IRAN'S NUCLEAR MENACE • MAPPING THE SEAFLOOR • FROM CATS TO QUANTUM PHYSICS

SCIENTIFIC AMERICAN

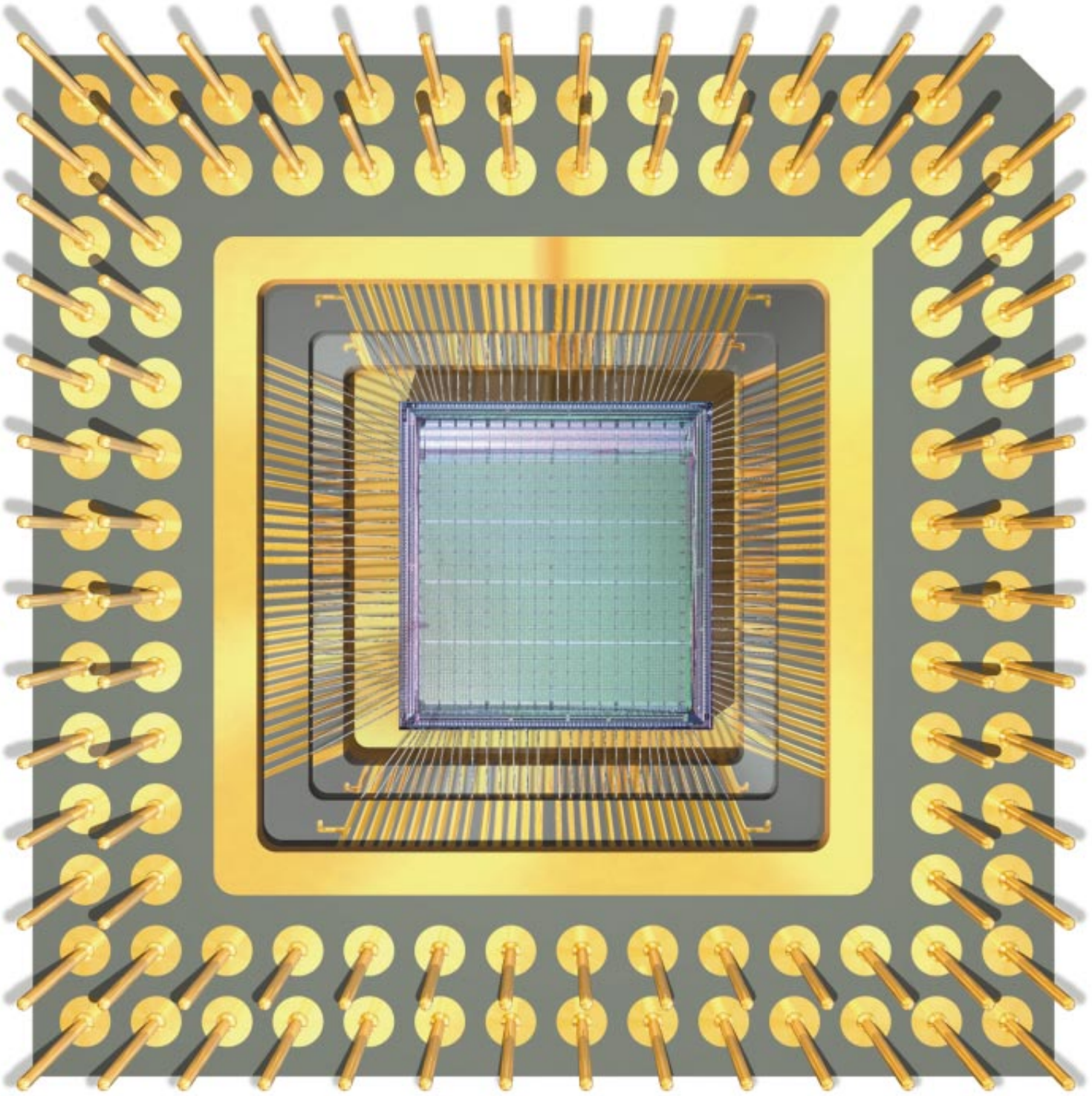
JUNE 1997 $4.95

SPECIAL REPORT

GENE THERAPY:

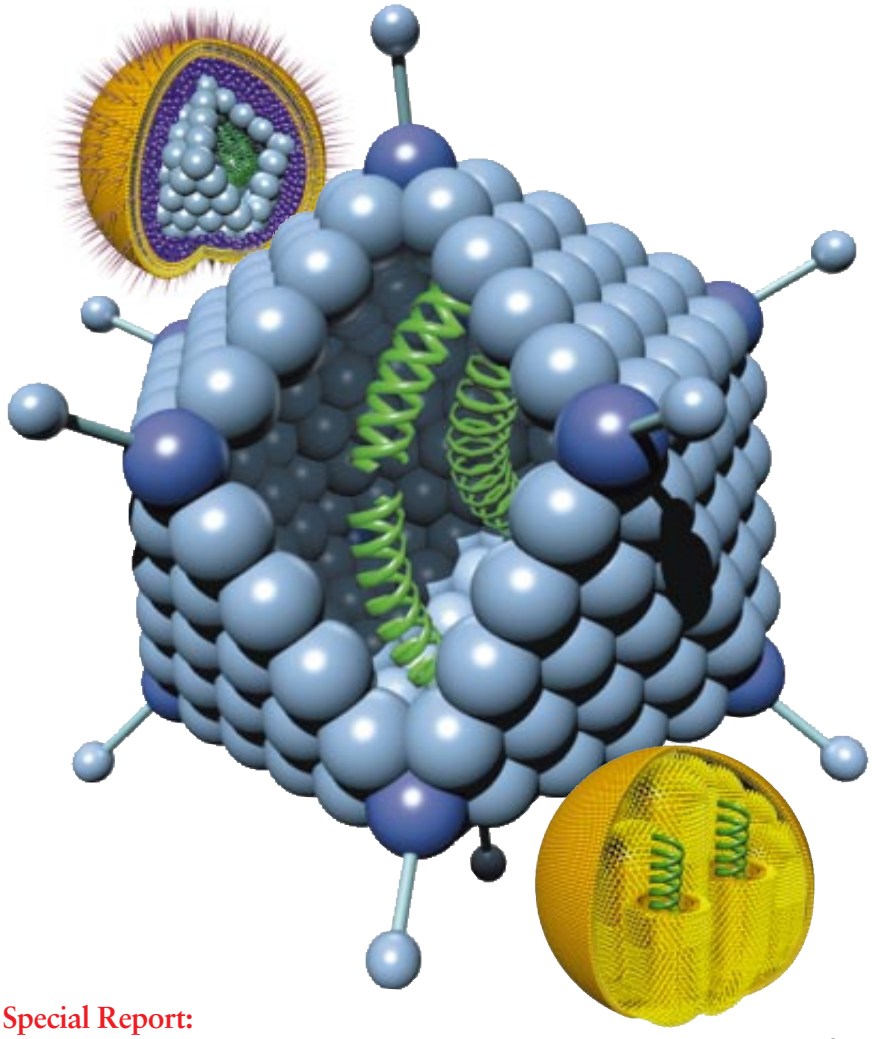
HOW IT WILL WORK AGAINST CANCER, AIDS, ALZHEIMER'S AND MORE

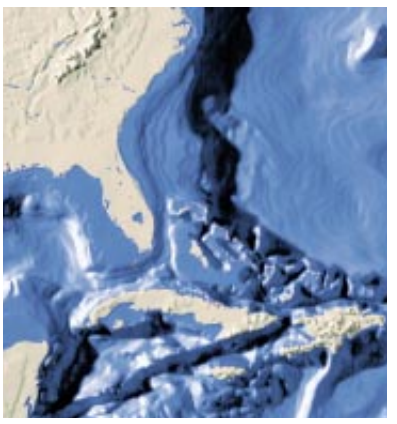
The microchip that rewires itself



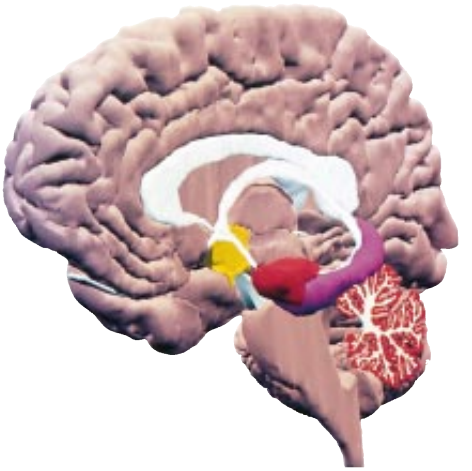
SCIENTIFIC AMERICAN

June 1997 Volume 276 Number 6





Scientific American (ISSN 0036-8733), published monthly by Scientific American, Inc., 415 Madison Avenue, New York, N.Y. 10017-1111. 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 (outside U.S. $47). Institutional price: one year $39.95 (outside U.S. $50.95). 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; fax: (212) 355-0408 or send e-mail to info@sciam.com Visit our World Wide Web site at http://www.sciam.com/ **Subscription inquiries: U.S. and Canada (800) 333-1199; other (515) 247-7631.**



About the Cover
The architecture of this integrated circuit includes field-programmable gate arrays (FPGAs), which can be physically modified during the chip's operation to improve its performance. Chip supplied by Xilinx; image by Slim Films.

Recognizing Technological Genius

Americans have always taken great, justified pride in their inventiveness. Even at the birth of this country, that tradition was in place: when the founding fathers weren't busy inventing the U.S., they were often inventing other useful things, too. Benjamin Franklin was the archetypal American Leonardo, a Renaissance man born two centuries too late. He invented bifocals, the lightning rod and his eponymous stove and pioneered the study of electricity. Thomas Jefferson is revered as an architect for having designed and built both his own home, Monticello, and the University of Virginia. But he was also an inveterate tinkerer and fan of new gadgetry and an ardent practitioner of scientific farming. His improvements to the moldboard of the common agricultural plow eventually led to that design becoming the standard for its time.

INVENTIVENESS builds national prosperity.

The roster of this country's technology innovators is long. To name only a few: Thomas Alva Edison. Alexander Graham Bell. Henry Ford. George Washington Carver. Eli Whitney. Orville and Wilbur Wright. Robert Fulton. Buckminster Fuller. Charles Goodyear. Samuel Morse. Elias Howe. George Eastman. Elmer Ambrose Sperry. Charles A. Lindbergh. Edwin H. Land. Grace Murray Hopper. Jack Kilby and Robert Noyce. Jonas Salk. Robert H. Goddard. Vannevar Bush. This country could barely have survived, let alone flourished, without their genius.

Against that backdrop of achievement, the National Medal of Technology stands as the preeminent honor that can be bestowed on any American for excellence in technological innovation. Since 1985 the president of the U.S. has annually awarded this recognition to individuals and corporate teams who, in the opinion of the independent steering committee, have made lasting contributions to American competitiveness and to standards of living.

SCIENTIFIC AMERICAN has of course always had its own strong interests in these areas, since it was founded in 1845 as "The Advocate of Industry and Enterprise, and Journal of Mechanical and Other Improvements." More than a few of the past and present winners have previously written about their work for this magazine. We are delighted to be associated with the National Medal of Technology and to join President Bill Clinton and the Department of Commerce in saluting this year's winners. A special bulletin describing them and their accomplishments appears this month, beginning on page 16. We commend them for their inspiration and for the real benefits they have brought to this republic.

JOHN RENNIE, *Editor in Chief*
editors@sciam.com

SCIENTIFIC AMERICAN®
Established 1845

John Rennie, EDITOR IN CHIEF

Board of Editors
Michelle Press, MANAGING EDITOR
Philip M. Yam, NEWS EDITOR
Ricki L. Rusting, ASSOCIATE EDITOR
Timothy M. Beardsley, ASSOCIATE EDITOR
Gary Stix, ASSOCIATE EDITOR
John Horgan, SENIOR WRITER
Corey S. Powell, ELECTRONIC FEATURES EDITOR
W. Wayt Gibbs; Kristin Leutwyler; Madhusree Mukerjee; Sasha Nemecek; David A. Schneider; Glenn Zorpette
Marguerite Holloway, CONTRIBUTING EDITOR
Paul Wallich, CONTRIBUTING EDITOR

Art
Edward Bell, ART DIRECTOR
Jessie Nathans, SENIOR ASSOCIATE ART DIRECTOR
Jana Brenning, ASSOCIATE ART DIRECTOR
Johnny Johnson, ASSISTANT ART DIRECTOR
Jennifer C. Christiansen, ASSISTANT ART DIRECTOR
Bridget Gerety, PHOTOGRAPHY EDITOR
Lisa Burnett, PRODUCTION EDITOR

Copy
Maria-Christina Keller, COPY CHIEF
Molly K. Frances; Daniel C. Schlenoff; Terrance Dolan

Administration
Rob Gaines, EDITORIAL ADMINISTRATOR
Sonja Rosenzweig

Production
Richard Sasso, ASSOCIATE PUBLISHER/ VICE PRESIDENT, PRODUCTION
William Sherman, DIRECTOR, PRODUCTION
Carol Albert, PRINT PRODUCTION MANAGER
Janet Cermak, MANUFACTURING MANAGER
Tanya DeSilva, PREPRESS MANAGER
Silvia Di Placido, QUALITY CONTROL MANAGER
Carol Hansen, COMPOSITION MANAGER
Madelyn Keyes, SYSTEMS MANAGER
Carl Cherebin, AD TRAFFIC; Norma Jones

Circulation
Lorraine Leib Terlecki, ASSOCIATE PUBLISHER/ CIRCULATION DIRECTOR
Katherine Robold, CIRCULATION MANAGER
Joanne Guralnick, CIRCULATION PROMOTION MANAGER
Rosa Davis, FULFILLMENT MANAGER

Advertising
Kate Dobson, ASSOCIATE PUBLISHER/ADVERTISING DIRECTOR
OFFICES: NEW YORK:
Meryle Lowenthal, NEW YORK ADVERTISING MANAGER
Randy James; Thomas Potratz,
Elizabeth Ryan; Timothy Whiting.
DETROIT: 3000 Town Center, Suite 1435,
Southfield, MI 48075; Edward A. Bartley, DETROIT MANAGER
WEST COAST: 1554 S. Sepulveda Blvd., Suite 212,
Los Angeles, CA 90025;
Lisa K. Carden, WEST COAST MANAGER; Tonia Wendt.
225 Bush St., Suite 1453,
San Francisco, CA 94104; Debra Silver.
CANADA: Fenn Company, Inc. DALLAS: Griffith Group

Marketing Services
Laura Salant, MARKETING DIRECTOR
Diane Schube, PROMOTION MANAGER
Susan Spirakis, RESEARCH MANAGER
Nancy Mongelli, ASSISTANT MARKETING MANAGER

International
EUROPE: Roy Edwards, INTERNATIONAL ADVERTISING DIRECTOR, London. HONG KONG: Stephen Hutton, Hutton Media Ltd., Wanchai. MIDDLE EAST: Peter Smith, Peter Smith Media and Marketing, Devon, England. PARIS: Bill Cameron Ward, Inflight Europe Ltd. PORTUGAL: Mariana Inverno, Publicosmos Ltda., Parede. BRUSSELS: Reginald Hoe, Europa S.A. SEOUL: Biscom, Inc. TOKYO: Nikkei International Ltd.

Business Administration
Joachim P. Rosler, PUBLISHER
Marie M. Beaumonte, GENERAL MANAGER
Alyson M. Lane, BUSINESS MANAGER
Constance Holmes, MANAGER, ADVERTISING ACCOUNTING AND COORDINATION

Chairman and Chief Executive Officer
John J. Hanley

Corporate Officers
Robert L. Biewen, Frances Newburg, John J. Moeling, Jr., Joachim P. Rosler, VICE PRESIDENTS
Anthony C. Degutis, CHIEF FINANCIAL OFFICER

Program Development
Linnéa C. Elliott, DIRECTOR

Electronic Publishing
Martin O. K. Paul, DIRECTOR

SCIENTIFIC AMERICAN, INC.
415 Madison Avenue • New York, NY 10017-1111
(212) 754-0550

PRINTED IN U.S.A.

LETTERS TO THE EDITORS

ANIMAL EXPERIMENTATION DEBATE, ROUND TWO

Since I was quoted by Neal D. Barnard and Stephen R. Kaufman in their debate article "Animal Research Is Wasteful and Misleading" [February] to justify an argument with which I disagree, I feel I must respond. My 1982 article, [which they cited to show the inadequacy of animal testing], was an evaluation of a specific biological assay for carcinogenicity that fails to meet the minimum standards of good scientific design. Just because some people do foolish things with animals is no reason to believe that all experiments using animals are worthless. The science of pharmacology has brought great understanding to the study of life, much of it through animal research and testing.

DAVID SALSBURG
New London, Conn.

Scientific American has a distinguished history of presenting readable, accurate articles from various scientific disciplines. That is why I am mystified that the editors chose to cast doubt on the vital contributions of animal research to medical progress by presenting the topic as a debate between two scientists and two animal rightists hiding behind M.D. degrees. What did you hope to gain by giving voice to two individuals who have previously been caught in gross distortions of medical history and publicly chastised?

Indeed, one even presented false credentials to your readers. Barnard has no track record in nutrition science; he is a psychiatrist. You have naively melded two issues into one: philosophical or religious beliefs concerning the relation of humans to animals have been allowed to hide behind allegedly scientific facts. A discussion of the former topic could have been presented in association with the scientific article by Jack H. Botting and Adrian R. Morrison ["Animal Research Is Vital to Medicine," February], but theirs should not have been paired with a masquerade.

FREDERICK K. GOODWIN
Director, Center on Neuroscience, Medical Progress and Society

I think it is time that the American myth that (to quote staff writer Madhusree Mukerjee) "in 1975 the animal-rights movement exploded onto the scene with the publication of *Animal Liberation*" should itself be exploded ["Trends in Animal Research," February]. Peter Singer's superb book was not about the idea of promoting animal rights but was an attack on speciesism—a concept I invented in Britain in 1970, as Singer has always acknowledged.

"I applaud those researchers who, like me, support both animal rights and animal research."

Singer was importing the concept from Oxford to New York. But it is praiseworthy indeed that Scientific American should seek to find some middle ground between animal welfarists and thoughtful scientists—something we have already partly succeeded in doing in Britain, where a civilized dialogue, helpful to animals and scientists alike, is well advanced.

RICHARD D. RYDER
Royal Society for the Prevention of Cruelty to Animals

Of course, I care about my daughter more than any other animal. Of course, I would choose to save her life at the expense of any other animal—human or nonhuman. But emotional bonds—to kin, countrymen, race or species—are no basis for a consistent morality. As the philosopher Arthur Schopenhauer said, "Universal compassion is the only guarantee of morality."

ANNE M. GREEN
Carnegie Mellon University

As a premedical student with an interest in neurology, I fully support animal research. But what is so difficult about admitting that, yes, we are making these animals suffer for our benefit and then trying to make their lives as comfortable as possible? Animal-rights extremists cannot expect animal research to cease, but they are not wrong in demanding that the treatment of animals be policed. Scientists cannot expect to go along without rules and governing bodies that represent both sides of this issue. I applaud those researchers who, like me, support both animal rights and animal research.

MARY SHAUGHNESSY
Houston, Tex.

The Editors reply:

Surely Goodwin does not think we invented the debate between scientists and animal rightists. What we hoped to gain was something that rarely occurs, an intelligent exchange between these sides that the public could judge on its own merits. Barnard is a psychiatrist, but he has also published numerous articles and books on nutrition. Ad hominem criticisms aside, the fact remains that Barnard and Kaufman head up two of the largest organizations of physicians critical of animal experimentation. However one feels about the views of those physicians, they are part of the biomedical community, and Barnard and Kaufman represent their position.

WHITHER BELL LABS?

I enjoyed the profile of our former colleague Ronald L. Graham of AT&T Labs ["Juggling Act," by John Horgan, March] with one small exception. Creation of AT&T Labs does not mean that Bell Labs disappeared. Bell Labs is still Bell Labs. We are the research and development arm of Lucent Technologies, the communications systems and software company spun off last year by AT&T.

ARUN N. NETRAVALI
Vice President, Research, Bell Labs

LARGE NUMBERS

A very amusing story by Richard E. Crandall ["The Challenge of Large Numbers," February]. But what's the point of numbers such as pi to a billion places? I've calculated that by the time you can tell me what it's good for, my beer can will have fallen over.

STEPHEN ZANICHKOWSKY
Brooklyn, N.Y.

Letters selected for publication may be edited for length and clarity.

50, 100 AND 150 YEARS AGO

SCIENTIFIC AMERICAN

JUNE 1947

COMPETITION FOR LEAD—"Plastics are eating their way into former lead applications. They can replace lead in tank linings and pipes in the chemical industry and in cable sheathing in the electrical industry. Substitute pigments are being developed for paints, to take the place of time-honored white lead. Glass and rubber offer many of the inert advantages of lead and are being used for jobs where lead was formerly the only material considered."

NEW FUNGICIDE—"An agricultural fungicide derived from petroleum and sulfur has been developed. The latex-like material holds chemicals so that they cannot be washed away by rain or dew, thus providing maximum killing action against blights and diseases. After spraying and drying, the material forms a microscopic web that can be removed only by scraping, decomposition of the materials, or expansion by growth."

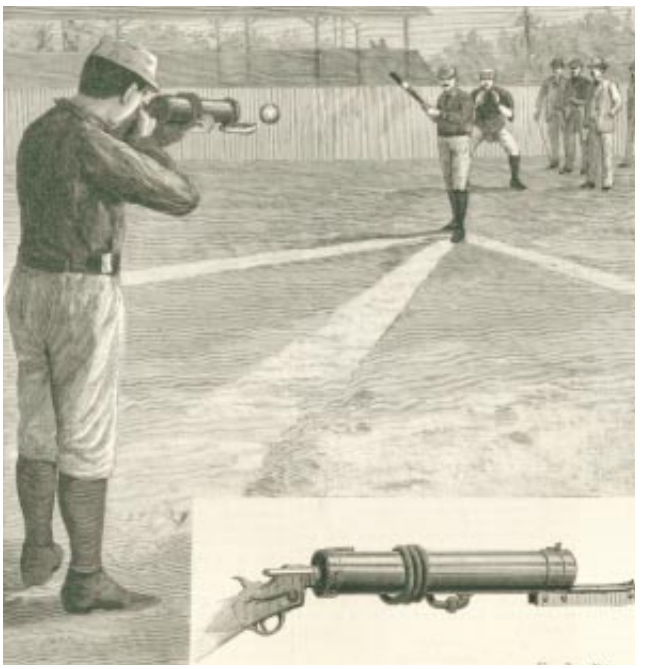
JUNE 1897

FLUORINE LIQUEFIED—"The distinguished chemist Prof. James Dewar has just succeeded in liquefying fluorine gas at a temperature of –185° C. The product was a yellow mobile liquid which had lost chemical activity. Great interest has been felt in the element fluorine since its isolation by M. Moissan in 1887. The efforts of chemists to investigate it in a satisfactory manner were baffled, because its chemical affinities were so numerous and acute that, when driven from one combination, it instantly combined with some other substance with which it came in contact. Owing to this difficulty, there had been some uncertainty as to its elementary nature."

THE CINEMATOGRAPH—"The popularity of the art of moving- or chrono-photography has led to the invention of numerous devices. One of the most recent cameras is that invented by the Lumière Brothers, of Paris, France. An ingenious device for producing an intermittent movement without sprocket wheels or cogs is one of the features of the camera, while its lightness and facility of operation make it adaptable for use in most any place. The same camera can be converted into a projecting apparatus for throwing moving pictures on the screen."

WHALE HUNTING—"Owing to the scarcity of right whales in northern waters, Newfoundland is about to follow the example of Norway in making humpbacks and fin whales, which are said to be found in immense numbers round the coast, the objects of systematic pursuit. The superintendent of fisheries has organized a fleet of small steamers, with harpoons and explosive lances, such as are used in Norway, to carry on the fishery. If the whalers of Newfoundland take many specimens, it might be worthwhile to try preparing its flesh for the market. If the prejudice against its use could be overcome, there is no reason why 'whale steak,' preserved and put up in tins, should not find ready sale."

MECHANICAL BASEBALL PITCHER—"We present some engravings of the new gunpowder gun for pitching a baseball, tried at the Princeton ball field on June 8, 9 and 10. A charge of powder in a tube coiled about the barrel is ignited, the gases are delivered behind the ball and it is flung from the barrel. Two 'fingers,' thin plates of metal curved and covered with rubber, project over the thickness of the barrel, and impart a velocity of spin to the ball; this spin gives it a curved path."



Baseball gun for delivering a curved ball

JUNE 1847

ASTROPHOTOGRAPHY—"Combining the telescope with the Daguerreotype in astronomy has lately occupied the attention of the Royal Society of Bohemia. Professor Christian Doppler says that the extreme susceptibility of the human eye for impressions is surpassed many thousands of times by an iodized silver plate. The diameter of one of the papillae of the retina is no more than 1/8000 of an inch, but on the space of a Daguerre plate equal to one retina papillae, more than 40,000 minute globes of mercury are to be met with. Therefore images of the smallest fixed stars can be obtained."

FIREPROOFING SAILING VESSELS—"A gentleman in Glasgow, Scotland, suggests a ready method to prevent sailing vessels from being consumed by fire. Every vessel should carry as ballast a quantity of chalk. In the event of fire in the hold, by pouring diluted sulphuric acid onto the chalk, such a quantity of carbonic acid gas [carbon dioxide] would be generated as would effectually put out the flames."

COURTESY OF U.S. DEPARTMENT OF COMMERCE

The 1997 National Medal of Technology

RAY M. DOLBY
Chairman
Dolby Laboratories

ROBERT S. LEDLEY
Professor of radiology, physiology and biophysics
Georgetown University Medical Center

NORMAN R. AUGUSTINE
Chairman and CEO
Lockheed Martin Corporation

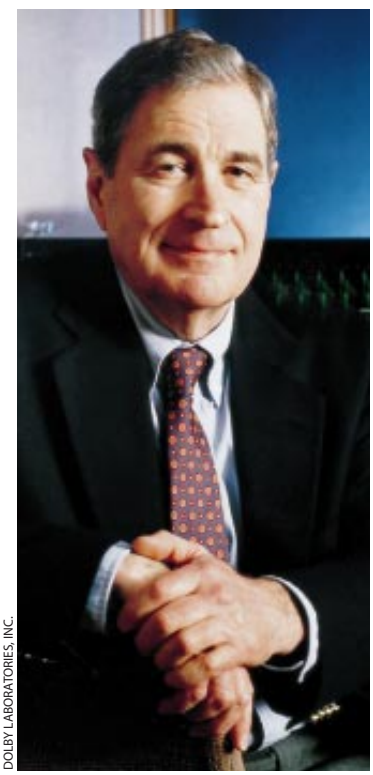
VINTON G. CERF
Senior vice president of Internet architecture
MCI Communications Corporation

ROBERT E. KAHN
President
Corporation for Research Initiatives

Among this nation's highest honors, this prize recognizes outstanding achievements in the innovation, development and commercialization of technology, as well as the human resource management that advances innovation. This year's winners include an audio pioneer, a biomedical inventor, an aerospace executive and two Internet designers

RAY M. DOLBY

Chairman
Dolby Laboratories
San Francisco, Calif.



