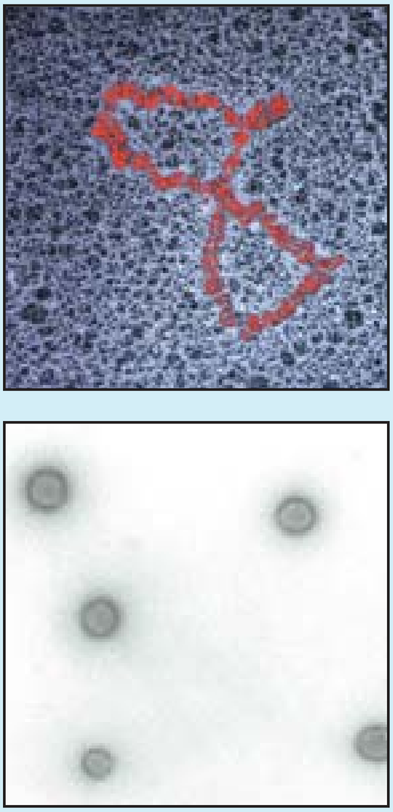
Davis chose CDs because they are relatively nontoxic, do not elicit an immune response and are soluble in water. He originally intended to package DNA for gene therapy into nanometer-size particles of CDs alone but found that this simple combination was unstable when administered to animals. So Davis and a graduate student, Suzie Hwang Pun, came up with the idea of altering the surface of the CD particles by adding adamantane-conjugated polyethylene glycol (PEG). The modification generates uniformly sized nanoparticles of CDs and DNA that resist clumping into useless aggregates with proteins in the blood serum.

"Decorating" the surfaces of the particles with the PEG compound also provided Davis and Pun with chemical hooks for attaching other molecules that could lead the particles to home in on, and deliver their genes to, specific types of cells. Insert Therapeutics in Pasadena, Calif.—which Davis founded and where Pun now works—is testing whether these targeted complexes can help treat various cancers and liver disease.

Specialized polymers are also being developed as gene therapy vectors by David M. Lynn of the University of Wisconsin–Madison. Lynn, who was a postdoctoral fellow in my laboratory at the Massachusetts of Technology, has synthesized sets, or libraries, of biodegradable cationic polymers called polyaminoesters. Along with Daniel Anderson, another postdoc in my lab, and David Putnam, who is now at Cornell University, Lynn has made hundreds of such polymers and has devised screening tests to identify the most useful ones according to how readily they can bind to DNA, how soluble they are in blood serum and how well they enter, or transfect, cells. Using this method, the researchers have found a number of polymers that can transfect cells more efficiently than the standard nonviral vectors lipofectamine and polyethylenimine.

Fred E. Cohen of the University of California at San Francisco has also employed the library approach. In collaboration with Ronald Zuckerman of Chiron in Emeryville, Calif., Cohen has synthesized a new class of polymers called peptoids or (more technically) cationic *N*-substituted glycine oligomers. Some of these can condense DNA into particles of between 50 and 100 nanometers in size that are capable of transfecting cells.

Lipids are also useful for delivering gene therapy, according to Sung Wan Kim of the University of Utah. Kim wraps a particular gene to be administered in a sheath of stearyl polylysine and then coats it with low-density lipoprotein. In studies using rabbits, he and his co-workers have employed this system to deliver the gene that encodes vascular endothelial growth factor (VEGF) to heart muscle damaged by a lack of oxygen (a condition called ischemia), as can occur when blood vessels are blocked. Kim plans to begin testing the approach next year in patients with ischemic heart disease. The hope is that the introduced VEGF gene will spur the growth of new blood vessels to bring oxygen and nutrients to starved areas of the heart muscle. *—R.L.*



CIRCLET OF DNA, called a plasmid (*top*), must enter a cell and begin functioning to be an effective gene therapy. Most gene therapy tests in people have used viruses to introduce genes into body cells, but they have raised safety questions. Researchers are now looking to package plasmids in cages of polymers (*bottom*) that are readily taken up by cells.

J. L. CARSON *Custom Medical Stock Photo* (*top*); MARK E. DAVIS AND SUZIE HWANG PUN *California Institute of Technology* (*bottom*)

“Smart” drug delivery systems would detect chemical signals in the body and release drugs in response to such signals.

whereas the latter cling together readily.

By making the aerosol particles larger, Edwards thought he could also decrease their uptake by lung macrophages, which tend to engulf smaller particles and to destroy drugs. Edwards and other scientists in our respective laboratories have shown that a single inhaled dose of insulin formulated in a large-particle aerosol can last up to four days in the lungs of animals. This approach is now being tested in humans by Alkermes, a company based in Cambridge, Mass.—in conjunction with pharmaceutical firms such as Eli Lilly—using several different drugs.

Intelligent Microchips

LOOKING FURTHER ahead, I see “smart” drug delivery systems as particularly exciting. These would detect chemical signals in the body and release drugs in response to such signals, keeping the concentration of a drug in the body at a desired therapeutic level.

A number of years ago, while I was watching a television program about how silicon chips for computers were made, I realized that the same technology might be applied to create smart systems for administering drugs. In conjunction with Michael J. Cima, an M.I.T. expert in ceramics processing, I found an undergraduate at the University of Michigan at Ann Arbor, John T. Santini, Jr., who agreed to explore the idea as a summer research project. Santini, who ended up doing his Ph.D. at M.I.T., worked out a way to create silicon microchips that contain a number of wells that can be loaded with drugs and covered with caps of thin gold foil. Applying a one-volt electrical signal to one or more of the wells dissolves the gold covers and releases the drug. Santini is now president of MicroCHIPS in Bedford, Mass., which is currently developing these systems for testing in human patients.

Microchips could be implanted under the skin or into the spinal cord or brain to deliver drugs ranging from pain medications to chemotherapies against cancer. Animal studies conducted by James Anderson of Case Western Reserve University and his colleagues have shown that the materials in the microchips are quite biocompatible and unlikely to cause side effects. Chip-based systems, which would include a small, wearable power source, would be easy to use and could keep an accurate record of how much drug a given patient is taking: data from the devices could be downloaded into a computer at home, at a doctor’s office, or in the hospital, providing a permanent record of a patient’s drug history. ChipRx, a drug delivery company based in Lexington, Ky., is also developing implantable devices to sense the level of a drug in the body and administer appropriate doses in response. It states that it is now preparing publications for scientific journals describing its devices.

We look forward to a day when any drug can be administered at the right time, in the right dosage, anywhere in the body with specificity and efficiency. SA

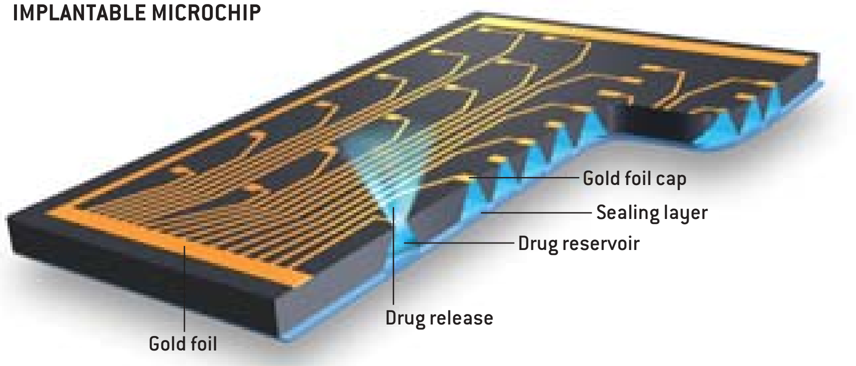
CONTROLLING RELEASE

HURDLE: Keeping the drug at a desired therapeutic level in the body while avoiding the need for frequent administration

SOLUTION: Implantable microchips with drug-filled reservoirs capped by gold foil that can be dissolved by an electrical charge to release the drug at appropriate times

SOME COMPANIES WORKING ON THE PROBLEM:
ChipRx, Lexington, Ky.
MicroCHIPS, Bedford, Mass.

IMPLANTABLE MICROCHIP



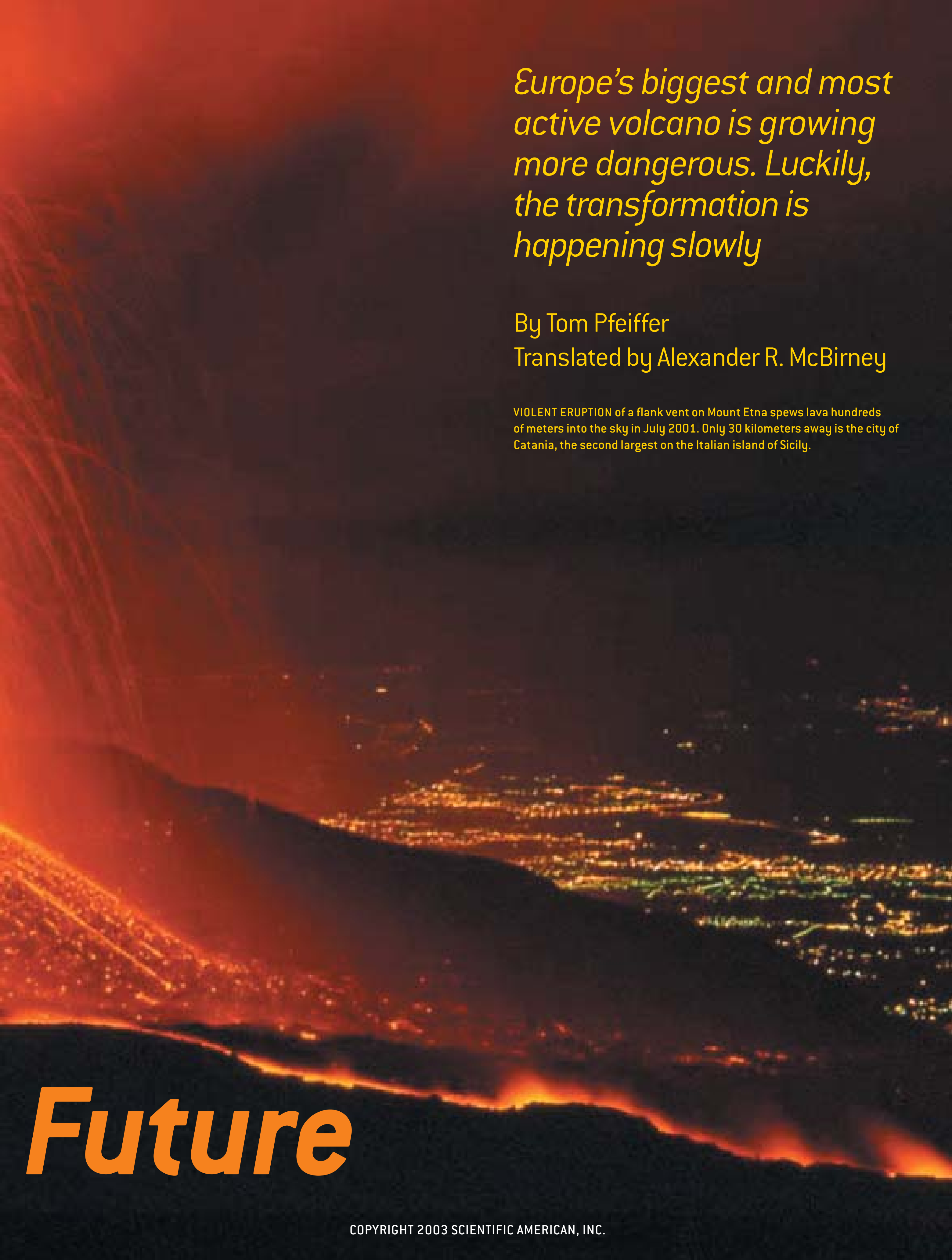
MORE TO EXPLORE

Drug Delivery and Targeting. Robert Langer in *Nature,* Vol. 392 (Supplement), pages 5–10; April 30, 1998.

Drug Delivery: Drugs on Target. Robert Langer in *Science,* Vol. 293, pages 58–59; July 6, 2001.

Ultrasound-Assisted Insulin Delivery and Noninvasive Glucose Sensing. Joseph Kost in *Diabetes Technology and Therapeutics,* Vol. 4, No. 4, pages 489–497; 2002.

Mount Etna's *Ferocious*



Future

Europe's biggest and most active volcano is growing more dangerous. Luckily, the transformation is happening slowly

By Tom Pfeiffer
Translated by Alexander R. McBirney

VIOLENT ERUPTION of a flank vent on Mount Etna spews lava hundreds of meters into the sky in July 2001. Only 30 kilometers away is the city of Catania, the second largest on the Italian island of Sicily.

LAVA FOUNTAIN erupts on Mount Etna's southern flank on October 30, 2002.

CORBIS (*preceding pages*); TOM PFEIFFER (*above*)

LAST OCTOBER ABOUT 1,000 ITALIANS FLED THEIR HOMES AFTER MOUNT ETNA, THE FAMOUS VOLCANO ON THE ISLAND OF SICILY, RUMBLED TO LIFE.

Shooting molten rock more than 500 meters into the air, Etna sent streams of lava rushing down its northeastern and southern flanks. The eruption was accompanied by hundreds of earthquakes measuring up to 4.3 on the Richter scale. As a huge plume of smoke and ash drifted across the Mediterranean Sea, residents of Linguaglossa (the name means "tongues" of lava) tried to ward off the lava flows by parading a statue of their patron saint through the town's streets.

Perhaps because of divine intervention, nobody was hurt and damage was not widespread. But the episode was unnerving because it was so similar to an erratic eruption on the volcano's southern flank in the summer of 2001 that destroyed parts of a tourist complex and threatened the town of Nicolosi. Some of the lavas discharged in both events were of an unusual type last produced in large amounts at the site about 15,000 years ago. At that time, a series of catastrophic eruptions led to the collapse of one of Etna's predecessor volcanoes.

The Sicilians living near Mount Etna have long regarded the volcano as a restless but relatively friendly neighbor. Though persistently active, Etna has not had a major explosive eruption—such as the devastating 1980 event at Mount Saint Helens in Washington State—for hundreds of years. But now some researchers believe they have found evidence that Etna is very gradually becoming more dangerous. It is unlikely that Etna will explode like Mount Saint Helens in the near future, but fierce eruptions may become more common.

Mountain of Fire

THE NAME "ETNA" is derived from an old Indo-Germanic root meaning "burned" or "burning." Extensive reports and legends record about 3,000 years of the volcano's activity, but a reliable chronicle has been available only since the 17th century. Most of the earlier accounts are limited to particularly violent eruptions, such as those occurring in 122 B.C. and A.D. 1169, 1329, 1536 and 1669. During the eruption in 1669, an enormous lava flow buried part of the city of Catania before pouring into the sea.

With a surface area of approximately 1,200 square kilometers, Etna is Europe's largest volcano [*see map on page 62*]. Its 3,340-meter-high peak is often covered with snow. Only the upper 2,000 meters consists of volcanic material; the mountain rests on a base of sedimentary rock beds. Blocks of this material are occasionally caught in the magma—the molten rock moving upward—and ejected at the surface. Numerous blocks of white sandstone were blown out during the 2001 and 2002 eruptions. This phenomenon occurs whenever magma must open new paths for its ascent, as is usually the case with lateral eruptions (those that occur on the volcano's flanks).

The volcano is more than 500,000 years old. Remnants of its earliest eruptions are still preserved in nearby coastal regions in the form of pillow lavas, which emerge underwater and do in fact look like giant pillows. At first, a shield volcano—so called because it resembles a shield placed face-up on the ground—grew in a depression in the area where Etna now stands. Today a much steeper cone rests on the ancient shield volcano. It consists of at least five generations of volcanic edifices that have piled up during the past 100,000 to 200,000 years, each atop the remnants of its eroded or partly collapsed predecessor. The present-day cone has been built in the past 5,000 to 8,000 years. Among Etna's special features are the hundreds of small cinder cones scattered about its flanks. Each marks a lateral outbreak of magma. One of the world's most productive volcanoes, Etna has spewed about 30 million cubic meters of igneous material each year since 1970, with a peak eruption rate of 300 cubic meters a second.

Etna is also one of the most puzzling volcanoes. Why has the magma that produced it risen to the surface at this particular spot, and why does it continue to do so in such large quantities? The an-

Overview/*Etna's Evolution*

- Long regarded as a relatively tame volcano, Mount Etna has rocked the Italian island of Sicily over the past two years. Eruptions on Etna's flanks have produced lava flows that have destroyed tourist facilities and threatened nearby towns.
- Researchers believe that some of Etna's molten rock is being generated by the collision of two tectonic plates. If this hypothesis is correct, the volcano may eventually become much more violent and explosive.

swers should be found in the theory of plate tectonics, which posits that the earth's outermost shell consists of about a dozen vast plates, each between about five and 150 kilometers thick. The plates constitute the planet's crust and the uppermost part of the mantle. Like pieces of ice floating on the ocean, these plates drift independently, sometimes moving apart and at other times colliding. The 530 active volcanoes of the world are divided into three major types according to their positions on or between these plates.

The first and most numerous type is found along the rift zones, where two plates are moving apart. The best examples are the long midocean ridges. Forces beneath the plates rip them apart along a fracture, and the separation causes an upwelling of hotter material from the underlying mantle. This material melts as it rises, producing basalt (the most common kind of magma), which contains large amounts of iron and magnesium. The basaltic melt fills the space created by the separating plates, thus continuously adding new oceanic crust.

The second type is located along the subduction zones, where two plates converge. Normally, a colder and heavier oceanic plate dives below a continental plate. The process that leads to the formation of magma in this environment is completely different: water and other fluids entrained with the sinking plate are released under increasing pressure and temperature, mainly at depths of about 100 kilometers. These fluids rise into the overlying, hotter mantle wedge and lower the melting temperature of the rocks. The resulting magmas, which are more viscous and gas-rich than the basaltic melts of the rift zones, contain less iron and magnesium and more silica and volatile components (mainly water and carbon dioxide).

These factors make the volcanoes in subduction zones far more menacing than volcanoes in rift zones. Because the viscous, gas-rich magma does not flow easily out of the earth, pressure builds up until the molten rock is ejected explosively. The sudden release of gases fragments the magma into volcanic projectiles, including bombs (rounded masses of lava), lapilli (small stony or glassy pieces) and

TOM PFEIFFER

MUSHROOM CLOUD of ash rises from Etna's northeastern flank on October 28, 2002.

ash. Such volcanoes typically have steep cones composed of alternating layers of loose airborne deposits and lava flows. Some of the best-known examples of subduction-zone volcanoes rise along the margins of the Pacific Ocean and in the island arcs. This Ring of Fire includes Mount Saint Helens, Unzen in Japan and Pinatubo in the Philippines, all of which have erupted in the past three decades.

The third type of volcano develops independently of the movements of the tectonic plates and is found above hot spots caused by mantle plumes, currents of unusually hot material that ascend by thermal convection from deep in the earth's mantle. As the mantle plumes approach the surface, decreasing pressure causes them to produce melts that bore their way through the crust, creating a chain of hot-spot volcanoes. Most hot-spot volcanoes produce highly fluid lava flows that build large, flat shield volcanoes, such as Mauna Loa in Hawaii.

At the Crossroads

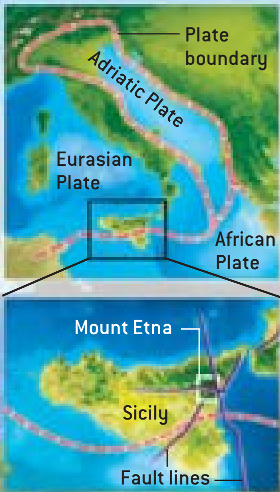
ETNA, HOWEVER, cannot be assigned to any of the three principal categories of volcanoes. It is located in a geologically complex area, which owes its current form to tectonic processes that have been active for the past 50 million to 60 million years. An ocean basin that formerly existed between Eurasia and the northward-moving African continent was swallowed to a large extent by the Eurasian plate. About 100 million years ago two smaller plates, Iberia and Adria, split off from the Eurasian and African plates because of enormous shearing stresses related to the separation of North America from Eurasia (and the opening of the Atlantic Ocean).

Mountain belts arose along the fronts where the plates collided. Italy's Apennines developed when the Iberian and Adriatic plates met. During this process, the Italian peninsula was rotated counterclockwise by as much as 120 degrees to its current position. Today Etna is sit-

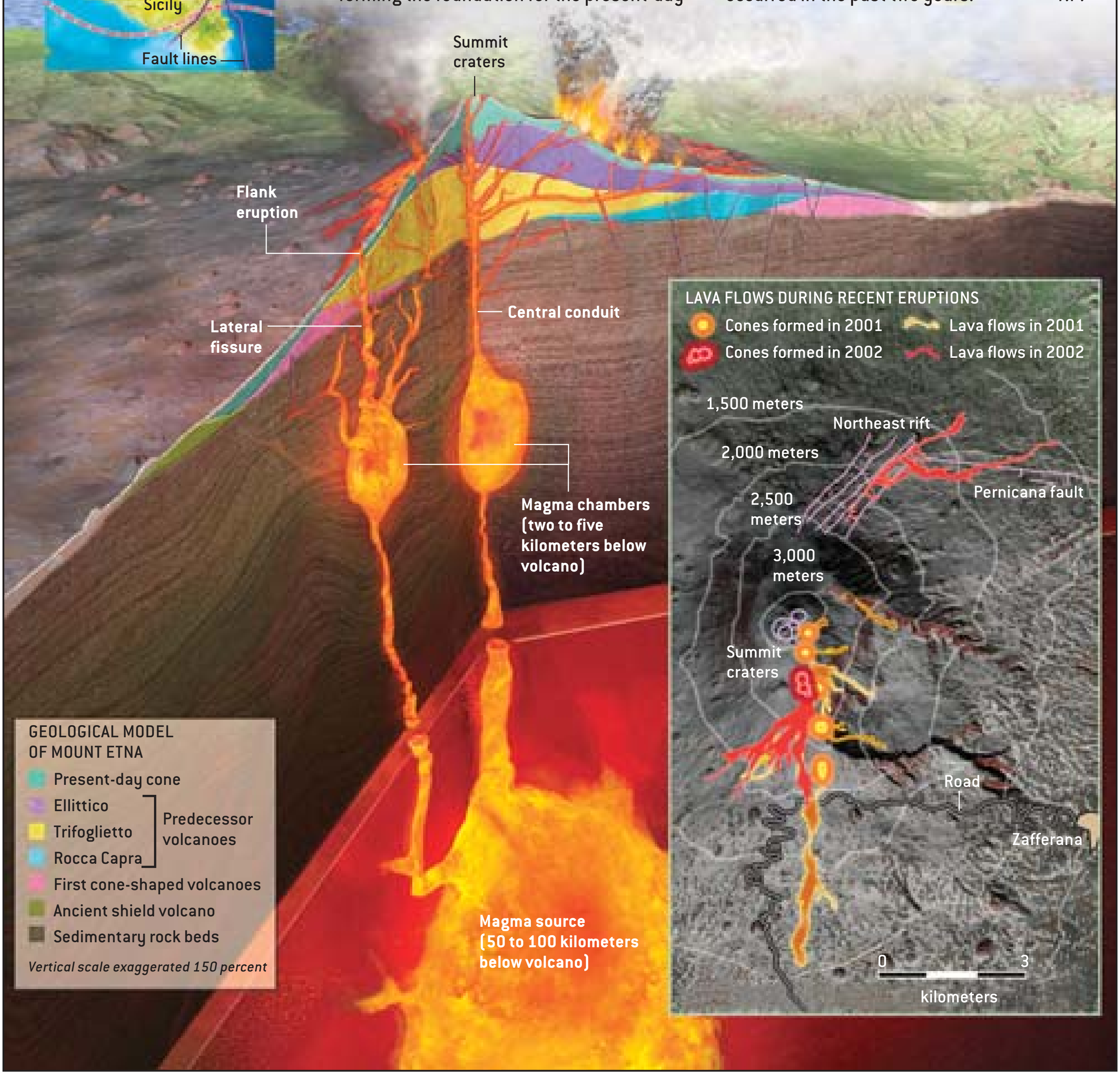
THE AUTHOR

TOM PFEIFFER has become very familiar with Mount Etna, photographing many of the volcano's recent eruptions. He is a Ph.D. student in the department of earth sciences at the University of Århus in Denmark. Pfeiffer has done research at the Hawaiian Volcano Observatory at Kilauea volcano and the Vesuvius Observatory in Naples. His dissertation is about the Minoan eruption on the Greek island of Santorini that devastated the eastern Mediterranean region around 1645 B.C. An earlier version of this article appeared in the May 2002 issue of *Spektrum der Wissenschaft, Scientific American*'s sister publication in Germany.

PORTRAIT OF A VOLCANO



MOUNT ETNA is situated close to the juncture of the Eurasian, African and Adriatic tectonic plates (*left*). The movements of these plates have fractured Sicily's crust along fault lines. A cross section of Etna (*below*) reveals much of the volcano's 500,000-year history. First, a flat shield volcano spread across the sedimentary rock beds; then a cone-shaped volcano rose above it. The succeeding generations of volcanic edifices—named Rocca Capra, Trifoglietto and Ellittico—piled atop their predecessors, forming the foundation for the present-day cone (dubbed Mongibello Recente). Recent eruptions on Etna's flanks seem to arise from a fissure that is not connected to the volcano's central feeding system. The two conduits appear to have separate magma chambers about two to five kilometers below the volcano's summit, although they share the same magma source 50 to 100 kilometers farther down. (This part of the cross section is not drawn to scale.) A contour map (*bottom right*) shows the locations of the flank eruptions and lava flows that have occurred in the past two years. *—T.P.*



DAVID FIERSTEIN

uated close to the junction of the African, Eurasian and Adriatic plates. Individual blocks from these plates have been superimposed and welded together on Sicily. Major tectonic faults cross the area around the volcano as a result of intense regional stresses within the crust.

For a long time researchers believed that Etna's position at the crossroads of these faults was the explanation for its volcanism. The presence of faults, however, accounts only for the ability of magma to reach the surface; it does not explain why the magma is produced in the first place. According to most theories, the prevailing forces in the Sicilian crust are similar to those in rift zones—extensional stresses that cause thinning of the crust and upwelling of the underlying mantle. But at Sicily the African and Eurasian plates are colliding, so one would expect the stresses to be compressive rather than extensional. Moreover, only about 20 percent of the magma erupted at Etna has a chemical composition similar to that of a rift-zone volcano.

ASH PLUME from Mount Etna is clearly visible in this image taken by NASA's Terra satellite.

HO, NASA, TERRA SATELLITE *AP Photo*

Judging from its magma and pattern of activity, Etna is most similar to hotspot volcanoes such as those in Hawaii. Recent theories suggest that it has developed above an active mantle plume, but no direct evidence for this plume has been detected. So far scientists have been unable to explain all the characteristics of this enigmatic volcano. For example, Etna is one of the few volcanoes in which magma is almost continuously rising. Its active periods can last for years or even decades and are interrupted only by short intervals of quietness. This pattern implies the existence of two things: first, a constant flow of magma from the mantle to the deep and shallow magma reservoirs beneath the volcano and, second, an open conduit through which magma can rise. In fact, the conduits between Etna's magma chambers and the summit craters seem to be very long lived structures. Seismic investigations have shown that the rising magma produces little noise and appears to move rather smoothly, without encountering major obstacles.

The kind of activity that prevails at Etna depends primarily on the level of magma inside its conduits. The low pressure in the upper part of the magma column allows the dissolved gases (mainly water and carbon dioxide) to escape. The resulting bubbles rise within the magma column and pop at the surface, throwing out liquid and solid fragments. When the level of the magma column is fairly deep inside the volcano, only gases and fine ash particles reach the crater rim. When it is closer to the surface, larger fragments (lapilli and bombs) are thrown out as well. In the rare cases when the magma column itself reaches the crater rim, the degassing magma pours over the rim or through a crack and forms a lava flow.

Besides lava flows, Etna produces an almost constant, rhythmic discharge of steam, ash and molten rock. Known as a strombolian eruption (named after Stromboli, a volcano on one of the Aeolian Islands about 100 kilometers north of Etna), this activity sometimes culminates in violent lava fountains jetting hundreds of meters into the air. During the spectacular series of eruptions at Etna's southeast crater in the first half of 2000, these fountains rose as high as 1,200 meters above the crater's rim—a stunning height rarely observed at any volcano.

To witness such an eruption from close range can be extremely dangerous, as I have learned from experience. In February 2000, violent eruptions at Etna's southeast crater were occurring at 12- or 24-hour intervals. On the evening of February 15, while I was observing the crater from about 800 meters away with a group of spectators, a white cloud of steam rose from the crater's mouth. It rapidly became thicker and denser. After a few minutes, the first red spots began dancing above the crater, rising and falling back into it. The explosions grew stronger, first slowly, then with breathtaking speed, throwing bombs more than 1,000 meters above the rim. Soon the volcanic cone surrounding the crater was covered with glowing rocks. At the same time, a fountain of lava started to rise from a fracture on the flank of the cone. Several other fountains rose from the crater and formed a roaring, golden curtain that illuminated

STROMBOLIAN EXPLOSIONS illuminate newly formed craters on Etna's northern flank.

TOM PFEIFFER

the scene like daylight. Some larger bombs crashed into the snow not far from us, but we felt secure in our viewing position. The fountain was nearly vertical, and a strong wind carried the mass of glowing lapilli and ash gently away from us.

Suddenly the lava fountain changed direction, sending a lateral outburst straight toward us. Just in time we reached the shelter of an abandoned mountain hut with a thick concrete roof. A heavy rain of incandescent stones fell around us; lava bombs of all sizes tumbled down, spraying thousands of sparks. Fortunately, our shelter was not hit by anything large, although a two-meter-wide bomb plunged into the snow nearby. After an endless two minutes, the lava fountain rose vertically again and stayed in this position for another 10 minutes. Then its supply of magma from below seemed to be exhausted. The fountain collapsed as if it were sucked back into the crater. The entire spectacle was finished 30 minutes after it began. In front of us, the 300-meter-high cone still glowed red but was completely silent.

Natural Air Polluter

ETNA'S REPUTATION as a relatively friendly volcano stems mainly from the fact that its lavas are very fluid. Such lavas are easily ejected to the surface, unlike the viscous magmas produced by subduction-zone volcanoes. But Etna's magmas also contain a great amount of gas, which can make eruptions much more explosive. During a particularly violent phase, Etna expels up to 20,000 tons of sulfur dioxide a day, making the volcano one of nature's worst air polluters. The high sulfur content of Etna's magma is hard to understand; this characteristic is more typical of subduction-zone volcanoes than of basaltic volcanoes.

What is more, Etna's composition indicates that the volcano has indeed experienced major explosive eruptions similar in size to those of Pinatubo in 1991 and Mount Saint Helens in 1980. Etna's last big explosion appears to have occurred in 122 B.C. During that event, more than one cubic kilometer of basaltic lava erupted in a giant column loaded with lapilli and ash. Deposits formed by this eruption are up to two meters thick on Etna's upper slopes and are still exposed in some areas. In Catania, about 30 kilometers from the summit, the deposits are between 10 and 25 centimeters thick. If such an event were to occur today, it would be a disaster. The roofs of many houses in the area would collapse from the weight of the ash.

The unusual flank eruptions of 2001 and 2002 made it clear that Etna is not tame. In 2001 as many as five fractures opened on both sides of the mountain, through which huge masses of lava started to pour. A new crater was born at an elevation of 2,500 meters. Extremely active, it spewed lava fountains and dense clouds of ash, growing within a few days to a cone about 100 meters high. Especially spectacular were the giant magma bubbles that rose within the new crater and detonated with awesome power. Even at a distance of several kilometers, the force of the explosions rattled doors and windows.

Researchers soon determined that two distinct eruptions were occurring simultaneously. The opening of the fractures near Etna's summit (between 2,700 and 3,000 meters above sea level) was a continuation of the volcanic activity that had been roiling the summit craters for years. But the eruptions at the lower fractures (at elevations between 2,100 and 2,500 meters) produced a more evolved type of magma that obviously had rested for a prolonged period in a separate chamber, where it could change its chemical composition. (A similar pattern was also evident in the 2002 eruptions.) This second kind of magma included centimeter-size crystals of the mineral amphibole, which is very rarely found in Etna's lavas. Besides iron, magnesium and silica, amphibole incorporates water in its crystal structure. The mineral can form only from a magma that contains sufficient amounts of water. Obviously, two different plumbing systems of the volcano were active at the same time: one associated with the central, more or less constantly active conduit and the other with an independent conduit off to the side.

The magmas ejected through this second conduit were last produced in large quantities at Etna about 15,000 years ago, when devastating eruptions caused the collapse of one of Etna's predecessors, the Ellittico volcano. Is their reappearance a sign that a catastrophic explosive eruption will happen in the near future? The answer depends on where Etna's magmas come from. Identifying their origins can be tricky: analyzing the erupted magma can be misleading, because the chemical composition of the original melt often changes during its ascent through the crust. Geologists have learned, however, that surface lavas sometimes contain crystals that preserve the composition of the original magma. If a crystal begins to form at an early stage in the life of a magma, it may include minuscule droplets of the primitive melt and grow around them. These melt inclusions are thus isolated from all subsequent chemical changes.

Analyzing such melt inclusions, though, is difficult. Until recently, almost no suitable data were available for Etna.

In 1996 a French-Italian research team consisting of Pierre Schiano (Blaise Pascal University in France), Roberto Clocchiatti (National Center for Scientific Research in France), Luisa Ottolini (National Research Council in Italy) and Tiziana Busà (then at the University of Catania in Italy) began a comprehensive investigation of the magmas of Etna and neighboring volcanoes. The researchers looked for glassy inclusions in olivine crystals, which are among the first to form from a primitive melt. The tiny inclusions they discovered, each less than two tenths of a millimeter in diameter, were remelted on a heating plate, then quenched to create a homogeneous glass. The team determined the chemical composition of the inclusions using a microprobe (which directs narrow beams of x-rays at a sample) and a secondary ion mass spectrometer (which employs ion beams).

Changing Character

THE SCIENTISTS PAID special attention to the trace elements, such as cesium and barium, which are rare in igneous rocks. When a melt forms deep underground, the trace elements in the source rock migrate almost completely to the magma. Because their relative concentrations remain nearly unchanged, the trace elements offer a geologic fingerprint of the origin of the melt. The magmas that erupted at Etna more than about 100,000 years ago had compositions similar to those from the older, now extinct volcanoes of the Iblean Mountains in southern Sicily. The trace-element patterns were also close to those found in magmas from hot-spot volcanoes in Hawaii and the Azores. The early volcanism at Etna was apparently fueled by a mantle plume, probably the same one that fed the Iblean volcanoes about 100 kilometers to the south.

But the analysis of the younger magmas (those that have been expelled within the past 100,000 years) revealed a much different picture. They have large concentrations of trace elements such as cesium, potassium, rubidium and barium, but they appeared to be depleted of elements such as titanium and zirconium. Remarkably similar patterns are found at the Aeolian Island volcanoes, which include Stromboli and Vulcano. This island arc most likely owes its existence to tectonic forces—specifically, the subduction of oceanic crust from the Ionian Sea under the Calabrian block (the southernmost part of the Italian mainland). Schiano and Clocchiatti are convinced that the similarity of the magmas is no coincidence. They believe that Etna has two sources of magma: the mantle plume that gave birth to the volcano and a second component that is identical to the magma feeding the Aeolian volcanoes. Furthermore, Etna's youngest magmas have the greatest amounts of this second component.

How does Etna produce its fiery mix of magmas? One possibility is that the two magmas form at different locations and mix somewhere within Etna's plumbing system. This hypothesis would imply that the magma below the Aeolian Islands travels more than 100 kilometers along a tectonic fault to Etna. It is considered highly unlikely, though, that such an underground magma passage exists. Researchers think it is more probable that the two magma *sources* are mixing. According to this model, part of the subducted slab of the Ionian plate has slowly migrated southward and come within reach of the plume beneath Etna. When the rising plume passes by the edge of the sinking slab, it creates the mix of magmas that emerges from the volcano.

Etna's activity has increased markedly since 1970, with more frequent eruptions and more volcanic material ejected. Researchers cannot be certain, however, whether this upsurge is caused by tectonic forces or by a fresh batch of magma rising from the mantle. If Etna is indeed transforming into an explosive subduction-zone volcano, the process will be a gradual one. As Schiano and Clocchiatti emphasize, "The observed change [from a hot-spot toward an island-arc volcano] is taking place in geological time and not in a human lifetime." Thus, Etna is unlikely to experience a catastrophic explosive eruption soon.

But if the researchers' hypothesis is correct, Etna's eruptions will grow increasingly violent. Some tens of thousands of years from now, Etna may well become as dangerous as Mount Saint Helens or Pinatubo. Fortunately, the Sicilians should have plenty of time to adapt to the new situation. SA

LAVA FLOWS block roads and cut through forests on Mount Etna's northern side in November 2002.

TOM PFEIFFER

MORE TO EXPLORE

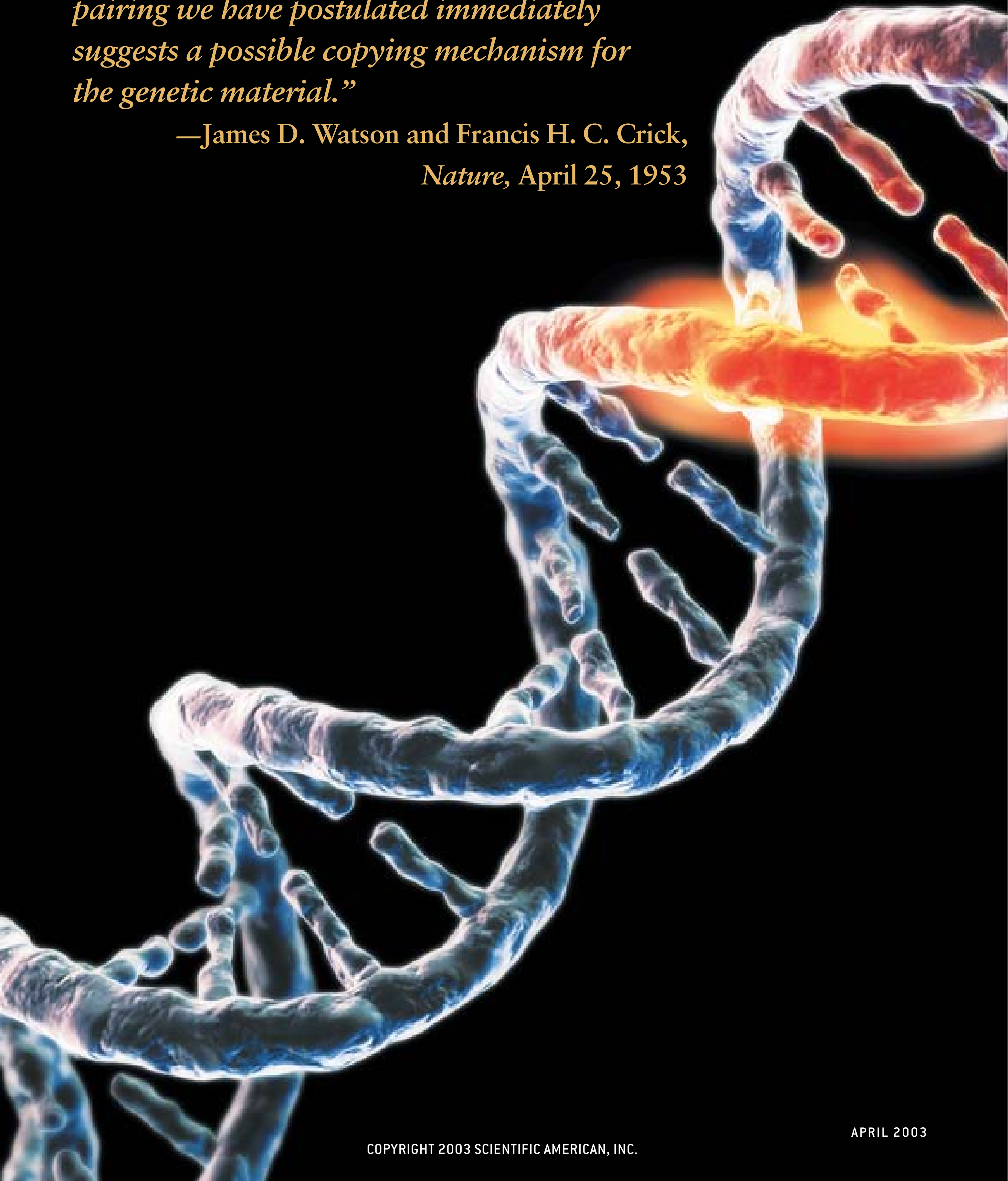
Mount Etna: Anatomy of a Volcano. B. K. Chester, A. M. Duncan, J. E. Guest and C.R.J. Kilburn. Chapman and Hall, 1985.

Transition of Mount Etna Lavas from a Mantle-Plume to an Island-Arc Magmatic Source. P. Schiano, R. Clocchiatti, L. Ottolini and T. Busà in *Nature,* Vol. 412, No. 6850, pages 900–904; August 30, 2001.

More information about Mount Etna, Stromboli and other volcanoes is available at **boris.vulcanoetna.com, www.stromboli.net** and **www.decadevolcano.net**

"It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material."