DOLBY LABORATORIES, INC.

The next time you are in a movie theater and the sound of a volcano erupting on screen sets your teeth to shaking or the pure sound of music on tape mentally transports you to a concert hall, spare a moment's thanks for Ray M. Dolby. Over the past 30 years he has profoundly influenced the science of sound recording and reproduction through his nearly ubiquitous Dolby technologies. Products incorporating his innovations range from the cassettes played in personal headsets and car stereos to the soundtracks of blockbuster films. This year he receives the National Medal of Technology for his inventions and for fostering their adoption worldwide through the products and programs of his company.

Dolby's involvement in sound engineering started early. While earning his undergraduate degree in electrical engineering at Stanford University, from which he graduated in 1957, he worked with the team at Ampex Corporation that produced the first practical video recorder. Dolby went on to receive his doctorate in physics from the University of Cambridge in 1963.

In 1965, shortly after his return to the U.S., he founded Dolby Laboratories to develop and commercialize his budding ideas for improving the ways in which sounds were recorded and reproduced from tape. Among his first inventions was a signal-processing method that eliminated the noise, usually noticeable as a hiss, inherent in most tape recordings. Unlike other, earlier antihiss techniques, Dolby's did not distort the underlying sound quality. His method involved separating and sorting the acoustic components of a given sound into different electronic channels according to their frequency and amplitude, eliminating those signals that contributed most to noise and then recombining the other components.

Dolby first marketed his technology to sound studios, where it helped to spawn an era of sophisticated multitrack recording that transformed the music industry. Later, Dolby Laboratories developed less costly noise reduction methods suitable for home audio systems and other consumer products.

Cassettes employing Dolby noise reduction quickly overtook long-playing records as the leading medium for prerecorded music; not until the early 1990s were cassettes surpassed by compact discs. Yet his technique has also evolved with the times: his original analog signal-processing methods have yielded to digital ones, which now shape the sound of audiocassettes as well as laser discs, video games and multimedia products. Overall, consumers have purchased more than 600 million products incorporating Dolby technologies.

Meanwhile other Dolby creations have gone Hollywood. In 1975 he introduced a multichannel soundtrack for optical films that produced higher-quality sound, at a lower cost, than previous multichannel methods. The Dolby soundtrack not only produced stereo sound but also provided extra channels for special effects, such as the low-frequency rumblings that make cinematic earthquakes and explosions more realistic.

EVERETT COLLECTION

***STARS WARS,* originally released in 1977, was one of the first major films to employ Dolby sound.**

George Lucas, one of the first directors to put Dolby sound into his films, credits the technology with having helped *Star Wars* become a hit in 1977. More than 6,000 feature films with Dolby-encoded soundtracks have been released since then, and Dolby-based playback equipment has been installed in more than 33,000 theaters worldwide.

Although Dolby Laboratories continues to manufacture sound equipment for professionals, it has disseminated its inventions primarily by licensing patents to other manufacturers, a strategy now common in the electronics industry. Although licensing fees are kept low to maximize market share, the company has so far earned a total of more than $250 million in royalties.

ROBERT S. LEDLEY

Professor of radiology,
physiology and biophysics
Georgetown University
Medical Center
Washington, D.C.

COURTESY OF ROBERT S. LEDLEY

The science of medical imaging has changed dramatically over the past quarter century. Until then, simple black-and-white x-ray photographs commonly represented the state of the art for peering into the body to diagnose what might be wrong with it. Today physicians can order up a variety of colorful imaging technologies—computed tomography (CT), magnetic resonance imaging (MRI), positron-emission tomography (PET) and more—that can virtually dissect out a troubled organ or tissue for scrutiny from any angle. That newfound diagnostic capability, along with many others, owes much to the work of Robert S. Ledley, whose contributions have now been honored with a National Medal of Technology.

Throughout his career, Ledley has excelled at applying advances in information processing to the field of medicine. He obtained a doctorate in dental surgery from New York University in 1948 and a master's degree in mathematical physics from Columbia University in 1949. After a stint in the U.S. Army, he went on to work for the National Bureau of Standards and at Johns Hopkins University before arriving at Georgetown University in 1970.

In 1959 Ledley and Lee B. Lusted of the University of Rochester co-authored "Reasoning Foundations of Medical Diagnosis," published in *Science*. That paper, along with Ledley's 1965 book, *Use of Computers in Biology and Medicine,* has been credited with sowing the seeds of the field of medical informatics, in which computers and other information technologies aid physicians in diagnosing and treating patients.

Ledley pioneered the creation of biomedical databases in the mid-1960s. Together with the late Margaret O. Dayhoff of the National Biomedical Research Foundation, he compiled a list of all known sequences of proteins and nucleic acids. Originally published as a book, *Atlas of Protein Sequence and Structure,* it was later released in electronic form under the name Protein Information Resource. (It is soon to be placed on the World Wide Web.) The success of these ventures encouraged the creation of similar databases, which have proved to be crucial for biomedical research.

Ledley was also a leader in the automation of prenatal

COURTESY OF ROBERT S. LEDLEY

FIRST FULL-BODY CT MACHINE, the automatic computed transverse axial (ACTA) scanner, shown here with its inventor, Ledley, is now on display at the Smithsonian Institution's National Museum of American History.

screening for birth defects. In the 1960s he developed algorithms and instruments, including a motorized microscope with pattern-recognition capability, that made it possible to scan chromosomes for abnormalities that cause Down's syndrome and other disorders. He has recently refined those techniques for the detection of more subtle genetic mutations associated with cancer and other diseases.

Perhaps Ledley's most prominent contributions, however, have been improvements in CT scanning technology. In the early 1970s Ledley invented the automatic computed transverse axial (ACTA) scanner, which was the first CT scanner capable of making cross-sectional images of any part of the human body. The device revolutionized the fields of radiology and medical imaging and set the standard for all subsequent CT scanners.

Previous CT scanners, which create three-dimensional images of the interior of the body by passing x-rays through it from various angles, required that the object being scanned be housed in a cumbersome, water-filled container that absorbed excess radiation. As a result, such scanners had been limited to studies of the human head. By redesigning the x-ray emitter and detector, the gantry on which they are mounted and the table on which the patient is placed, Ledley created a machine that dispensed with the need for a water container and could focus on any part of the patient's body.

The algorithms that Ledley devised for processing the signals from the x-ray detector also generated sharper images in much less time than previous CT scanners. The algorithms were later adapted for use in magnetic resonance imaging and positron-emission tomography. The prototype of Ledley's ACTA scanner is now on display at the Smithsonian Institution's National Museum of American History.

NORMAN R. AUGUSTINE

Chairman and CEO
Lockheed Martin Corporation
Bethesda, Md.

BARRY THUMMA *AP Photo*

The White House has awarded a National Medal of Technology to Norman R. Augustine for his visionary leadership of the aerospace industry, identifying and championing technical and managerial solutions to the challenges posed by civil and defense systems and helping to maintain U.S. preeminence in this crucial technology sector.

Augustine has spent more than 30 years as an engineer and manager in both the aerospace industry and the U.S. Department of Defense. After obtaining bachelor's and master's degrees in aeronautical engineering from Princeton University, he joined Douglas Aircraft Company in 1958, where he eventually became program manager and chief engineer.

In 1965 he took his first position in the Defense Department, serving as assistant director of research. After a stint at LTV Missiles and Space Company from 1970 to 1973, he returned to the Pentagon as assistant secretary and later as undersecretary of the army. Augustine joined Martin Marietta Corporation in 1977.

During the post–cold war era, which began in the late 1980s, defense spending fell by 60 percent, and Augustine set an example for the rational downsizing of a large defense contractor. He guided Martin Marietta through a series of mergers and acquisitions, culminating in the company's 1995 merger with the Lockheed and Loral corporations. As chairman and chief executive officer of Lockheed Martin, Augustine now oversees 190,000 people, 62,000 of whom are scientists and engineers.

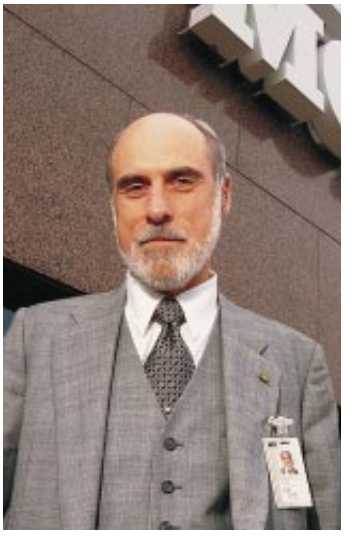
As early as the 1960s, Augustine championed taking military advantage of the U.S. superiority in high tech-

LOCKHEED MARTIN CORPORATION

TITAN MISSILE, the largest unmanned launch vehicle built in the U.S., is one of the most visible products of the newly formed Lockheed Martin Corporation.

nology by building "smart" weapons and other advanced equipment. He has also long advocated the cost-effectiveness of upgrading existing aircraft, ships, tanks and other platforms with more sophisticated electronics—such as radar, computers, communications and electronic-warfare gear—rather than developing new weapons systems from scratch.

Although he has rejoined the private sector, Augustine continues to advise the government on technology policy. In 1990 he headed a committee convened by the White House and the National Aeronautics and Space Administration to consider the future of the civilian space program. The so-called Augustine Report helped to shape NASA's plans for its shuttle program and the International Space Station. In 1986 Augustine chaired a Defense Department task force that recommended the U.S. take steps to bolster its domestic semiconductor industry. As a result, Congress provided funds for the establishment of SEMATECH, an institution that sponsors research on semiconductors by industry, academia and the government.



SHEPARD SHERBELL *SABA*

VINTON G. CERF

VINTON G. CERF

Senior vice president
of Internet architecture
MCI Communications
Corporation
Reston, Va.

ROBERT E. KAHN

President
Corporation for
Research Initiatives
Reston, Va.

Huge computer networks such as the Internet have become such an established part of business life, one might never know that they were once technically impossible. Vinton G. Cerf and Robert E. Kahn, joint recipients of a National Medal of Technology, created and sustained the protocols that made large-scale networks feasible.

Before the time of either personal computers or distributed computing, computer networks were few and isolated. Systems built around different types of hardware and software were essentially incompatible and could not communicate. Machines within an institution might be able to swap data, but they usually could not share it directly with outside machines.

In 1974, however, while Cerf was an assistant professor at Stanford University and Kahn was at the Defense Advanced Research Projects Agency (DARPA), they co-wrote a paper showing how diverse types of networks could be interlinked. They outlined an architecture that called for the creation of nodes, or "gateways," where data from different networks would be processed according to common protocols.

They also advocated a scheme called packet switching, in which messages are broken up into separate bundles of data, or packets. Each packet is assigned a code corresponding to its source and destination. Packets representing a single message can take different routes through a network and can be transmitted with packets from other sources before being reconstructed at the final destination. Communications are both faster and more robust, because calls can be more easily rerouted around areas where lines are congested or have failed.

Their concepts were incorporated into Arpanet, a network created by DARPA that allowed researchers around the U.S. and elsewhere to communicate. (The network was also supposed to serve as a prototype for a classified military network that could withstand a nuclear attack.) In subsequent years Cerf and Kahn steadfastly maintained that the methods they developed should be freely available and should not be associated with any particular vendor. In large part as a result of their efforts, Arpanet evolved into the Internet, which now has more than 30 million users and has spawned one of the nation's most rapidly growing industries.

COURTESY OF CORPORATION FOR RESEARCH INITIATIVES

ROBERT E. KAHN

Cerf earned a B.S. in mathematics and computer science from Stanford in 1965 and a Ph.D. in computer science from the University of California at Los Angeles in 1972. Kahn obtained a bachelor's degree in electrical engineering from the City College of New York in 1960 and a doctorate from Princeton University in 1964. They worked together both at DARPA, which Kahn joined in 1972 and Cerf in 1976, and at the Corporation for Research Initiatives, a nonprofit organization they founded in 1986. Cerf accepted his current position at MCI in 1994.

—Reporting by the Editors

MICHAEL GOODMAN

PACKET SWITCHING breaks messages into "packets" of data, each of which is tagged with a code (signified here by colors). The packets from a single message may take different routes through a network before being reconstructed at their destination.

NEWS AND ANALYSIS

IN FOCUS

PINNING DOWN INFLATION

Cosmologists strive to preserve a popular theory of creation

JESSICA BOYATT

***INFLATION'S CREATOR,** Alan H. Guth, hopes future observations will confirm his "vague idea."*

In 1979 a young physicist at the Stanford Linear Accelerator, while tinkering with some leading theories of elementary particles, reached a startling conclusion. Under the extreme conditions that might have prevailed in the primordial universe, gravity may have briefly become a repulsive rather than attractive force, causing the cosmos to undergo a stupendous growth spurt before subsiding to the relatively sedate expansion observed today. His excitement mounting, Alan H. Guth wrote "SPECTACULAR REALIZATION" in his notebook and set it off from the surrounding equations with two concentric boxes.

Guth's exhilaration turned out to be warranted. His theory, which he called inflation, explained some of the universe's fundamental features, such as the uniformity of the big bang's afterglow. The cosmological community immediately embraced inflation, as Guth himself recounts in his new book, *The Inflationary Universe*. More than 3,000 papers on the topic have been published in peer-reviewed scientific journals since Guth's original article in 1981. Many theorists would agree with Alan P. Lightman of the Massachusetts Institute of Technology that inflation is "the most significant new development in cosmological thinking" since the big bang theory itself.

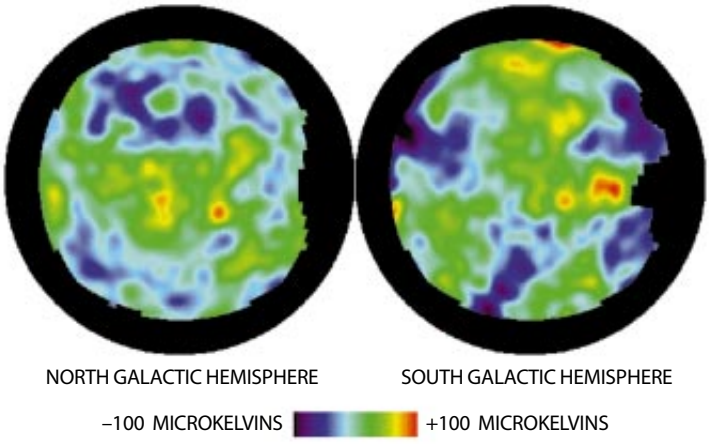
On the other hand, recent observations have contradicted one major prediction of inflation—or at least the version favored by most cosmologists. Some worry that even if inflation survives this challenge, it may never be confirmed in the same sense that the big bang theory has been. Guth, whose proposal helped to win him a full professorship at M.I.T., acknowledges that inflation remains a "vague idea" in need of substantiation. But he is confident that further observations and theoretical work will uphold the theory. "Inflation is here to stay," he declares.

Inflation's persistent popularity stems from its ability to resolve several cosmic conundrums, such as the apparent lack of curvature, or flatness, of space. According to general rela-

tivity, space could have assumed an infinite number of curvatures. Open-curvature universes expand forever; closed ones eventually collapse back on themselves. But inflation's exponential expansion of the universe would render it utterly flat, just as blowing up a balloon smoothes out its wrinkles. A flat cosmos keeps expanding but at an ever decreasing rate.

Similarly, inflation explained why the microwave radiation pervading the universe—thought to be the afterglow of the big bang—appears so homogeneous, or smooth, in all directions. Calculations based on preinflation models had suggested that there had not been time after the big bang for conditions to reach thermal equilibrium. But if inflation occurred, the visible universe emerged from a region so tiny that it had time to reach equilibrium before it inflated.

Guth's proposal also suggested why the universe is not a completely homogeneous consommé of radiation but contains lumps of matter in the form of stars and galaxies. Quantum mechanics suggests that even empty space contains energy and that this energy constantly fluctuates, like waves rippling on the surface of a windblown lake. The peaks generated by these quantum fluctuations in the nascent universe could have become large enough, after being inflated, to serve as the seeds from which stars and galaxies would grow.



MICROWAVE FLUCTUATIONS
revealed by COBE support inflation, but also other theories.

When the Cosmic Background Explorer (COBE) satellite discovered fluctuations in the microwave radiation in 1992, proponents of inflation crowed that it had been confirmed. Yet other theories—such as those involving cosmic strings, textures and other primordial "defects" in the fabric of space-time—also forecast such fluctuations. "Every prediction of inflation can be mimicked in other ways, albeit in contrived ways," says Neil G. Turok of the University of Cambridge. For that reason, Turok fears, "inflation can never be proved" beyond a reasonable doubt.

Moreover, what was once inflation's greatest asset, its resolution of the flatness problem, has now become its greatest liability. In recent years, measurements of the density of matter in the universe have consistently come up short of the amount needed to produce a flat universe. At a meeting held in March at the University of California at Irvine, several groups reported having found a mass density only 30 percent of what is required for flatness.

To account for the discrepancy, theorists have resuscitated an idea called the cosmological constant; first proposed by Albert Einstein, it assumes that empty space contains enough residual energy to exert an outward pressure on the universe. Although this "vacuum energy" has never been directly detected, it crops up in various models of particle physics and cosmology; in fact, an extremely dense speck of vacuum energy is thought to have triggered inflation. The cosmological constant may be sufficient, theorists suggest, to serve as the missing matter needed to make the universe flat.

Alternatively, P. James E. Peebles of Princeton University and others have showed how inflation might generate an open, rather than absolutely flat, universe. To be sure, open inflation and those versions employing the cosmological constant are more complicated than Guth's original formulation, but Peebles considers that to be a healthy development. "It would have been too remarkable," he says, "for a true model to be so elegant and simple."

Andrei D. Linde of Stanford University admits that he prefers the flat-universe version of inflation. "On the other hand," he points out dryly, "when the universe was created, we were not consulted." He has proposed variants of inflation that produce not only open universes and flat ones but even closed ones. He has also shown that inflation may stem not just from the so-called unified theories investigated by Guth almost two decades ago but from much more generic—albeit still hypothetical—quantum effects.

Observations may invalidate specific versions of inflation, Linde declares, but not the central idea. "If inflation is killed, it will be killed not by experimental data but only by a better theory." Others demur. "When a theory doesn't fit the data," says Charles L. Bennett of the National Aeronautics and Space Administration Goddard Space Flight Center, "it gets more complicated and convoluted until nobody believes it anymore except the founder."

More stringent tests of inflation may emerge from upcoming observations of the microwave background. "That's the cosmological mother lode," says Michael S. Turner of the University of Chicago. NASA's Microwave Anisotropy Probe, or MAP, scheduled for launching in the year 2000, should provide measurements more than 30 times more precise than COBE, and even more discerning observations will emerge from the European Space Agency's Planck Satellite starting in 2004. Several balloon-based missions are also being planned.

Theorists such as Marc Kamionkowski of Columbia University have produced detailed predictions of the imprint that inflation should have left on the microwave background. He concedes Turok's point that other theories can produce similar effects. "But if inflation keeps passing these tests," Kamionkowski says, "that will set it apart from other theories." No other model, he adds, has such explanatory power.

Guth thinks that "when the dust settles," observations may still support the simplest, flat-universe version of inflation. The theory's standing may also be bolstered, he says, by theoretical explorations of superstrings, black holes and other concepts from the frontier of physics, which could yield deeper insights into the universe's murky beginning. At the moment, he notes, such concepts are even more hypothetical than inflation is. "They're not describing the real world yet," he adds, "but that is a big hope for the future." It may take another spectacular realization for Guth's original vision to be completely fulfilled. *—John Horgan*

POLICY

WAR WITHOUT END?

Land mines strain diplomacy as technology advances

As an international push gathers force to ban antipersonnel land mines, new technologies show promise for speeding up humanitarian mine clearing. But negotiations in Geneva aimed at a global ban on the weapons are moving slowly, and a senior official of the Canadian Ministry of Foreign Affairs suggests that the U.S. has not negotiated seriously on an alternative, fast-track Canadian initiative—despite the declaration of 15 retired U.S. military officers that a ban on antipersonnel mines would be "militarily responsible."

At the moment, the U.S. is negotiating a ban through the United Nations Conference on Disarmament in Geneva. But Stephen Goose of Human Rights Watch says it is "increasingly clear to most observers" that the Conference on Disarmament will make little progress. Major mine producers such as China and Russia, as well as some developing countries, have shown scant interest in discussing land mines at the conference, which requires a step-by-step consensus.

In rallying international support for a fast-track treaty, Canada hopes to ban the transfer, production and use of antipersonnel mines, of which 100 million lie hidden in the ground in 64 countries. Most remain deadly for decades, killing and maiming 25,000 every year, mainly civilians. The treaty would most likely be open for signature in Ottawa this year. Although President Bill Clinton has said he seeks a ban, the U.S. has retained the right to use the weapons for now. (The U.S. has said that only in Korea will it use mines that remain dangerous indefinitely.) The U.S. is not participating in the "Ottawa process."

Canadian prime minister Jean Chrétien, who held talks with Clinton in April, spoke of "problems" the U.S. was experiencing at Geneva and asked Clinton to have negotiators state U.S. requirements for joining Canada. The country believes it can muster more than 50 signatories this year. Possibly, later negotiations in Geneva could extend an Ottawa treaty, a Canadian official indicates. But a spokesman for the U.S. National Security Council says the U.S. believes "it is important we try first" in Geneva.

Even after a treaty is signed, it will take decades to make such countries as Cambodia and Bosnia-Herzegovina safe. Worldwide, 20 mines are now emplaced for every one removed. Humanitarian mine clearing has stricter safety demands than military countermine operations, and most deminers still use simple metal detectors and handheld tools.

In the past two years, however, the Department of Defense has developed unclassified devices that could speed up the painfully slow process. The $14-million program was created at the urging of Senator Patrick Leahy of Vermont and is directed from Fort Belvoir near Washington, D.C., by Harry N. ("Hap") Hambric, a retired combat engineer. Commercially available infrared detectors reveal thermal anomalies that surround mines, says Hambric, who demined the Mogadishu bypass with them. And in combination with warming lightbulbs, such detectors can locate otherwise invisible trip wires in vegetation.

With a system that combines various detectors, "I have seen mines in grass," Hambric says. A remotely operated vehicle he is developing unearths the small bombs with an excavator and a supersonic "air knife" that removes soil. Once exposed, mines can be detonated with a spray-on explosive foam, a compound soon to be tested on mines in Cambodia. If an explosion is risky, special guns can inject chemicals that burn a mine's charge quietly. Crude but effective armored vehicles flail the ground or comb through it with tines. Other systems in development consist of probes and shielded vegetation-cutting machinery.

Remote-controlled vehicles equipped with arrays of metal detectors can find in the ground metal pieces weighing less than one gram, Hambric says, and he judges that ground-penetrating radar might in time have comparable potential. A study of humanitarian demining technologies led by Paul Horowitz of Harvard University for the Jason program, a defense advisory group, notes that techniques known as nuclear quadrupole resonance and x-ray backscatter can help in the crucial task of discriminating mines from debris or shrapnel, because they locate explosives.

Sweeps with acoustic and microwave energy sources and sensors might also be valuable, as could be probes that measure thermal conductivity. But for identifying TNT, the main charge in 80 percent of mines, systems that "sniff" vapors might offer the best solution, the Jason study concluded. A few companies are developing sniffer devices. And the Jason report suggests that biotech-

P. DURAND *Sygma*

MINE CLEARING,
shown here being taught by French soldiers to Afghan mujahiddins, could be made less hazardous with new, more sophisticated tools.

IN BRIEF

Topping Taxol

Last December, Samuel J. Danishefsky and his colleagues at Sloan-Kettering Institute for Cancer Research synthesized epothilone-A, an anticancer chemical produced by bacteria. Now they have artificially made its more potent cousin, epothilone-B. Both epothilones are natural products, as is taxol, the well-known cancer drug first derived from the yew tree. And like taxol, both kill tumor cells by stabilizing microtubules—organelles that help cells maintain normal shapes.

Screaming Leaves

Talking to plants seems reasonable, but listening to them? Physicists at the University of Bonn are doing just that to find out why so many geranium seedlings die in transit from the Mediterranean to Germany every year. To do so, they use a sensitive hearing aid: a laser excites ethylene molecules—gas that plants release when exposed to drought, cold or other forms of stress—and a resonance tube amplifies the ensuing shock waves. Higher ethylene emissions make for louder sounds. The device should help reveal what disturbs the green refugees.

DAVID CARRIERE *Tony Stone Worldwide*

Fatal Attraction

As many New York City women are killed by their partner at home as by strangers in robberies, sexual assaults and random attacks combined. (Only 6 percent of men murdered die by their partner's doing.) Health department researchers reexamined all murders of women aged 16 or older that occurred between 1990 and 1994—one of the first such surveys of its kind. Among women killed by their spouse, one third were trying to end the relationship. And in a quarter of the murders attributed to boyfriends, children watched the crime or were killed or injured themselves. The review also revealed that whereas men are typically killed by guns, women are more often beaten and burned.

More "In Brief" on page 24

nology may have a part to play: perhaps tagged bacteria or fruit flies could be engineered to locate TNT.

But as long as mines are being emplaced at current rates, the deminers can never catch up. In the meantime, though, anti-mine sentiment is growing in some corporate ranks. Motorola has said it will seek to ensure its products are not used in the devices, and in response to a campaign by Human Rights Watch, 17 other companies have followed suit. That organization is also naming companies that have made mine components but will not make such a commitment. Top of the list is Alliant Techsystems in Hopkins, Minn.

—Tim Beardsley in Washington, D.C.

ENTOMOLOGY

BEE BLIGHT

Looking for alternatives to the troubled honeybee

For several decades, honeybees in the U.S. have been dying off. The culprits are varied: pesticides, habitat loss and, most acutely, mites. Tiny tracheal mites and the larger varroa mite debilitate and ultimately destroy entire bee colonies. The number of managed colonies fell from roughly six million in the 1940s to three million in 1996. And as for wild honeybees, there are virtually none left, says Hachiro Shimanuki of the U.S. Department of Agriculture's Bee Research Lab. The honeybee's demise has led some entomologists to seek other kinds of bees to carry the pollen load.