—James D. Watson and Francis H. C. Crick,
Nature, April 25, 1953



CELEBRATING THE GENETIC JUBILEE:

A CONVERSATION WITH JAMES D. WATSON

"This structure has novel features which are of considerable biological interest" is one of science's most famous understatements. James D. Watson and Francis H. C. Crick penned it for their report in the April 25, 1953, issue of Nature, *in which they proposed the double helix model for the structure of DNA, the breakthrough that opened up our understanding of molecular biology and genetics.*

To mark the 50th anniversary, Scientific American *Editor in Chief John Rennie recently spoke with Watson in his office at Cold Spring Harbor Laboratory on Long Island, where he was director for 25 years. Watson reflected on the origins of the double helix discovery, the current state of molecular biology, and controversies surrounding genetic science. (Because of poor health, Crick was not available for comment.) Here is an abridged, edited version of that conversation.*

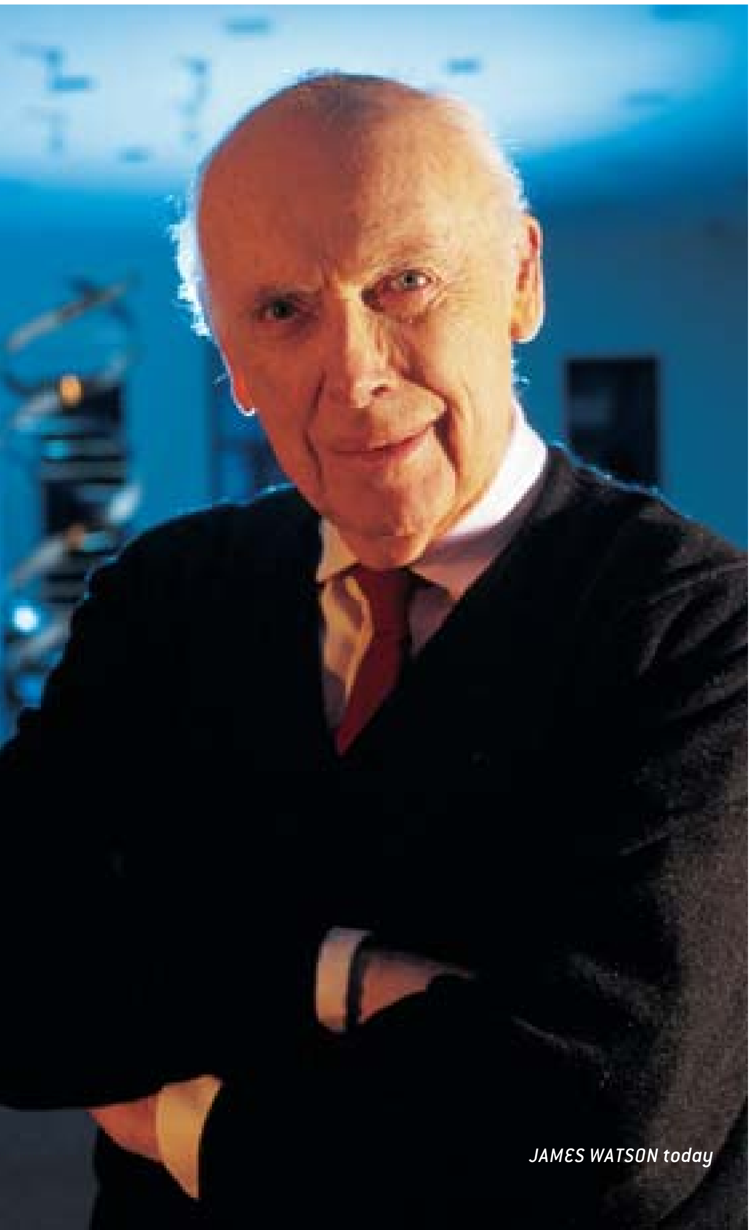
MIKE MEDICINE HORSE *Hybrid Medical Animation*

SCIENTIFIC AMERICAN: DNA is no longer just a scientific entity. It's erupted as this huge cultural phenomenon, as a metaphor for our natures. It's in our daily conversation, in art. When you were working on the double helix, did you foresee DNA ever becoming this well known?

JAMES WATSON: No, no, we couldn't. Because no one had ever sequenced DNA or amplified DNA.

The famous Australian immunologist [Frank Macfarlane] Burnet published this article in a medical journal that came out in 1961 or 1962, in which he said DNA and molecular biology will not have an influence on medicine. Because that's only possible when you can read the DNA. That's why the Human Genome Project is so important.

Back in 1953, all we wanted to do was find out how DNA provided the information and what the cellular machinery was for making proteins. That's really all; we didn't think about gene therapy. It took about 15 years before people began to think about that, around 1968—once restriction enzymes came along and, soon after, DNA sequencing.



JAMES WATSON today

SA: You've said that you first became involved with research on DNA because of your interests in evolution and information.

JW: [Physicist Erwin] Schrödinger probably wasn't the first, but he was the first one I'd read to say that there must be a code of some kind that allowed molecules in cells to carry information. By the time Schrödinger's book [*What Is Life?* (1944)] came out, a few people, such as [biologist J.B.S.] Haldane, were making that connection between genes and proteins. Back in those days the amino acid sequence for a protein hadn't been worked out. You knew there was some sequence, but that's all. It was only our getting the DNA structure and [chemist Frederick] Sanger sequencing the first polypeptide [protein] chain that let some air in.

SA: So would you say that your work was driven more by fascination with that idea than by ambition?

JW: I was born curious. I would rather read economic history than history, for example, because I liked explanations. And so if you wanted an explanation for life, it had to be about the molecular basis for life. I never thought there was a spiritual basis for life; I was very lucky to be brought up by a father who had no religious beliefs. I didn't have that hang-up. My mother was nominally a Catholic, but that's as far as it went.

SA: As we look back on the race for the double helix, it's obvious that individual personalities strongly influenced the specifics of who found it first and how. And yet the discovery at roughly that time also seems to have been inevitable. There were so many of you who were so close—you and Crick, Linus Pauling of the California Institute of Technology, Maurice Wilkins and Rosalind Franklin of King's College. If you and Crick had not made the discovery when you did—

JW: I can't believe a year would have passed.

SA: There has always been controversy about Wilkins showing you Franklin's crystallography images without her permission, giving you and Crick an important clue to DNA's structure. In retrospect, would it have been more appropriate for the Nobel committee to have given the prize to her, along with you and Crick, rather than to Wilkins?

JW: I think not. Wilkins gave us the crystalline photograph of the A form, and she gave us the B form. So you could have said that in an ideal, perfect society, they would have gotten the

I think there's something in me of that same weakness that is so apparent in John McEnroe. I just can't sit while people are saying nonsense.

MIRIAM CHUA *Cold Spring Harbor Laboratory (preceding page and left)*

People say, "Well, these would be designer babies," and I say, "What's wrong with designer clothes?"

chemistry prize, and Crick and I would have gotten the biology prize. That would have been a nice way to honor the four of us. But no one thought that way.

We're very famous because DNA is very famous. If Rosalind had talked to Francis starting in 1951, shared her data with him, she would have solved that structure. And then she would have been the famous one.

SA: In a century, we went from rediscovering Mendel's laws and identifying chromosomes as agents of heredity to having the human genome largely worked out. Finding the double helix drops neatly in the middle of that span. How much, with respect to DNA, is left for us to do? Are there still great discoveries to be made, or is it all just filling in details?

JW: The major problem, I think, is chromatin [the dynamic complex of DNA and histone proteins that makes up chromosomes]. What determines whether a given piece of DNA along the chromosome is functioning, since it's covered with the histones? You can inherit something beyond the DNA sequence. That's where the real excitement of genetics is now.

And it seems to be moving pretty fast. You don't really want to make a guess, but I'd guess that over these next 10 years, the field will be pretty played out. A lot of very good people are working on it. We have the tools. At some stage, the basic principles of genetics will be known in terms of gene functioning, and then we'll be able to apply that more to problems such as how the brain works.

SA: If you were starting out as a researcher now—

JW: I'd be working on something about connections between genes and behavior. You can find genes for behaviors, but that doesn't tell you how the brain works. My first scientific interest was in how birds migrated. Until you know how the bird brain works, you're not going to know how genes can tell that bird where to migrate. Because, you know, that mother bird isn't telling the young one where to go! So it's got to be inherited.

There are lots of other big behavioral things [to solve]. Some people say they're mystified that men can like men, but I say, "It's just as mysterious as why men like women!"

These things are so difficult. Francis insists that brain research doesn't have [the equivalent of] a DNA molecule. It doesn't have a central thing from which everything else flows.

SA: You have a reputation for being outspoken, and you get criticized for it. Do you have any regrets about things you've said?

JW: Occasionally. I think there's something in me of that same weakness that is so apparent in [tennis champion] John McEnroe. I just can't sit while people are saying nonsense in a meeting without saying it's nonsense!

WATSON and Francis Crick (right), *1953*

A. BARRINGTON BROWN *Photo Researchers, Inc.*

SA: On the subject of politics, many gene-related issues are in the public arena these days: genetically engineered foods, cloning, DNA fingerprinting and so on. How much confidence do you have in the political supervision of these?

JW: I think they're so contentious that the state shouldn't enter in. Yes, I would just stay out of it, the way [government] should stay out of abortion. Reproductive decisions should be made by women, not the state.

I mean, cloning now is the issue. But the first clone is not like the first nuclear bomb going off. It's not going to hurt anyone!

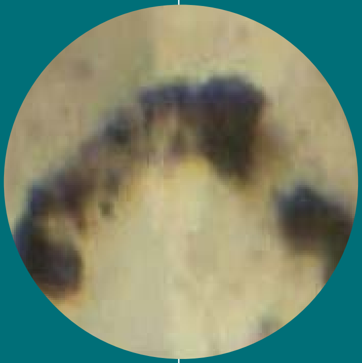
I know a famous French scientist who never had children because there was madness in his family. He didn't want to take a chance on more madness. That's what I mean. Cloning might mean you would know there wasn't going to be any more madness. I think the paramount concern should be the rights of the family, as opposed to the rights of the state.

People say, "Well, these would be designer babies," and I say, "Well, what's wrong with designer *clothes?*" If you could just say, "My baby's not going to have asthma," wouldn't that be nice? What's wrong with therapeutic cloning? Who's being hurt?

If you thought every plant was the product of a god who put it there for a purpose, you could say that you shouldn't change it. But America isn't what it was like when the Pilgrims came here. We've changed everything. We've never tried to respect the past, we've tried to improve on it. And I think any desire to stop people from improving things would be against the human spirit. SA

An extended version of this conversation and a list of suggested readings is available at www.sciam.com. Watson's latest book is DNA: The Secret of Life *(Knopf, 2003).*

In the past year scientists have been forced to reconsider how they identify life in the most ancient rocks on earth—and elsewhere in the solar system



ISUA, GREENLAND (*opposite page*), is home to rocks harboring well-accepted evidence that living cells existed on earth more than 3.7 billion years ago, only several million years after the planet itself formed. But other extremely old finds in Greenland and Australia—including microscopic squiggles thought to be fossil cells (*above*)—have recently been challenged.

Questioning the Oldest Signs of Life

By Sarah Simpson

Prehistoric creatures left seemingly endless proof of their existence fossilized in age-old rocks. Monstrous thighbones of lumbering dinosaurs that drowned in flooded rivers now lie encased in sandy mudstone. Jagged fronds of tropical ferns once growing in muddy swamps are pressed between jet-black layers of coal. Squiggly worm burrows, excavated within a slimy seabed, lace steel-gray limestone. These signs of life are unmistakably distinct from their stony tombs. But the more ancient the creature, the more obscure its grave.

Before life began walking, slithering or putting down roots, nothing much more than solitary microscopic cells populated the globe. Virtually all traces of those that lived before about 2.5 billion years ago—a period known as the Archean eon—have since become nearly indistinguishable from the rocks that entomb them. Millions of brutal years of burial and resurfacing, akin to repeated pressure cooking, permitted very few fossilized cells to survive in rocks accessible at the earth's surface today. Often geologists must instead rely on other signs of life, or biosignatures—including rather subtle ones, such as smudges of carbon with skewed chemical compositions unique to biology.

A decade ago high-powered microscopes enabled geologists to detect apparent fossils in earthly rocks an astonishing 3.465 billion years old—as close as tangible fossils may ever take us to the time when simple inorganic molecules first self-replicated and began selectively interacting with their environment, defining life's earliest moments. By 1996 new techniques for measuring slight variations in the chemical makeup of carbon samples seemed to solidify previous hints that life existed at least 365 million years earlier. That same year the stunning announcement that a meteorite found in Antarctica was carrying 3.9-billion-year-old biosignatures from Mars further energized scientists, who grew more confident in their ability to detect past life from faint traces in rocks.

But the uncontested glory did not last. Evidence from the Martian meteorite came under fire almost immediately, and only one life sign from Mars still clamors for acceptance [*see box on page 76*]. In early 2002 confidence over the widely accepted views on earthly evidence experienced a similar meltdown. A flurry of research has cast serious doubt on interpretations of the earth's two oldest geologic records of life, in Greenland and Australia. Some new evaluations of the geology suggest that the rocks formed in environments where life never could have thrived. Others question whether lifeless chemistry might have been able to mimic the special traces of carbon or even the shapes of microscopic fossils—rendering these clues useless as biosignatures.

The revisionist analyses are fueling an ongoing debate over how anyone can ever be sure of correctly identifying primitive life on earth—or elsewhere in the solar system. "If we can't get it right on earth, we can't get it right on Mars," says paleontologist Bruce Runnegar of the University of California at Los Angeles. That's a point that space scientists are sure to keep in mind when evaluating Martian rocks slated to be inspected by two NASA rovers early next year.

To Hell and Back

THE MOST ANCIENT—and easily the most controversial—evidence of life on the Blue Planet turned up seven years ago on a tiny nubbin of land at the southwest corner of Greenland's cold and barren island of Akilia. The island, which squats 30 kilometers south of the capital, Nuuk, extends a mere two kilometers across at widest reach; a hiker can traverse the ground in question in five minutes. There, underneath thick patches of arctic moss and lichens, all-important bands of milky, quartz-rich rock gleam among the darker volcanic slabs surrounding them. Using radioactive elements found only in volcanic minerals, scientists have dated one nearby slab to a whopping 3.83 billion years ago, ranking it among the oldest rocks preserved on the earth's surface today. And based on the position of the gleaming white rock (which contains no datable minerals), many geologists say that the white material is even older.

Akilia's stark landscape gives the impression that its rocks have looked this way ever since their primordial origins. But the earth is a dynamic place. These outcrops—like most from the Archean eon—have suffered one of the most torturous geologic processes that the earth has to offer: metamorphism. For 85 percent of the planet's lifetime, these rocks were buried, twisted, folded or pumped full of fluids; they were plunged to hellish depths nearly 70 kilometers underground and baked at 700 degrees Celsius before returning to the surface during at least two different episodes of uplift. If any of

Overview/*Early Life*

- Many traces of the earth's most ancient microscopic life—so-called biosignatures—are subtle and encased in rocks. Apparent biosignatures become most convincing as signs of life when geologists can confirm that the host rocks formed in a biologically friendly environment, such as a shallow sea.
- Several biosignatures can be mimicked by nonbiological chemical reactions that occur at high temperatures and pressures—reactions often experienced by deeply buried rocks or those that harden from a molten state.
- The ambiguity surrounding the origins of apparent biosignatures demonstrates that accurate interpretations of the geology are key to any search for signs of very ancient life, be it here on earth, on Mars or beyond.

J. WILLIAM SCHOPF *University of California, Los Angeles (page 70)*; JAMES L. AMOS *Corbis (preceding page)*

WHITE STRIPES in this rocky outcrop at the southern tip of Greenland's tiny island of Akilia hide flecks of carbon that were initially hailed as a signature of life older than 3.8 billion years.

CHRISTOPHER M. FEDO *George Washington University*

these rocks are the relics of the floor of an ocean once teeming with microscopic creatures, it could well be impossible to find traces of those organisms still intact.

Yet in 1996 geochemist Stephen J. Mojzsis, now at the University of Colorado at Boulder, glimpsed a suggestion of life inside that tortured white stone. Through the probing eye of a scanning electron microscope, he discerned black specks of graphite, a pure-carbon mineral that sometimes forms when organic matter is heated. He also noted that tough crystals of apatite encasing the graphite probably sheltered it from the harshest metamorphic transformations.

But what convinced Mojzsis were the uniquely skewed ratios of isotopes in the two dozen specks he analyzed in more detail. Each was diagnostically enriched in the lightest and most common isotope of carbon: carbon 12. Living organisms are frugal. So when they use carbon dioxide to fuel their activities, they exploit the light carbon isotope more efficiently than they do carbon 13, which is notably heavier because of an extra neutron in each atom's nucleus. This preference leaves them with a surplus of carbon 12 atoms—roughly 2 to 3 percent more than exists in carbon dioxide dissolved in the ocean.

This light carbon signature had been gaining support as an uncontestable marker of life for almost 60 years as researchers published thousands of similar measurements from younger rocks. Therefore, when Mojzsis's graphite samples clustered around 3.7 percent, it made perfect sense for him to declare them compelling evidence for the oldest known life on earth. This conclusion had an additional implication—that life got its start in a hostile period when devastating meteorite impacts were boiling off the oceans and turning the earth's atmosphere into a scorching mist of vaporized rock for millennia on end. Indeed, many scientists hailed Mojzsis's discovery as the key that would unlock a virtually unknown era of earth history, says geologist Christopher M. Fedo, now at George Washington University.

A year later, in 1997, Fedo accompanied Mojzsis and several other geologists to Akilia. Fedo recalls that at first it felt "like visiting hallowed ground." But almost immediately the two young researchers began seeing different pictures of the past—and different explanations for what the light carbon signal really means. From the makeup and structural relation of the rocks, Mojzsis and his colleagues had inferred that the graphite-bearing rock originated in a biologically friendly environment: an ocean basin where sand and other particles, including the cells of marine organisms, formed layers of quartz-rich sedimentary rock. On seeing the rocks for himself, Fedo, who had just spent a year mapping Archean rocks in Zimbabwe, became extremely skeptical. He knew that igneous rocks—those solidified from hot magma—can look sedimentary, and vice versa, once they have lost or gained key minerals during metamorphism. "If we're going to understand life, we'd better understand geology," Fedo says.

Subsequently, Fedo and geochronologist Martin J. Whitehouse of the Swedish Museum of Natural History in Stockholm returned to Akilia to make their own maps and chemical analyses. Their verdict, which they published last spring, was that the quartz-rich rocks that Mojzsis and others were calling old sediments were actually the progeny of igneous rocks that had endured a particular metamorphic process known to create graphite from nonbiological sources of carbon. Fedo and Whitehouse insisted that the graphite's light carbon signature might have to be explained by some process unrelated to biology. It is wrong, Fedo asserts, to believe that inorganic reactions cannot mimic light carbon signatures just because they haven't been proved to do that so far.

Geology Is Everything

ENORMOUSLY COMPLICATED geology is the reason that Mojzsis, Fedo and a slew of other investigators disagree about Akilia. It is also the reason that field geologist Minik T. Rosing of the University of Copenhagen's Geological Museum calls Akilia "utterly uninteresting." No combination of geologists can agree on the history of the rocks, he says, and so "we might never be able to resolve the problem." And that's from a Greenland native who has spent more than 20 years studying the icy island's geology.

But many investigators have not let go of settling the debate at Akilia. Indeed, Akilia is near the top of Bruce Runnegar's agenda as the new director of NASA's Astrobiology Institute, a consortium of 15 research teams across the U.S. dedicated to pursuing evidence for the origin and evolution of life on earth and beyond with an annual budget approaching $20 million. "Within a year or so we plan to take the people who really know the rocks and get some sort of consensus on what we're sampling in the field so that everybody knows what everybody else is talking about," he says.

Yet Rosing and other researchers—including paleontologist J. William Schopf of U.C.L.A.—point out that even if scientists can agree that the rocks at Akilia are former sediments, they still won't be able to prove where or when the carbon originated. They insist that light carbon, and indeed graphite, in such highly metamorphosed rocks can only suggest the possibility of life. By itself, it cannot constitute proof. When sediments are pressure-cooked, fluids might carry in carbon from other, younger sources. Also, the carbon bonds of any organic matter within them start to break and can be reset, even if the carbon is sheltered by tougher crystals of apatite. "There really is no good way, in my opinion, to go from [the measured] value back to what the original values were," Schopf says. "There's a big difference between knowing it and having a hint."

Akilia's limelight has diverted attention from a much more convincing hint of early life that exists about 180 kilometers northeast of that island, in a part of Greenland called Isua. There Rosing recently detected the light carbon biosignature in rocks that he argues experienced a mere migraine compared with Akilia's hellish past. This pocket of relatively gentle metamorphism was not easy to find. Isua's four-kilometer-wide belt of Archean rocks stretches 35 kilometers along the western edge of the bluish-gray monolith of the Greenland ice cap. The distinctive shimmer of the rocks and their marble-size crystals of red garnet, black

Biosignatures at a Glance

BEFORE PLANTS AND ANIMALS AROSE, single-celled microbes populated the earth. Scientists gather physical signs of these primitive organisms by combing ancient rocks for subtle traces of their existence, called biosignatures. But such finds can be questioned as evidence for life if their presence can be plausibly explained by nonbiological processes—as in the disputes described below. *—S.S.*

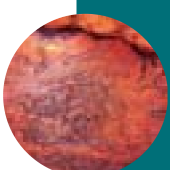
LIGHT CARBON

DEFINITION: Carbon having a higher ratio of carbon 12 to carbon 13 than occurs in nonbiological materials; the higher ratio reflects organisms' preference for using carbon 12 as they convert carbon dioxide into cellular material.
OLDEST DISCOVERY: Tiny specks of carbon found in Akilia, Greenland, in rocks more than 3.8 billion years old. Recent research contradicts assertions that the host rocks came from an environment that could have supported life. The debate leaves carbon (*black dots, left*) in rocks that formed more than 3.7 billion years ago in Isua, Greenland, as the oldest uncontested relic of life on earth.

LIGHT SULFUR

DEFINITION: Sulfur having a higher ratio of sulfur 32 to sulfur 34 than does sulfur that has not been processed by the microscopic organisms that use this element as a source of energy.
OLDEST DISCOVERY: Flecks of the iron-sulfide mineral pyrite within 3.5-billion-year-old rocks from northwestern Australia. Some researchers question whether the spiky gray crystals (*right*) that harbor the pyrite truly formed in an environment that could have sustained life. Records of this biosignature become unambiguous by about 2.5 billion years ago.

STROMATOLITES

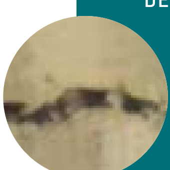


DEFINITION: Layered, domelike formations constructed by colonies of microbes.
OLDEST DISCOVERY: Fossilized mounds discovered in northwestern Australia and dating to about 3.5 billion years ago; these are the oldest known macroscopic representatives of life on earth (*left*). Most other stromatolites of this age are disputed as evidence for life because their simpler structures strongly resemble mineral layers that can be produced nonbiologically.

MOLECULAR FOSSILS

DEFINITION: Complex organic molecules resembling those in living cells today.
OLDEST DISCOVERY: Hydrocarbons found in Australian rocks some 2.7 billion years old. These molecules, derived from fossilized cell membranes, are the oldest undisputed evidence for eukaryotic cells (those containing a true nucleus) and for oxygen-producing cyanobacteria, possibly even the lineage that led to *Eoentophysalis* (*right*) 700 million years later.

MICROFOSSILS



DEFINITION: Remains of once living cells.
OLDEST DISCOVERY: Microscopic, carbon-rich squiggles (*left*) found in 3.5-billion-year-old rocks in northwestern Australia. These finds were originally interpreted as remnants of ancient microbes, but recent work suggests lifeless chemistry could have created them. Younger microfossils, including a two-billion-year-old cyanobacterium from Canada, are widely accepted.

BIOLOGICAL MINERALS

DEFINITION: Mineral grains produced by living cells.
OLDEST DISCOVERY: Unique forms of the magnetic mineral magnetite (*right*)—nearly identical to those known to occur in certain modern bacteria on earth—found in the Martian meteorite ALH84001. The Martian minerals are thought to be 3.9 billion years old; similar magnetite crystals have been detected in Australian rocks nearly two billion years old. Both finds remain in question.

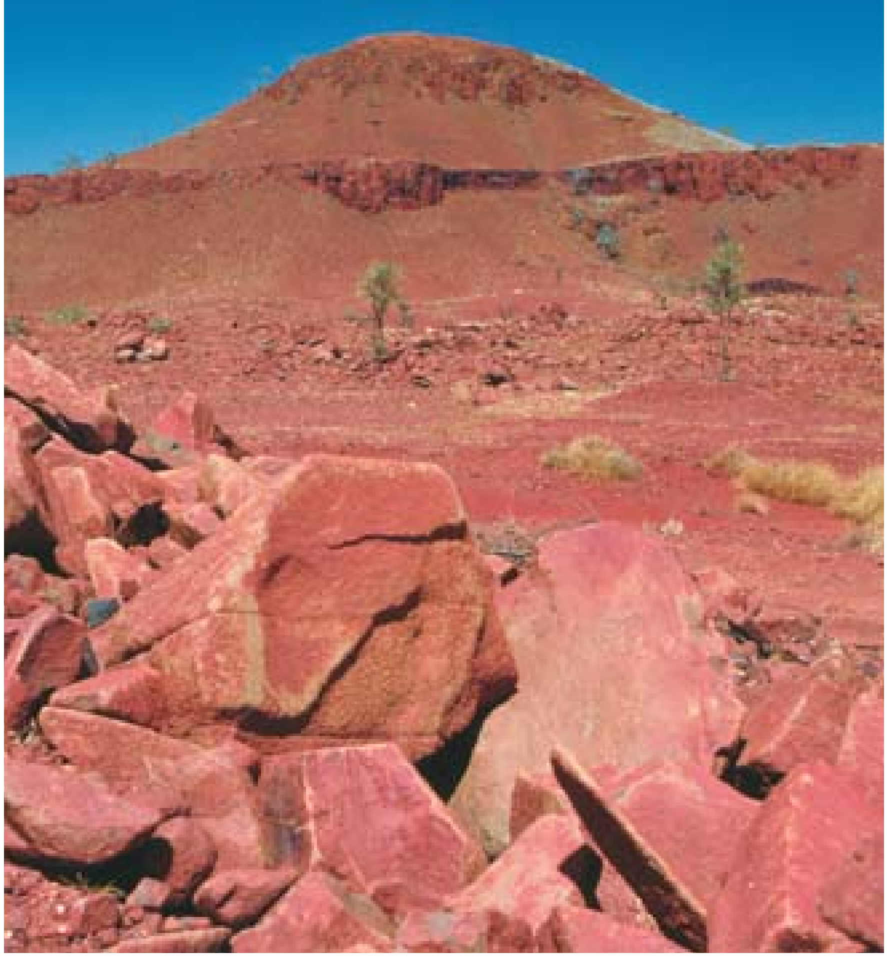
hornblende and sage-green diopside in most places attest to metamorphic tortures nearly as intense as those experienced at Akilia.

"I didn't go there to look for life, and I've always been skeptical about it," says Rosing, who initially visited Isua to understand how hot fluids transform the rocks. Since his first research visit in 1980, he has traipsed repeatedly among caribou, white polar hare and ptarmigan in this remote area—accessible only by helicopter—for one to three months at a shot. This intensive examination eventually enabled him to set aside rocks that were igneous or otherwise too complicated to harbor clear signs of life. Then, in 1999, he described a promising outcrop of old sediments at Isua's western edge that other geologists agree must be older than 3.7 billion years, making them potentially as old as Akilia's. And that's where he found the light carbon biosignature.

The relatively undisturbed geology in this part of Isua revealed important details about the sedimentary environment that were simply impossible to decipher at Akilia. The light carbon was tied up in layers of clay, which would have trapped organic particles when they settled to the bottom of an ocean. The carbon is also abundant in the rocks, and it persists through a thick pile of ancient sediments that represents as much as a million or more years of slow deposition. Every day, every night, for all that time, carbon with exactly the same composition of that in present-day microorganisms rained down to the bottom of a deep ocean, Rosing explains.

So far no one has seriously challenged his basic interpretations, and he has been working to characterize the microbes that might have produced light carbon. "That's what everyone thinks is the best bet for evidence of biology in Greenland," Runnegar says.

But even the best evidence in Greenland offers only a solitary sign of life. Surely a multiplicity of clues from a single site would be even more convincing. That was precisely the strength of Schopf's landmark interpretation of rocks on the other side of the world, one that went uncontested for nearly a decade.



MARS-LIKE LANDSCAPE of northwestern Australia's Pilbara region is the source of 3.5-billion-year-old microscopic structures that some interpret as fossilized strings of ancient cells.

Surf or Sizzle?

OF ALL POTENTIAL biosignatures, what pleases scientists most is a bona fide fossil of an organism's body, even if it is only one or two cells in size. In this category, Australia's timeless landscape holds most of the records. Uncontestable microfossils—including those of oxygen-producing cyanobacteria—exist in rocks about two billion years old. And convincing examples of so-called molecular fossils—relics of complex organic molecules that were once fatty constituents of cell membranes—turned up three years ago in rocks 2.7 billion years old. But neither of these exciting finds pushed back the fossil evidence as far as the pioneering work of Schopf, who, after devoting three decades of study to ancient microscopic fossils, launched a new wave of early-life research in 1993.

As in Greenland, the oldest signs of life in the land Down Under are exposed in a remote, desolate area—an ancient landscape about 1,200 kilometers north of Australia's west coast port of Perth. If it weren't for the wallabies hopping among silica-tipped spines of spinifex grass or the occasional roadhouse marking the spot where one dirt track meets another, the dusty, low hills of northwestern Australia might be mistaken for Mars. Near Marble Bar, a tiny watering hole in this sublime sea of red, geologists long ago described the Apex chert—the final home of Schopf's tiny, famous fossils—as a mixture of sand and small pebbles once churned by waves along a shallow seaway flanked by volcanoes. The chert (which, like the graphite-bearing rocks of Greenland, cannot be dated directly) is conveniently sandwiched between two lava flows, which have been dated precisely at 3.46 billion and 3.47 billion years old. Of the half a dozen claims of Archean microfossils—including ones from four locations in South Africa—this age made Schopf's the oldest. Further analysis would reveal that his cache, if truly biological, was also the most diverse, with 11 new species of microorganisms identified.

According to Schopf, the chert contained telltale smudges of graphite that he

OPPOSITE PAGE: MINIK T. ROSING *Geological Museum, University of Copenhagen* (*light carbon*), TIM LYONS *University of Missouri* (*stromatolites*), J. WILLIAM SCHOPF *U.C.L.A.* (*microfossils*), ROGER BUICK *University of Washington* (*light sulfur*), H. J. HOFMANN *McGill University* (*molecular fossils*), KATHIE L. THOMAS-KEPRTA *Lockheed Martin* (*biological minerals*); THIS PAGE: PENNY TWEEDIE *Corbis* calls threadlike strands of once living cells. His biological interpretation of these smudges was backed by the chert's distinct enrichment in light carbon and the nearby presence of fossilized stromatolites, mineralized mounds of bacterial mats that serve as the only sign of Archean life visible to the naked eye. Based on this trio of evidence, everything from textbooks to television, even the *Guinness Book of World Records*, touted this landmark finding as the earth's oldest fossil evidence for life.

But early last year Schopf's celebrated vision of the cradle of life was upended by a reinterpretation of the local geology—and of the fossils themselves. In March micropaleontologist Martin D. Brasier of the University of Oxford and seven of his colleagues published the first robust reanalysis of Schopf's 1993 conclusions. Among several challenges, Brasier asserted that the chert harboring the presumed fossils was not deposited on the sunny floor of a shallow sea but rather deep within the dark subsurface plumbing of seafloor hot springs. This distinction is critical because Schopf had proposed that many of his fossils may have been light-loving cyanobacteria. Even more damaging, however, was Brasier's suggestion that many of the microscopic structures that evoke life for Schopf may be nothing more than lifeless artifacts.

Today Schopf admits that it was a mistake to rely on the geologic mapping of others in the construction of his original story, conceding, albeit reluctantly, that his famed fossils may not have been photosynthesizers. Instead they may have been the forebears of heat-loving bacteria that color the steamy geyser pools in Yellowstone National Park and populate submarine volcanoes. Schopf has recently reevaluated the carbon smudges using a technique known as Raman imagery, which employed scattered light to probe their internal three-dimensional structures. The results indicate that many of the structures in question look like cell walls.

Still, Brasier asserts that the light carbon enrichments may well be able to form through lifeless chemical reactions—much as Fedo and others have argued could have occurred at Akilia. Little definitive research has been done in this area, but a handful of experiments do suggest that the right combination of metals and other chemicals—like those abundant in seafloor hot springs—might ignite reactions that could mimic biology's preference for the light isotope of carbon. And although Brasier does admit that some of the carbon could represent scattered remains of microbes, he insists that "you must falsify [nonbiological] origins for these materials before you accept biological ones."

The same policy must be applied, Brasier and others would argue, to another biosignature—light sulfur—preserved within 3.47-billion-year-old Australian rocks that sit a two-hour, backbreaking car ride northeast of the controversial Apex chert. Sulfur-rich rocks from the ironically named North Pole district contain a surplus of sulfur 32 (relative to heavier sulfur 34) that is characteristic of waste materials produced by bacteria known to use sulfur as an energy source. Light sulfur signatures, like those of carbon, unambiguously record life over much of the earth's history. But in these very old rocks in western Australia,

Martian Magnets

METEORITE ALH84001 hit Antarctica 13,000 years ago.

PULLING CONVINCING SIGNS OF LIFE out of a single space rock isn't easy. But one of the last surviving claims that biosignatures endured meteorite ALH84001's trip from Mars has weathered a recent barrage of criticism.

For the past seven years microscopist Kathie L. Thomas-Keprta of Lockheed Martin in Houston and her colleagues have kept alive the idea that minuscule grains of the mineral magnetite within the potato-size meteorite—bits widely accepted to have formed on Mars some 3.9 billion years ago—are indistinguishable from the tiny magnets made by some aquatic bacteria on earth.

Skeptical researchers have pointed out that, as is true of certain putative biosignatures on earth, a lifeless chemical process could have created the lifelike material, in this case a heat-induced transformation of iron-rich minerals during a collision the rock suffered while still on Mars. But such an impact would have created magnetite with impurities (such as magnesium and manganese), Thomas-Keprta notes, and the grains she calls biosignatures are 100 percent pure—a finding she and her colleagues recently confirmed with new three-dimensional tomography scans of the Martian magnetite.

In addition to the lack of chemical impurities, about 25 percent of the magnetite in ALH84001 shares at least five other distinct characteristics with magnetite made by the bacterial strain MV-1 here on earth. The Martian and MV-1 magnetite grains are in the same size range, lack significant structural defects and share an unusual elongated crystal shape that strengthens the mineral's magnetic properties, for instance. If any single criterion is absent from a Martian grain, it is excluded from further consideration as a possible biosignature. The test is so stringent that nearly one third of the bacterial crystals would fail to pass the test, Thomas-Keprta points out.

MARS

The new evidence makes her almost certain that the tiny magnets are true signs of past life on Mars, but many other scientists remain unconvinced. Even Thomas-Keprta says she won't be satisfied with finding only one likely trace. "Defining a biosignature is almost as difficult as defining life itself," she admits. In the next year she and her colleagues will begin to scour the meteorite for a particular iron sulfide mineral that bacteria on earth are known to produce. *—S.S.*

SETI LEAGUE, USED BY PERMISSION *www.setileague.org* (*top*); NASA AND THE HUBBLE HERITAGE TEAM (*bottom*)

STROMATOLITES, towering structures built by colonies of microorganisms, live today in such places as Shark Bay, about midway along the western coast of Australia. Ancient versions dating to about 3.5 billion years ago are among the few putative signs of life that old that are still largely unchallenged.

MITSUAKI IWAGO *Minden Pictures*

the controversy is the same: Did the rocks form in low-temperature environments inhabited by bacteria or in higher-temperature locales where nonbiological reactions could have mimicked the bacterial isotope patterns? Australian geologist Roger Buick of the University of Washington and his collaborators argue the former, stating that sulfur-bearing crystals formed in an evaporating lagoon. But not everyone, most notably Runnegar, agrees with this interpretation of the region's geology.

Despite the controversies, Schopf maintains his basic position. He counters that although an individual biosignature can be cast in an uncertain light, the uncertainty does not render the evidence useless. A suite of biosignatures from a single location—even if disputable when viewed individually—packs a powerful punch. He is fond of saying, "If it looks like life, has the ecology of life, has the isotopes of life and fits with all other evidence of life, then most likely it's life."

Earth and Beyond

STILL, IN THE END, the interpretations of the oldest rocks in both Greenland and Australia are unavoidably complicated by the possibilities of both biological and nonbiological origins. Lest you worry that scientists are losing their ability to recognize early life, though, remember that the brouhaha over Akilia and Apex are just about being the very oldest signs of life. That's important but not the be-all and end-all. Rocks at Isua and South Africa's Transvaal Basin are just a tad younger and, some would argue, much less controversial. Although many scientists quibble over the details, the great antiquity of life is generally accepted.

Perhaps the most important conclusion that has crystallized from these arguments is this: whether you are investigating ancient rocks on earth or potato-size meteorites from other planets, don't count on a single smoking gun to be your proof of life. That conclusion has serious implications for further hunts for early life on earth and for how future evidence from Mars is interpreted. Brasier cautions that without having "the criteria for the detection of early life clear in our minds *before* we have a robotic or manned mission to Mars ... we will end up having profitless debates that may simply demoralize the scientific community."

NASA scientists share that concern, which is why they are determined to decipher the geology of Mars before tackling a search for past life. "When you consider the fact that field geologists have been crawling over the earth for 200 years and are still having problems [agreeing on reliable biosignatures], we're way far away from being able to do a credible search for biosignatures on Mars," says planetary scientist Steven W. Squyres of Cornell University.

Since 1997 Squyres has been working as chief scientist for NASA's upcoming Mars landing missions, which will scour the planet's surface for hints that its past environment might have been biologically friendly. Two remote-controlled, robotic geologists, called the Mars Exploration Rovers, are scheduled for launch in late May and June and should begin their fieldwork on the Red Planet in January 2004. The landing sites, to be announced sometime this month, will target spots where orbiting spacecraft have discovered tantalizing hints that liquid water—a requirement for all known forms of life—once existed.

As experiences from Greenland and Australia illustrate, it is difficult to find readable records here on earth because the constant motion of tectonic plates have so chewed them up over the past four billion years. But because such a global process may never have existed on Mars, researchers predict that its surface has remained intact—except for a few meteorite impacts—during that same period. Squyres notes the irony: "If life did come to being on Mars, evidence for it might be much easier to find." SA

Sarah Simpson is a contributing editor at Scientific American.

MORE TO EXPLORE

Evidence for Life on Earth before 3,800 Million Years Ago. Stephen J. Mojzsis et al. in *Nature,* Vol. 384, pages 55–59; November 7, 1996.

^{13}C-Depleted Carbon Microparticles in >3700-Ma Sea-Floor Sedimentary Rocks from West Greenland. Minik T. Rosing in *Science,* Vol. 283, pages 674–676; January 29, 1999.

Cradle of Life: The Discovery of Earth's Earliest Fossils. J. William Schopf. Princeton University Press, 2001.

Questioning the Evidence for Earth's Oldest Fossils. Martin D. Brasier et al. in *Nature,* Vol. 416, pages 76–81; March 7, 2002.

Metasomatic Origin of Quartz-Pyroxene Rock, Akilia, Greenland, and Implications for Earth's Earliest Life. Christopher M. Fedo and Martin J. Whitehouse in *Science,* Vol. 296, pages 1448–1452; May 24, 2002.

Magnetofossils from Ancient Mars: A Robust Biosignature in the Martian Meteorite ALH84001. Kathie L. Thomas-Keprta et al. in *Applied and Environmental Microbiology,* Vol. 68, No. 8, pages 3663–3672; August 2002.

By linking digital processors, storage systems and software on a global scale, grid technology is poised to transform computing from an individual and corporate activity into a general utility

THE GRID: Computing without Bounds

By Ian Foster

Early humans ate only the food they caught and used only the tools they made.

Millennia of effort to overcome the coupling of production and consumption eventually led to agriculture, mass manufacturing and electric power distribution. The resulting specialization of work, economies of scale and novel technologies define our modern world—and allow me to sit in a café sipping an espresso and writing this article on a laptop computer, with no thought to the ultimate sources of the water, coffee, electricity and wireless-network bandwidth that I consume.

The ready availability of these resources exemplifies the concept of virtualization, which (to computer scientists) refers to hiding useful functions behind an interface that conceals the details of how they are implemented. When the café's *barista,* for example, turns on a water spigot, it is as if he taps a bottomless barrel. The same phenomenon occurs when I plug my laptop into a wall socket. Given the huge unseen electric grid beyond the plug, who knows how and where that power was generated?

Computing itself, however, is anything but virtualized. My laptop, a home computer and even a sophisticated corporate data center are largely self-contained systems running locally stored software programs. We would not accept a situation in which every home and business had to operate its own power plant, library, printing press and water reservoir. Why should we do so for computers?

This less-than-optimal state of affairs has motivated computer scientists to look for more efficient alternatives. Digital networks are getting faster and faster, so why not assemble "computers" dynamically from distributed pieces so that users can call on resources—processing, storage, data and software—regardless of location and from any suitable supplier? That is, why not virtualize general computational services? Such a computing "grid" could be as ubiquitous as the electric grid—and just as useful.

Grid Business

LIKELY BENEFICIARIES of a pervasive computing grid would include e-commerce enterprises, which could tailor their information and computing systems to changes in demand while simultaneously linking those systems to partners, suppliers and customers to expand the services they receive. Imagine an adventure travel firm that provides virtual-reality explorations of its dive locations. As a potential customer starts to navigate the virtual waters off Cozumel, the company needs to access databases and software to retrieve relevant geographic information, render data into three-dimensional graphics, superimpose the appropriate commercial information and integrate real-time video from on-site cameras. No travel company could afford to operate these capabilities itself. But with grid technologies, the necessary resources can be assembled from multiple suppliers, each serving many clients and thus achieving favorable economies of scale [*see box below*].

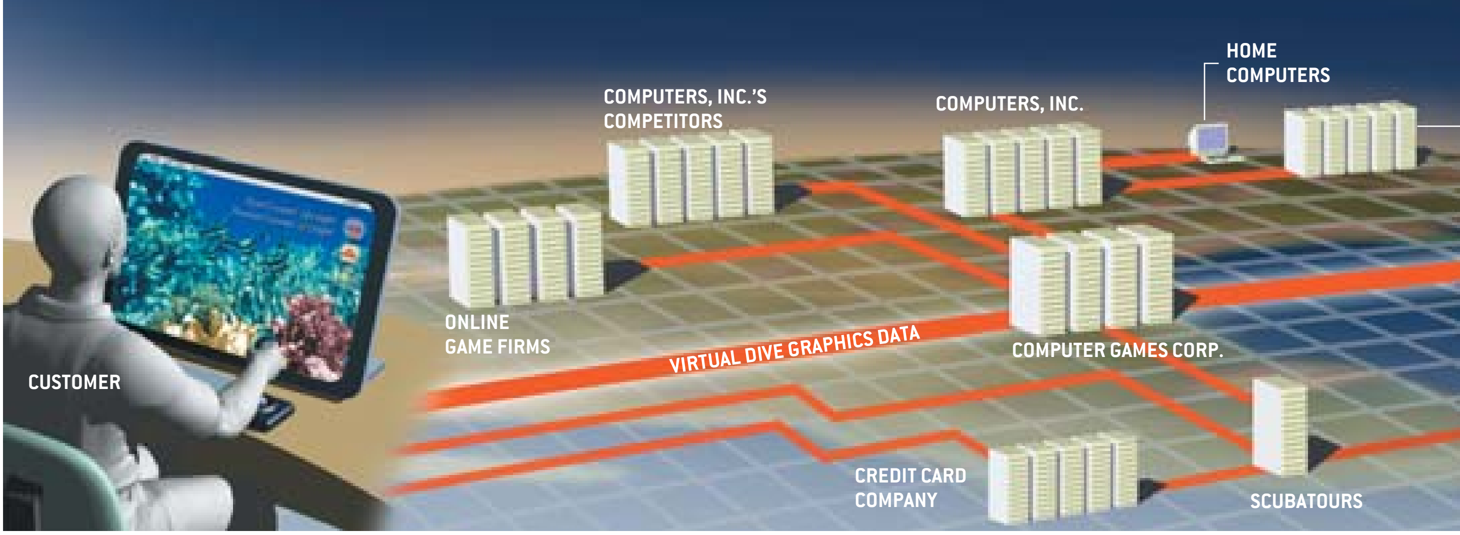
Grid systems hold promise for more esoteric domains as well. In medicine, access to remote computers and databases could allow my physician to compare my x-rays with those of a million other patients, perhaps enabling him to catch a

SLIM FILMS (*preceding pages*); GEORGE RETSECK (*below*)

GRID-ENABLED VIRTUAL-REALITY DIVE TOUR

TO VISUALIZE what a full-fledged grid-computing system could do in the commercial arena, imagine an adventure travel firm (ScubaTours) that provides virtual-reality explorations of its diving locations. This service allows a potential customer armed only with a home computer and joystick to "fly through" a detailed three-dimensional representation of a coral reef—in this case, one lying off the Mexican resort of Cozumel.