Honeybees (*Apis mellifera*) are responsible for pollinating up to $10-billion worth of apples, almonds and other crops every year, a far more valuable service than their simultaneous production of $250-million worth of honey. So far there has been no major shortfall in crops, as large-scale producers rent hives from migratory beekeepers—who move their charges north for the summer and south for the winter. (But migration, which enables bees from different hives to mix, probably helped the mite epidemic to spread: in Canada, where bee transport is limited, there is minimal infestation.) The true sufferers will be those with small orchards and backyard vegetable plots, who rely chiefly on pollination by wild honeybees.

Beekeepers currently use chemicals to treat mites, but "we'd like to get away from using pesticides," comments the USDA's William Bruce. Overuse could lead to the mites developing resistance to the pesticides, as they already have in parts of Europe.

An alternative is to breed honeybees that are naturally resistant to mites. Such insects groom and pick mites off one another, toss out mite-infested pupae from the nests or have short growth cycles that allow fewer mites to breed. Roger Hoopingarner of Michigan State University, who selects for such traits, has produced some colonies that have no mites at all. Still, such "hygienic" honeybees are a long way from the market, because the behavior does not reproduce reliably.

More immediate success might come from native pollinators, which do not get varroa mites. *Apis* is in reality an import, brought over by early settlers. (The mites arrived in the 1980s.) For fruition of some crops such as cranberries, bumblebees are far more efficient, says Suzanne W. T. Batra of the USDA.

SCOTT CAMAZINE *Photo Researchers, Inc.*

VARROA MITE,
which rides on the backs of honeybees, has devastated the bees' colonies.

In Brief, *continued from page 22*

British Blues

Prozac—and other drugs like it, collectively referred to as selective serotonin reuptake inhibitors (SSRIs)—swept the States in the 1980s. Now their popularity has reached Europe's shores. The *British Medical Journal* recently reported that between 1990 and 1995, new prescriptions for antidepressants in England rose by 116 percent; new prescriptions for SSRIs soared by 732 percent.

Gorilla Warfare

The war in Zaire continues to claim scores of innocent victims—including humankind's closest relatives, the apes.

ART WOLFE *Tony Stone Images*

Indeed, a new report from the World Wildlife Fund warns that land mines, forest stripping, random shootings and disease are decimating ape populations in Rwanda, Zaire and Uganda. The western lowland gorilla, for instance, is now extinct in Zaire. And other species, such as chimpanzees and mountain gorillas, are disappearing fast.

MAP Kinase Confusion

It all started in April when Craig Malbon and his colleagues from the State University of New York at Stony Brook announced that they had found *the* switch—an enzyme called mitogen-activated protein (MAP) kinase—behind breast cancer. Malbon tested 30 women, 11 of whom had the disease, and found that MAP kinase levels were five to 20 times higher in tumor cells than in normal breast cells. Shortly after Malbon presented his bold conclusions at a press conference, however, many experts expressed doubt: most dividing cells—be they cancerous or not—have elevated MAP kinase levels.

Tracks or FAQs

Fearing that readers will abandon train sets for computer nets, *Railway Modeller* magazine has taken an unusual stand. The editors have refused to publish URLs, even in ads. Needless to say, more than a few loyal hobbyists are blowing off steam on-line, threatening to cancel their subscriptions. Lucky for them, they can still access the magazine through its publisher's Web site: http://www.mmcltd.co.uk/peco/rm/rm_home.htm

More "In Brief" on page 26

"They don't fool around and go to other plants," she explains, whereas honeybees fly far and wide, dispersing their favors loosely.

The efficacy of other insects is, however, unclear. Very little is known about the 3,500 or more species of native pollinators or their distribution. In the Northeast, with its meadows and forests, they are abundant, but in the Midwest, with its vast expanses of monoculture crops, they are scarce. Marion Ellis of the University of Nebraska suggests dedicating stretches of roadside and wasteland to prairie wildflowers to encourage native bees to come back.

A few entomologists are hoping the honeybee hole is a window of opportunity for native pollinators. Most local bees are solitary; as a result, they may have a harder time finding food than the social honeybee. "They're not told, 'Fly two kilometers north by northeast, and there's the good stuff,'" points out Stephen L. Buchmann of the Arizona-Sonora Desert Museum in Tucson. No one has actually proved that competition with honeybees has caused a species to decline; yet many observers believe honeybees have a negative impact. In Australia, for example, beekeeping is banned near national parks for fear that the foreign *Apis* will outcompete local insects and birds.

"We've not had a diversified pollination portfolio," insists Buchmann, who co-directs the Forgotten Pollinators Campaign, which spreads the word about wild local bees. But only time will tell if native pollinators will rebound after 350 years of the extraordinarily busy honeybee. —*Madhusree Mukerjee*

ENVIRONMENT

WHEN NUTRIENTS TURN NOXIOUS

A little nitrogen is nice, but too much is toxic

If global warming seems ominous, consider this new assessment of how humans have disrupted the natural cycling of nitrogen. By using fertilizers, burning fossil fuels and cultivating crops that convert nitrogen into forms plants can use, humankind has over the past century doubled the total amount of atmospheric nitrogen that is converted, or fixed, every year on land. The nitrogen glut is already causing "serious" loss of soil nutrients, acidification of rivers and lakes, and rising atmospheric concentrations of the greenhouse gas nitrous oxide. Moreover, the oversupply probably explains decreases in the number of species in some habitats, as well as long-term declines in marine fish catches and, in part, the algal blooms that are an unwelcome spectacle in many coastal areas.

That alarming evaluation, to be formally published this summer in *Ecological Applications,* is the work of eight senior ecologists chaired by Peter M. Vitousek of Stanford University. Their study identifies as the chief culprit the industrial fixation of nitrogen gas to make fertilizer. "The immediacy and ra-

RICHARD PACKWOOD *Oxford Scientific Films/Earth Scenes*

DEMAND FOR FERTILIZER has led to excessive nitrogen fixation on land.

In Brief, *continued from page 24*

Oil's Lasting Effects

Cleaning oil-slicked seabirds may help people heal, but not birds, several recent studies say. Daniel W. Anderson of the University of California at Davis tracked 112 treated pelicans, as well as 19 uncontaminated birds. After two years, only 10 percent of the oiled birds could be found, compared with 55 percent of the unaffected animals. So, too, marsh coots exposed to oil often die prematurely, presumably because the oil has left them immunosuppressed. And independent ornithologist Brian E. Sharp has shown that common murres, western grebes and white-winged scoters fare worse. The researchers suggest that money used to rehabilitate birds might be better spent on finding ways to prevent oil spills in the first place.

BEN OSBORNE *Tony Stone Images*

Flunking Genetic Tests

Patients beware: a recent study in the *New England Journal of Medicine* warns that many doctors may not be ready for genetic testing. The authors, led by Francis Giardiello of Johns Hopkins University, found that among 177 people screened for a rare inherited colon cancer in 1995, 30 did not actually need the test, based on their family history. Only 33 of these people received any counseling to help them interpret the test results. And more troubling still, physicians misinterpreted the results in 56 of the cases.

Jumpin' Jupiter

On February 20, Galileo cruised in closer to Jupiter's moon Europa than ever before, and again the probe returned telling photographs of that icy orb. First, these pictures reveal what look like floating blocks of ice, not unlike the icebergs seen on Earth during springtime thaws. Their presence lends further support to the idea that Europa sports subterranean seas. In addition, the images capture several crater-free patches on the moon's surface, prompting some scientists to suggest that Europa is in fact much younger than originally thought. *—Kristin Leutwyler*

SA

pidity of the recent increase of nitrogen fixation is difficult to overstate," the researchers say. More than half the nitrogen fertilizer ever made before 1990 was used during the single decade of the 1980s, they note.

Industry now fixes 80 million metric tons of nitrogen every year to make fertilizer. Leguminous crops, which harbor nitrogen-fixing bacteria, and fossil fuels, which liberate nitrogen compounds when burned, together make another 60 million tons of nitrogen available to living things. The natural global rate of nitrogen fixation on land is between 90 and 140 million metric tons, and the excess stimulates plant growth. Moreover, by clearing forests and draining wetlands humans make the situation worse, because those activities liberate nitrogen that would otherwise be stored.

The Environmental Protection Agency, recognizing the damage caused by nitrogen oxides from combustion, has introduced regulations to limit by several million tons emissions from power stations and other industrial plants. And it is negotiating further limits on the already tightly controlled amounts emitted by vehicles. But there are no effective federal controls on the amount of fertilizer a farmer can use. "It is my feeling that this is an emerging issue," says Gary T. Gardner of the Worldwatch Institute. Gardner asserts that demand for industrially produced fertilizer could be reduced if farmers instead put on their fields recovered municipal food and yard waste, rich sources of nitrogen that together make up a third of the waste volume.

Employing fertilizers more efficiently might be "our best hope for doing something," Vitousek suggests. The Sierra Club Legal Defense Fund is pressuring the EPA to limit runoff into the Mississippi, which the litigation group contends is responsible for a 7,000-square-mile "dead zone" that appears every summer in the Gulf of Mexico. Reductions are possible: some states, including Arizona, have initiated successful incentive programs to lower fertilizer runoff. And some U.S. farmers have reduced fertilizer consumption voluntarily.

A spokesman for the Fertilizer Institute in Washington, D.C., a manufacturers organization, says industry is already developing ways of getting more growth from less fertilizer. But assessments such as Vitousek's report should, the institute argues, acknowledge the rapid escalation in the human population's demand for food. It points out that global nitrogen fertilizer use in 1995 was down 3 percent from the peak year of 1988—although it apparently is on the rise again. Those numbers may need closer scrutiny as the global population zooms to an estimated 10 billion during the next century.

—Tim Beardsley in Washington, D.C.

GENETICS

SEX, FLIES AND VIDEOTAPE

A mutant gene alters the sexual behavior of fruit flies

Science has taken an important step forward in the effort to expose the genetic underpinnings of sexual predilection—in fruit flies. A group led by Michael McKeown of the Salk Institute for Biological Studies in La Jolla, Calif., has found that a single mutant gene, called *dissatisfaction*, makes female flies too choosy and male flies not choosy enough.

Previous research had shown that females carrying *dissatisfaction* never lay eggs, but the precise causes of the infertility remained unknown. McKeown and his three colleagues put the mutant females in transparent chambers with normal males and videotaped their shenanigans. Normal females copulate after several minutes of male courtship, which includes poking, licking and vibrating a wing—or "singing," as the investigators describe it in the February 4 issue of *Proceedings of the National Academy of Sciences.*

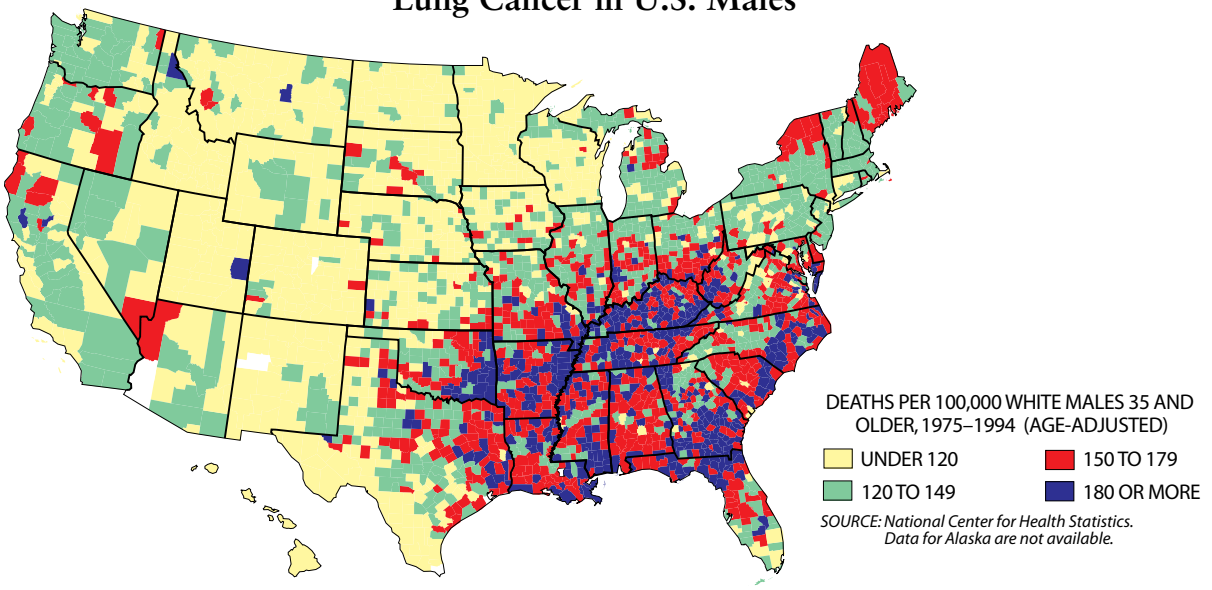
Continued on page 31

MICHAEL MCKEOWN

SEXUAL BEHAVIOR OF FRUIT FLIES is disrupted by* dissatisfaction *gene. The flies shown here are normal.

BY THE NUMBERS

Lung Cancer in U.S. Males



SOURCE: National Center for Health Statistics. Data for Alaska are not available.

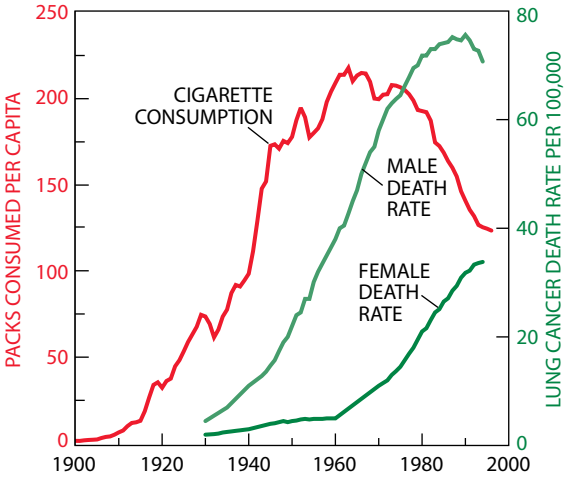
The U.S. is now in the seventh decade of a lung cancer epidemic that started with the introduction of milder, more inhalable cigarettes near the turn of the century. Because of the disease's long incubation period, lung cancer mortality did not rise until the 1930s, but as early as 1912, critics were claiming that cigarettes caused cancer. There was, however, no strong evidence until 1950, when published reports showed smoking to be far more common among those with the disease. Later research confirmed beyond a reasonable doubt that smoking not only caused most lung cancer—more than 80 percent—but also contributed to a variety of other diseases. The Centers for Disease Control and Prevention estimates that 420,000 Americans died of a smoking-related disease in 1990 and that, of these, 28 percent died of lung cancer, 24 percent of coronary heart disease, 19 percent of other forms of cardiovascular disease and 15 percent of obstructive lung diseases, such as emphysema. About half of those who start smoking regularly as teenagers can expect to die before their time from a smoking-related disease.

Geographically, lung cancer mortality among men follows roughly the prevalence of cigarette smoking but also reflects local habits and practices, such as those of the Cajun population of Louisiana, who are heavy users of hand-rolled cigarettes. Another factor influencing the pattern on the map is occupational exposure to carcinogens, for example, in the shipyards of the Gulf and South Atlantic coasts. Other factors include availability of high-quality medical care (low in many parts of the South), religion (Mormons, among others, proscribe smoking), air pollution, radiation and possibly even diet. Smoking tends to be higher among blue-collar workers and the less well educated. Between 1987 and 1990 about 40 percent of blue-collar men smoked, compared with 24 percent of white-collar men. A 1993 survey showed that among those with less than a high school education, 37 percent smoked, compared with 14 percent of college graduates and, from a different survey, 6 percent among the most highly educated—doctors, dentists and clergy, for instance.

(The geographical pattern among women differs markedly from that of men, being highest on the West Coast and Florida. Among blacks, lung cancer is highest in a broad band stretching from Pennsylvania through Nebraska and south to Kentucky and Tennessee.)

The decline among men in the lung cancer death rate since 1990 (*chart*) reflects the falling prevalence of smoking among males from about 60 percent in the 1950s or early 1960s to 25 percent in 1995. The lung cancer death rate among women now appears to be leveling off and will presumably decline, reflecting a later peaking of cigarette use. Smoking prevalence among women reached a high of about 35 percent in the mid-1960s, followed by a decline to 21 percent by 1995.

Compared with other developed countries, lung cancer death rates among American males are in the middle range. Rates for American women, however, are the highest among developed countries, which is not surprising considering that the American tobacco industry pioneered mass advertising of cigarettes to women as early as the 1920s. *—Rodger Doyle*



SOURCE: Mortality data are age-adjusted rates per 100,000 total population and are from the National Center for Health Statistics. Cigarette consumption data are for people over 18 and are from the Centers for Disease Control and Prevention.

RODGER DOYLE

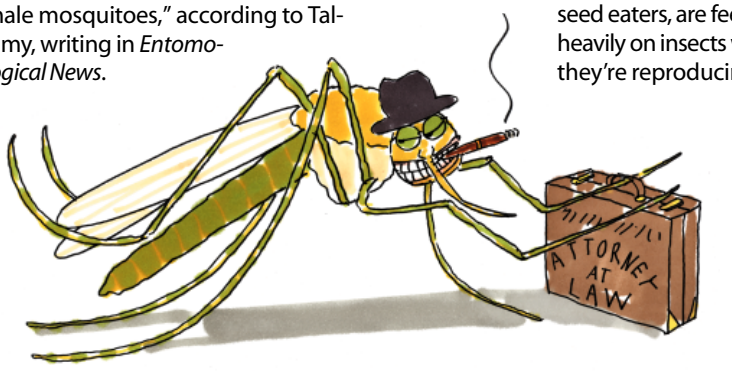
ANTI GRAVITY

Small Fry

Douglas Tallamy returned from an insect-collecting trip in April to find a message on his answering machine. Being a professor of entomology at the University of Delaware, Tallamy often goes bug hunting. Being a human being, he often gets phone messages. The messages, however, never before included threats from lawyers.

How the lawyers discovered Tallamy dates back to another trip, taken three years ago. "I had driven through the tollgates after the Delaware Memorial Bridge, and they had bug zappers," he said. "And I remember sitting there and watching the bugs get zapped."

Tallamy suspected that the crackling bits of biomass were most likely not mosquitoes because of a 1983 study by entomologist Roger Nasci, now at the Centers for Disease Control and Prevention. That report revealed that "the average zapper in South Bend, Ind., killed more than 3,000 insects per day, 96.7 percent of which were not female mosquitoes," according to Tallamy, writing in *Entomological News*.



MICHAEL CRAWFORD

(Only female mosquitoes bite; males spend their lives eating flowers, drinking nectar and generally being merry.) Those low death rates were not all that surprising, given that female mosquitoes are attracted to carbon dioxide—the better to find exhalers lousy with blood—and not to the ultraviolet light these devices use to attract insects. "I said," he recalls, " 'I wonder what really is being killed.' "

The next day Tallamy was approached by high school student Tim Frick, who was looking for research experience. The fateful words, "I wonder what really is being killed," hung in the air like Obi-Wan's reminder to use the Force. Before he knew what he had gotten into, he and Tallamy were perusing thousands of dead insects collected from six insect-electrocution devices hanging from people's houses in suburban Newark, Del.

All the homes were near water, one only 65 meters from a stream with plenty of stagnant pockets. Mosquitoes should have been dropping like flies in a Little League game. The final tally, however, was stunning, even, brace for impact, shocking: of 13,789 dead, female mosquitoes accounted for 18. Of course, other biting flies were killed, too. Thirteen others. The grand total: 31.

Assuming his numbers are representative, Tallamy figures that four million insect-electrocution devices running for 40 summer nights could be blasting to chitin bits some 71 billion insects, most of which wouldn't hurt a fly. Now, even that vast number may be only a drop in the bucket and may not upset delicate food chains. "But 71 billion insects is a lot of insects," Tallamy says, "and we do know all the things that feed on them. There are an awful lot of bird-watchers, and they love the birds. And after they watch them, they go home and put up their bug zappers. These birds, even the seed eaters, are feeding heavily on insects when they're reproducing."

Tallamy notes that without a survey of the mosquito population near his study sites, he can't be certain that 18 dead females isn't all of them. But he doubts it. "It is highly unlikely," he wrote, "that our lowland, wooded sites which were rich in aquatic breeding habitats, produced so few adult mosquitoes in the course of nine weeks that 18 electrocuted females would represent adequate control of these flies."

These numbers drove Tallamy to conclude that "electric insect traps are worthless for biting fly reduction." And apparently, the word "worthless" served as the flame that attracted the lawyer. "I may be sued," says Tallamy, who also studies social parasitism among insects—good preparation for any legal actions.

—*Steve Mirsky*

Continued from page 26

But the videotapes revealed that *dissatisfied* females rarely assume the proper position for copulation. Often the females bolt from their suitors or kick those who don't get the message. When the reluctant females are inseminated, they still remain infertile, because they lack the nerves that ordinarily signal the uterus to expel the eggs.

Male bearers of *dissatisfaction*, on the other hand, court all their fellow flies, male and female, indiscriminately. Desire is not matched by performance, however: neural defects hamper the efforts of *dissatisfied* males to curl their abdomens into the proper mating posture. The males "attempt copulation," notes Barbara J. Taylor of Oregon State University, a member of McKeown's team. "They're just not very good at it."

A gene called *fruitless* has been shown to affect males in a similar manner, McKeown notes. *Fruitless* males court both males and females indiscriminately but cannot mate (hence the gene's name) because of malformations of their abdomen. *Fruitless* has no apparent effect on females.

The investigators hope to determine whether variants of *dissatisfaction* occur in other species. So-called homologues of other fruit fly genes, including some that control the development of eyes, have been found in various organisms, humans among them. A *dissatisfaction* homologue would not necessarily be directly linked to sexual behavior in other species, Taylor notes; it could have a more generalized role, such as regulating the development of neural synapses.

The investigators do not shy away from the implications of their research. It raises the possibility, Taylor says, that the sexual behavior of more complex species—including humans—may be regulated not by hundreds of genes (each of which has a minute effect) but by relatively few genes.

Four years ago a team led by Dean Hamer of the National Cancer Institute claimed to have found a gene associated with male homosexuality. That result has not been independently corroborated. Given that at least two genes—*dissatisfaction* and *fruitless*—can affect the behavior of fruit flies, McKeown adds, it may be "naive" to expect to find a single "master sex gene" controlling the behavior of *Homo sapiens*. Let's hope that if such a gene is found, it merits a name more promising than *dissatisfaction*. —*John Horgan*

SCIENTIFIC AMERICAN CORRESPONDENCE

Reprints are available; to order, write Reprint Department, Scientific American, 415 Madison Avenue, New York, NY 10017-1111, or fax inquiries to (212) 355-0408. E-mail: info@sciam.com

Back issues: $8.95 each ($9.95 outside U.S.) prepaid. Most numbers available. Credit card (Mastercard/Visa) orders for two or more issues accepted. To order, fax (212) 355-0408.

Index of articles since 1948 available in electronic format. Write SciDex®, Scientific American, 415 Madison Avenue, New York, NY 10017-1111, fax (212) 980-8175 or call (800) 777-0444.

***Scientific American*-branded** products available. For free catalogue, write Scientific American Selections, P.O. Box 11314, Des Moines, IA 50340-1314, or call (800) 777-0444. E-mail: info@sciam.com

Editorial correspondence should be addressed to The Editors, Scientific American, 415 Madison Avenue, New York, NY 10017-1111. Manuscripts are submitted at the authors' risk and will not be returned unless accompanied by postage. E-mail: editors@sciam.com

Advertising correspondence should be addressed to Advertising Manager, Scientific American, 415 Madison Avenue, New York, NY 10017-1111, or fax (212) 754-1138. E-mail: advertising@sciam.com

Subscription correspondence should be addressed to Subscription Manager, Scientific American, P.O. Box 3187, Harlan, IA 51537. The date of the last issue of your subscription appears on each month's mailing label. For change of address notify us at least four weeks in advance. Please send your old address (mailing label, if possible) and your new address. We occasionally make subscribers' names available to reputable companies whose products and services we think will interest you. If you would like your name excluded from these mailings, send your request with your mailing label to us at the Harlan, IA address. E-mail: customerservice@sciam.com

Visit our Web site at http://www.sciam.com/

PROFILE: RAYMOND V. DAMADIAN

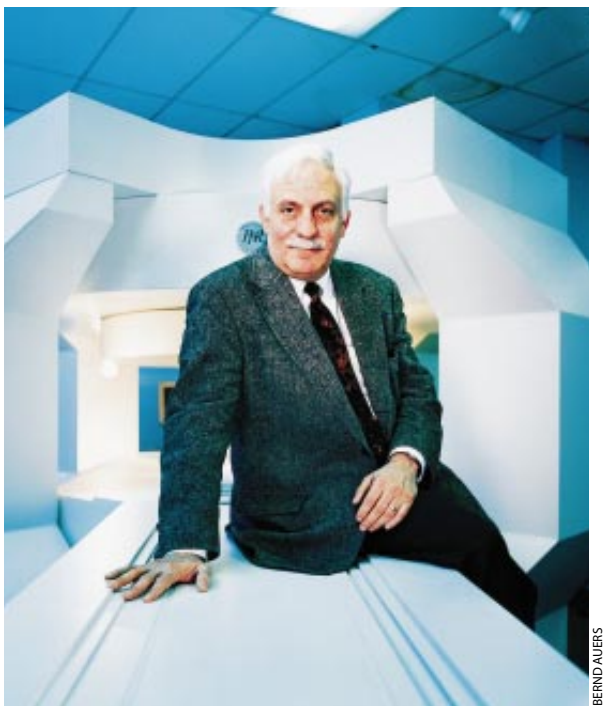
Scanning the Horizon

On a chartered bus somewhere outside Washington, D.C., Raymond V. Damadian lifts a megaphone to his mouth and addresses his fellow passengers, as if acting as a tour guide. Instead of describing the historic attractions in the city they are about to visit, he reviews why they have been traveling the interstate since the wee hours of the morning and what they might say when they arrive at the Capitol and meet with their elected representatives. Most of his audience probably need little reminder, but this scientist, inventor and entrepreneur wants there to be no doubt about the seriousness of their mission. To his mind, they are there to avert a national disaster.

The catastrophe he foresees is the demise of effective patent protection for the country's inventors. And Damadian is certainly one to speak for that group. Twenty years ago, in a basement laboratory at the Downstate Medical Center in Brooklyn (part of the State University of New York), Damadian designed and built a machine he had conceived—and patented—some six years earlier: a medical scanner that could probe the body using the phenomenon of nuclear magnetic resonance. This first prototype for magnetic resonance imaging, which he dubbed "Indomitable," is now held by the Smithsonian Institution, along with Edison's lightbulb and the Wright flyer.