The small company can manage to deliver this supercomputerlike capability by employing the grid to contract out the task of assembling the far-flung computational resources needed to produce a realistic simulation. When a potential customer wants to navigate the virtual Caribbean waters, ScubaTours's computers contact remote servers that handle the details of accessing remote databases and software. These



- Grid computing refers to the large-scale integration of computer systems (via high-speed networks) to provide on-demand access to data-crunching capabilities and functions not available to one individual or group of machines.
- Using shared languages and interaction protocols, grid systems reach out across the globe to access the computing resources, information and services required to satisfy local user needs.
- To users, the highly integrated networks that embody grid systems are transparent—it seems as if the services furnished from afar are provided by local computers.
- Grid technology enables large-scale scientific and business collaboration among members of virtual organizations, remote experimentation, and high-performance distributed computing and data analysis.

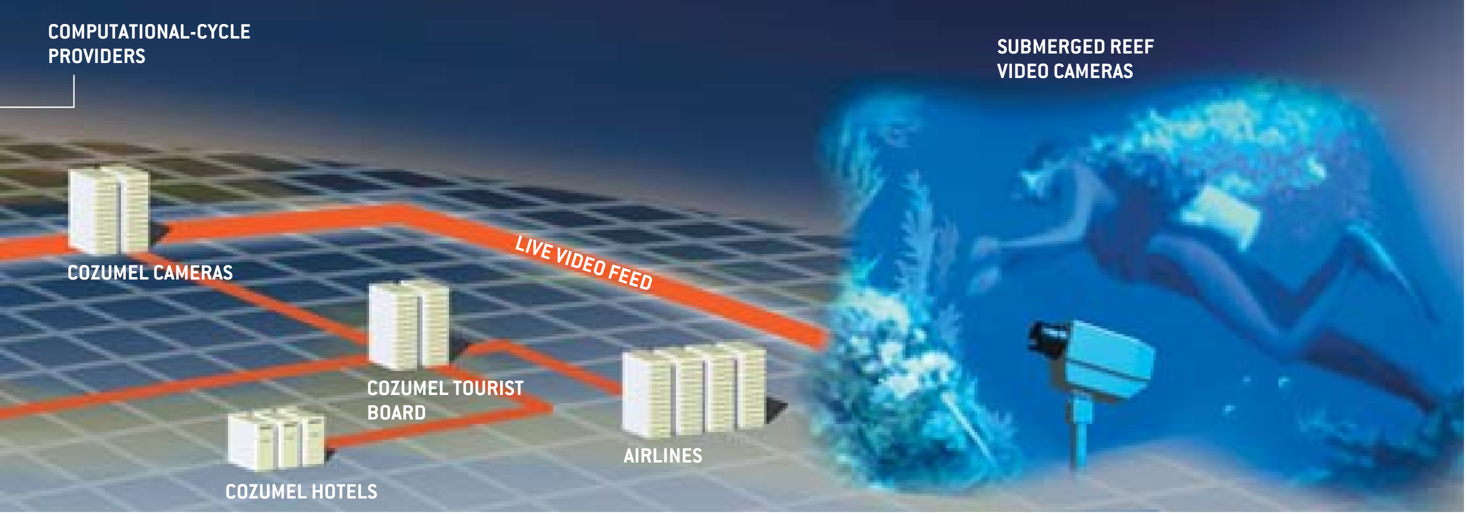
malignancy earlier. Elsewhere it might let biochemists screen 10,000 drug candidates in an hour rather than in a year or allow civil engineers to test an earthquake-resistant bridge design in minutes rather than months.

This concept of globally virtualized grid computing is a natural extension of today's Internet. The Internet virtualizes communications, permitting any person or device to connect with any other person or device, regardless of location or the means used to do so. The result has been an explosion of innovative functions: e-mail, the World Wide Web, peer-to-peer applications, including file-sharing systems such as KaZaA, and simple distributed-computing schemes such as SETI@home, FightAIDS@home and the Smallpox Research Grid. The goal of those of us now developing the grid is to virtualize computing and information so that any person or device can furnish software services to any other and—equally significant—so that access to disparate collections of such services will be secure and reliable.

My interest in grid computing arose through a fascination with how computational tools can enhance human intellect. In the early 1990s I was working on tools for scientific computing at the U.S. Department of Energy's Argonne National Laboratory. High-speed networks were starting to appear, and it became clear that if we could integrate digital resources and activities across networks, it could transform the process of scientific work. Lab instruments, for instance, could be linked with remote computers to enable real-time data analysis. And multiple distributed databases could be combined to search for interesting correlations.

Thus, in 1994 I decided to refocus my research on distributed computing. With partners Steven Tuecke of Argonne and Carl Kesselman, then a research fellow at the California Institute of Technology and now director of the Center for Grid Technologies at the University of Southern California's Information Sciences Institute, I established a project to create a software system that would enable scientific collaboration on a worldwide scale. We gave this effort the somewhat grandiose name the "Globus Project."

servers thus retrieve relevant cartographic and image information, render data into three-dimensional graphics, integrate real-time video from on-site cameras, and superimpose appropriate commercial information. All these assets, as well as the software that generates a compelling virtual experience, are obtained via a mix of long-term contracts and competitive bidding from multiple suppliers and brokers (such as Computer Games Corp., and Computers, Inc.), each serving many clients and thus achieving favorable economies of scale. Other providers on the grid (*not shown*) offer reputation management, billing, and similar functions. [*See page 85 for a more detailed explanation of how grid systems could enable this sophisticated business application.*]



SELECT INTERNATIONAL GRID PROJECTS

NAME	WEB ADDRESS	SPONSORING INSTITUTIONS	FOCUS
Biomedical Informatics Research Network	www.nbirn.net	U.S. National Institutes of Health	Integration and analysis of biomedical data
Department of Energy Science Grid	doesciencegrid.org	U.S. DOE Office of Science	Links resources and applications at DOE science laboratories
Earth System Grid	www.earthsystemgrid.org	U.S. DOE Office of Science	Analysis of large climate-model data sets
European Grid of Solar Observatories	www.mssl.ucl.ac.uk/grid/egso	European Union	Integration of research by European solar-data centers
European Union DataGrid	www.eu-datagrid.org	European Union	Applications in high-energy physics, environmental science and bioinformatics
GridLab	gridlab.org	European Union	Grid technologies and applications
Grid Physics Network	griphyn.org	U.S. National Science Foundation	Data analysis for large-scale physics experiments
Information Power Grid	www.ipg.nasa.gov	NASA (U.S.)	Aerospace sciences research support
MyGrid	www.mygrid.org.uk	U.K. e-Science program	Workbench for bioinformatics applications
National Fusion Grid	www.fusiongrid.org	U.S. DOE Office of Science	Computational fusion research
National Research Grid Initiative	www.naregi.org	Japanese Ministry of Education	Japanese national research grid project
National Virtual Observatory	www.us-vo.org	U.S. National Science Foundation	Integration and analysis of astronomical data
Network for Earthquake Engineering Simulation	www.neesgrid.org	U.S. National Science Foundation	Earthquake engineering research
Particle Physics Data Grid	ppdg.net	U.S. DOE Office of Science	Data analysis in high-energy and nuclear physics experiments
Singapore BioGrid	www.bic.nus.edu.sg/biogrid	Singapore government	Bioinformatics
TeraGrid	teragrid.org	U.S. National Science Foundation	High-speed infrastructure links among four major science-resource sites
WorldGrid	ivdgl.org/demo/worldgrid	U.S. DOE, National Science Foundation, European Union	International infrastructure for data-intensive science

Distributed computing is by no means a new idea: many concepts that underlie today's grid systems predate even the Internet. Back in the mid-1960s, for example, Fernando Corbato, the father of time-sharing operating systems, described the then revolutionary Multics system as a "computing utility." The banking and airline industries have run sophisticated distributed systems for decades. We, however, approached the problem from a different perspective, one defined by the needs of scientific research communities. In our experience, the often extreme requirements and only partially controlled chaos of scientific investigation can be strong drivers for innovation. It seems no coincidence that the World Wide Web was invented by Tim Berners-Lee, a computer scientist whose work with high-energy physicists inspired him to create a universal system for sharing information.

In our case, we saw that scientific communities require technologies that allow groups of collaborators located at different institutions—what we called virtual organizations (VOs)—to share resources in a controlled and well-managed manner. Rather like professional soccer players pulled from separate clubs to train together as a national team to compete in the World Cup, such scientific VOs cut across conventional institutional boundaries, creating, of course, myriad organizational and policy problems. VO participants in particular need to share information, computers, storage and software in a controlled fashion as they engage in collaborative work, but unlike the soccer players, they might not wear the same uniforms, speak the same language, follow the same rules, or perhaps even play the same game!

Kesselman, Tuecke and I envisioned a new class of integrating software that would enable this sharing, not by replacing systems at participating sites but rather by linking them into a VO structure. The software would accomplish this task by standardizing a variety of functions—most important, the ability to authenticate identity, to authorize a requested activity, to define and access available resources, and to control the movement of data. Because these VO utilities would operate on top of existing computational structures, they could achieve these goals at an acceptable cost.

Early Grid Work

OUR IDEAS WERE PUT to the test sooner than we anticipated. In late 1994 Rick L. Stevens, director of the mathematics and computer science division at Argonne, and Thomas A. DeFanti, director of the Electronic Visualization Laboratory at the University of Illinois at Chicago, proposed establishing temporary links among 11 high-speed research computer networks to create a national grid (the "I-WAY") for the two weeks before and during an industry conference called Supercomputing '95. Their call for applications for the system generated more than 60 responses from scientists.

Stevens and DeFanti also persuaded me and a small team of co-workers at Argonne to develop the software that knitted the 17 participating I-WAY sites into a single virtual system. These protocols allowed users to run applications on computers across the country. Users could log on once, locate suitable computers, reserve time, load application codes, and then monitor their execution. In one example, a group led by Lori Freitag of Argonne constructed a network that allowed combustion engineers across the U.S. to work collaboratively on improvements to industrial incinerators.

The I-WAY experiment was a great success and inspired much that followed. The U.S. Defense Advanced Research Projects Agency gave the Globus Project money for more research. In 1997 we unveiled the first version of our grid software system, the Globus Toolkit, and demonstrated its operation across 80 sites worldwide. Meanwhile the National Science Foundation funded the creation of the National Technology Grid to connect university scientists with high-end computers, NASA started work on its Information Power Grid (which had similar goals), and the DOE initiated pioneering work on applying grids to science research.

The remainder of the 1990s saw growing implementation of grid concepts and technology. The high-energy physicists designing the Large Hadron Collider at CERN, the European organization for nuclear research, for instance, recognized that a grid system would be essential to their analysis of the huge quantities of data the new accelerator was to produce. Projects were subsequently established to develop the needed grid technologies, notably the European Data-Grid, as well as the U.S. Grid Physics Network and Particle Physics Data Grid efforts. These and similar undertakings spurred the development of grid-related infrastructure, user communities and new software applications.

The international nature of science makes nearly seamless interoperability important to any resource-sharing infrastructure. It has helped that almost all efforts to develop grid systems build on the Globus Toolkit, but it was clear that well-defined, community-based technical standards would be essential if grid computing were to expand beyond its initial users. Therefore, in 1998 several of us convened the first meeting of what two years later became the Global Grid Forum, an international user-community and standards organization.

Exploiting the Grid

GRID TECHNOLOGIES are infrastructure, which doesn't make for compelling reading—after all, when infrastructure works, it is as invisible as the sewage system. The grid infrastructure has made possible some virtualized projects that *are* compelling, however.

Ever since Galileo pointed his telescope at Saturn, long, cold nights in the observatory have been the astronomer's lot. Yet advances in computing and sensor technologies are now enabling a new generation of "armchair astronomers" to make significant advances. Rather than freezing in nighttime isolation, these astronomers sit in comfortable offices during the day, instructing computers to mine digital sky surveys. For these New Age astronomers, a principal obstacle to progress is not a shortage of telescope time but the difficulty in finding the software, storage, network and computing resources they need to analyze the vast amounts of data these surveys produce.

A team at the University of Chicago, Fermilab and the University of Wisconsin–Madison has been looking for ways to apply grid technologies to the analysis of data from one of these new archives, the Sloan Digital Sky Survey. This extensive effort to map a quarter of the night sky will determine the positions and absolute brightnesses of more than 100 million celestial objects. By harnessing computational resources at labs across the nation, we aim to allow scientists to perform what are now weeklong scans in the time it takes to brew a cup of coffee.

This group has already generated a database of galaxy clusters, which interests cosmologists working on theories describing the birth of the universe. Next is a search for near-Earth objects, a topic that concerns us all, given that asteroids will one day strike our planet. This effort feeds into a far more ambitious international undertaking focused on linking many such astronomical databases into a global virtual observatory.

Grid technologies are also making their mark in clinical medicine, where digital imaging has made it easier to compare mammograms and other body scans both across time for individual patients and across case populations. But the improved scanning systems mean contending with a deluge of visual data. Currently experts estimate that one fifth of first-time mammograms are misdiagnosed. Further, in 20 percent of all cases, previously prepared images cannot be located. Digital image libraries now being created by projects such as the National Digital Mammography Archive in the U.S. and eDiamond in the U.K. are addressing these problems. These grid systems also link physicians with advanced analytical tools for automated diagnosis and permit researchers to conduct in-depth studies to determine the effects of environment and lifestyle choices on disease [*see box on next page*].

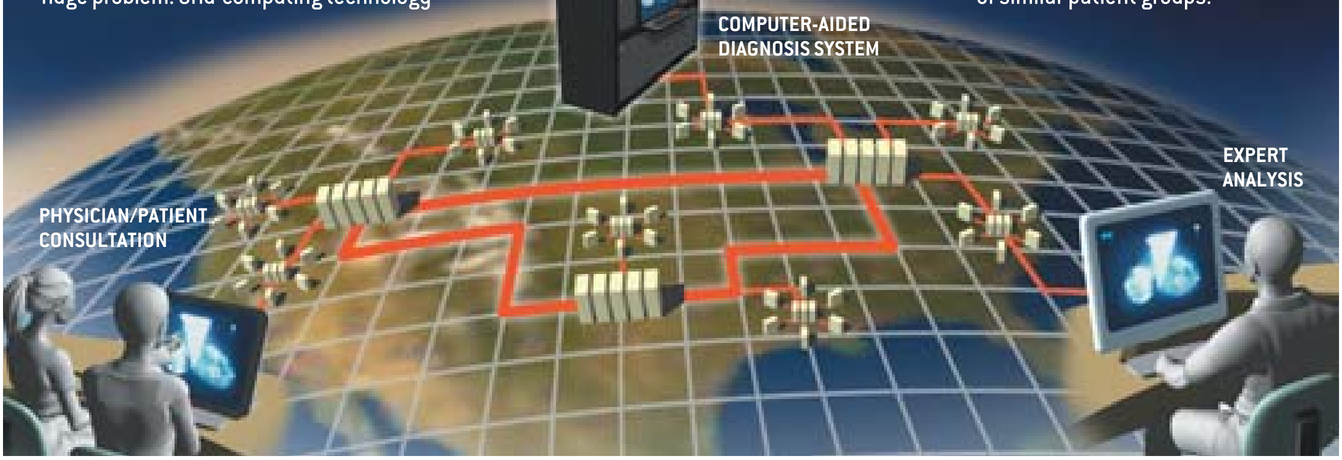
Another health-related imaging grid, the U.S. Biomedical Informatics Research Network, is integrating body-scan data from various sources to allow comparative studies of brain images from many databases. This grid may help answer questions such as "How are brain structures in Alzheimer's patients different from those organs in healthy people?"

THE AUTHOR

IAN FOSTER is originally from New Zealand. He now resides in Chicago, where he is associate director in the mathematics and computer science division at Argonne National Laboratory and professor of computer science at the University of Chicago. Foster also serves on the advisory boards of several companies, including Entropia and Insors. His many honors include the Global Information Infrastructure Next Generation award and the British Computer Society's Lovelace Medal.

DISTRIBUTED ARCHIVE OF DIGITAL BREAST IMAGES

BY ENABLING ACCESS to all kinds of computational resources from afar, grid computing is transforming the way enormous processing tasks are accomplished. Digital mammography is a case in point. Computer-based breast imagery yields more accurate diagnoses than do film-based photographs. But storing and accessing the estimated 28 terabytes of image data that can be generated every day in the U.S. is a huge problem. Grid-computing technology being implemented for the National Digital Mammography Archive is creating a vast virtual file cabinet that can not only reliably store and transfer these images with little loss of detail but can make possible real-time expert consultation and the use of computer-aided diagnosis techniques from across the country. Thus, a patient's mammogram can be readily compared both with her previous images and with those of similar patient groups.



The U.S. Network for Earthquake Engineering Simulation (NEES) shows how grid technologies can support experimentation. Civil engineers designing earthquake-resistant structures test their concepts in experimental facilities housing equipment such as shake tables and centrifuges. Part of the $82-million program focuses on assembling a NEESgrid that links existing and newly constructed test facilities with data archives, computational resources and the nationwide community of users. NEESgrid makes it possible for researchers at remote sites to participate in experiments taking place elsewhere. Planned for this summer, for example, is a study of bridge columns that will be tested simultaneously on shake tables in California and Illinois. Networks will communicate the relevant force-feedback levels, generating data about how structures behave in earthquakes.

Laying Down the Grid

SCIENTISTS ARE NOT the only ones who are excited by the potential of grid computing. Since 2000, more and more companies have pursued commercial applications for grid technologies, including Avaki, DataSynapse, Entropia, Fujitsu, Hewlett-Packard, IBM, NEC, Oracle, Platform, Sun and United Devices. Phrases such as "utility computing," "e-business on demand," "planetary computing," "autonomic computing," and "enterprise grids" have become increasingly common in marketing literature and business plans.

To understand this growing interest, we must look at the high costs and endless frustration associated with operating computers, an endeavor that has never progressed much beyond what is essentially still a cottage industry. We all know that users spend far too much time installing, maintaining, troubleshooting, fixing and upgrading their private systems, which furthermore do not work well together. Grid computing offers the promise of overcoming these problems by decoupling production and consumption, thereby enabling specialized functions and common services to be turned into commodities.

Progress toward this goal will be anything but straightforward. Similar transitions, however, have occurred in fields such as manufacturing, electric power and telephony. The automobile, for example, was initially a luxury plaything of expert hobbyists. Today cars and trucks are driven by nearly everyone in the developed world, with support from a vast global production and service network.

Equivalent success in the computing industry could lead to a computational ecosystem made up of manufacturers, distributors and consumers. Producers could gain economies of scale and improve security and reliability. Distributors could focus on brokering consumer demand with available resources and on finding novel forms of value-added services. Consumers could avail themselves of a rapidly developing array of applications and devices powered by new remote services.

Many innovative technologies are required to make this vision of grid-enabled computing a reality. The Internet and the World Wide Web allow us to send messages and access Web pages from just about anywhere, but much more is required if users are to share resources on a large scale. In effect, we have speech but not yet a common language. Collaborators need to agree on how services are described, how parties establish one another's identity, how access rules are described and verified, and how activities are conducted.

The grid community took a significant step toward addressing these issues with the launch in 2002 of an initiative to define the Open Grid Services Architecture. This system architecture integrates

GEORGE RETSECK

Web services technologies that have become popular in industry with the grid techniques pioneered by the scientific community.

Plugged into the Grid

TO ILLUSTRATE what a full-fledged open architecture can do, let's return to our adventure travel company, ScubaTours [*see box on page 80*]. What happens when a customer asks for a virtual dive off Cozumel? ScubaTours first establishes that the customer has a broadband connection, determines what the customer wants to see, and then hands responsibility for fulfilling the request to a specialized virtual-reality service provider, Computer Games Corp. (Computer Games Corp. also supports multiparticipant Internet gaming companies, hence its name.) Computer Games Corp. in turn assembles the data and software required to deliver the desired "fly-through"—streaming video generated in response to the user's navigational commands. Its computers reach out into the grid to locate the most attractive supplier of computational cycles for the rendering of the video stream. Computer Games Corp. then installs this software, and the virtual dive gets under way. In effect, Computer Games Corp. acts as a middleman, aggregating services from several sources to afford a specialized capability to the user.

Although these negotiations should appear transparent to the customer, making them work is far from easy. Computer Games Corp. engineers must have already built specialized software able to run on multiple computers, and they must know how to estimate and then locate the computing power and network capacity they need. To this end, they contact a trusted "broker" who serves as a clearinghouse for computing services. Having located a suitable candidate, say, Computers, Inc., they can contact it directly using some common language. Then they negotiate to establish mutually acceptable bona fides ("I am not a hacker" and "I am able to pay" and "You are an established company") and to determine the terms of the transaction (the number of computers, performance, cost and so forth).

Once the engineers have negotiated access to equipment at Computers, Inc., Computer Games Corp. must load its software and manage the virtual dive experience. Perhaps the customer will head in an unexpected direction, increasing computational requirements, or perhaps some computers or networks will fail. In either case, Computer Games Corp. has to incorporate additional computers or reduce the fidelity of the simulation.

Notice that these interactions require standard languages for describing, advertising and querying the capabilities supplied by resource operators and for making conversations among the participants possible. Mechanisms for establishing identity, accounting for usage, organizing payment, and so on are also needed. Software for accomplishing many of these tasks is included today in the Globus Toolkit, and more general standards are being disseminated by Global Grid Forum working groups.

I expect standardization to spur innovation and competition in computing services, just as it has in other industries. It could become feasible to use aggregated desktop or even home computers as a significant computing and storage resource. We find, for instance, that about half the computing resources employed in the Sloan Digital Sky Survey were from desktop systems aggregated by a system called Condor. Companies such as DataSynapse, Entropia, Platform and United Devices are implementing grid protocol interfaces to desktop computer pools to enable virtualized computing.

The success of grid computing depends on its broad adoption. One proven strategy for overcoming this classic catch-22 is to provide free access to core technologies. Hence, it is important that not only all key grid specifications be freely available to everyone but that an open, easy way to implement those specifications exists. The Globus Toolkit fills this requirement. Further progress depends on contributions from academia and industry and from commercial support vendors, as well as on integration of grid concepts into commercial software and third-party training.

Where Next?

HAVING ATTAINED success within the relatively friendly confines of scientific collaborations and corporate intranets, grid technologies are expanding to larger communities, including those based on purely commercial relationships. Standardization is spurring investment and innovation. During the next few years, this work will establish the basis for a global grid system.

Meanwhile grid researchers are tackling the next set of challenges: How can we manage large, distributed infrastructures so that they deliver reliable service in the face of failures? How can we enable users to exploit the availability of on-demand resources and services? How must grid concepts and technologies evolve as the number and power of computers rise by orders of magnitude? The answers will emerge from both research and practical experience and will surely draw on ideas being pioneered in the related fields of autonomic, ubiquitous and peer-to-peer computing [see "The Worldwide Computer," by John Kubiatowicz and David P. Anderson; SCIENTIFIC AMERICAN, March 2002]. These disciplines are all converging on the same vision of the computing environment of tomorrow. SA

MORE TO EXPLORE

The Grid: Blueprint for a New Computing Infrastructure. Edited by I. Foster and C. Kesselman. Morgan Kaufman, 1999. (Second edition in press.)

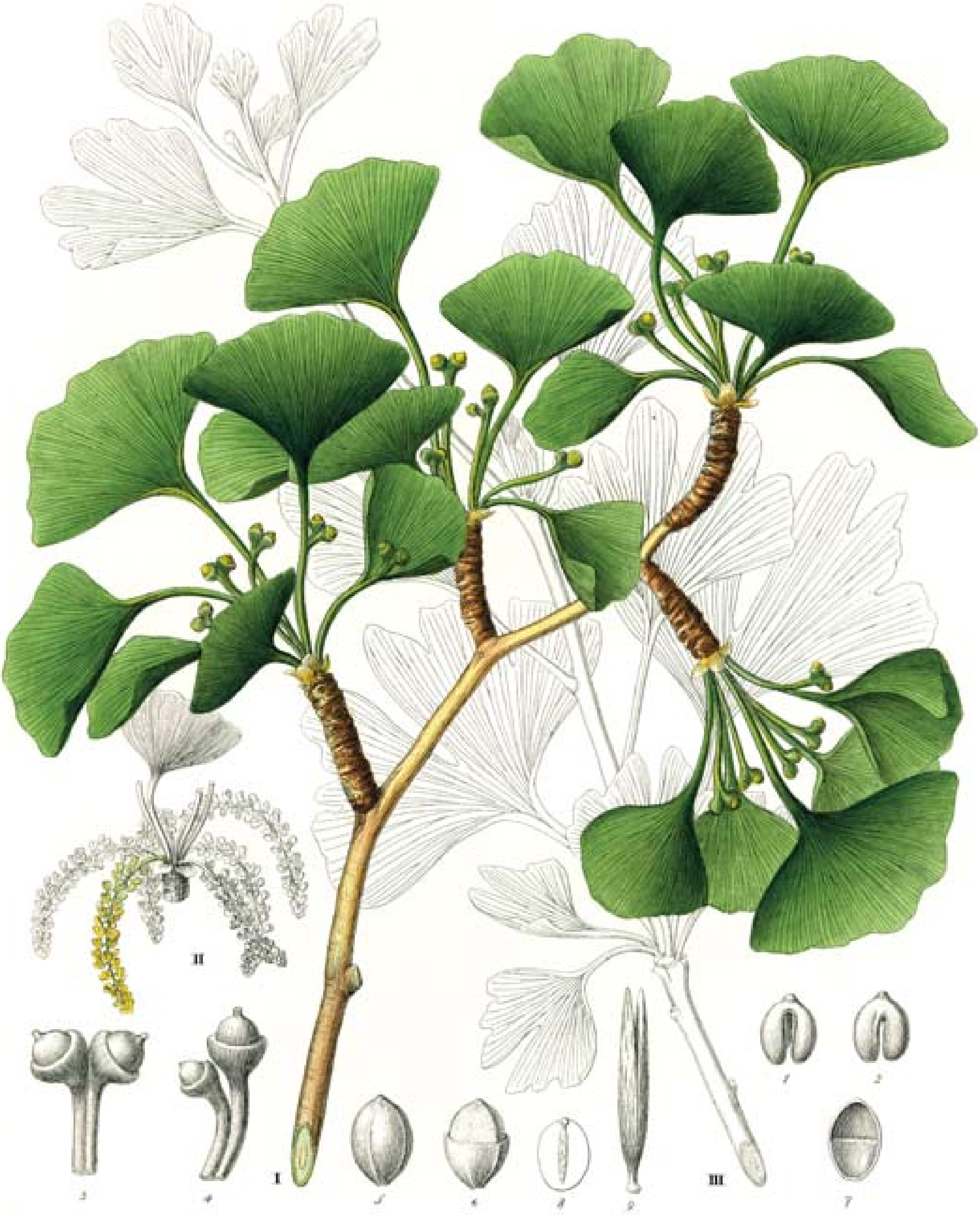
The Grid: A New Infrastructure for 21st Century Science. I. Foster in *Physics Today,* Vol. 55, No. 2, pages 42–47; 2001.

The World-Wide Telescope. A. Szalay and J. Gray in *Science,* Vol. 293, pages 2037–2040; 2001.

The Physiology of the Grid: An Open Grid Services Architecture for Distributed Systems Integration, 2002. I. Foster, C. Kesselman, J. Nick and S. Tuecke. **www.globus.org/research/papers/ogsa.pdf**

Grid Computing: Making the Global Infrastructure a Reality. Edited by F. Berman, G. Fox and T. Hey. John Wiley & Sons, 2003.

More information can be found at **www.mcs.anl.gov/~foster**



FAN-SHAPED LEAVES of the ginkgo tree are shown in this drawing from *Flora Japonica,* a book written by 19th-century German physician Philipp Franz von Siebold. The extract obtained from the leaves (*inset on opposite page*) is one of the most widely used herbal treatments aimed at improving memory.

THE LOWDOWN ON GINKGO BILOBA

This popular herbal supplement may slightly improve your memory, but you can get the same effect by eating a candy bar

By Paul E. Gold, Larry Cahill and Gary L. Wenk

FROM *FLORA JAPONICA*, BY SIEBOLD AND ZUCCARINI, LEIDEN 1835/42, IN THE LEIDEN UNIVERSITY BRANCH OF THE NATIONAL HERBARIUM OF THE NETHERLANDS (*opposite page*); GETTY IMAGES (*above*)

The ginkgo tree (*Ginkgo biloba*) is remarkable in many ways. Although indigenous to Korea, China and Japan, the tree can be found in parks and along city sidewalks around the world. It may grow as high as 40 meters and live for more than 1,000 years. Ginkgo fossils have been dated as far back as 250 million years ago, and Charles Darwin referred to the tree as "a living fossil." Nowadays, however, the ginkgo's primary claim to fame is the extract obtained from its fan-shaped leaves.

The use of ginkgo leaf extracts can be traced back for centuries in traditional Chinese medicine. Today ginkgo biloba is perhaps the most widely used herbal treatment aimed at augmenting cognitive functions—that is, improving memory, learning, alertness, mood and so on. Ginkgo is especially popular in Europe; officials in Germany recently approved the extract for treating dementia. In the U.S. the National Institute on Aging is currently supporting a clinical trial to evaluate the efficacy of ginkgo in treating the symptoms of Alzheimer's disease.

But is there any evidence that ginkgo biloba can really improve cognitive functions? Information on most dietary supplements is based far more on folklore than on experimental findings. Because the U.S. Food and Drug Administration does not regulate herbal treatments, the manufacturers are not required to test the effectiveness or safety of their products. More attention to supplements such as ginkgo biloba is clearly warranted; even if the products do not cause medical problems, they can be costly and may prevent patients from seeking more pragmatic treatments. In an attempt to close the gap in our knowledge, we have reviewed the experimental evidence both for and against the usefulness of ginkgo biloba in enhancing brain functions.

Many Studies, Few Answers

THE TYPICAL DAILY DOSE of ginkgo biloba—and the one used in many of the experiments described in this article—is 120 milligrams of dried extract in two or three oral doses. The extract contains several flavonoids, a large group of natural plant products that are characterized by a specific chemical structure containing a series of carbon rings. Ginkgo extract also contains some biflavonoids, a related group of compounds, and two different types of terpenes, a class of naturally occurring chemicals that includes the active ingredients in catnip and marijuana.

Ginkgo's Effects on the Brain

RESEARCHERS CANNOT SAY FOR CERTAIN WHETHER GINKGO BILOBA CAN IMPROVE COGNITIVE FUNCTIONS, BUT THEY HAVE FOUND THAT THE EXTRACT CAN AFFECT THE BRAIN IN SEVERAL WAYS

CIRCULATORY

- Stimulates widening of the blood vessels, which leads to increased blood flow to the brain and lowered blood pressure (perhaps reducing the risk of stroke).
- Reduces cholesterol levels in the blood (excessive cholesterol is correlated with an increased risk of Alzheimer's disease).
- Inhibits the aggregation of blood platelets and the formation of clots. This may lower the risk of an occlusive stroke (caused by a clot blocking a blood vessel in the brain) but raise the chance of a hemorrhagic stroke (caused by bleeding in the brain).

ANTIOXIDANT

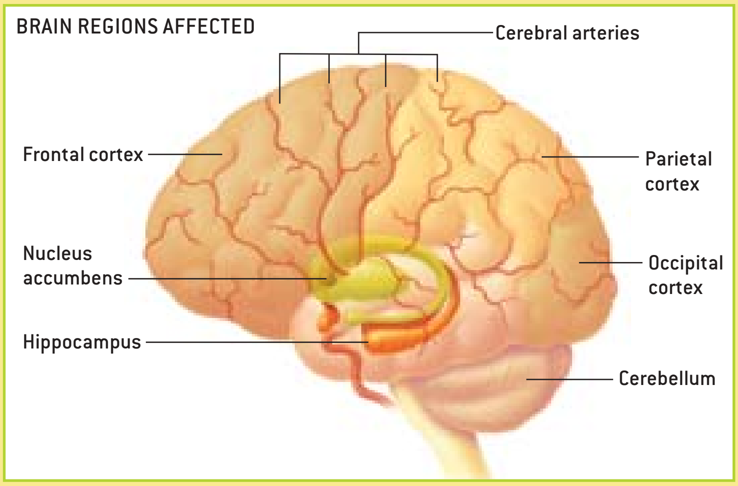
- Curbs the creation of free radicals, highly reactive oxygen molecules that may injure neurons and cause age-related changes in the brain.
- Alleviates the effects of cerebral ischemia—the loss of blood flow to the brain—by inhibiting the production of toxic free radicals after an ischemic episode.

GLUCOSE UTILIZATION

- Boosts the absorption of glucose, the body's primary fuel, in the frontal and parietal cortex, areas of the brain important for processing sensory information and for planning complex actions.
- Also increases glucose absorption in the nucleus accumbens and the cerebellum, brain regions involved in experiencing pleasure and controlling movement, respectively.

NEUROTRANSMITTER SYSTEMS

- Appears to help neurons in the forebrain absorb the nutrient choline from the blood. Choline is one of the components of acetylcholine, a brain chemical that transmits signals between certain neurons.
- Slows the attrition of neuron receptors that direct the response to serotonin, a neurotransmitter that reduces stress and anxiety.
- Enhances the release of gamma-amino butyric acid (GABA), another neurotransmitter that can relieve anxiety. Lowering stress may reduce the level of glucocorticoid hormones in the blood, which in turn may protect the hippocampus, a brain structure critical to normal learning.
- Raises the production of norepinephrine, yet another neurotransmitter. Enhanced activation of the norepinephrine system by certain antidepressants has been shown to reduce the symptoms of depression.



To date, dozens of investigations have examined the cognitive effects of ginkgo in humans, but many of the research reports are in non-English publications or in journals with very restricted distribution, making assessment of the findings difficult. The great majority of studies have involved subjects with mild to moderate mental impairment, usually a diagnosis of early Alzheimer's. Most of the experiments that show evidence of cognitive enhancement in Alzheimer's patients have used a standardized ginkgo extract known as EGb 761.

The ginkgo researchers have usually employed tests of learning and memory; less attention has been paid to other mental functions such as attention, motivation and anxiety. Moreover, because most of the investigators introduced the tests to the subjects after long-term use of ginkgo biloba (typically several months), it is hard to identify which cognitive abilities have been affected. For example, higher scores on the memory and learning tests might stem from the fact that subjects who used ginkgo paid better attention to the initial instructions. To get more specific data on ginkgo's effects, researchers need to administer the tests both before and after the subjects take the supplement.

Because the studies have varied so greatly in the numbers of subjects and the control over experimental conditions, it is useful to focus on only the most rigorous investigations. In 1998 Barry S. Oken of Oregon Health Sciences University and his colleagues considered more than 50 studies involving subjects with mental impairment and selected four that met a conservative set of criteria, including sufficient characterization of the Alzheimer's diagnosis, use of a standardized ginkgo extract, and a placebo-controlled, double-blind design (in which neither the subjects nor the investigators know until the end whether a given patient is receiving the extract or the placebo). Each of these studies showed that the Alzheimer's patients who received ginkgo performed better on various cognitive tests than did patients who received a placebo. Improvements were evident in standardized tests measuring attention, short-term memory and reaction time; the average extent of improvement resulting from ginkgo treatment was 10 to 20 percent.

Oken and his colleagues reported that ginkgo's effect was comparable to that of the drug donepezil, which is currently the treatment of choice for Alzheimer's. Donepezil enhances brain

ALICE Y. CHEN

activity by inhibiting the breakdown of acetylcholine, a brain chemical that transmits signals between certain neurons. Despite these apparently encouraging findings, though, another recent, large and well-controlled trial of EGb 761 (sponsored by its manufacturer, Dr. Willmar Schwabe Pharmaceuticals in Karlsruhe, Germany) involving patients with a mild or moderate stage of dementia reported no "systematic and clinically meaningful effect of ginkgo" on any of the cognitive tests employed.

A critical question concerns whether the ginkgo treatment in studies showing positive effects actually improved cognitive abilities in Alzheimer's patients or merely slowed their deterioration. Two different answers to this key question have come from a 1997 investigation led by Pierre L. Le Bars of the New York Institute for Medical Research. In this study, which was one of the four analyzed by Oken, the results varied according to the cognitive test that was employed. Measured by the Alzheimer's Disease Assessment Scale Cognitive Subscale, the performance of the patients treated with the placebo slowly deteriorated over a year, whereas the performance of patients treated with ginkgo remained stable. But according to a second test—the Geriatric Evaluation by Relative's Rating Instrument—ginkgo-treated subjects improved by about the same amount that placebo-treated subjects deteriorated.

Furthermore, at least one study has reported positive effects on mentally impaired subjects after just a single treatment of ginkgo. Herve Allain of the University of Haute Bretagne in Rennes, France, gave one fairly high dose of ginkgo—320 or 600 milligrams—to a small group of elderly people with mild, age-related memory impairment. An hour after the treatment, Allain tested the subjects' memory by rapidly presenting short lists of words or drawings and then asking the patients to recall the lists immediately afterward. Their ability to recall the rapidly presented material increased significantly after ingestion of ginkgo. This finding raises the possibility that short-term, rather than long-term, biological actions provide the basis for ginkgo's reported effects on cognition.

It should be noted that ginkgo has also been shown to impair performance. For example, in a small study of elderly people with mild to moderate memory impairment, Gurcharan S. Rai of Whittington Hospital in London and his team found that after 24 weeks of treatment, patients who took ginkgo could not recall digits as well as patients who received a placebo.

Help for the Healthy?

UNFORTUNATELY, FAR FEWER studies have examined the cognitive effects of ginkgo biloba on healthy young adults. In one small study during the mid-1980s, Ian Hindmarch of the University of Leeds in England administered a battery of tests to eight healthy subjects (ages 25 to 40) after they took the ginkgo extract EGb 761. Hindmarch reported that the highest dose tested (600 milligrams) improved performance in only a short-term memory test. More recently, two reports from Cognitive Drug Research, a laboratory in Reading, England, provide minor support for the view that ginkgo may enhance cognitive functions in young people. One study reported that subjects who took a dose of ginkgo performed better on tasks assessing attention than did subjects who took a placebo. The other study showed an improvement in memory among middle-aged subjects (ages 38 to 66) who were treated with a combination of ginkgo and ginseng, another herbal remedy touted as a memory aid. The effects seen in the latter study, however, could not be attributed to ginkgo alone and did not increase with the dosage, which would be expected of a truly effective substance.

Information on most dietary supplements is based far more on folklore than on EXPERIMENTAL FINDINGS.

For most pharmaceuticals, researchers conduct a large number of studies with lab animals before they test the drugs in humans. Such experiments can be useful in determining a drug's safety and effectiveness. But because ginkgo is unregulated, its manufacturers have not been required to perform animal tests. As a result, there are relatively few research reports in refereed journals examining the efficacy of ginkgo in improving learning and memory in animals. Perhaps the most notable example is a 1991 study of young adult mice that were trained to press a lever to receive food. Mice treated with ginkgo for four to eight weeks learned the task slightly more quickly than did the control mice. As with humans, though, it is difficult to pin down whether ginkgo actually enhances the learning process or whether it has other effects that improve the animals' performance at a specific task. For instance, investigators have reported that repeated administration of ginkgo reduced stress in rats, and altered stress responses can themselves have a great influence on learning and memory.

If ginkgo can really enhance mental processes, how does it work? Studies of humans and lab animals have indicated several classes of biological effects that might contribute to ginkgo's putative improvement of cognitive functions [*see box on opposite page*]. Whatever its effects, ginkgo appears to pose few health

THE AUTHORS

PAUL E. GOLD, LARRY CAHILL and *GARY L. WENK* are leading authorities on the enhancement of brain functions. Gold is professor of psychology and neuroscience at the University of Illinois at Urbana-Champaign. Cahill is assistant professor of neurobiology and behavior and fellow of the Center for the Neurobiology of Learning and Memory at the University of California, Irvine. Wenk is professor of psychology and neurology at the University of Arizona.

MONICA STEVENSON *FoodPix*

The Other "Brain Boosters"

By Mark A. McDaniel, Steven F. Maier and Gilles O. Einstein

OLDER ADULTS HAVE shown a strong interest in over-the-counter "brain boosters," many of which are marketed with grand claims touting their benefits. There are sound biochemical reasons for expecting some of these nutrients to be effective. In a review of published research, we found studies showing that some of these substances had robustly enhanced memory in lab animals and occasionally produced impressive improvements in humans as well. Nevertheless, there are numerous questions about the sample sizes in the studies, the generality of the results across different memory tests and populations, and other aspects of the procedures and data. These problems, in conjunction with a general lack of research demonstrating that the results can be replicated, dampen enthusiasm for the effectiveness of these nutrients in substantially arresting or reversing memory loss. Here is an abbreviated summary of the findings for six kinds of nonprescription compounds that are claimed to be memory enhancers and treatments for age-related memory decline.

PHOSPHATIDYLSERINE (PS)

A naturally occurring lipid, PS has been shown to reduce many consequences of aging on the neurons in older rats and mice and to restore their normal memory in a variety of tasks. Research on the impact in humans is limited, though. For older adults with moderate cognitive impairment, PS has produced modest increases in recall of word lists. Positive effects have not been as consistently reported for other memory tests.

CHOLINE COMPOUNDS

Phosphatidylcholine, which is typically administered as lecithin, has not proved effective for improving memory in patients with probable Alzheimer's disease. Research on citicoline is practically nonexistent, but one study reported a robust improvement in story recall for a small sample of normally aging older adults.

PIRACETAM

Developed in 1967, Piracetam has not been approved by the U.S. Food and Drug Administration, but it is sold in Europe and Mexico under several names (including Nootropil and Pirroxil). Animal studies suggest that the drug may improve neural transmission and synaptic activity and also combat age-related deterioration of neuronal membranes. But there is no clear sign of any cognitive benefits in patients with probable Alzheimer's or adults with age-associated memory impairment.

VINPOCETINE

An alkaloid derived from the periwinkle plant, vinpocetine increases blood circulation in the brain. In three studies of older adults with memory problems associated with poor brain circulation or dementia-related disease, vinpocetine produced improvements in performance on cognitive tests measuring attention, concentration and memory.

ACETYL-L-CARNITINE (ALC)

An amino acid included in some "brain power" supplements sold in health food stores, ALC participates in cellular energy production, a process especially important to neurons. Animal studies show that ALC reverses age-related decline in the number of receptor molecules on neuronal membranes. But studies of patients with probable Alzheimer's have reported only nominal advantages for ALC in a range of memory tests.