Nuclear magnetic resonance (NMR) is the phenomenon by which atomic nuclei placed in a moderately large magnetic field will absorb and emit radio waves at certain well-defined frequencies. Its discovery was first reported in 1938 by the physicist Isidor I. Rabi and his colleagues at Columbia University. Since the close of World War II, physicists and chemists have routinely used nuclear magnetic resonance in their laboratories to probe the nature of various substances. But before Damadian's bold innovation, none of these scientists had considered scanning the human body using this method. Magnetic resonance imaging (MRI), of course, has since become an indispensable medical tool.

"He's like the missing link," quips David Terry, Damadian's brother-in-law and secretary-treasurer of FONAR Corporation, the company Damadian founded in 1978 to commercialize his invention. Terry rightly points out that Damadian bridged the gap between the many research scientists familiar with nuclear magnetic resonance and the many doctors desperate for better ways to detect cancerous tumors in the body.



BERND AUERS

MRI SCANNER
surrounds its inventor, Raymond V. Damadian.

The key was Damadian's background. After winning a scholarship from the Ford Foundation, Damadian entered the University of Wisconsin as a 16-year-old freshman in 1952. His major area of study was mathematics, but he then chose to go to medical school. "The one thing I found appealing about medicine was that it didn't seal your options," Damadian notes. He earned a medical degree at the Albert Einstein College of Medicine in Bronx, N.Y., and completed his internship and residency at S.U.N.Y.'s Downstate Medical Center. After a couple of postgraduate stints, Damadian assumed a professorship at Downstate in 1967.

Knowing his subsequent accomplishments, one suspects that Damadian presents false modesty when he reports his initial reticence to follow a career in laboratory research. "I lacked confidence. I was always one of those guys who dropped the crucible," he proclaims. He nonetheless found encouragement at Downstate, where he engaged in studies of the balance of electrolytes in the body. And it was his investigation of sodium and potassium in living cells that led him in 1969 to experiment with nuclear magnetic resonance using borrowed time on the latest equipment.

Damadian swiftly began to appreciate what NMR physicists had known for some time. The dominant NMR signal from cells comes from the hydrogen atoms in water they contain. What is more, the signal varies with the configuration of that liquid—for example, whether the water molecules are bound tightly to various cell structures or more loosely held. Damadian then asked himself a crucial question: How might the NMR signal change between healthy cells and cancerous ones? The answer, he was soon to discover, was that the differences were dramatic.

After testing normal mouse tissues against tumors extracted from the animals, Damadian determined that NMR signals persisted for much longer in cancerous cells than in healthy ones. He published these results in 1971 in a paper entitled "Tumor Detection by Nuclear Magnetic Resonance." This scholarly report only hinted at what he would outline much more fully in the patent application for his pioneering invention, which he filed the following year. There he described how with magnetic fields and radio waves doctors could scan the human body for cancerous tumors.

Damadian had completed his initial experiments on mice without any sort of research grant at all, so his first task was to look for funds to build a human-size scanner. But the idea of probing the body in this way for cancer was almost unimaginable in 1971, and he was laughed at by many of his academic col-

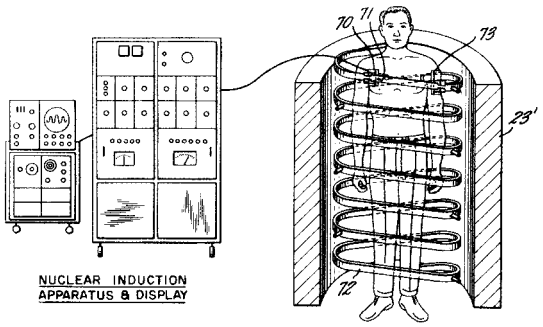
leagues. His grand claims for the potential of the technique did not help, either. In fact, his brash assertions simply alienated many of his conservative peers, and as a result, he failed to find the needed money through the usual channels. After one particularly disappointing refusal from the National Institutes of Health, Damadian resorted to a more direct approach for garnering government aid: he wrote to then president Richard M. Nixon, who had just declared a multibillion-dollar war on cancer, explaining the value of nuclear magnetic resonance and boldly asking him to intercede.

"I was young and not understanding of the way things worked," Damadian admits. Curiously, his letter to the White House did not disappear into bureaucratic oblivion. Damadian soon received a telephone call from an administrator at the NIH ostensibly reprimanding him for writing directly to the president. "The thing that amazed me was that he did something," says Damadian, who subsequently received a modest grant.

Damadian's early political activism on behalf of his research project did not end there. In December 1976, with funds evaporating as rapidly as the precious liquid helium he used to cool the cryogenic magnet he and his graduate students constructed for their prototype scanner, Damadian decided to visit Plains, Ga., home of the president-elect, Jimmy Carter. There, in rural Georgia, this charismatic New York researcher quickly became friendly with Jimmy's cousin Hugh, who made a living raising earthworms on a nearby farm. And when Hugh's son later joined the Carter administration, Damadian made an appeal for research funds using this rather unconventional point of contact (an effort that brought him no great gain).

The sense of urgency with which Damadian sought funds was heightened by the knowledge that a few who were swayed by his success distinguishing tumors were beginning to compete with him in building NMR imaging equipment, hoping themselves to perform the first human scan. So Damadian pushed himself and his students relentlessly and found private backers to keep the research going on a shoestring budget.

And in the summer of 1977 Damadian finally stepped into Indomitable, his ungainly metal creation, to make the first magnetic resonance scan of the human chest. The attempt failed, in essence because of Damadian's heft. His girth was too ample to fit within the largest radio pickup coil he and his motley crew could get to work. But after it was clear that sitting for hours in the intense magnetic field of the machine produced absolutely no ill effects, one of Damadian's graduate students, Larry Minkoff, volunteered his younger and slimmer frame. Minkoff thus became the first person to have his torso revealed by a magnetic resonance scan.



PATENT DRAWING from Damadian's 1972 filing illustrates human scanning.

The initial picture obtained using Indomitable was quite crude. But Minkoff's heart, lungs and chest wall could all be clearly discerned. And that success brought Damadian a certain amount of popular notoriety. Television news crews visited his Brooklyn laboratory to report about his work. His gargantuan magnetic apparatus appeared prominently in *Popular Science* magazine. But the publicity proved a mixed blessing. Some of the coverage, particularly from the influential *New York Times*, cast doubt on Damadian's claims that the newly demonstrated technology would eventually be able to find hidden tumors. Many scoffed at the thought, and when Damadian went on to try to commercialize the invention, venture capitalists were nowhere to be found.

Other scientists with a patented discovery of this magnitude would have probably chosen to license the technology to an established manufacturer of medical equipment. Damadian flirted with that option, but he ultimately decided that to bring magnetic resonance scanning to the world, he needed to do more—and he needed to do it himself. So Damadian and a small group of committed friends, students and family members began a grassroots campaign to start a new industry.

Damadian named his fledgling company FONAR, using the first and second letters of the words he used to describe the seemingly magical new technique: Field fOcused Nuclear mAgnetic Resonance. Although this phrase remains obscure, the use of nuclear magnetic resonance in medicine is anything but forgotten. With countless improvements and embellishments from researchers around the globe, magnetic resonance imaging soon evolved to a point that physicians could see the interior of the body in minute detail and were able to diagnose everything from brain tumors to slipped disks. In 1988 Damadian received the National Medal of Technology for his innovation. And thousands of MRI scanners can now be found at hospitals and clinics in the U.S. alone, most produced by such industrial giants as General Electric, Toshiba and Siemens.

Indeed, the manufacture of magnetic resonance imaging machines by other companies and the years of legal wrangling required to defend his patent convinced Damadian that the lone inventor rarely fares well when forced to confront huge corporations. That his company has only recently been awarded some $100 million in damages from General Electric confirms for him the hurdles inventors face. He becomes particularly animated in discussing the current proposal in the U.S. Congress to privatize the patent office—a move he believes will let big businesses exert undue influence and profit at the cost of smothering technological innovation. "The other charming feature of this bill, which I'm sure will delight you, is the gift clause," he explains as he reads a provision that would appear to sanction monetary gifts to a newly constituted government patent corporation. "It's astonishing."

It should probably be no surprise then that he is ready to lobby Congress as fervently as he has confronted his scientific critics and his business competitors. What is startling is that he pursues each of these activities with such intense conviction and energy. Twenty years later he seems able to muster the same enormous drive that allowed him to prove that NMR scanning of the body would, after all, work. One wonders whether the most indomitable thing to emerge from that dingy laboratory in Brooklyn was a novel machine or Damadian himself. —*David Schneider*

TECHNOLOGY TRANSFER

SELLER BEWARE

German high-tech sales to Iran provoke concerns in the U.S.

Several years ago, in the wake of the war in the Persian Gulf, a "new world order" was proclaimed. Nations would work together to isolate and contain rogue countries that flouted international treaties or standards of decency. It was a good if obvious idea. But it was difficult to reconcile with the fact that numerous Western companies—with the tacit approval of their governments—had supplied the high-tech equipment and materials that enabled various rogue countries, such as Iraq, to embark on programs to produce weapons of mass destruction.

Six years after the war ended in the Gulf, some observers claim that the same pattern of technology acquisition that enabled Iraq to sustain nuclear- and chemical-weapons programs is occurring in Iran. Moreover, German high-tech companies—whose products turned up in abundance in the Iraqi nuclear- and chemical-weapons programs—are once again at the center of controversy. "The U.S. has been widely concerned over the past five years with what our Western allies, particularly the Germans, have been doing with the Iranians," says David A. Kay, a national security expert in the McLean, Va., office of Science Applications International Corporation.

Among those fretting about the German-Iranian links, apparently, is the U.S. Central Intelligence Agency. This past March a man using the name of Peyton K. Humphries, who was an official in the U.S. embassy in Bonn, was expelled by the German government. *Der Spiegel,* the German news magazine, identified the man as a CIA employee. According to *Der Spiegel*'s article, Humphries had tried to recruit an employee of the German Ministry of Economics to provide information on sales to Iran of German high-tech goods and services.

A subsequent article in the U.S. newsletter *Nucleonics Week,* citing unnamed U.S. and German government sources, indicated that Humphries's particular interest was in so-called dual-use technologies. Such technologies have both military and nonmilitary uses. Dual-use items fall into a vast category, including everything from supercomputers to certain high-strength materials. (The U.S. State Department, the U.S. embassy in Bonn, the CIA and the German embassy in Washington all declined to comment on the Humphries case.) Although Iranian officials have steadfastly denied that they have a military nuclear program, virtually all Western analysts believe the country is trying to build a nuclear weapon [see "Iran's Nuclear Puzzle," by David A. Schwarzbach, on page 62].

Germany has been Iran's largest trading partner in recent years. According to the German economics ministry, German companies sold $736-million worth of electrotechnical, chemical and optical products, machinery and precision tools to Iran in the 11 months ending November 1996. The proportion of these goods that could be considered dual use was not clear.

Controls on German exports were strengthened considerably after the war in the Persian Gulf, when it was discovered that many Western companies, including dozens of German ones, had helped Iraq build poison-gas factories and had supplied critical equipment for the country's atomic bomb project. At least two German firms, Leybold AG and Karl Schenck, have supplied both the Iraqi and Iranian nuclear programs.

At present, German companies cannot export dual-use items without a license and must inform the German government if they plan to export any item to an arms manufacturer in Iran, Iraq, Libya, Syria or several other countries. In 1994 the German economics ministry and various industrial organizations lobbied unsuccessfully for the removal of Iran and Syria from the list.

Even under the tightened restrictions, millions of dollars' worth of control computers and tunnel-digging machinery from German companies wound up in recent years near Tarhuna, Libya, where construction is ongoing, if intermittent, on a large underground factory. The plant is expected to produce chemical weapons, such as mustard gas and nerve agents. The German equipment got into Libya illicitly through phony companies in Belgium and Thailand. (In the 1980s some 30 German companies were involved in the construction of a chemical-weapons plant at Rabta, Libya.)

Phony companies located in the exporting country itself can also be useful for circumventing export controls. "What I have seen is that the Iranians are following the examples of Pakistan and Iraq," says Andrew Koch, an analyst at the Center for Nonproliferation Studies at the Monterey Institute of International Studies in California. "They set up a network of front companies, through which they import dual-use technologies." Evidence of this strategy, according to Koch, is Iran's ongoing effort to buy Sket Magdeburg GmbH, a machine-tool manufacturer in the former East Germany. The proposed deal

UNITED NATIONS/SYGMA

CHEMICAL WEAPONS COMPONENTS, such as these empty bomb casings, were destroyed in Iraq in 1991. Some analysts fear that Iran is following Iraq's example.

calls to mind Matrix Churchill Ltd., a British-based machine-tool company that Iraq purchased in 1987 and subsequently used as a front for exports to the country's weapons plants.

An official in the German embassy in Washington responded that "if an Iranian company did acquire Sket, German export laws and the entire control system would still apply. The exports would still have to be approved, and there is a pretty tight system for that."

The sentiment offers little comfort to Koch. "The point is, the export-control system is based on the exporter taking some of the responsibility for determining where its exports are actually going," he notes. "If Iran is determining that for itself, where is the check?"

—Glenn Zorpette

CIVIL ENGINEERING

FLOATING GIANTS

Sea-based platforms eyed for launch sites and airstrips

Offshore oil rigs are feats of modern engineering, able to weather monster waves and hurricane-force winds while producing the lifeblood of modern society. In coming years, the technology that mines black gold from under the sea may be deployed for other uses, from launching rockets to landing airplanes.

Beginning in 1998, a converted offshore oil-drilling platform is slated to become the launching site for rockets that will take satellites into orbit from a location more than 1,000 miles southeast of Hawaii. Sea Launch, consisting of four companies led by Boeing Commercial Space, plans to take advantage of the additional rotational speed at the equator to give rockets more momentum for lifting satellites into a fixed geostationary orbit. Launching from the equator also means that a satellite is already aligned with its orbital path and does not have to be repositioned from another latitude.

Kvaerner, Europe's largest shipbuilder, and one of the Sea Launch partners, is refurbishing an oil-drilling platform that had been damaged by an explosion in the North Sea, adding a launchpad, a hangar for storage of a Russian-Ukrainian rocket, and facilities for rocket fuels. The platform, measuring 430 feet by 220 feet and weighing 31,000 tons, will rest on a series of columns attached to two submerged pontoons. Construction costs for the entire project, which also includes a specially outfitted command ship, are expected to reach $500 million. The first mission is planned for June 1998.

The wherewithal of mobile, semisubmersible oil platforms has not gone unnoticed by the U.S. Department of Defense. An ongoing series of studies using scale models is trying to determine whether a set of interlocked platforms could be used as an offshore military base. A multibillion-dollar sea base would eliminate the difficulty of finding near a battle theater a friendly country from which to resupply troops. Self-propelled platforms, each at least as large as a standard oil rig, could move close to a conflict area and then link together. The resulting several-thousand-foot runway could accommodate C-130 transport aircraft. Underneath would remain millions of square feet of storage space.

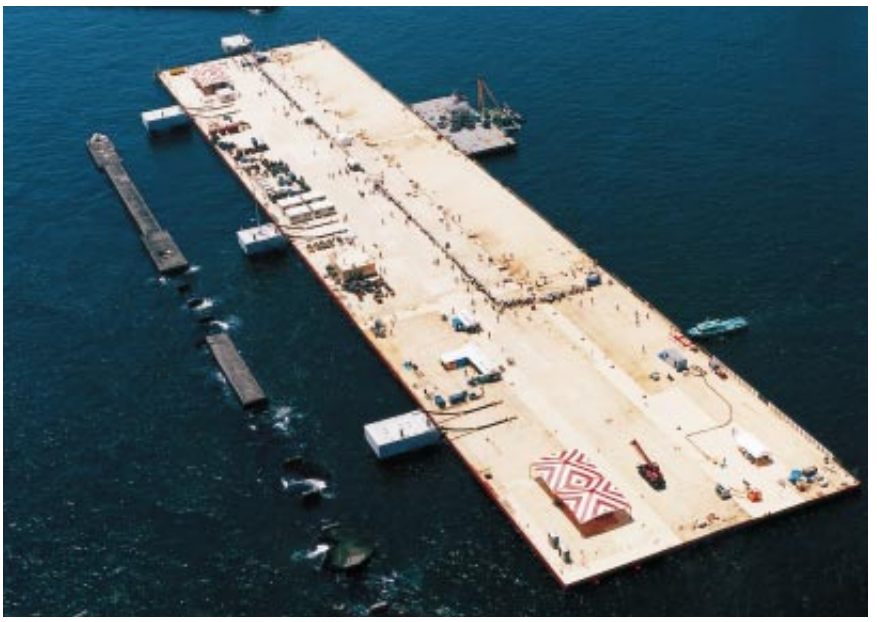
Still unanswered is whether gargantuan platforms could be coupled in a tossing sea using giant male-female connectors, hinges or bridgelike structures. "The forces that get generated between the modules are huge," says Albert J. Tucker, a division director at the Office of Naval Research, which is overseeing a $16-million research program.

Not everyone has given the concept an ardent welcome. Factions within the defense community believe a mobile base could undermine the case for continued spending on aircraft carriers.

Japan, given its limited land area, has shown interest in floating platforms, although the structures are not derived from offshore oil platforms. A group of Japanese shipbuilders and steel companies, the Technological Research Association of Mega-Float, recently built a nearly 1,000-foot-long experimental floating platform in Tokyo Bay. The so-called mega-float technology, which might be used for floating airports, pier facilities or power plants, consists of a series of hollow steel blocks welded together to create a pontoonlike structure that is moored to the sea bottom with pilings. In September the U.S. and Japan agreed to investigate the technology as one option for restationing the U.S. Marine helicopter wing now at the Futenma Air Station on Okinawa.

Despite the flurry of proposals, engineers caution against waterworld fantasies, which foresee cities occupying the high seas. "You really have to have a very compelling reason to be out in the ocean—this is a very expensive technology," says William C. Webster, associate dean of the College of Engineering at the University of California at Berkeley. Still, the experience in building floating "islands" may make deep ocean waters accessible for a few clearly defined endeavors.

—Gary Stix



TECHNOLOGICAL RESEARCH ASSOCIATION OF MEGA-FLOAT

"MEGA-FLOAT" EXPERIMENTAL PLATFORM in Tokyo Bay is nearly 1,000 feet long, 200 feet wide and seven feet high. The technology might eventually be used for airport runways.

APPLIED MATHEMATICS

MEDICAL MISMATCH

When hospitals act like men

On the ides of March, 13,554 graduating medical students in the U.S. each opened an envelope. The dreaded missives named the hospitals where they would go for their residencies. "It's serious business," says Kevin J. Williams of Jefferson Medical College. If he has his way, the 1997 rite of passage will be the last to have the odds stacked against the students.

Williams himself went through the National Resident Matching Program (NRMP) in 1980. The program, subscribed to by most medical colleges, requires students to first apply to the hospitals. After interviews, the students rank the hospitals in order of preference, while the hospitals similarly rank the students. The NRMP then matches the participants via an algorithm, with the final results being binding. On examining the formula, Williams, Victoria P. Werth (his wife), and classmate Jon A. Wolff discovered that contrary to the NRMP's claims, it was biased against the students.

The very first students to go through the NRMP program, in 1952, appear to have protested this bias as well. It was not until a decade later, however, that mathematicians analyzed the general problem. David Gale and Lloyd S. Shapley, then at Brown University and the Rand Corporation, respectively, considered a (heterosexual) society in which the men and women rank to whom they want to be married and are then paired off. The matches must be stable: a man and woman should not prefer each other over their assigned partners.

For the simplest case—two men and two women—three kinds of conflicts can arise. Both men want the same woman, in which case she chooses; both women want the same man, and so he chooses; or Alice wants Bob, Bob wants Mary, Mary wants Dan and Dan wants Alice. There are two ways to slash through this last tangle: either the men get to choose, or the women do. Because men propose (in most cultures), women tend to lose. So Bob proposes to Mary, Dan proposes to Alice, and because under the rules of the game getting married is better than staying single, both women accept, neither getting the man she preferred.

Williams argues that the NRMP matching algorithm sets up a virtual universe in which the hospitals act as men,

Spying Saucer

Among the hundreds of experimental machines built to go where humans cannot (or should not), there have been rollers, crawlers, fliers, orbiters and undersea cruisers. Now there is a flying saucer, and it is boldly going where no flying drone has gone before. It is meandering down urban streets, peeping in windows and setting down gently on the roofs of buildings.

Appropriately enough, demonstrations of the saucer's capabilities are coinciding with the 50th anniversary of the notorious incident in Roswell, N.M. In that event, which occurred during the evening of July 2, 1947, a downed balloonlike device, part of a secret U.S. Air Force project, caused an enduring sensation when it was mistaken for a flying saucer of extraterrestrial origin. Ironically, the real flying saucer, which is called Cypher, has not yet provoked any similar episodes, partly because timely articles in the local press at some of the places where the saucer has been flown have explained its earthly origins and missions. (This article is not part of an insidious cover-up conspiracy. Honest!)

Though not otherworldly, Cypher is at least revolutionary. Built by a small team at Sikorsky Aircraft in Stratford, Conn., the two-meter-diameter flier is a rotary-wing aircraft, similar in some respects to a helicopter. Unlike a helicopter, however, the aircraft is propelled by two rigid rotors, one above the other, which spin in opposite directions. Cypher is not the first experimental vehicle to exploit this propulsion scheme, which eliminates the need for a tail rotor. But it is the first pilotless craft configured in this manner that shrouds the rotors with its fuselage.

SIKORSKY AIRCRAFT

This shrouding allows the saucer to bump into tree branches, buildings or other objects without causing a catastrophe. The 110-kilogram aircraft can stay in the air for about two and half hours, covering a range of 30 kilometers. Its diminutive rotary engine—the size and weight of a lawnmower engine—puts out an astounding 52 horsepower.

SIKORSKY AIRCRAFT

FLYING OBJECT ***can carry a camera*** **(lower left photo)** ***for peering in windows.***

Advanced software grants the flier an unusual degree of autonomy. In tests last autumn, the saucer used software from Lockheed Martin and Northrup Grumman to find and trail a solitary soldier walking in a field. During the 25-minute flight, operators sent only two orders to the craft, according to James Cycon, who leads the project at Sikorsky. One command instructed the machine to take off; the other, issued after it had found the soldier and had followed him for a short while, told it to return.

Another notable test was carried out this past January at Fort Benning in Georgia. The army is looking for ways of making sure that troops are not ambushed in urban settings by snipers. At Fort Benning, where a mock town has been used to test anti-sniper concepts, the saucer cruised up and down streets only six meters wide, searching for hostile sharpshooters, and landed on the roof of one of the buildings. It looked inside some buildings by aiming a video camera through their windows. "The beauty of Cypher," Cycon says, "is that it can fly low and slow."

Cycon and company are now experimenting with new rotors and, in general, ascertaining the capabilities of their strange little saucer. "We're trying to show people what the aircraft can do," Cycon explains. "At the same time, we're trying to understand what it can do."

—Glenn Zorpette

proposing to the students. It is thus inherently biased against the students. In practice, the bias is small: 0.1 percent, affecting a few hundred students over the years. That may be because students tend to agree on which the best programs are, and vice versa, leading to the first two types of conflict rather than the third. Or it may be because the number of hospitals is enormous compared with the 15 or so that students actually rank.

Nevertheless, the American Medical Students Association (AMSA) has objected strongly to the bias, which impacts students' lives and careers. So did the hospitals' program directors at an NRMP meeting in November 1996. The NRMP board planned to meet in May to vote on the question. "They'll have very little choice but to change," predicts Andrew J. Nowalk of the AMSA. Although too late for this year's graduates, the sex switch may sweeten Match Day 1998, for some.—*Madhusree Mukerjee*

MEDICINE

ATTACKING ARTHRITIS

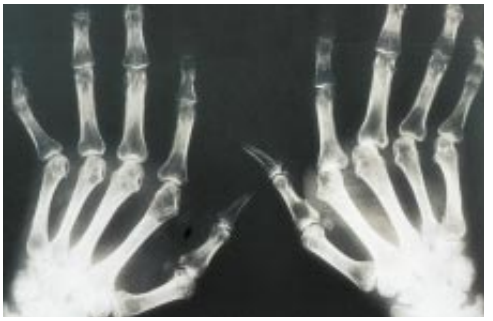
New treatments seek to rebalance the immune system

For some people, aches and pains in the joints flare up with bad weather. But for the more than two million Americans suffering from rheumatoid arthritis, stiff and swollen joints are the result of an internal storm in the immune system. Chemicals that the body normally releases to fight off infections flood the tissues in the joints, attacking them as though they were foreign invaders, eventually eroding the cartilage and bone. Over the past several decades, doctors have had few options for treatment. As knowledge of the immune system has expanded, however, researchers have developed various new drugs that aim to knock the body's defense system back in line.

Last fall, at a meeting of the American College of Rheumatology (ACR), several groups presented results on three novel therapies, all of which work by interfering with the deluge of chemicals released by the immune system in the course of rheumatoid arthritis. (The illness is distinct from the more common osteoarthritis, which stems from a lifetime of wear and tear on the joints.) Researchers at Amgen described their initial trials of a drug that inhibits the activity of interleukin-1, the naturally occurring protein that induces inflammation by activating the cells lining the blood vessels.

Workers at IDEC Pharmaceuticals, in collaboration with scientists at SmithKline Beecham Pharmaceuticals, have continued testing the drug they described at the ACR meeting. The compound is a monoclonal antibody that works by binding to the surface of immune system cells known as *T* cells. These cells direct the functioning of other parts of the immune system; overactive *T* cells, however, can provoke the body's natural defenses to destroy healthy tissue. When these monoclonal antibodies attach to *T* cells, they slow the immune response and seem to protect against joint damage. Phase III trials, which, as efficacy tests on humans, are the last and most crucial aspect of drug development, should begin later this year.

A third novel class of rheumatoid arthritis drugs targets the molecule called tumor necrosis factor, or TNF. This hormonelike substance, known as a cytokine, appears early in the chain reaction leading to joint destruction and has a wide range of functions—in particular, promoting the release of other inflammatory cytokines and enzymes that damage cartilage and bone. After successful early trials, Immunex Corporation recently started Phase III trials of its drug Enbrel, which soaks up TNF in the blood, thereby preventing it from causing further damage.



COURTESY OF THE AMERICAN COLLEGE OF RHEUMATOLOGY

***RAVAGES OF RHEUMATOID ARTHRITIS** may be eased by new drugs.*

In February, Immunex announced the discovery of an enzyme known as TACE, which acts even earlier in the inflammation cascade by stimulating the initial release of TNF. Michael Widmer, vice president of biological sciences at Immunex, indicates that the company, along with Wyeth–Ayerst Research, is now investigating how to block the release of TACE with a compound that could potentially be administered in pill form, rather than by injections, as required for Enbrel and other therapies.