ANTIOXIDANTS

Antioxidants such as vitamins E and C help to neutralize tissue-damaging free radicals, which become more prevalent with age. But studies have found that vitamin E does not offer significant memory benefits for patients with Alzheimer's or early Parkinson's disease. A combination of vitamins E and C did not improve college students' performance on several cognitive tasks.

A more detailed version of this article appeared as "Brain-Specific Nutrients: A Memory Cure?" by Mark A. McDaniel, Steven F. Maier and Gilles O. Einstein in Psychological Science in the Public Interest *(May 2002). (Available at www.psychologicalscience.org/journals/pspi/3_1.html) McDaniel is chair of the department of psychology at the University of New Mexico. Maier is director of the Center for Neuroscience at the University of Colorado at Boulder. Einstein is chair of the department of psychology at Furman University.*

BUYERS BEWARE: Researchers have not found convincing evidence that dietary supplements can enhance memory.

CHUCK SAVAGE *Corbis*

risks, particularly at the typical doses of 120 to 240 milligrams a day. Some complications have been noted, though, including subdural hematomas (clots between the skull and brain) and gastrointestinal problems. As is the case with most plant extracts and medications, ingestion of ginkgo is at times associated with nausea and vomiting. In addition, some users experience increased salivation, decreased appetite, headaches, dizziness, tinnitus (ringing in the ears) and skin rash. Large doses may lead to orthostatic hypotension, a condition of low blood pressure sometimes seen after abrupt postural changes, such as standing up after being seated. Still, the general impression is that the incidence of serious adverse consequences after use of ginkgo is quite low. Also, this incidence may be reduced further if and when optimal individual dose regimens for ginkgo are established.

Because of the differences in experimental designs used to test ginkgo and other treatments, it is difficult to make direct comparisons of efficacy. For instance, on a test involving memory of a brief story, glucose enhanced performance in young adult and healthy elderly subjects by about 30 to 40 percent. In people with Alzheimer's, the improvement on a similar memory test approached 100 percent, with smaller enhancements seen on other measures. The extent of improvement in these experiments is much larger than the 10 to 20 percent gain shown with ginkgo. But most of the experiments testing the effects of glucose have used short-term treatments and compared performance before and afterward, whereas most of the experiments testing ginkgo have used long-term treatments and compared ginkgo-treated subjects with a control group.

Does ginkgo biloba in fact enhance cognitive function? The proof for even a MILD BENEFIT is weak.

But to return to our original question, does ginkgo biloba in fact enhance cognitive function? In general, the reported effects are rather small. The number of experiments is also small, and they are of mixed quality, so the proof for even a mild benefit is weak. In humans, ginkgo may slow cognitive decline during dementia. It is possible that ginkgo's main effects on memory kick in after one dose and are relatively short-lived, but the research literature is so limited that significant issues such as this one cannot be adequately evaluated at this time.

The Bottom Line

GIVEN THE AVAILABLE EVIDENCE, is ginkgo biloba the best therapy for improving memory? Other supplements have been found to influence cognitive function in humans and lab animals [*see box on opposite page*]. Pharmaceuticals such as donepezil can strongly enhance learning and memory in rodents and induce modest though significant improvements in humans. But relatively simple interventions can produce some of the same results. For example, hearing an exciting story apparently releases epinephrine from the adrenal glands into the circulation, thereby enhancing one's memory without any drugs at all. One mechanism by which epinephrine might enhance memory is by liberating glucose from stores in the liver, thereby increasing the circulating glucose available to the brain.

Eating a simple sugar can also improve one's memory. Considerable evidence indicates that glucose administered systemically (to humans by ingestion and to rodents by injection) enhances cognitive performance in young and elderly rats, mice and humans, including people with Alzheimer's. Like most treatments that improve memory, glucose's effects follow a dose-response curve in the shape of an inverted U. Only intermediate doses improve memory; low doses are ineffective, and high doses may actually impair memory.

Establishing direct comparisons of efficacy is vital to identifying which treatments improve cognition the most. This is one of many instances in which further studies of rodents would be useful because of the researchers' ability to control all the variables in the experiment. Only one study has directly compared the effects of ginkgo with those of other treatments. This study showed that the peak enhancement observed with ginkgo was about half of that seen with other drugs. More direct comparisons, in both humans and lab animals, are clearly needed.

We began our survey of the research literature with healthy skepticism but with a commitment to avoid prejudging the findings. We found evidence supporting the view that ginkgo enhances cognitive functions, albeit rather weakly, under some conditions. Our overriding impression, however, is that we do not have enough information to say conclusively whether ginkgo does or does not improve cognition. There are too few experiments on which to base clear recommendations, and most of the studies showing benefits have involved too few subjects.

But there are enough positive findings, perhaps just enough, to sustain our interest in conducting further research on ginkgo. Many years of experience with investigations of new drugs have demonstrated that the initial positive results from studies involving a small number of subjects tend to disappear when the drugs are tested in larger numbers of subjects from diverse populations. So the true test of ginkgo's efficacy lies ahead. SA

MORE TO EXPLORE

A more detailed version of this article appeared as **"Ginkgo Biloba: A Cognitive Enhancer?"** by Paul E. Gold, Larry Cahill and Gary L. Wenk in *Psychological Science in the Public Interest,* Vol. 3, No. 1, pages 2–11; May 2002. Available at www.psychologicalscience.org/journals/pspi/3_1.html

More information about ginkgo biloba and other dietary supplements can be found at **dietary-supplements.info.nih.gov/** and **www.cfsan.fda.gov/~dms/supplmnt.html**

TRANSDERMAL DRUG DELIVERY

Potent Patches

Swallowing a pill is simple. Yet people still forget to take their medication, and drug levels in the bloodstream surge and sink with each dose. Medicated patches that stick to the skin prevent such problems.

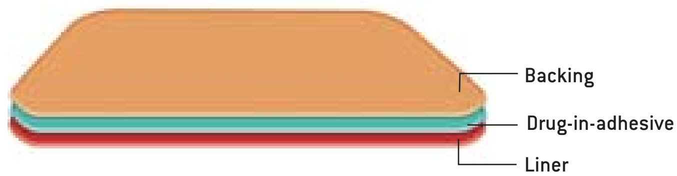
"The patch" became popular around 1990, as smokers used it to kick the habit. A series of patches worn over several weeks' time supplied decreasing amounts of nicotine, gradually weaning people from their addiction. Today transdermal patches deliver estrogen for hormone replacement therapy, nitroglycerin for angina, scopolamine for motion sickness and seasickness, fentanyl for pain control, clonidine for hypertension and, recently, ethinylestradiol plus norelgestromin for contraception.

In most cases the patch is saturated with the drug, which steadily diffuses through microscopic gaps between skin cells and through the skin's pores. The trick is designing the appropriate polymer adhesive that will retain a given drug molecule yet also allow it to diffuse, a delicate balancing act, notes Sharon Grosh of 3M Drug Delivery Systems.

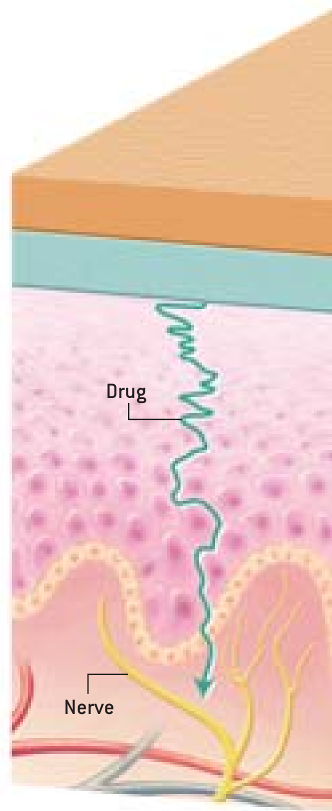
The skin's outer layer of dead cells—the stratum corneum—is a good barrier, so it controls the diffusion rate. Drug molecules must be small enough to sneak between the cells and must dissolve in the oily lipids there, so they can reach capillaries deeper in the skin. Furthermore, the stratum corneum varies in thickness and porosity from person to person, so it is advantageous if a drug has a big therapeutic window—not too toxic at a high concentration yet still potent at a low one, says Mark R. Prausnitz of the Georgia Institute of Technology. Because of these vagaries, only a handful of drugs are effective in patch form.

Researchers are trying to modify the passive patch with active drivers that force molecules that are larger or water-soluble, or both—such as insulin and vaccines—through the skin. Electric current from a tiny battery within the patch could widen cell gaps and push ionized drug molecules through them. Or an array of hollow microneedles on the patch could provide microscopic funnels through the skin. The active designs are costlier and bulkier than passive patches, but several are already in advanced clinical trials.

—Mark Fischetti



SINGLE-LAYER PATCH has an adhesive polymer layer that sticks to the skin and also holds the drug. Like all currently approved patches, it has a protective backing and a liner that seals in the drug until a person peels it off for application. The diffusion rate depends on the drug concentration in the patch, the lipid concentration in the skin, and the skin area covered.



DRUG MOLECULES from a patch diffuse through the lipids between dead cells in the skin's stratum corneum and past living cells in the epidermis to capillaries in the dermis, where they enter the bloodstream.

ALICE Y. CHEN

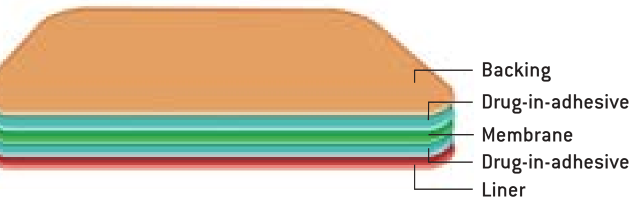
DID YOU KNOW . . .

➤ THROW IT AWAY: A patch lasts only so many hours or days before its drug delivery falls below an effective threshold. The patch must then be removed, but it still contains medication. Users must therefore be certain that children or pets don't accidently ingest the patch, which could be toxic or even fatal.

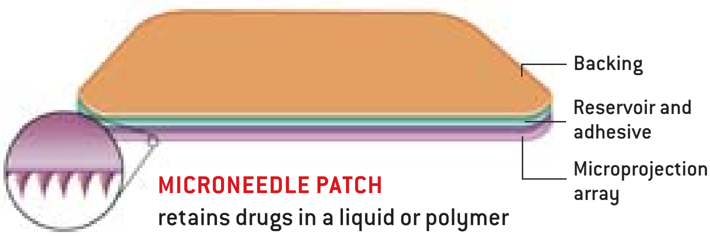
➤ MAX VACCINE: Certain vaccines—including those for HIV and influenza—may elicit a strong immune response in Langerhans immune cells in the epidermis. But hypodermic needles and pills miss this target, sending vaccine to the bloodstream. If patches can be developed to handle vaccine molecules, which are large, they could bathe Langerhans cells, maximizing the vaccines' effectiveness.

➤ THICK-SKINNED: The stratum corneum is thickest on the palms and soles; it is thinnest behind the ear, in the armpit and on the scalp. The region behind the ear is also rich with blood vessels, so it is an excellent place for a patch, although only small designs will fit there. The armpit and scalp are generally unsuitable because hair interferes. Most patches end up on the arm, inner thigh, lower back or chest, where they can remain hidden and are unlikely to be rubbed off.

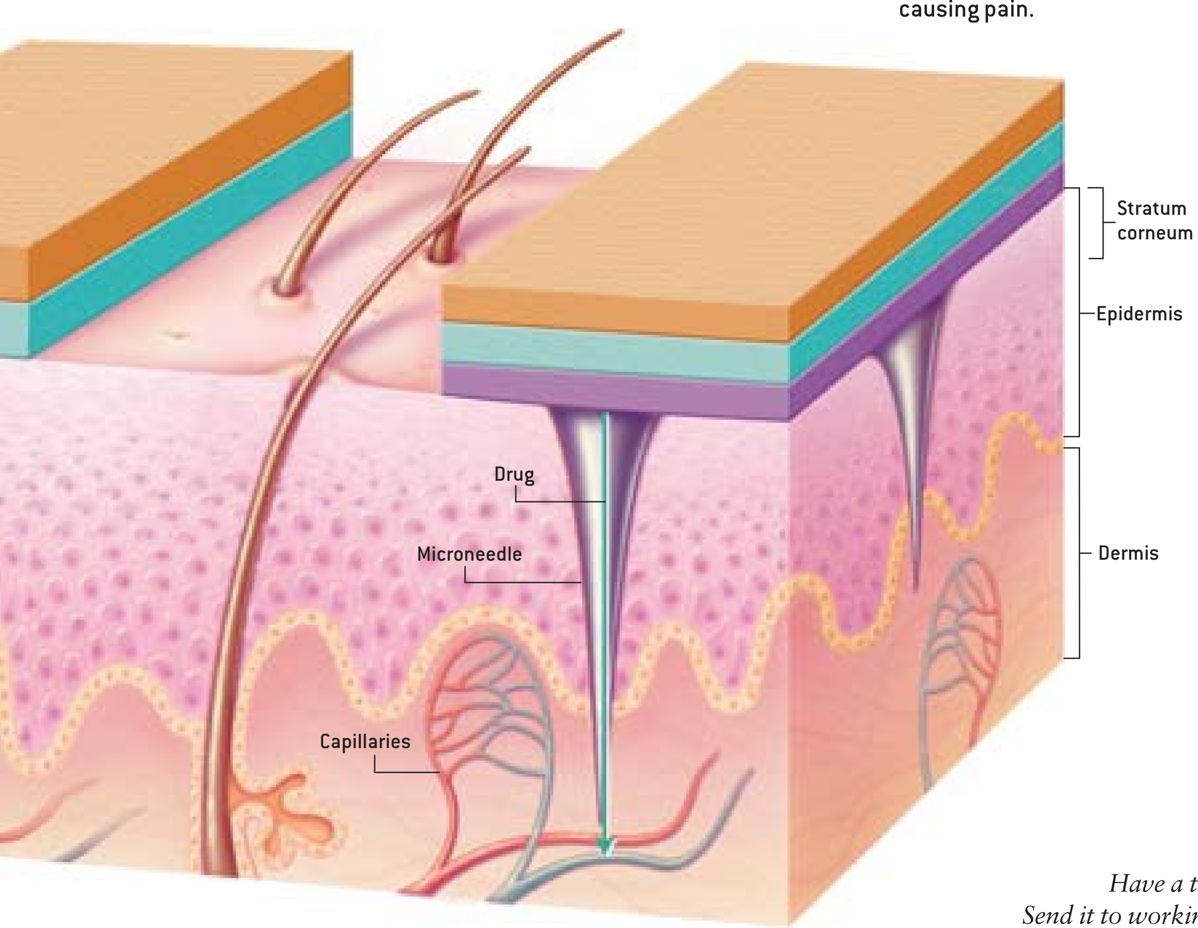
➤ NO PAIN? The microneedles on experimental patches don't hurt, yet developers are worried that the name will scare away customers. Their synonyms—microprojections, microblades—sound no better. A euphemism is sure to be found once trials are completed.



MULTILAYER PATCH has a membrane that controls the delivery rate of drugs, especially those that would otherwise diffuse too quickly into the skin.



MICRONEEDLE PATCH retains drugs in a liquid or polymer reservoir. Hollow, micromachined needles penetrate the top edge of the dermis—deep enough to bring drugs to capillaries but not so deep that they touch nerve endings, causing pain.



Have a topic for a future column? Send it to workingknowledge@sciam.com

Screen Writing

THE TABLET PC IS A HIGH-TECH TOOL FOR SCRIBBLERS **BY MARK ALPERT**

DIGITAL DOODLES can be drawn on the Toshiba Portégé 3500. You can rotate the screen 180 degrees and fold it back over the keyboard.

COURTESY OF TOSHIBA

When I was starting out in the journalism business 19 years ago, I got a piece of good advice from my first boss, Poody Walsh, who was the editor of the *Eagle Times* in Claremont, N.H. Poody (yes, believe it or not, that's what everyone called him) saw that I was handwriting my notes while I interviewed sources on the telephone, and he urged me to get accustomed to typing the notes instead. But I just couldn't do it—typing seemed to require more brainpower than handwriting, making it difficult to concentrate on the interview. Two decades later I'm suffering the consequences: I am surrounded by looming stacks of legal pads that cover nearly every square inch of my disordered office.

Since the early 1990s computer makers have tried to wean me from my dependence on paper. Their solution was called pen computing: I could use a stylus to jot my notes on the screen of a portable PC, and the device's software would convert my chicken scratches to crisp, clear text files that could be stored on a hard drive instead of cluttering my office. Unfortunately, none of the first-generation pen computers—Go Corporation's EO, Apple's Newton and Microsoft's WinPad, among others—was successful. The biggest problem was that the devices' handwriting-recognition software did not work as well as advertised. Cartoonist Garry Trudeau lampooned the technology in a *Doonesbury* strip; when Mike Doonesbury scribbles "I am writing a test sentence" on the screen, the computer translates it to "Ian is riding a taste sensation."

But PC hardware and software have advanced dramatically in the past 10 years, and last November Microsoft took another stab at pen computing when it introduced a software platform for a device called the Tablet PC. Competing models of this portable machine are being manufactured by Toshiba, Fujitsu, Compaq, Acer and other companies, but they all share an enhanced version of Microsoft's Windows XP operating system. As a result, the Tablet PC can incorporate handwriting into a wide variety of Windows-compatible programs; for example, you can mark up a Microsoft Word file with scrawled comments, then e-mail the revised document via Microsoft Outlook. Intrigued, I asked several of the manufacturers to let me try out their Tablet PCs, and two machines soon came to my office.

The first to arrive was the Toshiba Portégé 3500, a 4.1-pound, $2,300 device that at first glance looks like an ordinary laptop. Conforming to the traditional clamshell design, it opens to reveal a 12-inch screen and a full-size keyboard. When you boot it up, you see a standard Windows display. But you can turn the Portégé into a tablet by rotating the screen 180 degrees and folding it back over the keyboard. The image on the screen suddenly shifts from a horizontal to a vertical orientation. If you pick up the device, it feels like an unusually heavy clipboard. All you need do now is pop out the tablet pen, a thin black cylinder with a white plastic tip, from its convenient niche.

The tablet pen acts as the system's mouse. You point the pen's tip to the part of the screen where you want the cursor to go, and you tap on the icons to open files and programs. The really cool feature is that you don't have to actually touch the screen to move the cursor. As long as the pen's tip is within about an inch of the screen's surface, it can control the cursor's movements. The explanation for this "spooky action at a distance" (to borrow a term from quantum mechanics) is that a grid of wires underneath the screen emits an electromagnetic field that causes a coil inside the pen to oscillate at a resonant frequency. These oscillations produce signals that are detected by the grid. The electromagnetic resonance technology allows you to rest your hand on the screen without affecting it. Because the system knows where the tip of the pen is, it can distinguish between a pen stroke and the pressure of a finger or thumbnail. (You'd better not lose that pen, though.)

The program that best demonstrates the capabilities of the Tablet PC is Windows Journal. When you open this application, an image of a lined sheet of notebook paper appears on the screen. Writing on this surface with the tablet pen feels very much like writing on paper with a fine-point marker. But this marker is incredibly versatile. You have a wide choice of colors for the ink. A simple tap on an icon turns the pen into a highlighter. The other end of the pen is an electronic eraser—you can rub it on the screen to remove your slipups. And you can cut and paste the document by drawing rings around the parts you want to move.

The machine was great fun to play with (I drew some elaborate doodles), but

I wondered how well it would work during an actual interview. As a test, I jotted notes on the Portégé's screen while I spoke on the telephone with Craig Marking, a senior product manager at Toshiba. It took me a moment to adjust, but after a few minutes I was scribbling like a fiend. I discovered, to my delight, that if you cross out a misspelled word, it automatically disappears. And there are various shortcuts that make reviewing the document just as easy as flipping through the pages of a legal pad.

COURTESY OF FUJITSU

TAKING NOTES on the Fujitsu Stylistic ST4000 is fun, but it's hard to beat the convenience of pen and paper.

But being just as good as a legal pad is hardly good enough to justify a $2,300 purchase. The real test was how well the Portégé could convert my notes to text. Earlier pen computers such as Apple's Newton were designed to recognize a user's handwriting by learning the quirks of his or her penmanship. The Tablet PC, though, takes a brute-force approach: it compares each scrawl with a database containing hundreds of thousands of handwriting samples. For example, the software can interpret a sloppily scribed "it" because many poorly written "it"s are stored in the database.

This technique worked quite well in translating my notes. The Portégé made a few mistakes—turning "frequency" into "regency," for instance—but overall the recognition was surprisingly accurate. The only drawback was that the machine could convert just one page of handwriting at a time, so creating a text file for a long interview was frustratingly laborious.

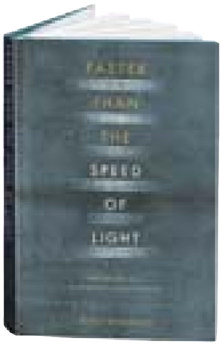
The other Tablet PC that I evaluated was the $2,200 Fujitsu Stylistic ST4000, which works much like the Portégé but looks less conventional. Instead of being incorporated into a folding clamshell, the screen and computer are contained in a single slab. Because the Stylistic is about a pound lighter than the Portégé, it's easier to carry in one hand like a legal pad or clipboard. But the Stylistic's screen is a bit smaller than the Portégé's, its microprocessor is less powerful, and it doesn't have a built-in keyboard (it comes instead with a keyboard that can be plugged in).

Tablet PCs are perhaps most useful for people who take a lot of notes while on the go: journalists, college students, insurance adjusters and so on. Although I nominally fall into the first category, I doubt that I'll actually buy one of the devices. The problem is, I'm just as disorganized in the electronic realm as I am in the world of pen and paper. After only a few days of using the Toshiba and Fujitsu machines, I'd cluttered up their hard drives with a bewildering array of files. Finding my interview notes (either the handwritten Windows Journal files or the converted Microsoft Word documents) took about as long as rummaging through the mounds of folders scattered across my desk.

And did the digital note-taking somehow improve the quality of this column? I don't know. Judging from my editor's scrawled comments ("NO, NO, NO! REWRITE!"), I guess not. SA

Was Light Faster in the Past?

A MAVERICK PHYSICIST POSITS THE POSSIBILITY—AND MUCH MORE **BY PHILIP MORRISON**



FASTER THAN THE SPEED OF LIGHT: THE STORY OF A SCIENTIFIC SPECULATION
by João Magueijo
Perseus Publishing, Cambridge, Mass., 2003 ($26)

Breaking the old speed limit posted by one Albert Einstein in his 20s, this book deploys a racy and provocative text to convey its popularized content of a new cosmology. Jocular, ironic, witty, self-centered, even indignant, Magueijo is all too ready to castigate his adversaries, those comfortable gatekeepers of learning. The author is no aspiring youth but a tenured professor of theoretical physics, age 35. In spite of his own stature within learned gates—University of Lisbon, then Cambridge on a prime fellowship, now enjoying tenure at great Imperial College in London—his voice is embittered. This journey of youthful success is recalled in complaint about the idiots, the sexually deficient, the money wasters. The thin volume is studded with familiar four-letter words, invoked with rude claims about the motives of colleagues, shadowy referees, editors and others encountered.

Our current scenario for cosmology clearly opened its second act among the high simplicities of the 1970s with two visible puzzles. Why is 3-D cosmic space accurately flat (like old Euclid's own), although it lies within Einstein's universal 4-D curved spacetime? Why is its content so uniform on large scale?

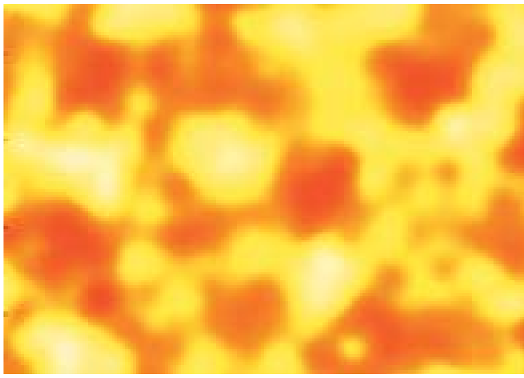
In 1980 Alan H. Guth of the Massachusetts Institute of Technology found a unitary explanation for both riddles. Named inflation, it postulates a minute interval of unusually sudden spatial expansion immediately before the slow, steady expansion of space carried all matter outward. That transient field eventually decayed to yield the complex mix of particles (including radiations) that still move through space. The early push is maintained in the Hubble expansion observably under way, now quite likely speeding up.

This very cosmos was in fact described well before any of its complex contents were known. In 1918 Einstein and his friend the Dutch astronomer Willem de Sitter found the broad space and time properties we now believe. Inflation is the repulsive side of gravity's attraction, a kind of matter that stretches cosmic space so fast and far that almost every flaw has been ironed out to approximate local flatness. Our current particle physics allows such behavior, making such a surprise acceptable. Today we freely use what seemed unrealizable in those days.

Nobody would have believed the account Einstein and de Sitter arrived at had it not fitted so neatly what we observe. Before and beyond all the starry galaxies, we see a distant uniform surface, the origin of almost all cosmic photons, pure thermal radiation with utter conformity to the established spectrum shape of old Max Planck's. The same temperature is seen at every point of the sky to better than one part in 100,000. Your coffee cream confirms: uniformity in fluids comes from stirring. It is easy to believe that those photons broke free of the expanding opaque plasma, to stream along while much slower action built the lumpy, gravitating assemblages we call galaxies. The time of that breakout was a rough half a million years after the inflationary flash.

It is the minor deviations from simplicity that give us any early detail. For the past half a dozen years, the task has been to analyze all those minor flaws as hints of the earliest matter and of its changes and motions as our present cosmos grew. There are no new real puzzles, although certainly a great deal remains to be learned—most importantly, the dark, enigmatic legacy of AE: his cosmological "constant."

The book at hand is a People's Manifesto by an articulate and inventive op-



COSMIC MICROWAVE BACKGROUND RADIATION can provide observational tests of João Magueijo's theory that the speed of light is variable.

position to the complacent consistency I have just expressed. The author and his colleagues are now skeptical of inflation: it is a tale much too pat, an expansion at unlimited speed. To stir the dense, hot mix in the early epochs, you have to race and beat light itself out to the remote boundaries of inflation. Faster than light? Einstein and his partner admitted only one way this could happen: with repulsive gravity. It is in their theory!

Perhaps there is another way, suggests Magueijo. If matter in motion is too slow for light, why not make the speed of light faster and faster into the past? Throwing out heavyweight Einstein and his near constant speed of light is no easy task. Yet that is the burden of the new iconoclasts. Maybe they can make a cosmos with wildly varying speeds of light, and maybe they can keep the gas uniform, but they give no clear reward for so denying our well-tested Einstein on this theorist's journey into the past. Their strongest argument is the very flatness of space: it turns out that a cosmos with a changing speed of light must be a flat one and a uniform one as well, if energy is to be conserved. There is much more to be said about the untested physics of these variable vacuum light speeds. More than one form of theory is out there, to say nothing of the myriad options opened by multiple dimensions.

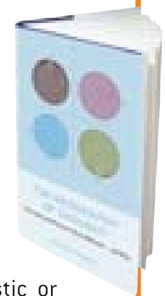
Magueijo sums up with the view that the AE establishment "think they own us; we think … they are just a bunch of squares.… We have all the fun in the universe." I hope my comments demonstrate that his last remark is wrong; there is fun with Einstein, too, plus plenty of impressive experimental support. As for the true prize, the grandeur of cosmology, neither the Academy nor its clever hecklers have yet grasped its origins. SA

Philip Morrison, professor of physics (emeritus) at the Massachusetts Institute of Technology, wrote the book reviews and the Wonders column for this magazine for 35 years.

THE EDITORS **RECOMMEND**

THE GEOGRAPHY OF THOUGHT: HOW ASIANS AND WESTERNERS THINK DIFFERENTLY … AND WHY

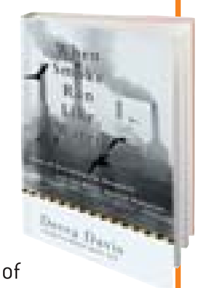
by Richard E. Nisbett. Free Press, New York, 2003 ($24)



Nisbett, a psychologist and Distinguished University Professor at the University of Michigan at Ann Arbor, used to believe that "all human groups perceive and reason in the same way." A series of events and studies led him gradually to quite another view, that Asians and Westerners "have maintained very different systems of thought for thousands of years." Different how? "The collective or interdependent nature of Asian society is consistent with Asians' broad, contextual view of the world and their belief that events are highly complex and determined by many factors. The individualistic or independent nature of Western society seems consistent with the Western focus on particular objects in isolation from their context and with Westerners' belief that they can know the rules governing objects and therefore can control the objects' behavior." Nisbett explores areas that manifest these different approaches—among them medicine, law, science, human rights and international relations. Are the societal differences so great that they will lead to conflict? Nisbett thinks not. "I believe the twain shall meet by virtue of each moving in the direction of the other."

WHEN SMOKE RAN LIKE WATER

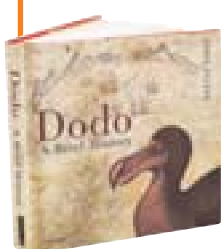
by Devra Davis. Basic Books, New York, 2002 ($26)



Davis tackles the subject of environmental pollution on two fronts, one personal and one professional. The first contains insight into her own life, starting from her roots in the metalworking town of Donora, Pa.—where the smog from pollution killed 20 outright in October 1948 and had lasting ill effects in townspeople, some fatal, in the months and years that followed. Her vivid descriptions of deadly smog in London as recent as the mid-1950s give the reader perspective about the inherent perils of industrial pollution to the public at large. An epidemiologist by training, Davis also chronicles the growing awareness of the spread of breast cancer (and pollution as a possible cause) in the 1990s, sterility and testicular cancer in men, and the impact of pollution on climate change. Although her prose relies heavily on statistics and historical accounts of pollution, Davis's personal narrative ties the story together nicely.

DODO: A BRIEF HISTORY

by Errol Fuller. Universe Publishing/Rizzoli, New York, 2002 ($22.50)



Sailors from a group of Dutch ships that reached the Indian Ocean island of Mauritius in 1598 found dodo birds in considerable numbers along the coast. Forty years later the species was all but gone there, and by 1690 it was extinct everywhere. Fuller, a British painter interested in curiosities of natural history, tells the dodo's story with grace and many intriguing illustrations. Hard facts about the dodo are minimal. The bird was "a gigantic, flightless pigeon" with a massive beak, a large head and a tuftlike tail. But its exact appearance is uncertain because drawings made while it still lived are contradictory. Written accounts are "as tantalizing as the pictures." Having set out the few facts, Fuller goes on to describe the dodo's popularity in literature—and to give it an epitaph as "the ultimate symbol of what can go wrong when man and nature come together."

All the books reviewed are available for purchase through www.sciam.com

The Graph of Life

BY DENNIS E. SHASHA

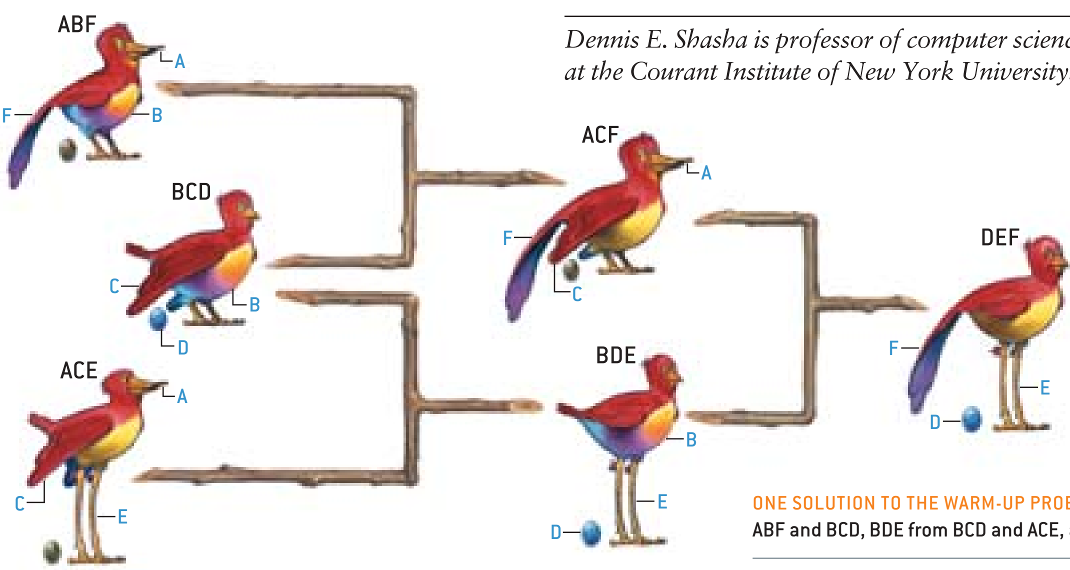
The history of life is often represented by a tree, with each existing species evolving directly from a single progenitor species. The newer species arises after some population of the original species adapts to different conditions and diverges genetically so much from the original group that the two can no longer mate. That's the standard story, but biology eschews absolutes. Many studies suggest that although individuals from different species *normally* do not interbreed, they *may* interbreed. What is more, such interbreeding, or hybridization, occasionally yields new species.

Suppose you are trying to trace the origins of a collection of species, some of which are known to be source, or founder, species and others of which are known to be derived. For this puzzle, we want each derived species to have as few direct evolutionary progenitors as possible. The direct evolutionary progenitors of species *X* are those species in the collection that combine to supply *X*'s traits without the involvement of any other species. Note that three species may combine to form a new species, *Y*, by first producing unseen intermediate species. Species *Y* would then be considered to have three direct progenitors, though, because we count only the species we see. One more constraint: no species can arise before its progenitors appear.

Warm-up: Assume we have three source species with traits ABF, BCD and ACE, respectively, and three derived species with traits ACF, DEF and BDE. Construct a derivation scheme that gives each derived species at most two direct progenitors (one or both of which may be derived). One solution is depicted below. In that solution, ACF and BDE had to arise before DEF could appear. According to the rules above, a direct derivation of DEF from the *source* species would require all three of those species—ABF, BCD and ACE—to cover all the traits.

Puzzle: Suppose four source species have the traits AB, DH, EF and CG, and the derived species have traits BEF, DEF, ADE, ACD, ACF, ADG, BEG and FGH. Construct a derivation scheme in which each derived species, except one, has only two direct evolutionary progenitors. (As in the warm-up, one or both may be derived.) Can you find four alternative source species, each having two traits, such that (*a*) every derived species here has only two direct progenitor species and (*b*) the path from a source species to a derived species goes through at most one other derived species? SA

Dennis E. Shasha is professor of computer science at the Courant Institute of New York University.



TRAITS OF INTEREST
A = Long beak
B = Rainbow belly
C = Long wings
D = Blue eggs
E = Long legs
F = Long tail

ONE SOLUTION TO THE WARM-UP PROBLEM: ACF derives directly from ABF and BCD, BDE from BCD and ACE, and DEF from ACF and BDE.

MATT COLLINS

Answer to Last Month's Puzzle

To guarantee opening the 10-switch safe, you need to try only the combinations below. The first letter in each row is the setting of the first switch, the second letter is the setting of the second switch, and so on. The letter A stands for high setting, B for middle and C for low. X means any setting is okay.

1. A A A A A A A A A A
2. A B B B B B B B B B
3. A C C C C C C C C C
4. B A B C A B C A B C
5. B B C A B C A B C A
6. B C A B C A B C A B
7. C A C B A C B A C B
8. C B A C B A C B A C
9. C C B A C B A C B A
10. X A A A B B B C C C
11. X A A A C C C B B B
12. X B B B A A A C C C
13. X B B B C C C A A A
14. X C C C A A A B B B
15. X C C C B B B A A A

Web Solution

For a full explanation and the answer to this month's puzzle, visit www.sciam.com

Truth in Advertising

THERE ARE BURGER JOINTS, AND THEN THERE ARE BURGERS AND JOINTS **BY STEVE MIRSKY**

The fault, Shakespeare once almost said, lies not in our stars but in our stuffing our faces. That sentiment is basically the reasoning behind a federal district court judge's January dismissal of a recent lawsuit against McDonald's, brought by two obese New York City teenagers who claimed that the fast food was at fault for their fat. "Common sense has prevailed," read a statement issued by a no doubt relieved McDonald's, which had probably contemplated a future where "over 99 billion served" would include the word "subpoenas."

"The plaintiffs have alleged that the practices of McDonald's in making and selling their products are deceptive and that this deception has caused the minors who have consumed McDonald's products to injure their health by becoming obese," observed Judge Robert W. Sweet in his ruling. In other words, the kids asked, how were we to know that a steady diet of hamburgers and french fries was going to make us fat? And the judge's response was, well, they *should* know and they therefore "cannot blame McDonald's if they, nonetheless, choose to satiate their appetite with a surfeit of supersized McDonald's products."

Sweet clearly came down on the side of personal responsibility, a stance he has long taken. "In the interest of consistency and integrity," he wrote in a footnote, "it should be noted that the author of this opinion publicly opposed the criminalization of drugs.... This belief is based upon the notion that, as long as consumers have adequate knowledge about even harmful substances, they should be entitled to purchase them, and that the issue should be one of health, not the criminal law."

Which brings us to Item Two, namely, the current odd ad campaign sponsored by the Office of National Drug Control Policy. My favorite in this series of public-service spots features two teenage boys smoking marijuana in the den of what appears to be an upper-middle-class home. After some insipid pot-inspired conversation, one dumb kid finds a handgun. The other dumb kid says, "Cool," and asks if it's loaded. The first dumb kid points the gun at the second, says, "No," and pulls the trigger. We hear a gunshot as the screen goes dark and then read the following: "Marijuana can distort your sense of reality. Harmless?"

This is an antidrug commercial? Because it sure looks to me like an antigun commercial. I grew up in a house with a rifle, which was dismantled and in a locked case, and I *know* that no card-carrying member of the National Rifle Association would keep a handgun loaded and easily accessible, especially in a home with children. Furthermore, I put forth the proposition that I would much rather find myself in a room full of stoned teenagers without guns than in a room full of straight teenagers carrying loaded weapons. (I actually did the first part of this experiment in the 1970s, but I categorically refuse to do the second part, unless the teenagers in question are in boot camp at Parris Island.)

Now, before I get audited, no one is suggesting that teenagers use drugs, just as no one is suggesting that they drink beer until they puke, except for companies that do extensive advertising to college kids, most of whom are below the legal drinking age. But I digress. Why not let teenagers know about the dangers of drug use in a way that would actually get their attention?

For example, a TV commercial that might very well drive many teenage boys away from marijuana could say: "Smoking a lot of pot may give men breasts worthy of the *Sports Illustrated* swimsuit issue." I can envision another commercial that ends with a sober guy telling his stoned buddy, "Look, there's not enough dope in the world to get her to sleep with you." And finally, how about one in which two fat stoners sue their dealer because the munchies made them eat too much McDonald's? SA

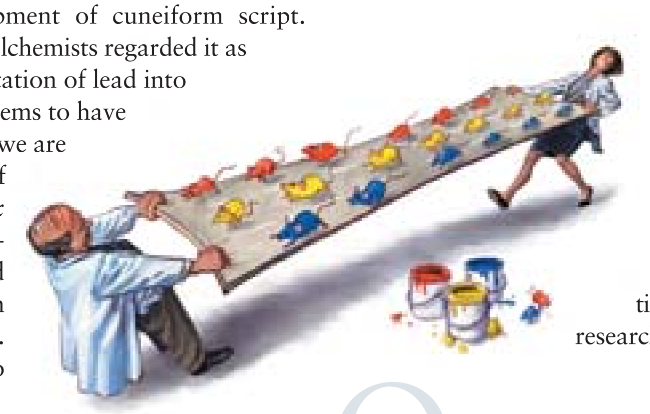
ILLUSTRATION BY MATT COLLINS; PHOTOGRAPH BY FRANK VERONSKY

What is the importance of this recent scientific discovery?

—Mervin Stykes, Cretin-on-Hudson, N.Y.

Norman Sansperson, professor of applied heuristics at the University of Winnepassaic, offers this perspective:

Since the dawn of time, people have wondered about this seemingly insoluble problem. Clay tablets unearthed during the 1920s in Mesopotamia indicate that the Sumerians expended considerable effort on it in 2500 B.C.; it might even have indirectly influenced their development of cuneiform script. Paracelsus and other European alchemists regarded it as a problem equal to the transmutation of lead into gold. Even Leonardo da Vinci seems to have dabbled in finding an answer, if we are to believe pages 31 through 35 of his notebook known as the *Codex Worcester* (he disguised his speculations with lemon juice and mirror writing to keep them from the prying eyes of the Church). Clearly, we have not been able to find an answer fast enough!



The recent announcement marks a clear breakthrough in our understanding. Although not all the pieces have yet dropped into place, scientists are now cautiously optimistic that a final answer may soon be within reach. More research will of course be needed, but this basic research discovery holds the potential to someday transform a wide variety of essential technologies. Some analysts predict that it could give birth to a major growth industry for the 21st century.

To understand what has recently been learned, an analogy may be useful. Picture the interstate highway system of the U.S. but, in place of automobiles, imagine that the roads are covered with mice. Some of these mice are painted red, others blue, others yellow, but they are all running in straight lines at 100 miles an hour. If some force were to stretch sections of the road like taffy, the mice would move farther apart but would not lose any of their information-carrying capacity.

Our new grasp of the situation might have proved impossible were it not for a series of advances in imaging over the past 20 years. (Poor Leonardo! Even 200 hand lenses could not have given him sufficient magnification!) Atomic force microscopy has improved so dramatically that we can now not only visualize individual atoms but also move them like pieces on a chessboard. By a happy coincidence, NASA investigators had been developing a new set of algorithms for use with data from the Hubble Space Telescope. The combination of these techniques in the mid-1990s gave us the serendipitous key we needed.

The new discovery is not without controversy. Protesters—particularly in Europe—have announced their opposition to any application of this work to the genetic modification of food crops. Industry lobbyists argue that the work is too preliminary for Congress to pass the proposed regulations. And the global-warming implications are especially troubling. More research is needed.

Hey, what was that?

—Miller Huntington, Afteryew, Miss.

Irene Adler Holmes, manager of the Mycroft Coffee Shop, explains:

Air masses warmed by the sun rise above colder air masses, leaving zones of low pressure. In the Northern Hemisphere, lateral currents of air spiraling counterclockwise fill the resulting depression. It was probably the wind.

Which month is cruelest?

—I. Newton, Cambridge, England

T. S. Eliot, treasurer of the Dead Poets Society, replies:

April is the cruelest month, breeding lilacs out of the dead land, mixing memory and desire, stirring dull roots with spring rain. November, also very bad. SA

For genuinely authoritative answers to much better questions than these, visit www.sciam.com/askexpert

ILLUSTRATION BY MATT COLLINS


EVOLUTION OF THE SPECIES
HOMO SAPIENS SAPIENS, 40,000 years ago
HOMO REGULARIS, ca. 1947
HOMO NONELECTRONICUS, ca. 2003
HOMO ANALOGICUS, ca. 2003
HOMO DIGITALICUS, ca. 2003