In a slightly different approach, Christopher Evans and his colleagues at the University of Pittsburgh have experimented with injections of therapeutic genes into joints affected by rheumatoid arthritis. The group inserts genes that trigger the production of a protein that in turn reduces the activity of interleukin-1, the familiar substance involved in inflammation and joint destruction. Gene therapy for treating rheumatoid arthritis must be tested further; so far only two patients have been treated, but Evans does regard the results as encouraging. "Patients accepted [the procedure] well—there were no safety or tolerability issues. And there is evidence that the gene transfer worked."

Other advances in the diagnosis and treatment of rheumatoid arthritis await further testing, including genetic screening, stem cell or bone marrow transplants, and vaccinations (some researchers have speculated that rheumatoid arthritis results from a viral or bacterial infection). Edward Keystone, professor of medicine at the University of Toronto, comments that "rheumatologists had six drugs to test in the past 50 years. We have 12 to 14 agents in testing right now, all of which have been developed in the past five or so years." William Koopman, president of the ACR, explains that researchers now have a better grasp of the biochemistry of rheumatoid arthritis, providing more options for treatment. "We now have more opportunities to target the molecules involved in pathogenesis," he says.

With this better understanding of how the immune system behaves in rheumatoid arthritis should come additional weapons against other diseases characterized by faulty immune systems, such as inflammatory bowel disease, multiple sclerosis, scleroderma and systemic lupus erythematosus. Keystone concludes on a confident note: "Remember that there are 50 or so autoimmune diseases that affect 20 million Americans. We're applying our new knowledge to these other diseases as well." —*Sasha Nemecek*

CYBER VIEW

Disliking the Internet

We are seeing a growing disenchantment with the Internet, nowhere more strikingly than in media coverage of the cult Heaven's Gate and the mass suicide of 39 of its members. Countless accounts blamed the World Wide Web for the tragedy. Cable television's CNN led the attack, presenting a view of the Internet teeming with mad Web page workers (several members of the cult designed cut-rate Web sites) and lonely, vulnerable surfers (one recruit may have seen the group's Web page before joining). Newspapers, magazines and radio stations followed suit, from the *New York Times* editorial page to *Newsweek*—its cover story ran with the headline "Web of Death" and opened, "They were watching the skies—and the Internet—for a sign."

Actually, the sign most likely came not from the Internet but from the radio. The cult probably learned about the object they believed was a "mothership" waiting to take them away by means of a November radio broadcast. Amateur astronomer Chuck Shramek of Houston, who had recorded an image of Comet Hale-Bopp that showed a "self-luminous Saturn-like object" nearby, talked about the object on *Coast to Coast AM with Art Bell,* a late-night radio talk show popular with UFO believers. (Shramek bills himself as a "noted expert in 41 fields not currently recognized as science by Harvard, M.I.T. or Yale.") The show evoked a massive call-in, and the host repeatedly told worried listeners that the object, which became known as an SLO, was very real.

"The story took flight on the radio. The Web chased behind," comments Paul Saffo of the Institute for the Future in Menlo Park, Calif. In this case, the Web debunked the mythmakers barely 24 hours later: on November 15, Russell Sipe, who maintains a leading home page on Comet Hale-Bopp, posted information showing that the SLO was not a mysterious entity but in fact an eighth-magnitude star, SAO 141894.

Sipe, who patiently fields about 200 Hale-Bopp-related e-mail messages a day, is appalled by the reporting of the cult's ties to the Internet. "Heaven's Gate did a number of things to raise money," he says. "They bought and restored old cars, but no one is questioning the role of restoring '57 Chevys in creating a cult." He also notes that the group's home page—a dense, unattractive Web site about as readable as the Unabomber's manifesto—became popular only after the suicides. "The group's rise and fall was tangential to the Net, but you'd never know it from the coverage."

If you don't regard press treatment of Heaven's Gate as revelatory of shifting public attitudes toward the Internet, then consider the many other signs. Har-

DAVID SUTER

old Sjursen, who directs a university program to deliver technical courses via the Net, says the Internet is no longer seen as a serious place to teach. "Saying that classes will be conducted on the Internet these days is like saying the classes are being offered on *The X-Files,*" he argues. "Many universities, well aware of this distaste, would prefer to be associated with Internet II [the high-speed successor now in development]. It's not merely a matter of bandwidth."

What has paved the way for this disenchantment? Part of the feeling stems from the invasion of the Internet by the masses: when Charles Manson talks of setting up a Web site, it's time for a new network. Playing a part, too, is fear of the Internet's efficiency at communication—the same fear directed in the past at television, public libraries and, in its time, the Gutenberg press.

There are other, more prosaic reasons for the erosion of optimism regarding the Net. For many, particularly those connected by modem, congestion and glitches make use of the Internet a time-wasting, unreliable pursuit. Many businesspeople, too, have soured on the Net, for they cannot figure out how to make money on it. "I think the Web is marvelous," says Richard van Slyke, a computer scientist and professor at Polytechnic University in Brooklyn, N.Y. "But if I were an investor, I'd feel quite different. After all, I didn't pay for any of it, and there's no prospect that I will."

Among the technically minded, distaste for the Net focuses sometimes on the commercialism, sometimes on its incessant hype, as each new development is touted by a computer press floundering to pick and promote winners. The latest is "push technology," the term for Internet broadcasting that delivers information, such as stock quotes, without prompting. The data just inch automatically across a small window on your screen. PointCast is the leader in this technology, but soon many other push media will be arriving. And they will appear not only on our desktops but also on our wrists, on our walls and on the dashboards of cars we rent in strange cities—street maps and locators will stream past, along with accompanying commercials. It's not clear, though, whether push will ever replace the browser, as some press pundits are claiming. A hybrid is more likely.

If hype causes people to devalue the Net, so, too, does its relentlessly teenage diction: the reflexive, omnipresent use of "cool," "wow," "killer app" and other youthful expressions. Couple the breathless word choice and hyperbole with juvenilia—find a home page for noted sourpuss St. Augustine of Hippo, and it will probably have a joke menu—and you have the Net's distinctive cadence, one that trivializes with every word.

Despite it all, most of us will probably keep right on using the Internet to stay in touch with colleagues, buy airline tickets or download tax forms instead of searching through dog-eared piles at the post office. "It takes time," Saffo says, "to take a raw, untamed technology and turn it into a compelling medium. All media go through adolescence; the Web happens to be going through a particularly rough one. For now, if people don't want to be associated with the Net, fine. I'll happily take their bandwidth." —*Anne Eisenberg*

Iran's Nuclear Puzzle

Rich in fossil-fuel resources, Iran is pursuing a nuclear power program difficult to understand in the absence of military motives

by David A. Schwarzbach

When the Iranian government announced in 1995 that it had signed contracts totaling $940 million with the Russian Ministry of Atomic Energy to complete a commercial nuclear power plant near the town of Bushehr, the U.S. response came immediately. U.S. Secretary of State Warren Christopher campaigned to convince the Russians that the proposed sale would contribute to the proliferation of nuclear weapons by helping Iran assemble an atomic arsenal.

Although Christopher's entreaties were rebuffed, little progress has been made over the past two and a half years on the ambitious project, which many experts believe will ultimately cost far more than $940 million. Nevertheless, Bushehr, on the Persian Gulf, is emblematic of Iran's baffling foray into nuclear technology. At the heart of this puzzle is a question: Why would a country with enormous reserves of natural gas and other fossil fuels, and with a gross domestic product of only $62 billion, commit itself to spending perhaps billions of dollars on a nuclear plant that could not possibly generate electricity as cost-effectively as a natural-gas plant? The question is difficult to answer realistically unless the Bushehr plant is viewed as a foothold from which Iran could climb toward an atomic bomb.

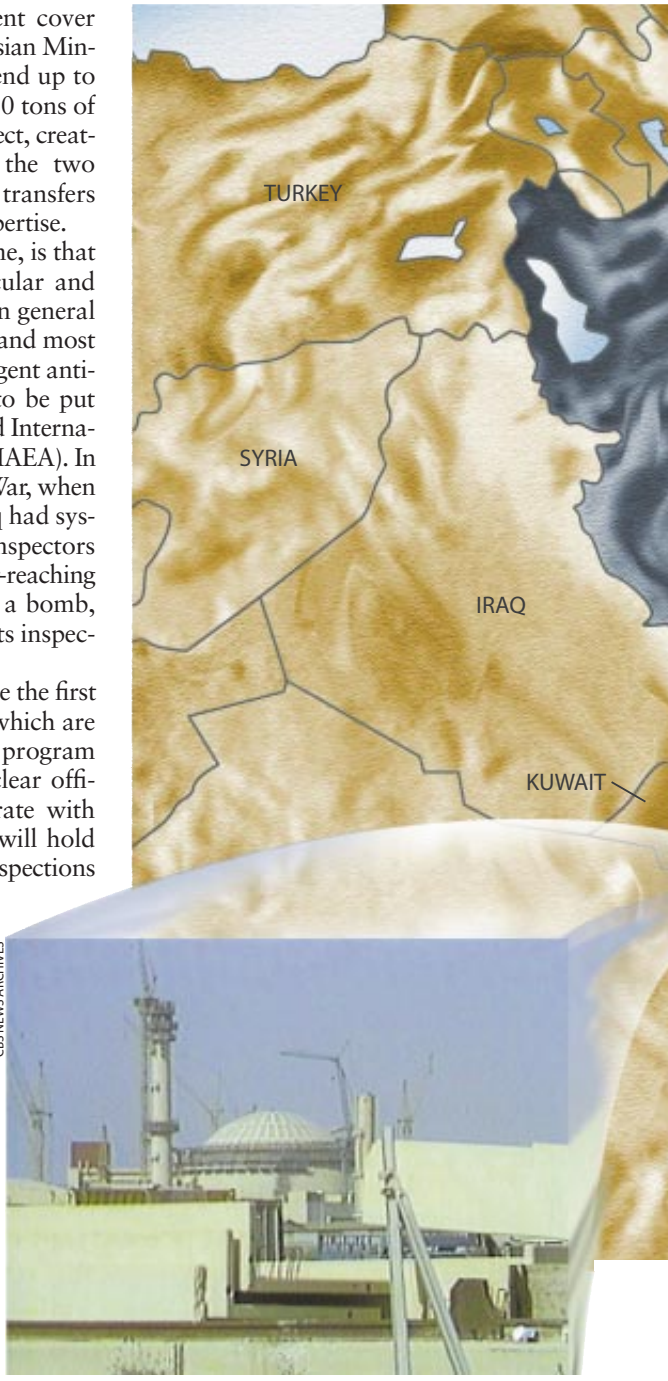
It is true that Bushehr could in some minor way help alleviate the country's serious shortage of electrical generating capacity. At the same time, the project would enable Iran to train a generation of engineers in the operation of a nuclear reactor—the basics of which apply equally whether the reactor's primary purpose is the production of electricity or of plutonium, one of the two standard fissile materials that can be used in the construction of nuclear weapons. More immediately, the huge Bushehr project could provide excellent cover for smuggling efforts. The Russian Ministry of Atomic Energy will send up to 3,000 workers and bring 7,000 tons of equipment to Iran for the project, creating enough traffic between the two countries to shield any covert transfers of equipment, materials or expertise.

The bright side, if there is one, is that the Bushehr project in particular and the Iranian nuclear program in general will constitute one of the first and most challenging tests of more stringent antiproliferation measures soon to be put into effect by the Vienna-based International Atomic Energy Agency (IAEA). In the wake of the Persian Gulf War, when the world discovered that Iraq had systematically deceived IAEA inspectors and managed to assemble a far-reaching clandestine program to build a bomb, the IAEA began overhauling its inspection and monitoring efforts.

Iran's nuclear projects will be the first test of the IAEA procedures, which are being formulated under a program known as 93+2. Iranian nuclear officials have pledged to cooperate with the IAEA; whether this vow will hold up under the more intrusive inspections remains to be seen.

It is not known conclusively whether Iran now has an active military nuclear program, although evidence gathered by several intelligence services tends to support the notion. Moreover, in the recent past the country launched an acquisition effort that was undoubtedly directed at producing a bomb.

According to knowledgeable observers, Iraq's devastating defeat in early 1991 convinced Iranian government officials that their coun-

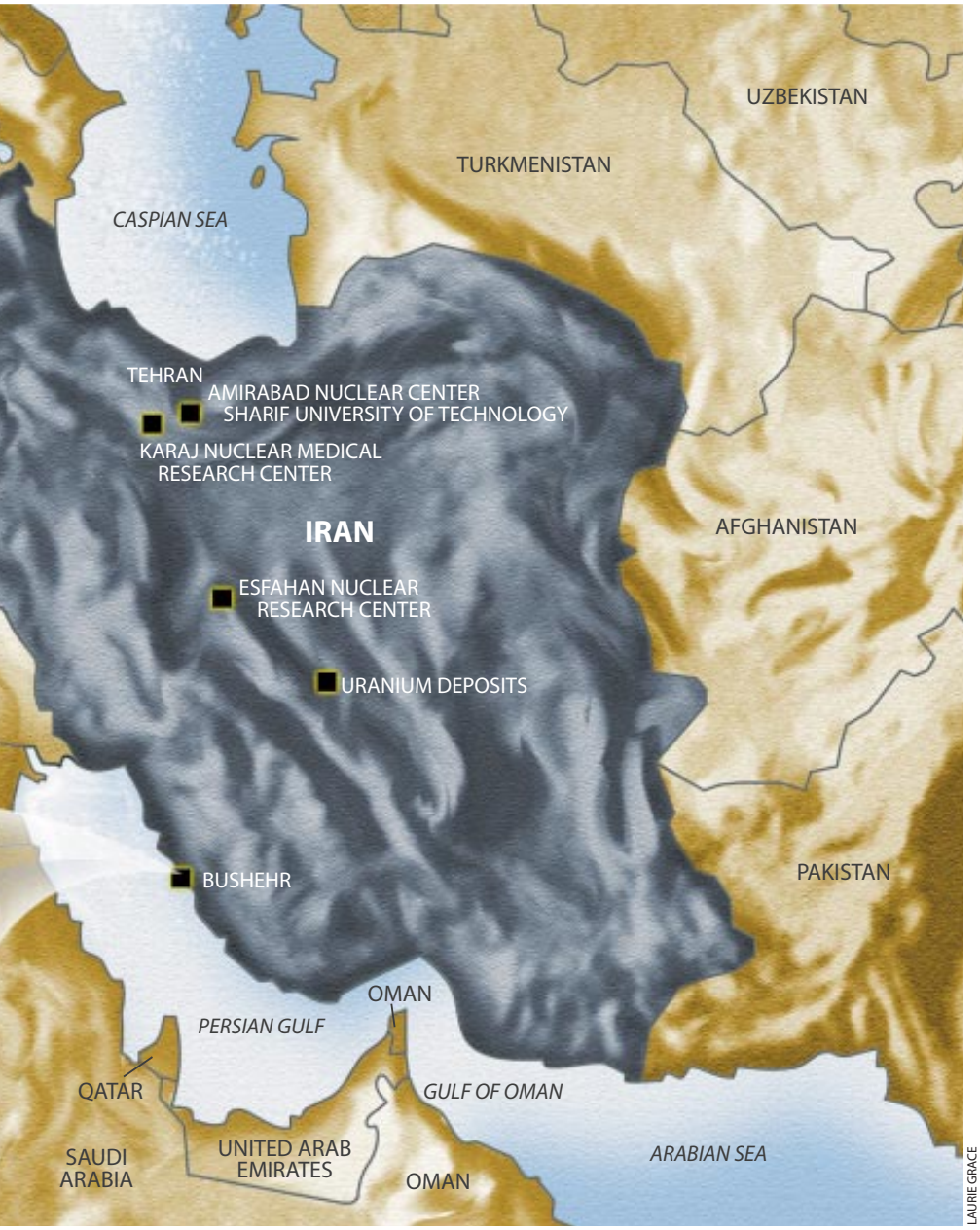


CBS NEWS ARCHIVES

try could not hope to rely on conventional forces alone to deter any future Western intervention in the Gulf region. In nuclear weapons Iranian officials saw a way not only to handle the West but also to counter the threat of chemical or biological weapons from regional enemies such as Iraq. In 1991 and 1992 Iran tried to purchase a variety of equipment from Argentine, Chinese, European and Indian sources; with the appropriate expertise, the sought-after components would have provided Tehran with the means to build a small atomic arsenal. Pressure from the U.S. blocked the deals, but intelligence reports spanning the past several years indicate that Iran's global smuggling network remains intact.

Having failed in its major procurement attempts, Iran maintains a relatively basic nuclear infrastructure. Two Iranian reactors are capable of producing plutonium today. One is a research reactor at the Tehran University Amirabad Nuclear Center, established in the 1960s under the Shah and equipped by the U.S. government. In addition to the reactor, the center has a small laboratory where plutonium can be separated from spent reactor fuel. The laboratory's limited resources permit the separation of only about 0.6 kilogram of plutonium annually; roughly five to seven kilograms of plutonium would be needed to construct a bomb, depending on the expertise of the bomb builders. (In comparison, the Bushehr reactors would be capable of producing upward of 180 kilograms of plutonium a year.) Nuclear scientists could in time secretly accumulate enough material from Amirabad for a weapon. Such a diversion would not be straightforward, though, because Amirabad, like all Iranian nuclear installations, is under IAEA safeguards designed to deter proliferation.

Iran's only other reactor with the ability to produce plutonium is capable of making only trivial quantities. The reactor is at the Esfahan Nuclear Research Center, which was begun in the mid-1970s by a French nuclear concern and completed, with help from the Chinese, after the overthrow of the Shah. Iran has plans to expand Esfahan, and extensive activities there prompted the IAEA to conduct inspections late last year. None of those inspections, however, turned up firm evidence that a clandestine weapons program was under way.



LAURIE GRACE

IRANIAN NUCLEAR ESTABLISHMENT includes several research facilities in or near Tehran and one at Esfahan, in addition to sizable deposits of uranium. A Russian electric power reactor is to be installed at Bushehr (*photograph at left*) in a facility built by the German firm Kraftwerk Union, which halted work after the Islamic revolution in 1979.

Better Way to a Bomb

Although most modern nuclear weapons are based on plutonium, it is also possible to build a device based on highly enriched uranium (HEU). In fact, for a developing country seeking to build a weapon surreptitiously, this option is in many ways the more desirable one, even though it takes roughly twice as much HEU to produce a weapon. Not only is it simpler to produce a weapon with HEU, but, more important, following this route eliminates the need for a kind of industrial-chemical plant, known as a reprocessing facility, to separate plutonium from spent reactor fuel.

Of course, producing HEU has its own set of industrial requirements, chief among them the equipment needed to enrich the uranium by boosting the concentration of the highly fissile uranium 235 isotope from its natural 0.7 percent to the 93 percent found in weapons-grade HEU. Enrichment is a rather exacting process in its own right; the necessary facilities, though, are more easily disguised as ordinary industrial plants than are reprocessing facilities, which

release various unique isotopes whose presence can instantly give away the facilities' true purpose.

The standard uranium-enrichment technique today uses hundreds of centrifuges that spin uranium hexafluoride gas at very high speeds, enabling centrifugal forces to separate the lighter uranium 235 hexafluoride molecules from the heavier uranium 238 ones. Various reports indicate that Iran has aggressively, though unsuccessfully, sought this enrichment technology on the black market and from the Russian Ministry of Atomic Energy.

Another, antiquated method of enriching uranium dates to the U.S. Manhattan Project during World War II. With this technique, known as electromagnetic isotope separation, a stream of uranium ions is deflected by electromagnets in a vacuum chamber. The heavier uranium 238 ions are deflected less than the uranium 235 ions, and this slight difference is used to separate out the uranium 235. The chamber and its associated equipment are actually a special type of cyclotron called a calutron (for "California University cyclotron"). The calutron requires much greater amounts of energy than the centrifuge approach, but the necessary components are more easily imported or manufactured domestically.

Analysts suspect that research into uranium enrichment has been carried out at three Iranian nuclear facilities: the Esfahan Nuclear Research Center, the Sharif University of Technology (at the University of Tehran) and the Karaj Agricultural and Medical Research Center. What little public knowledge there is regarding these efforts concerns the presence of a calutron and a cyclotron at Karaj.

A cyclotron purchased from the Belgian firm Ion Beam Applications was installed in 1991 at Karaj, leading French analysts to suspect that Iran was launching a uranium-enrichment research program. Karaj also has a small, Chinese-supplied calutron. Neither of these accelerators could produce militarily significant quantities of HEU, but both could be useful for research as well as for training in isotope separation.

Enrichment technology, such as centrifuges, is not used uniquely for making weapons; it is also needed to produce, among other things, power-reactor fuel. Yet given the glut of low-enriched uranium fuel available worldwide in the wake of the cold war, it is hard to understand why Iran would want to make its own fuel even if it does succeed in getting the Bushehr plant up and running. To develop such a capability would be extremely costly for a developing country, and Iran's energy security is assured by its vast reserves of fossil fuels.

Indeed, Iran's known reserves of natural gas are the world's second largest and by conservative estimates could easily accommodate all the country's electrical energy needs for the next 50 to 100 years.

Electricity Shortage

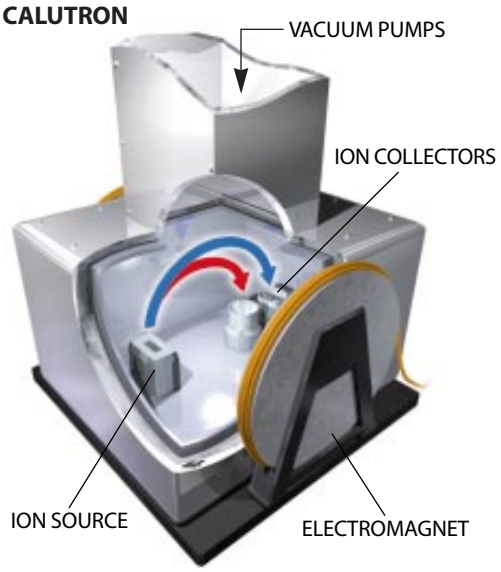
It is true, though, that Iran is starved for electricity. Present installed electrical capacity is about 20,000 megawatts. Growth in electrical demand is difficult to estimate but seems to be roughly 6 to 8 percent a year. (Although not unheard of in developing countries, this kind of growth is far in excess of the 2 or 3 percent typical of developed nations.) Notwithstanding its large oil reserves, Iran is unlikely to tap them for its own use. Oil is one of the very few means Iran has for amassing foreign currency, and oil sales account for upward of 85 percent of the country's trade earnings.

The economics of finishing Bushehr are similarly unencouraging. The project dates to the mid-1970s, when the Shah contracted with Kraftwerk Union in Germany to build two Siemens 1,200-megawatt electric reactors at Bushehr. The project was 70 percent complete in 1979 when the Islamic revolution halted work. The engineers from the Russian Ministry of Atomic Energy face enormous challenges in modifying the existing structures to accommodate a Russian VVER-1000 reactor and its support systems. Just the alterations needed for the Russian steam generators—which play the key role of converting heat from the reactor into steam to drive the turbines—are daunting. Six horizontal VVER steam generators will have to be installed in place of the four vertical Siemens units the structure was built for. The Russians will have to accomplish this feat without any technical documents or blueprints, because the Germans did not provide them to the Iranians in the 1970s.

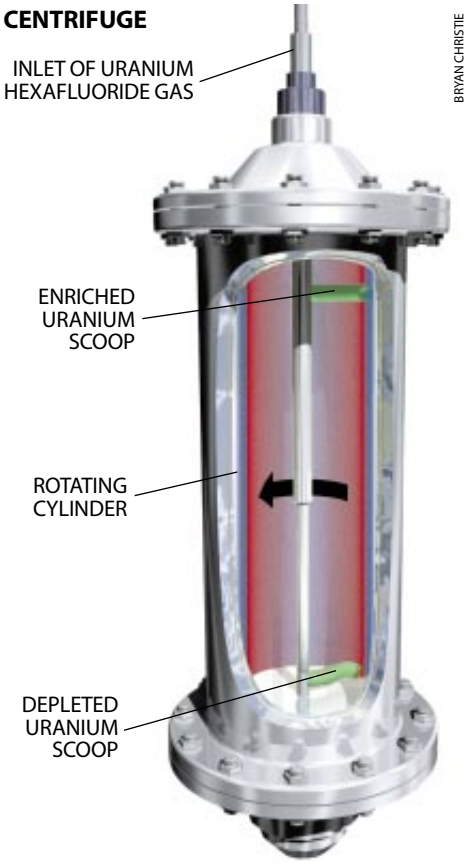
Realistic estimates of the cost to finish Bushehr run well over $1,000 per installed kilowatt. With cost overruns or construction delays caused by the unique difficulties of the project, the price could go much higher. In comparison, natural-gas-fired power plants run about $800 per kilowatt on average in the Middle East. With large-scale development the price would most likely fall to the $600 or less found in the West.

Nuclear power often makes up for being capital intensive by having lower fuel costs. But natural gas is extremely cheap in Iran and is likely to stay that way for the foreseeable future. Iran may learn the hard way—as other nations have—that nuclear power is uneconomic. In the meantime the country's dogged devotion to Bushehr in spite of more reasonable alternatives should make the international community suspicious.

As a signatory to the Non-Proliferation Treaty, Iran must put Bushehr under observation by the IAEA, greatly complicating any efforts to divert the plant's spent fuel into a secret weapons-making effort. In fact, Iran has stated that it has neither the need nor the desire to keep the spent fuel produced by the reactor once it goes into operation and would be willing to return it to Russia. Though preferable, this course would not eliminate completely the threat of diversion, because the spent fuel will pass several years cooling off in ponds



URANIUM ENRICHMENT is typically accomplished in a centrifuge (*right*), in which uranium hexafluoride gas is spun at extremely high speeds. Molecules based on the heavier uranium 238 (*blue*) atoms migrate closer to the cylinder wall than do those based on uranium 235 (*red*). With electromagnetic isotope separation, slight differences in the deflections of two streams of uranium ions in a device called a calutron (*above*) are able to distinguish the two isotopes.



at Bushehr before leaving the country.

Because of weak controls on nuclear materials and technology in Russia today, the traffic in equipment and workers will increase the risk that Iran could successfully obtain significant quantities of uranium or plutonium either directly on the black market or, like Pakistan, through the acquisition of uranium-enrichment technology. In addition, the project will provide legitimate grounds for Iran to expand nuclear-related research and training, making a military program easier to conceal.

The U.S. and the international community can work together to prevent such an outcome by taking an active role in influencing the conditions under which Bushehr will operate. Besides a Russian commitment to take back the spent fuel from Bushehr, Washington should bargain in particular for broad use by the IAEA of new methods that can greatly enhance the detection of a domestic Iranian effort to separate plutonium, enrich uranium and even construct nuclear weapons components. In 1993, as part of an effort to enhance nuclear safeguards after the Iraqi experience, the IAEA began to implement new nuclear inspection procedures, which it hoped at the time to have in place two years later (hence the name "93 + 2"). Under part one of the program, which was approved last year, inspections of declared nuclear facilities will now include the use of powerful, isotopic detection techniques that were previously barred.

An Engaging Proposition

Iranian officials recently agreed in principle to these measures, although implementation is still a matter of difficult negotiation. More important, the IAEA board of governors was expected to take up a proposal in May to expand the monitoring system to one that could be applied at any facility in any country that is a member of the IAEA—including facilities that were not declared as nuclear sites. Iran's willingness to allow this kind of broad monitoring, under part two of 93 + 2, will be a critical indicator of that country's intentions.

The isotopic detection techniques on which 93 + 2 will rely come under the heading of environmental monitoring. They take advantage of an inability to prevent minute quantities of a material from escaping an industrial plant or process. By using spectrometry, for example, a lab can accurately identify the isotopic ratio of a sample containing less than a billionth of a gram of material. Because the ratio of uranium 235 to uranium 238 in natural uranium is the same almost everywhere, samples with higher or lower ratios would most likely indicate that illegal enrichment had occurred.

With regard to plutonium, the presence of the element at levels in excess of those expected would suggest the existence of a reprocessing program. More likely to be detected are products of nuclear fission, such as radioactive iodine and krypton isotopes. Wipes from the surface of walls and equipment, along with soil, air, vegetation and water samples taken from suitably chosen locations, can help provide early warning of a nuclear weapons effort.

Because Iran has already agreed to allow inspections of any location by the IAEA, Tehran should have no objections to the establishment of a countrywide environmental monitoring program. In turn, the IAEA should push for the broadest possible implementation of these techniques, which proved so successful in revealing what had gone undetected in Iraq. All the same, it is important to remember what the 93 + 2 regime would not be able to discern: the acquisition of weapons-grade material on the black market or the theft of a "loose nuke."

Looking at the broader context, Iran, with its huge natural-gas reserves and proximity to emerging and established markets in Asia and eastern Europe, could become a major player in the next century as natural gas begins to eclipse oil as a primary form of energy. Pipelines from Iran could temper the country's nuclear ambitions by strengthening its economic ties to other countries and by promoting economic development at home.

This kind of investment would make much better use of Iran's limited capital than would venturing down a very expensive atomic trail. That is, of course, if acquiring nuclear weapons is not the real aim. SA

The Author

DAVID A. SCHWARZBACH is a former staff member of the Natural Resources Defense Council, a public-interest organization with offices in New York City, Washington, D.C., San Francisco and Los Angeles. An affiliate of the Center for Energy and Environmental Studies at Princeton University, his primary research interests center on ongoing efforts to restructure the North and South American electricity industries.

Further Reading

IRAN AND IRAQ: THE THREAT FROM THE NORTHERN GULF. Anthony Cordesman. Westview Press, Boulder, Colo., 1994.

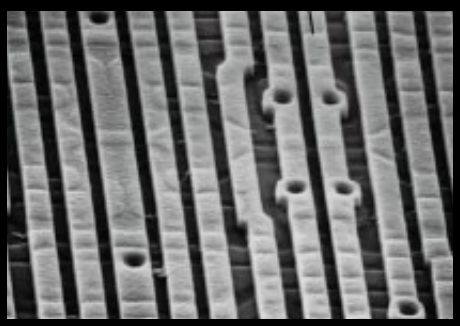
DOES IRAN WANT NUCLEAR WEAPONS? Shahram Chubin in *Survival,* Vol. 37, No. 1, pages 86–104; Spring 1995.

ENVIRONMENTAL MONITORING FOR NUCLEAR SAFEGUARDS. U.S. Congress Office of Technology Assessment. U.S. Government Printing Office (OTA-BP-ISS-168), Washington, D.C., September 1995.

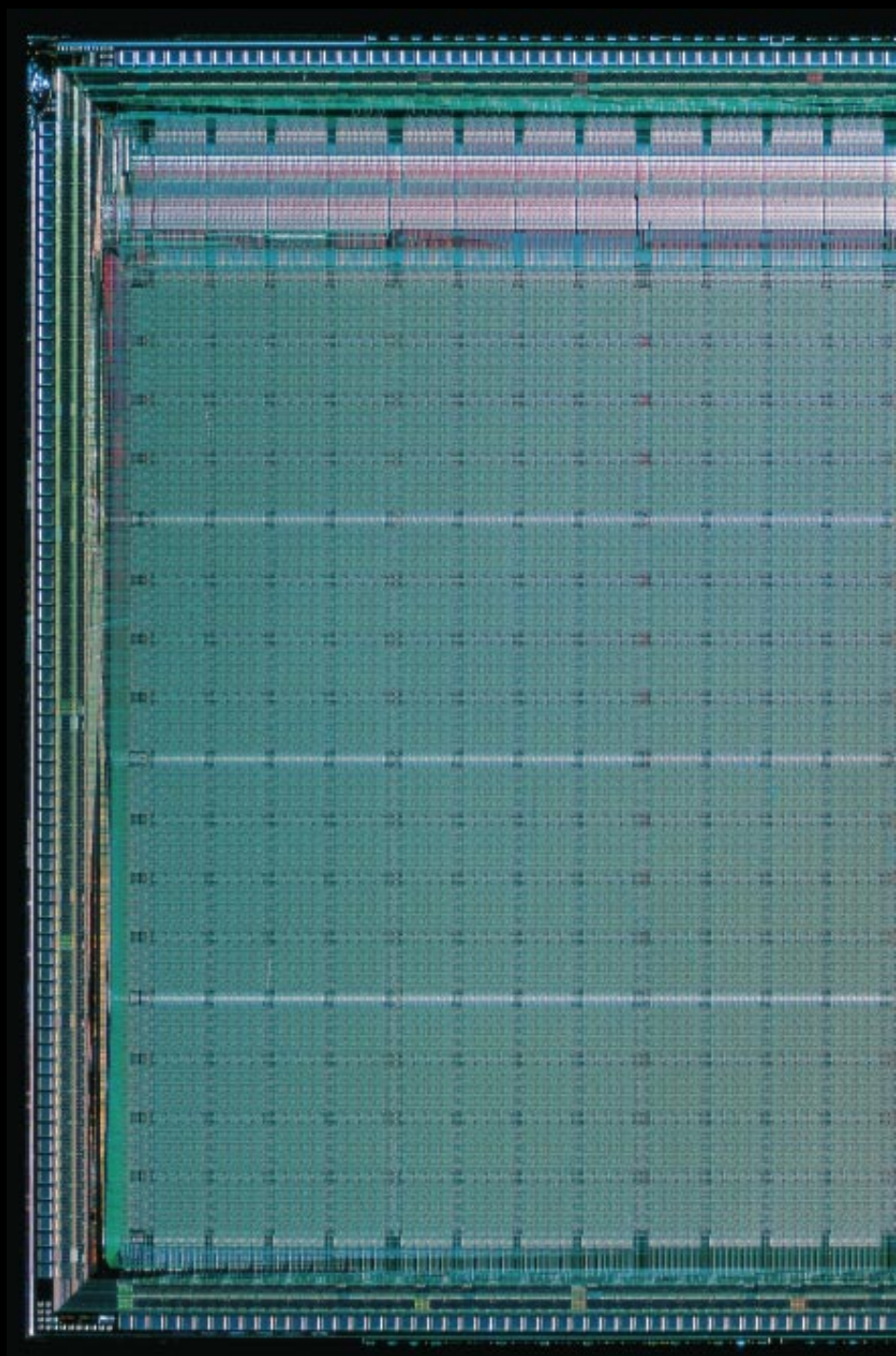
IRAN'S NUCLEAR PROGRAM: ENERGY OR WEAPONS? David Schwarzbach. A publication of the Natural Resources Defense Council, Nuclear Weapons Databook series, 1995.

Configurable Computing

Computers that modify their hardware circuits as they operate are opening a new era in computer design. Because they can filter data rapidly, they excel at pattern recognition, image processing and encryption



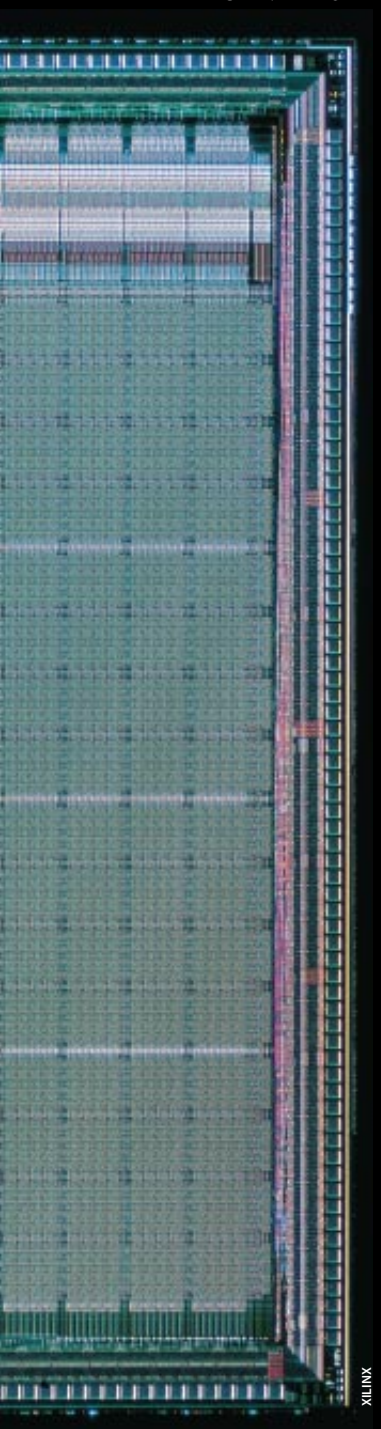
INTEGRATED CIRCUIT ENGINEERING



RECONFIGURABLE LOGIC DEVICE (*right*) **has circuits that can be partially or completely changed while it is operating. Scanning electron micrographs above (magnification 3,000 ×) show three fixed-circuit metal layers of a flip-flop cell in a field-programmable gate array. The flip-flop cell contains four memory cells with switches (*bottom, colored areas*) that can be opened or closed electronically to produce a conductive path.**

by John Villasenor and William H. Mangione-Smith

APPROXIMATE SIZE



Computer designers face a constant struggle to find the right balance between speed and generality. They can build versatile chips that perform many different functions relatively slowly, or they can devise application-specific chips that do only a limited set of tasks but do them much more quickly. Microprocessors (such as the Intel Pentium or Motorola PowerPC chips commonly found in personal computers) are general purpose: programming instructions encoded in binary format can lead a microprocessor through virtually any logical or mathematical operation a programmer can conceive. The Intel Pentium, for example, was never designed specifically to execute either Microsoft Word or the computer game DOOM, but it can run both.

In contrast, custom hardware circuits, often known as application-specific integrated circuits (ASICs), provide precisely the functionality needed for a specific task. By carefully tuning each ASIC to a given job, the computer designer can produce a smaller, cheaper, faster chip that consumes less power than a programmable processor. A custom graphics chip for a PC, for instance, can draw lines or paint pictures on the screen 10 or 100 times as quickly as a general-purpose central processing unit can.

As designers make their choices between versatility and speed, they must also confront the issue of cost. A well-designed ASIC will solve the specific problem for which it was designed, but probably not a slightly modified problem introduced after the ASIC design is finished. Furthermore, even if a modified ASIC can be developed for the new problem, the original hardware circuits may be too highly customized to be reused in successive generations. As a result, the engineering effort required to design and build an ASIC must be amortized over a relatively small number of units.

A new development in integrated circuits offers a third option: large, fast, field-programmable gate arrays, or FPGAs—highly tuned hardware circuits that can be modified at almost any point during use. FPGAs consist of arrays of configurable logic blocks that implement the logical functions of gates. Logic gates are like switches with multiple inputs and a single output. They are used in digital circuits to perform basic binary operations such as AND, NAND, OR, NOR and XOR. In most hardware that is used in computing today, the logical functions of gates are fixed and cannot be modified. In FPGAs, however, both the logic functions performed within the logic blocks and the connections between the blocks can be altered by sending signals to the chip. These blocks are structurally similar to the gate arrays used in some ASICs, but whereas standard gate arrays are configured during manufacture, the configurable logic blocks in FPGAs can be rewired and reprogrammed repeatedly, long after the integrated circuit has left the factory.

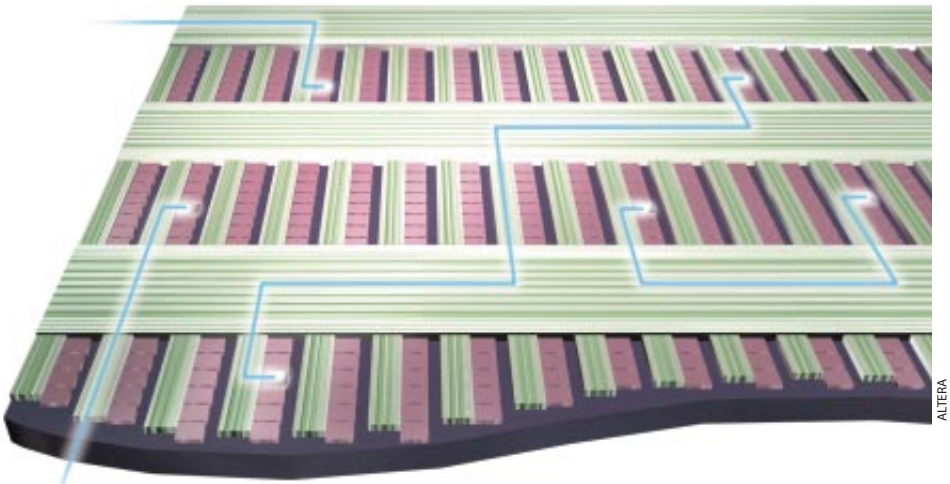
The key that has opened the door to configurable computing is the design of new FPGAs that can be configured extremely quickly. The earliest field-programmable arrays required several seconds or more to change their connections—perfectly suitable for engineers who wanted to test alternative circuit designs or for companies that sold devices that might need occasional upgrading. Newer FPGAs can be configured in one millisecond, and we expect to see devices with configuration times as low as 100 microseconds within two years. Ultimately, computing devices may be able to adapt their hardware almost continuously in response to changes in the input data or processing environment.

There are many variations on FPGA design, but the basic structure consists of a large number of configurable logic blocks and a programmable grid of connections that can link those blocks in any pattern the designer chooses. Those FPGAs that are coarse grained have a small number of powerful configurable logic blocks; those with a finer-grained structure have many simple blocks. A single element in a coarse-grained FPGA might be capable of adding or comparing two numbers. One block in a fine-grained device might be capable only of comparing two binary digits—in effect, it would be a single logic gate. A designer might choose to start with either a coarse- or fine-grained chip depending on the application at hand and the amount of time available for building complex subsystems from scratch.

Computing devices can make use of configurable elements in many different ways. The least demanding technique is to switch between functions on command—the hardware equivalent of quitting one program

Programmable Circuitry

Programmable circuits (*blue lines*) in a field-programmable gate array (FPGA) can be created or removed by sending electronic signals to gates in the logic elements. A built-in grid of circuits arranged in columns and rows allows the designer to connect a logic element to other logic elements or to an external memory or microprocessor. The logic elements are grouped in blocks that perform basic binary operations such as AND, OR and NOT; several firms, including Xilinx and Altera, have developed devices with the capability of 100,000 equivalent gates. *—J. V. and W. H. M. S.*



ALTERA

and then running another. Slow reconfiguration, on the order of several seconds, may well be acceptable in such an application. Faster programming times permit dynamic design swapping: a single FPGA performs a series of tasks in rapid succession, reconfiguring itself between each one. Such designs operate the chip in a time-sharing mode and swap between successive configurations so rapidly that it appears the FPGA is performing all its functions at once.

Using this approach, we have built a single-chip video transmission system that reconfigures itself four times per video frame. It thus requires only a quarter of the hardware that would be needed for a fixed ASIC. The FPGA first stores an incoming video signal in memory, then applies two different image-processing transformations and finally transforms itself into a modem to send the signal onward.

The most challenging and potentially most powerful form of configurable computing involves the hardware reconfiguring itself on the fly as it executes a task, refining its own programming for improved performance. An image-recognition chip might tune itself in response to a tentative identification of the object it is looking at: if an image contained a car or a truck, parts of the circuitry originally intended for tracking high-speed aircraft or slow-moving people could be reconfigured to focus instead on land vehicles. For some applications, such a radical departure from traditional computer design, in which the hardware is specified at the outset, could make for much faster and more versatile machines than are possible with either general-purpose microprocessors or custom chips.

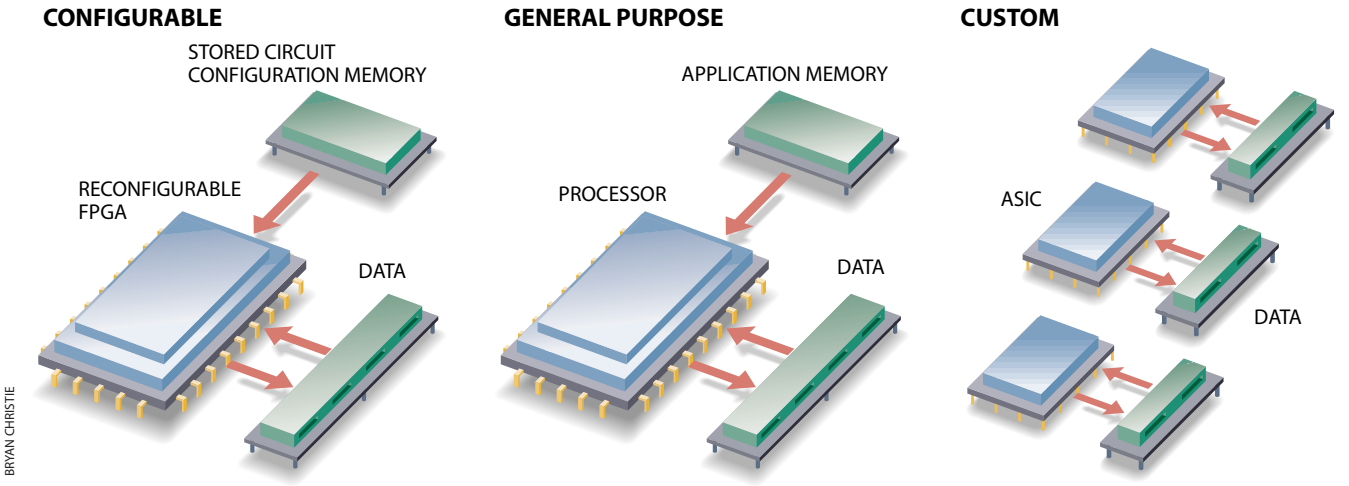
Cutting Critical Hardware

One of the most promising applications for configurable computing involves pattern matching. Pattern matching is used in tasks such as handwriting recognition, face identification, database retrieval and automatic target recognition. A fundamental operation of pattern matching involves comparing an input set of bits (representing an image, a string of characters or other data) with a set of templates corresponding to the possible patterns to be recognized. The system declares recognition when the number of input bits that match bits in the template exceeds some threshold.

In the case of target recognition—a military application that drove some of our initial work—the greatest challenge is the rapid comparison of an input image to thousands of templates. A template could represent, for example, a front or side view of a specific type of vehicle. Each image typically contains thousands of pixels (picture elements), and a target could appear at any position within an image. To recognize targets fast enough for military applications, a system needs to perform comparisons at the rate of several trillion operations per second, because all the pixels in the input image must be compared with all the pixels in many templates.

With support from the Defense Advanced Research Projects Agency (DARPA), we have built a prototype recognition system with configurable hardware that achieves significant hardware savings by tuning itself to each template in turn. Many of the pixels in a typical template do not contribute to the matching results [*see bottom illustration on opposite page*], and so the configurable computing machine could simply omit them from its calculations. A conventional system could not easily pare itself down in a similar way, because the pixels to be ignored differ from template to template. One can go further in exploiting the flexibility of configurable machines by tuning the hardware to take advantage of similarities among templates. The configurable hardware can process a set of templates in parallel, using only one comparison unit for each pixel whose value is the same for templates in that set. For example, rather than having eight separate hardware circuits consider a certain pixel for eight different templates, a single circuit can consider the pixel and then propagate its outcome to the seven other templates.

Most recently, we have built a prototype encryption system (also funded by DARPA) that takes advantage of configurable hardware. An FPGA implements the Data Encryption Standard (DES), which uses 56-bit-long keys to encrypt 64-bit-long blocks of data. (A key in encryption is a number used to scramble or unscramble a confidential message.) DES encryption usually proceeds in two steps: subkey scheduling and data processing. In the first step, a set of rotations and permutations translates the 56-bit encryption key into a series of 16 subkeys. Each subkey then processes the data in a separate round; a full set of 16



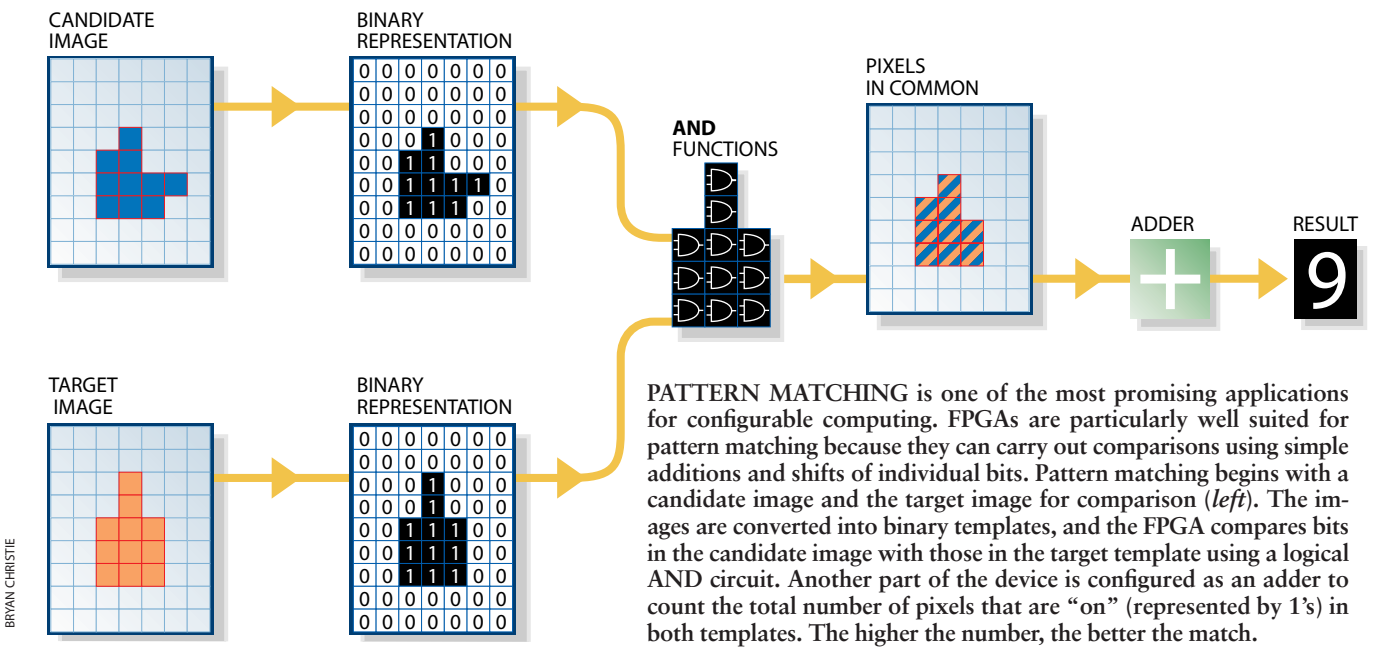
CONFIGURABLE COMPUTING architectures combine elements of general-purpose computing and application-specific integrated circuits (ASICs). The general-purpose processor operates with fixed circuits that perform multiple tasks under the control of software. An ASIC contains circuits specialized to a particular task and thus needs little or no software to instruct it. The configurable computer can execute software commands that alter its FPGA circuits as needed to perform a variety of jobs.

rounds encrypts or decrypts each 64-bit block. When the computer deals concurrently with multiple users, each dialogue between users must have a distinct key, and the encryption hardware will change keys as parts of messages arrive for different users.

In many applications of DES, the encryption key remains constant while a long block of data passes through the data path. For example, if two people are communicating over a secure network, they exchange a secure encryption key once and then use that key throughout the duration of their dialogue to generate the subkeys for each round of encryption or decryption. Some ASICs are designed to handle only one kind of encryption algorithm, such as DES; others—such as programmable digital signal processors—are capable of implementing many encryption algorithms.

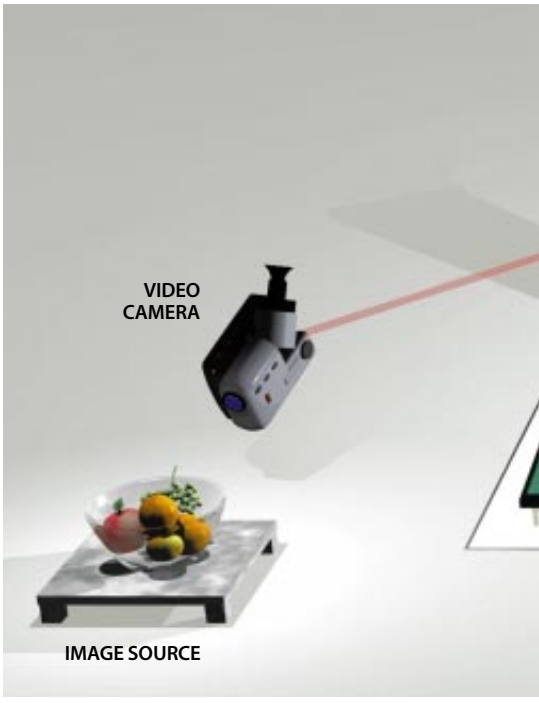
With a configurable chip, the software can calculate the subkey values once, and the data-processing circuitry can be optimized for those specific subkeys. This approach allows the subkey-scheduling hardware to be completely removed from the system. These savings have allowed us to implement the DES algorithm in a 13,000-gate FPGA, instead of the 25,000-gate circuit previously required. When the encryption key must be changed, the user can quickly specify a new circuit, customized to the new key, and download it to the FPGA.

The target-recognition and encryption prototypes we have built help illustrate the enormous flexibility that arises when the hardware in a computer can be customized to a diverse and changing set of external data. There are many other applications that could benefit from the ability to modify the computation



PATTERN MATCHING is one of the most promising applications for configurable computing. FPGAs are particularly well suited for pattern matching because they can carry out comparisons using simple additions and shifts of individual bits. Pattern matching begins with a candidate image and the target image for comparison (*left*). The images are converted into binary templates, and the FPGA compares bits in the candidate image with those in the target template using a logical AND circuit. Another part of the device is configured as an adder to count the total number of pixels that are "on" (represented by 1's) in both templates. The higher the number, the better the match.

PROTOTYPE VIDEO COMMUNICATIONS SYSTEM uses a single FPGA to perform four functions that typically require separate chips. A memory chip stores the four circuit configurations and loads them sequentially into the FPGA. Initially, the FPGA's circuits are configured to acquire digitized video data. The chip is then rapidly reconfigured to transform the video information into a compressed form and reconfigured again to prepare it for transmission. Finally, the FPGA circuits are reconfigured to modulate and transmit the video information. At the receiver, the four configurations are applied in reverse order to demodulate the data, uncompress the image and then send it to a digital-to-analog converter so it can be displayed on a television screen. The prototype is capable of processing eight frames per second.



hardware in this manner, including digital communications, design automation and digital filtering for radar.

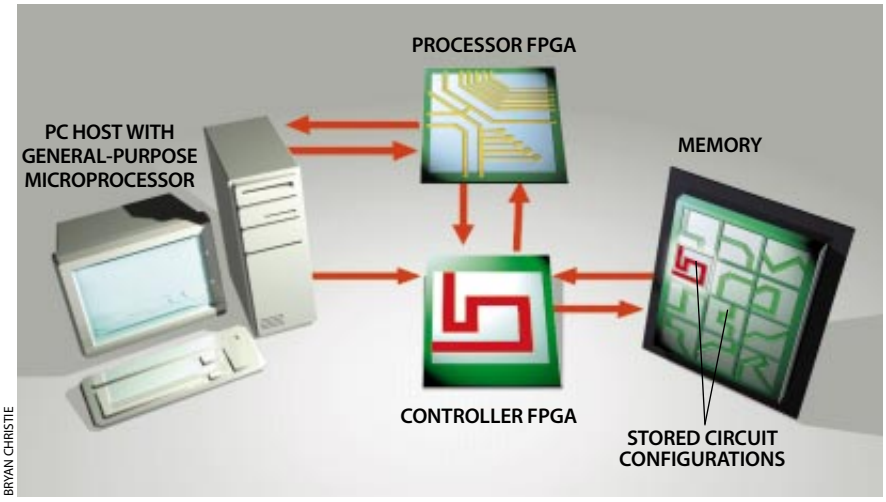
The Future of Configurable Computing

Configurable computing is still an extremely young field. Although Gerald Estrin of the University of California at Los Angeles proposed configurable computing in the late 1960s, the first demonstrations did not occur until a few years ago, and current FPGAs, with up to 100,000 logic elements, still do not come close to exploiting the full possibilities of the technique. Future FPGAs will be much larger; as with many other integrated circuits, the number of elements on a single FPGA has doubled roughly every 18 months. Before the decade is out, we expect to see FPGAs that have a million logic elements. Such chips will have much broader application, including highly complex communications and signal-processing algorithms.

Academic researchers and manufacturers are overcoming numerous other design limitations that have hindered the adoption of configurable computing. Not all computations can be implemented efficiently with today's FPGAs: they are well suited to algorithms composed of bit-level operations, such as pattern matching and integer arithmetic, but they are ill suited to numeric operations, such as high-precision multiplication or floating-point calculations. Dedicated multiplier circuits such as those used in microprocessor and digital signal chips can be optimized to perform more efficiently than multiplier circuits constructed from configurable logic blocks in an FPGA. Furthermore, FPGAs currently provide very little on-chip memory for storage of intermediate results in computations; thus, many configurable computing applications require large external memories. The transfer of data to and from the FPGA increases power consumption and may slow down the computations.

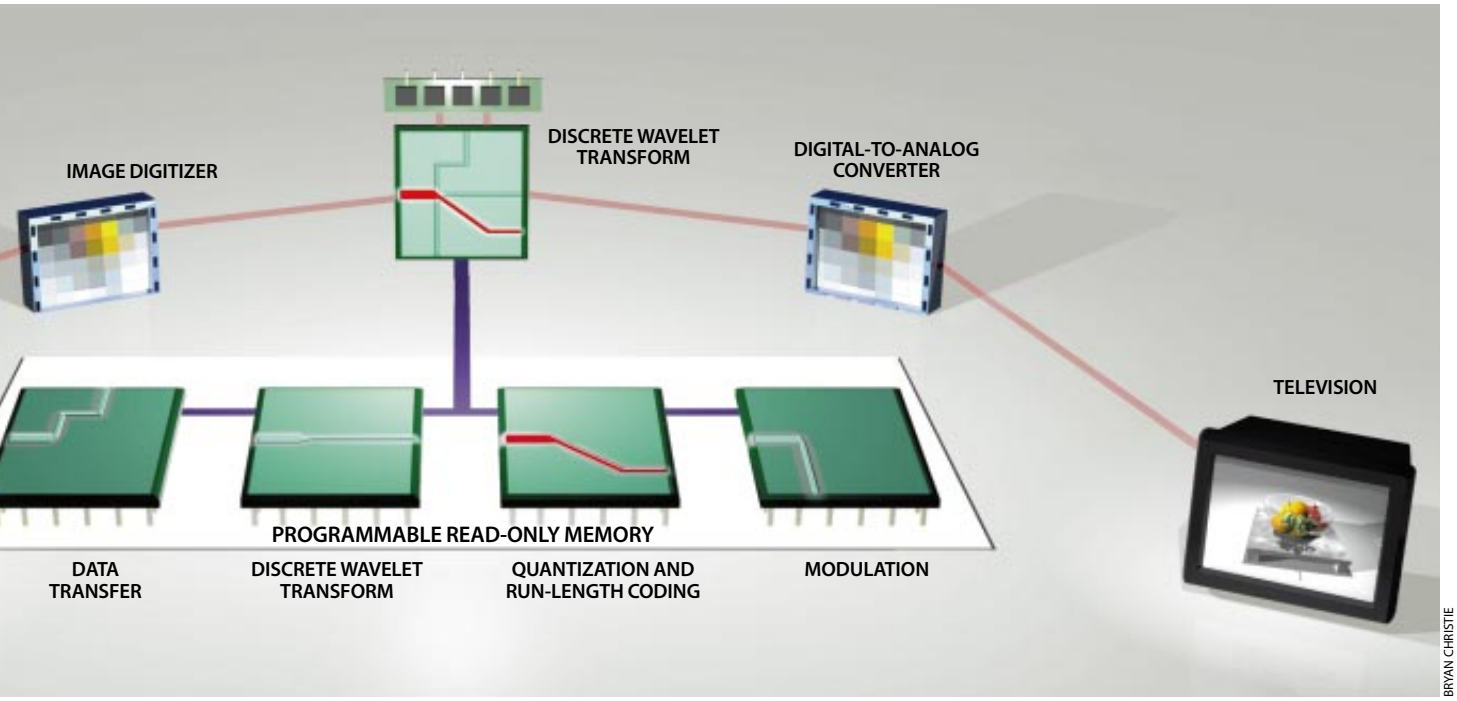
Fortunately, researchers are developing advanced FPGAs that contain memory, arithmetic processing units and other special-purpose blocks of circuitry. André DeHon and Thomas F. Knight, Jr., of the Massachusetts Institute of Technology have proposed an FPGA that stores multiple configurations in a series of memory banks. In a single clock cycle, which is on the order of tens or hundreds of nanoseconds, the chip could swap one configuration for another configuration without erasing partially processed data.

Meanwhile Brad L. Hutchings of Brigham Young University has used configurable computing to build a dynamic instruction set computer (DISC), which effectively marries a microprocessor to an FPGA and demonstrates the potential of automatic reconfiguration using stored configurations. As a program runs, the FPGA requests reconfiguration if the designated circuit is not resident. DISC allows a designer to create and store a large number of circuit configurations and activate them much as a programmer would initiate a call to a software subroutine in a microprocessor.



BRYAN CHRISTIE

HYBRID-ARCHITECTURE COMPUTER combines a general-purpose microprocessor and reconfigurable FPGA chips. The dynamic instruction set computer (DISC), built at Brigham Young University to demonstrate mixed architecture advantages, consists of two FPGAs and memory on a circuit board connected to a personal computer. The controller FPGA loads circuit configurations stored in the memory onto the processor FPGA in response to the requests of the operating program. If the memory does not contain a requested circuit, the processor FPGA sends a request to the PC host, which then loads the configuration for the desired circuit.



BRYAN CHRISTIE

The Colt Group, led by Peter M. Athanas of Virginia Polytechnic Institute and State University, is investigating a run-time reconfiguration technique called Wormhole that lends itself to distributed computing. The unit of computing is a stream of data that creates custom logic as it moves through the reconfigurable hardware.

John Wawrzynek and his colleagues at the University of California at Berkeley are developing systems that combine a microprocessor and an FPGA. Special compiler software would take ordinary program code and automatically generate a combination of machine instructions and FPGA configurations for the fastest overall performance. This approach takes advantage of opportunities to integrate a processor and an FPGA on a single chip.

FPGAs will never replace microprocessors for general-purpose computing tasks, but the concept of configurable computing is likely to play a growing role in the development of high-performance computing systems. The computing power that FPGAs offer will make them the devices of choice for applications involving algorithms in which rapid adaptation to the input is required.

In addition, the line between programmable processors and FPGAs will become less distinct: future generations of FPGAs will include functions such as increased local memory and dedicated multipliers that are standard features of today's microprocessors; there are also next-generation microprocessors under development whose hardware supports limited amounts of FPGA-like reconfiguration. Indeed, just as computers connected to the Internet can now automatically download special-purpose software components to perform particular tasks, future machines might download new hardware configurations as they are needed. Computing devices 10 years from now will include a strong mix of software-programmable hardware and hardware-configurable logic. SA

The Authors

JOHN VILLASENOR and WILLIAM MANGIONE-SMITH are on the faculty of the electrical engineering department at the University of California, Los Angeles. Villasenor received a Ph.D. in electrical engineering from Stanford University in 1989 and was with the Jet Propulsion Laboratory from 1990 to 1992. He moved to U.C.L.A. in 1992 and currently conducts research on image and video communications and on configurable computing. Mangione-Smith received a Ph.D. in computer engineering from the University of Michigan in 1991 and then worked for Motorola until 1995, becoming the systems architect for a second-generation personal wireless communicator. His research interests include low-power and high-performance computing engines, processors, systems architecture and systems software.

Further Reading

A Dynamic Instruction Set Computer. M. J. Wirthlin and B. L. Hutchings in *Proceedings of the IEEE Symposium on FPGAs for Custom Computing Machines,* April 1995. IEEE Computer Society Press, 1995.

A First Generation DPGA Implementation. E. Tau, D. Chen, I. Eslick, J. Brown and A. DeHon in *Proceedings of the Third Canadian Workshop on Field-Programmable Devices,* pages 138–143; May 1995.

Configurable Computing Solutions for Automatic Target Recognition. John Villasenor, Brian Schoner, Kang-Ngee Chia, Charles Zapata, Hea Joung Kim, Chris Jones, Shane Lansing and Bill Mangione-Smith in *4th IEEE Symposium on FPGAs for Custom Computing Machines (FCCM '96) April 1996.* Available from http://www.computer.org/cspress/catalog/pr07548.htm

Wormhole Run-Time Reconfiguration. Ray Bittner and Peter Athanas in *FPGA '97: ACM/SIGDA International Symposium on Field Programmable Gate Arrays.* Special Interest Group on Design Automation (SIGDA), ACM, 1997.

Early Hominid Fossils from Africa

A new species of Australopithecus, *the ancestor of* Homo, *pushes back the origins of bipedalism to some four million years ago*

by Meave Leakey and Alan Walker

The year was 1965. Bryan Patterson, a paleoanthropologist from Harvard University, unearthed a fragment of a fossil arm bone at a site called Kanapoi in northern Kenya. He and his colleagues knew it would be hard to make a great deal of anatomic or evolutionary sense out of a small piece of elbow joint. Nevertheless, they did recognize some features reminiscent of a species of early hominid (a hominid is any upright-walking primate) known as *Australopithecus,* first discovered 40 years earlier in South Africa by Raymond Dart of the University of the Witwatersrand. In most details, however, Patterson and his team considered the fragment of arm bone to be more like those of modern humans than the one other *Australopithecus* humerus known at the time.

The age of the Kanapoi fossil proved somewhat surprising. Although the techniques for dating the rocks where the fossil was uncovered were still fairly rudimentary, the group working in Kenya was able to show that the bone was probably older than the various *Australopithecus* specimens previously found. Despite this unusual result, however, the significance of Patterson's discovery was not to be confirmed for another 30 years. In the interim, researchers identified the remains of so many important early hominids that the humerus from Kanapoi was rather forgotten.

Yet Patterson's fossil would eventually help establish the existence of a new species of *Australopithecus*—the oldest yet to be identified—and push back the origins of upright walking to more than four million years (Myr) ago. But to see how this happened, we need to trace the steps that paleoanthropologists have taken in constructing an outline for the story of hominid evolution.

Evolving Story of Early Hominids

Scientists classify the immediate ancestors of the genus *Homo* (which includes our own species, *Homo sapiens*) in the genus *Australopithecus*. For several decades, it was believed that these ancient hominids first inhabited the earth at least three and a half million years ago. The specimens found in South Africa by Dart and others indicated that there were at least two types of *Australopithecus—A. africanus* and *A. robustus*. The leg bones of both species suggested that they had the striding, bipedal locomotion that is a hallmark of humans among living mammals. (The upright posture of these creatures was vividly confirmed in 1978 at the Laetoli site in Tanzania, where a team led by archaeologist Mary Leakey discovered a spectacular series of footprints made

MATT MAHURIN (*illustration*); ROBERT CAMPBELL (*left*); ALAN WALKER; © NATIONAL MUSEUMS OF KENYA (*center* and *right*)

AUSTRALOPITHECUS ANAMENSIS **(*right*) lived roughly four million years (Myr) ago. Only a few *anamensis* fossils have been found—the ones shown at the left include a jawbone and part of the front of the face (*left*), parts of an arm bone (*center*) and fragments of a lower leg bone (*right*)—and thus researchers cannot determine much about the species' physical appearance. But scientists have established that *anamensis* walked upright, making it the earliest bipedal creature yet to be discovered.**

3.6 Myr ago by three *Australopithecus* individuals as they walked across wet volcanic ash.) Both *A. africanus* and *A. robustus* were relatively small-brained and had canine teeth that differed from those of modern apes in that they hardly projected past the rest of the tooth row. The younger of the two species, *A. robustus,* had bizarre adaptations for chewing—huge molar and premolar teeth combined with bony crests on the skull where powerful chewing muscles would have been attached.

Paleoanthropologists identified more species of *Australopithecus* over the next several decades. In 1959 Mary Leakey unearthed a skull from yet another East African species closely related to *robustus*. Skulls of these species uncovered during the past 40 years in the northeastern part of Africa, in Ethiopia and Kenya, differed considerably from those found in South Africa; as a result, researchers think that two separate *robustus*-like species—a northern one and a southern one—existed.

In 1978 Donald C. Johanson, now at the Institute of Human Origins in Berkeley, Calif., along with his colleagues, identified still another species of *Australopithecus*. Johanson and his team had been studying a small number of hominid bones and teeth discovered at Laetoli, as well as a large and very important collection of specimens from the Hadar region of Ethiopia (including the famous "Lucy" skeleton). The group named the new species *afarensis*. Radiometric dating revealed that the species had lived between 3.6 and 2.9 Myr ago, making it the oldest *Australopithecus* known at the time.

This early species is probably the best studied of all the *Australopithecus* recognized so far, and it is certainly the one that has generated the most controversy over the past 20 years. The debates have ranged over many issues: whether the *afarensis* fossils were truly distinct from the *africanus* fossils from South Africa; whether there was one or several species at Hadar; whether the Tanzanian and Ethiopian fossils were of the same species; whether the fossils had been dated correctly.

But the most divisive debate concerns the issue of how extensively the bipedal *afarensis* climbed in trees. Fossils of *afarensis* include various bone and joint structures typical of tree climbers. Some scientists argue that such characteristics indicate that these hominids must have spent at least some time in the trees. But others view these features as simply evolutionary baggage, left over from arboreal ancestors. Underlying this discussion is the question of where *Australopithecus* lived—in forests or on the open savanna.

By the beginning of the 1990s, researchers knew a fair amount about the various species of *Australopithecus* and how each had adapted to its environmental niche. A description of any one of the species would mention that the creatures were bipedal and that they had ape-size brains and large, thickly enameled teeth in strong jaws, with nonprojecting canines. Males were typically larger than females, and individuals grew and matured rapidly. But the origins of *Australopithecus* were only hinted at, because the gap between the earliest well-known species in the group (*afarensis,* from about 3.6 Myr ago) and the postulated time of the last common ancestor of chimpanzees and humans (between 5 and 6 Myr ago) was still very great. Fossil hunters had unearthed only a few older fragments of bone, tooth and jaw from the intervening 1.5 million years to indicate the anatomy and course of evolution of the very earliest hominids.

Filling the Gap

Discoveries in Kenya over the past several years have filled in some of the missing interval between 3.5 and 5 Myr ago. Beginning in 1982, expeditions run by the National Museums of Kenya to the Lake Turkana basin in

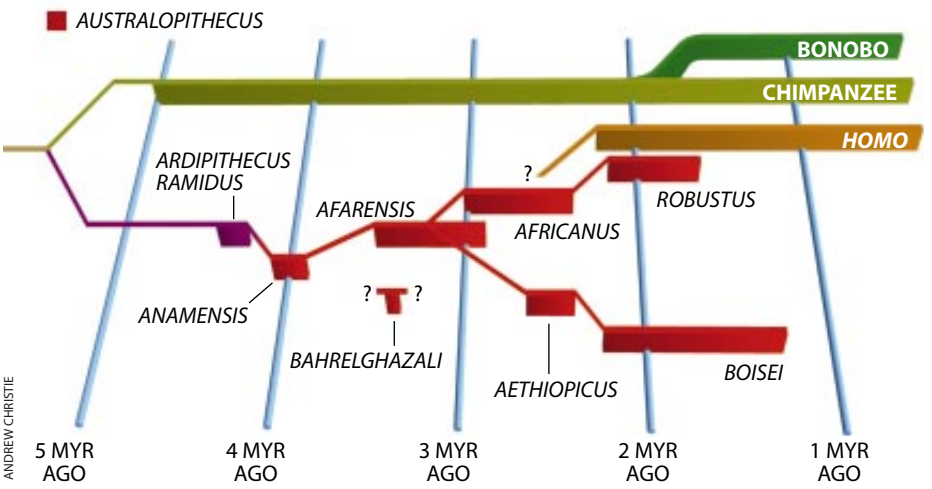
northern Kenya began finding hominid fossils nearly 4 Myr old. But because these fossils were mainly isolated teeth—no jawbones or skulls were preserved—very little could be said about them except that they resembled the remains of *afarensis* from Laetoli. But our recent excavations at an unusual site, just inland from Allia Bay on the east side of Lake Turkana [*see maps on page 78*], yielded more complete fossils.

The site at Allia Bay is a bone bed, where millions of fragments of weathered tooth and bone from a wide variety of animals, including hominids, spill out of the hillside. Exposed at the top of the hill lies a layer of hardened volcanic ash called the Moiti Tuff, which has been dated radiometrically to just over 3.9 Myr old. The fossil fragments lie several meters below the tuff, indicating that the remains are older than the tuff. We do not yet understand fully why so many fossils are concentrated in this spot, but we can be certain that they were deposited by the precursor of the present-day Omo River.

Today the Omo drains the Ethiopian highlands located to the north, emptying into Lake Turkana, which has no outlet. But this has not always been so. Our colleagues Frank Brown of the University of Utah and Craig Feibel of Rutgers University have shown that the ancient Omo River dominated the Turkana area for much of the Pliocene (roughly 5.3 to 1.6 Myr ago) and the early Pleistocene (1.6 to 0.7 Myr ago). Only infrequently was a lake present in the area at all. Instead, for most of the past four million years, an extensive river system flowed across the broad floodplain, proceeding to the Indian Ocean without dumping its sediments into a lake.

The Allia Bay fossils are located in one of the channels of this ancient river system. Most of the fossils collected from Allia Bay are rolled and weathered bones and teeth of aquatic animals—fish, crocodiles, hippopotamuses and the like—that were damaged during transport down the river from some distance away. But some of the fossils are much better preserved; these come from the animals that lived on or near the riverbanks. Among these creatures are several different species of leaf-eating monkeys, related to modern colobus monkeys, as well as antelopes whose living relatives favor closely wooded areas. Reasonably well preserved hominid fossils can also be found here, suggesting that, at least occasionally, early hominids inhabited a riparian habitat.

Where do these *Australopithecus* fossils fit in the evolutionary history of hominids? The jaws and teeth from Allia Bay, as well as a nearly complete radius (the outside bone of the forearm) from the nearby sediments of Sibilot just to the north, show an interesting mixture of characteristics. Some of the traits are primitive ones—that is, they are ancestral features thought to be present before the split occurred between the chimpanzee and human lineages. Yet these bones also share characteristics seen in later hominids and are therefore said to have more advanced features. As our team continues to unearth more bones and teeth at Allia Bay, these new fossils add to our knowledge of the wide range of traits present in early hominids.



ANDREW CHRISTIE

FAMILY TREE of the hominid species known as *Australopithecus* includes a number of species that lived between roughly 4 and 1.25 Myr ago. Just over 2 Myr ago a new genus, *Homo* (which includes our own species, *Homo sapiens*), evolved from one of the species of *Australopithecus*.

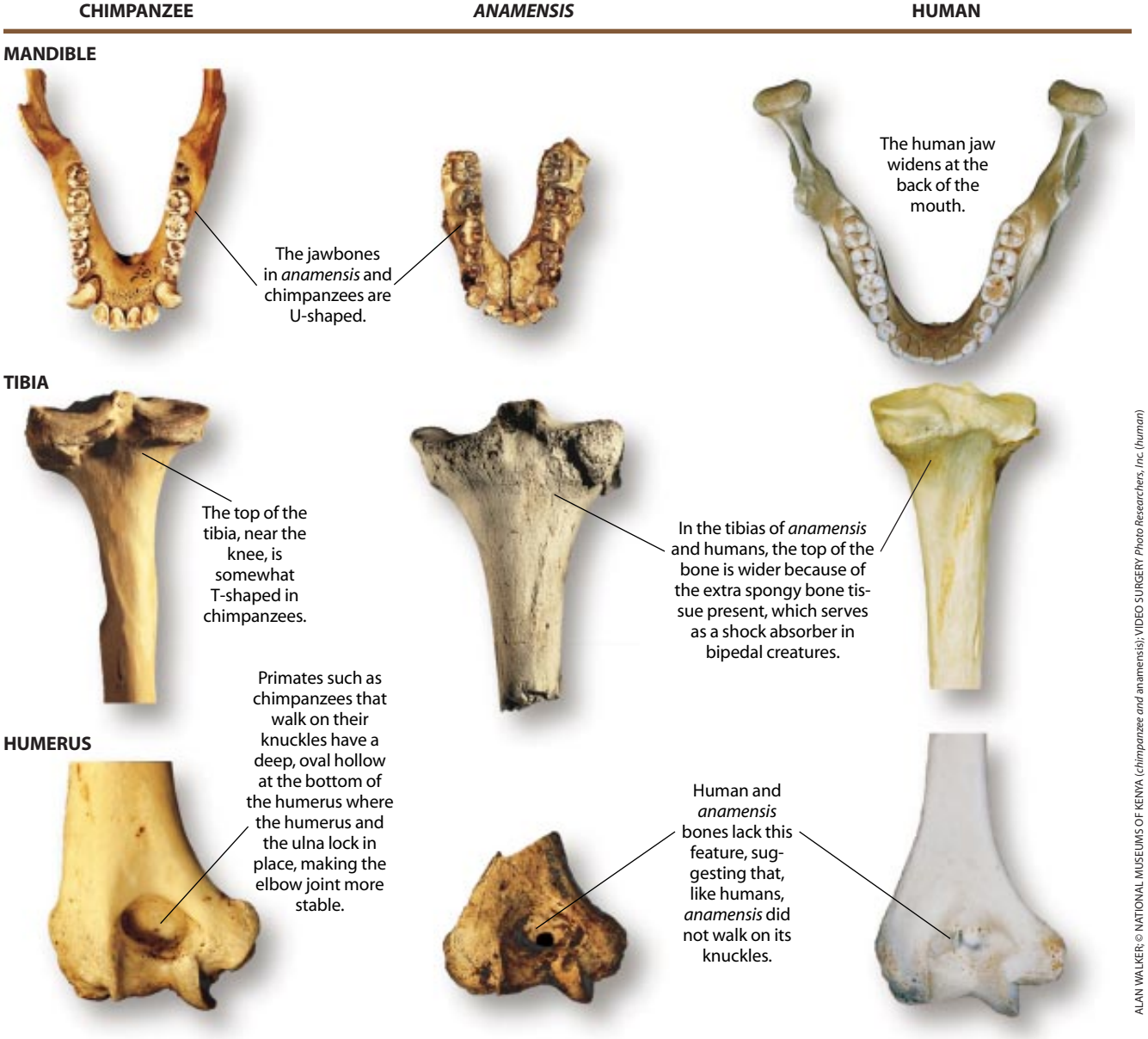
Return to Kanapoi

Across Lake Turkana, some 145 kilometers (about 90 miles) south of Allia Bay, lies the site of Kanapoi, where our story began. One of us (Leakey) has mounted expeditions from the National Museums of Kenya to explore the sediments located southwest of Lake Turkana and to document the faunas present during the earliest stages of the basin's history. Kanapoi, virtually unexplored since Patterson's day, has proved to be one of the most rewarding sites in the Turkana region.

A series of deep erosion gullies, known as badlands, has exposed the sediments at Kanapoi. Fossil hunting is difficult here, though, because of a carapace of lava pebbles and gravel that makes it hard to spot small bones and teeth. Studies of the layers of sediment, also carried out by Feibel, reveal that the fossils here have been preserved by deposits from a river ancestral to the present-day Kerio River, which once flowed into the Turkana basin and emptied into an ancient lake we call Lonyumun. This lake reached its maximum size about 4.1 Myr ago and thereafter shrank as it filled with sediments.

Excavations at Kanapoi have primarily yielded the remains of carnivore meals, so the fossils are rather fragmentary. But workers at the site have also recovered two nearly complete lower jaws, one complete upper jaw and lower face, the upper and lower thirds of a tibia (the larger bone of the lower leg), bits of skull and several sets of isolated teeth. After careful study of the fossils from both Allia Bay and Kanapoi—including Patterson's fragment of an arm bone—we felt that in details of anatomy, these specimens were different enough from previously known hominids to warrant designating a new species. So in 1995, in collaboration with both Feibel and Ian McDougall of the Australian National University, we named this new species *Australopithecus anamensis*, drawing on the Turkana word for lake (*anam*) to refer to both the present and ancient lakes.

To establish the age of these fossils, we relied on the extensive efforts of Brown, Feibel and McDougall, who have been investigating the paleogeographic history of the entire lake basin. If their study



FOSSILS from *anamensis* (*center*) share features in common with both humans (*right*) and modern chimpanzees (*left*). Scientists use the similarities and differences among these species to determine their interrelationships and thereby piece together the course of hominid evolution since the lineages of chimpanzees and humans split some five or six million years ago.

of the basin's development is correct, the *anamensis* fossils should be between 4.2 and 3.9 Myr old. Currently McDougall is working to determine the age of the so-called Kanapoi Tuff—the layer of volcanic ash that covers most of the fossils at this site. We expect that once McDougall successfully ascertains the age of the tuff, we will be confident in both the age of the fossils and Brown's and Feibel's understanding of the history of the lake basin.

A major question in paleoanthropology today is how the anatomic mosaic of the early hominids evolved. By comparing the nearly contemporaneous Allia Bay and Kanapoi collections of *anamensis,* we can piece together a fairly accurate picture of certain aspects of the species, even though we have not yet uncovered a complete skull.

The jaws of *anamensis* are primitive—the sides sit close together and parallel to each other (as in modern apes), rather than widening at the back of the mouth (as in later hominids, including humans). In its lower jaw, *anamensis* is also chimplike in terms of the shape of the region where the left and right sides of the jaw meet (technically known as the mandibular symphysis).

Teeth from *anamensis,* however, appear more advanced. The enamel is relatively thick, as it is in all other species of *Australopithecus;* in contrast, the tooth enamel of African great apes is much thinner. The thickened enamel suggests *anamensis* had already adapted to a changed diet—possibly much harder food—even though its jaws and some skull features were still very apelike. We also know that *anamensis* had only a tiny external ear canal. In this regard, it is more like chimpanzees and unlike all later hominids, including humans, which have large external ear canals. (The size of the external canal is unrelated to the size of the fleshy ear.)

The most informative bone of all the ones we have uncovered from this new hominid is the nearly complete tibia—

ROBERT I.M. CAMPBELL, *National Geographic Image Collection*

DESERT CONDITIONS now prevail in the region near Lake Turkana. When *anamensis* lived here some 4 Myr ago, this region was for a time covered by the ancient Lake Lonyumun, and lush forests would have bordered the rivers that fed into the lake.

the larger of the two bones in the lower leg. The tibia is revealing because of its important role in weight bearing: the tibia of a biped is distinctly different from the tibia of an animal that walks on all four legs. In size and practically all details of the knee and ankle joints, the tibia found at Kanapoi closely resembles the one from the fully bipedal *afarensis* found at Hadar, even though the latter specimen is nearly a million years younger.

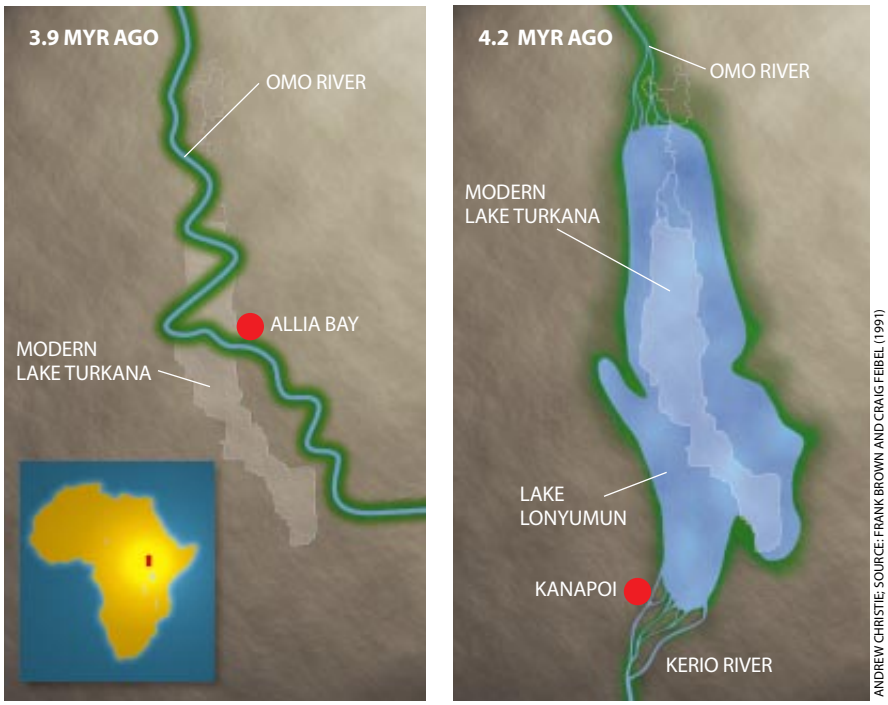
Fossils of other animals collected at Kanapoi point to a somewhat different paleoecological scenario from the setting across the lake at Allia Bay. The channels of the river that laid down the sediments at Kanapoi were probably lined with narrow stretches of forest that grew close to the riverbanks in otherwise open country. Researchers have recovered the remains of the same spiral-horned antelope found at Allia Bay that very likely lived in dense thickets. But open-country antelopes and hartebeest appear to have lived at Kanapoi as well, suggesting that more open savanna prevailed away from the rivers. These results offer equivocal evidence regarding the preferred habitat of *anamensis*: we know that bushland was present at both sites that have yielded fossils of the species, but there are clear signs of more diverse habitats at Kanapoi.

An Even Older Hominid?

At about the same time that we were finding new hominids at Allia Bay and Kanapoi, a team led by our colleague Tim D. White of the University of California at Berkeley discovered fossil hominids in Ethiopia that are even older than *anamensis*. In 1992 and 1993 White led an expedition to the Middle Awash area of Ethiopia, where his team uncovered hominid fossils at a site known as Aramis. The group's finds include isolated teeth, a piece of a baby's mandible (the lower jaw), fragments from an adult's skull and some arm bones, all of which have been dated to around 4.4 Myr ago. In 1994, together with his colleagues Berhane Asfaw of the Paleoanthropology Laboratory in Addis Ababa and Gen Suwa of the University of Tokyo, White gave these fossils a new name: *Australopithecus ramidus*. In 1995 the group renamed the fossils, moving them to a new genus, *Ardipithecus*. Other fossils buried near the hominids, such as seeds and the bones of forest monkeys and antelopes, strongly imply that these hominids, too, lived in a closed-canopy woodland.

This new species represents the most primitive hominid known—a link between the African apes and *Australopithecus*. Many of the *Ardipithecus ramidus* fossils display similarities to the anatomy of the modern African great apes, such as thin dental enamel and strongly built arm bones. In other features, though—such as the opening at the base of the skull, technically known as the foramen magnum, through which the spinal cord connects to the brain—the fossils resemble later hominids.

Describing early hominids as either



ANDREW CHRISTIE; SOURCE: FRANK BROWN AND CRAIG FEIBEL (1991)

TURKANA BASIN was home to *anamensis* roughly 4 Myr ago. Around 3.9 Myr ago a river sprawled across the basin (*left*). The fossil site Allia Bay sat within the strip of forest (*green*) that lined this river. Some 4.2 Myr ago a large lake filled the basin (*right*); a second site, Kanapoi, was located on a river delta that fed into the lake.

primitive or more advanced is a complex issue. Scientists now have almost decisive molecular evidence that humans and chimpanzees once had a common ancestor and that this lineage had previously split from gorillas. This is why we often use the two living species of chimpanzee (*Pan troglodytes* and *P. paniscus*) to illustrate ancestral traits. But we must remember that since their last common ancestor with humans, chimpanzees have had exactly the same amount of time to evolve as humans have. Determining which features were present in the last common ancestor of humans and chimpanzees is not easy.

But *Ardipithecus,* with its numerous chimplike features, appears to have taken the human fossil record back close to the time of the chimp-human split. More recently, White and his group have found parts of a single *Ardipithecus* skeleton in the Middle Awash region. As White and his team extract these exciting new fossils from the enclosing stone, reconstruct them and prepare them for study, the paleoanthropological community eagerly anticipates the publication of the group's analysis of these astonishing finds.

But even pending White's results, new *Australopithecus* fossil discoveries are offering other surprises, particularly about where these creatures lived. In 1995 a team lead by Michel Brunet of the University of Poitiers announced the identification in Chad of *Australopithecus* fossils believed to be about 3.5 Myr old. The new fossils are very fragmentary—only the front part of a lower jaw and an isolated tooth. In 1996, however, Brunet and his colleagues designated a new species for their specimen: *A. bahrelghazali.* Surprisingly, these fossils were recovered far from either eastern or southern Africa, the only areas where *Australopithecus* had been found until now. The site, in the Bahr el Ghazal region of Chad, lies 2,500 kilometers west of the western part of the Rift Valley, thus extending the range of *Australopithecus* well into the center of Africa.

The *bahrelghazali* fossils debunk a hypothesis about human evolution postulated in the pages of *Scientific American* by Yves Coppens of the College of France [see "East Side Story: The Origin of Humankind," May 1994]; ironically, Coppens is now a member of Brunet's team. Coppens's article proposed that the formation of Africa's Rift Valley subdivided a single ancient species, isolating the ancestors of hominids on the east side from the ancestors of modern apes on the west side. In general, scientists believe such geographical isolation can foster the development of new species by prohibiting continued interbreeding among the original populations. But the new Chad fossils show that early hominids did live west of the Rift Valley. The geographical separation of apes and hominids previously apparent in the fossil record may be more the result of accidental circumstances of geology and discovery than the species' actual ranges.

KENNETH GARRETT *National Geographic Image Collection*

FOSSIL HUNTER Alan Walker (*foreground*) and two colleagues excavate the bone bed at Allia Bay, where several *anamensis* fossils have been recovered. The bone bed appears as a dark band about 18 inches thick at the top of the trench.

The fossils of *anamensis* that we have identified should also provide some answers in the long-standing debate over whether early *Australopithecus* species lived in wooded areas or on the open savanna. The outcome of this discussion has important implications: for many years, paleoanthropologists have accepted that upright-walking behavior originated on the savanna, where it most likely provided benefits such as keeping the hot sun off the back or freeing hands for carrying food. Yet our evidence suggests that the earliest bipedal hominid known to date lived at least part of the time in wooded areas. The discoveries of the past several years represent a remarkable spurt in the sometimes painfully slow process of uncovering human evolutionary past. But clearly there is still much more to learn. SA

The Authors

MEAVE LEAKEY and ALAN WALKER, together with Leakey's husband, Richard, have collaborated for many years on the discovery and analysis of early hominid fossils from Kenya. Leakey is head of the division of paleontology at the National Museums of Kenya in Nairobi. Walker is Distinguished Professor of anthropology and biology at Pennsylvania State University. He is a MacArthur Fellow and a member of the American Academy of Arts and Sciences.

Further Reading

AUSTRALOPITHECUS RAMIDUS, A NEW SPECIES OF EARLY HOMINID FROM ARAMIS, ETHIOPIA. Tim D. White, Gen Suwa and Berhane Asfaw in *Nature,* Vol. 371, pages 306–312; September 22, 1994.

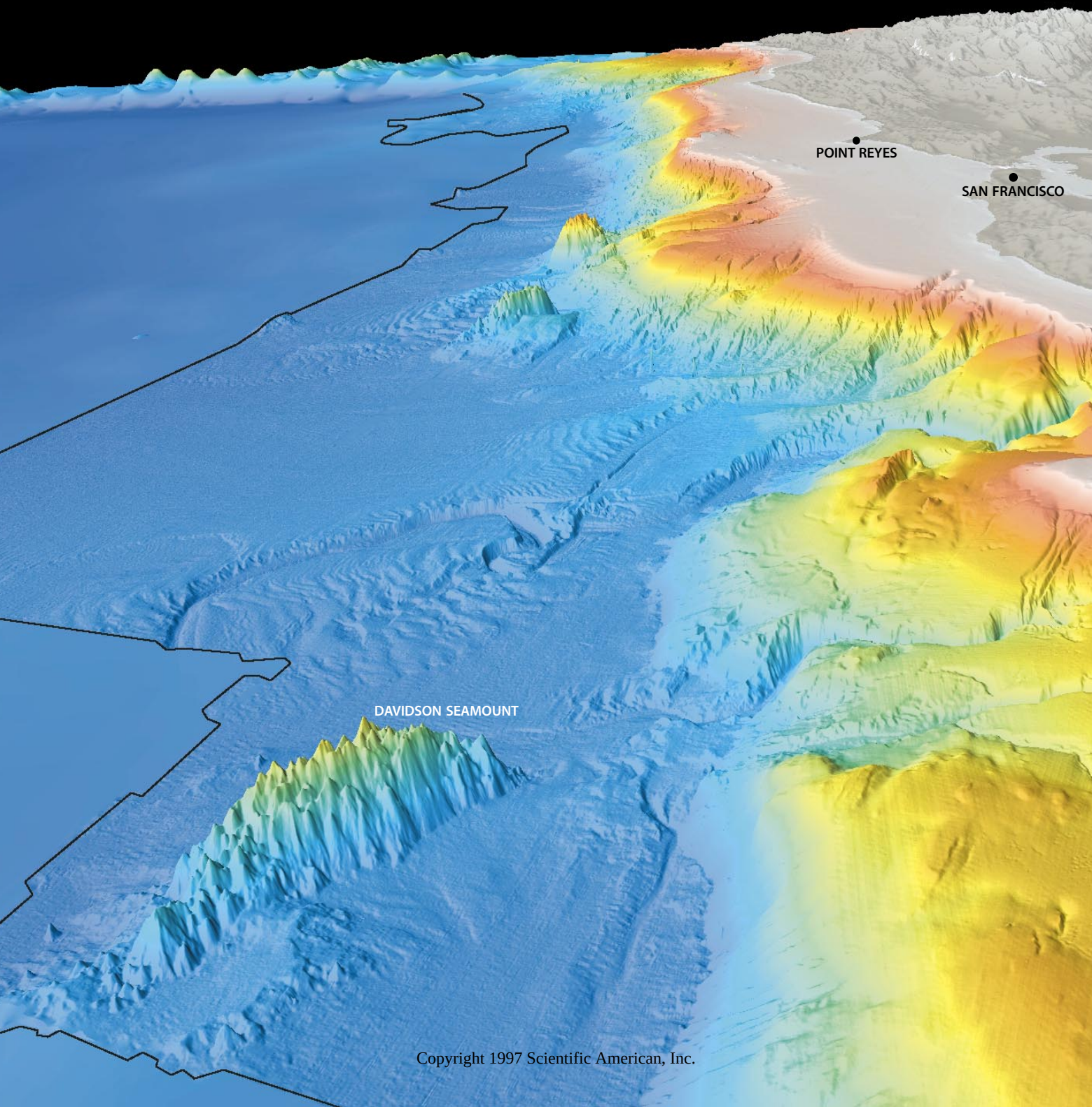
NEW FOUR-MILLION-YEAR-OLD HOMINID SPECIES FROM KANAPOI AND ALLIA BAY, KENYA. Meave G. Leakey, Craig S. Feibel, Ian McDougall and Alan Walker in *Nature,* Vol. 376, pages 565–571; August 17, 1995.

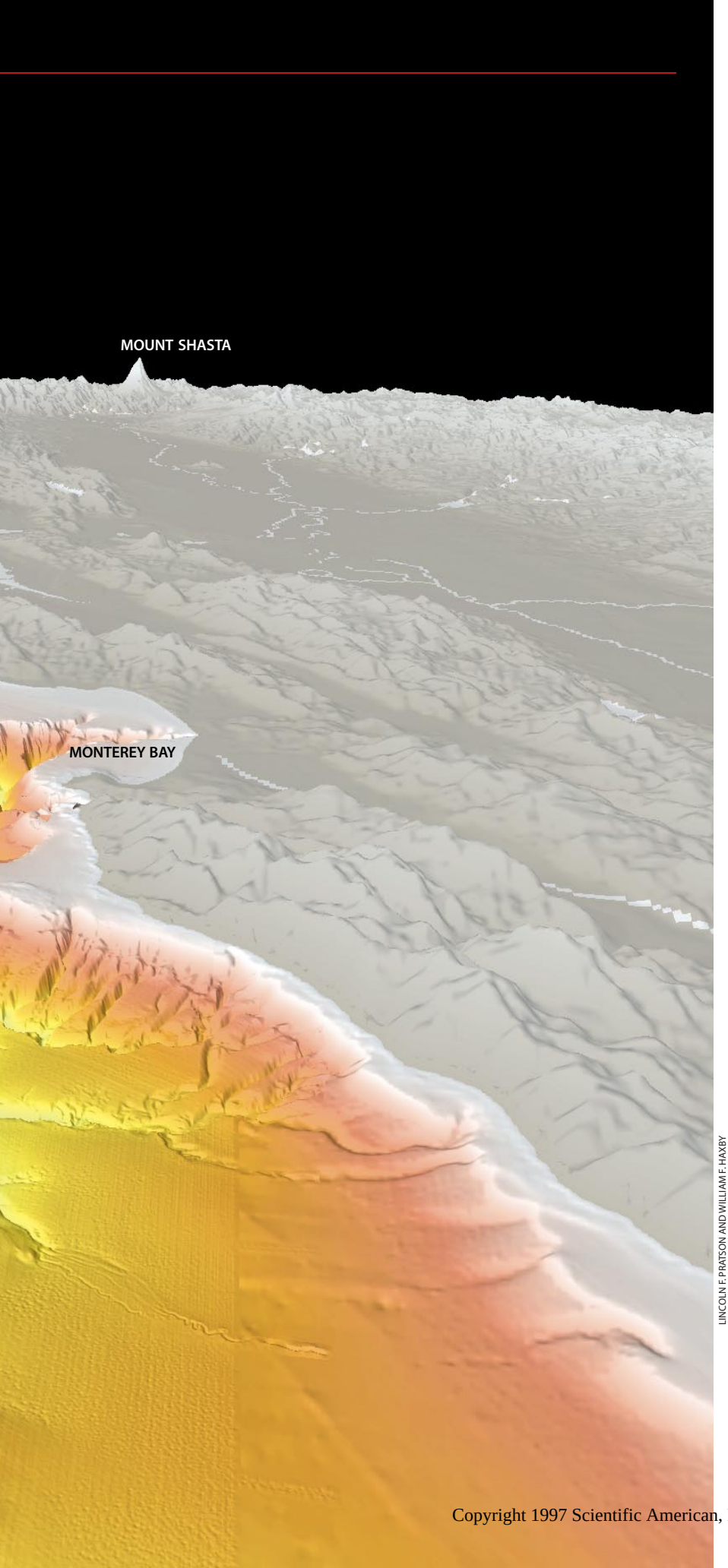
FROM LUCY TO LANGUAGE. Donald C. Johanson and Blake Edgar. Peter Nevraumont, Simon & Schuster, 1996.

RECONSTRUCTING HUMAN ORIGINS: A MODERN SYNTHESIS. Glenn C. Conroy. W. W. Norton, 1997.

Panoramas of the Seafloor

Modern sonar techniques map the continental margins of the U.S. and reveal the richly varied scenery usually hidden underwater





LINCOLN F. PRATSON AND WILLIAM F. HAXBY

by Lincoln F. Pratson
and William F. Haxby

In 85 B.C. or thereabouts, a Greek named Posidonius set sail on a curious mission. He was not carrying freight or passengers, nor was he engaged in war. He simply wanted to answer an age-old question: How deep is the ocean? Halting his vessel in the middle of the Mediterranean Sea, Posidonius coaxed his ship's crew to let out nearly two kilometers of rope before a large stone attached to the end of the line finally hit bottom. He and his men must have been jubilant—at least until they realized that they then had to haul the great weight back on board.

For the next 2,000 years, naval surveyors and oceanographers continued to use exactly the same laborious line-and-sinker method to probe the ocean's depths. It is not surprising that they made scant progress. Then, during the 1920s, oceanographers developed the first echo sounders—instruments that could measure the water's depth by bouncing sound waves off the bottom. With the wealth of measurements these devices provided, scientists got their first glimmers of the true shape of the ocean basins.

In the past few decades engineers have constructed ever more sophisticated acoustic devices to speed the mapping of this hitherto hidden part of the earth. The major impetus for these developments initially came from concerns about national defense, but more recently economic considerations have taken precedence.

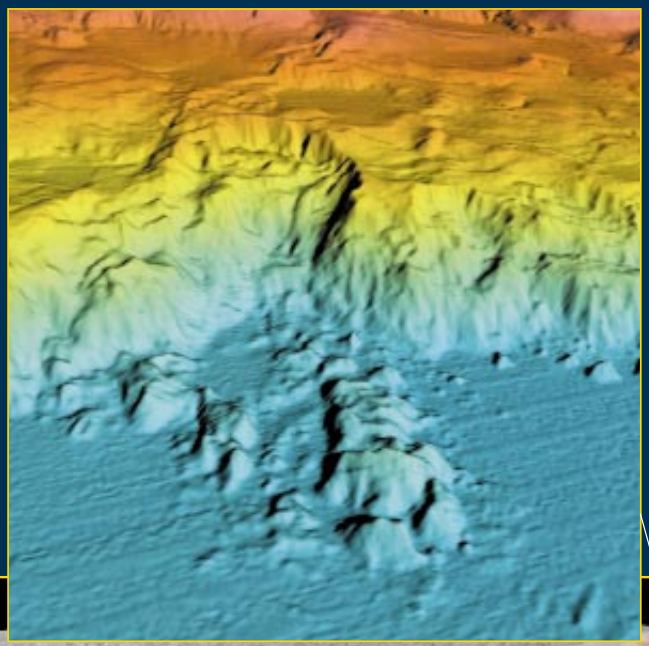
Beginning with the U.S. in 1981, the world's maritime nations declared the waters and seafloor within 200 miles of their shores to be "Exclusive Economic Zones." To help assess the value of the vast undersea expanse claimed by the U.S., the National Oceanic and Atmospheric Administration began surveying parts of the newly annexed area in 1983. That effort (which continued until 1993) mapped more than 200,000 square kilometers of the seafloor off the coasts of the Atlantic and Pacific oceans and the Gulf of Mexico.

Over this same period, the National Science Foundation funded two smaller sur-

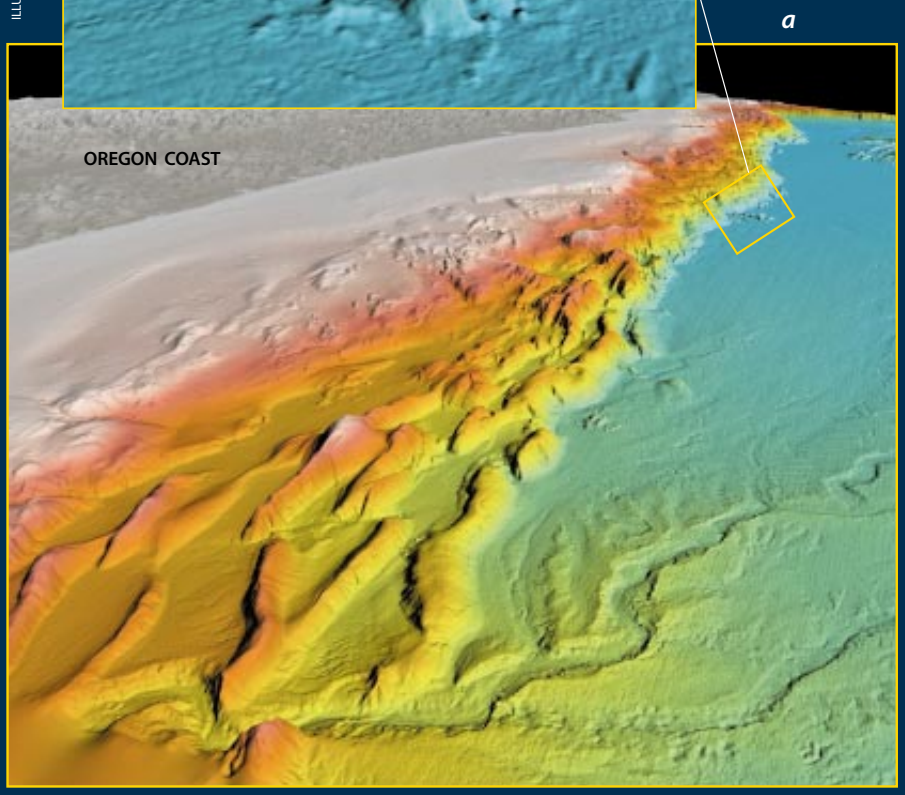
Continued on page 86

COMPUTER-GENERATED IMAGES of the seafloor surrounding the U.S. show geologic features in great detail in regions where specialized sonar mapping has been done (*right of black line*).

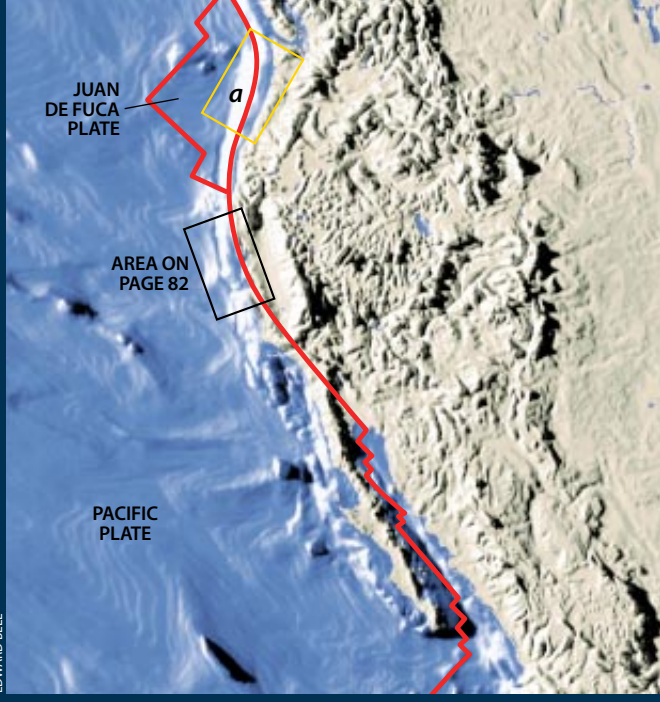
SEAFLOOR FAILURE carved a pocket in the continental slope offshore from central Oregon. The trail of debris extending outward from the crescent-shaped embayment in the continental slope marks the path the material traveled. The excavated pocket is six kilometers wide—about the width of the island of Manhattan. Some of the dislodged blocks that now rest on the continental rise are as tall as a modern skyscraper. The collapse that sent these huge chunks tumbling down the slope was most likely triggered by an earthquake. Such dramatic failures of the continental slope can generate violent tsunamis that may inundate the coast nearby or travel across the Pacific and create havoc on distant shores.



ILLUSTRATIONS BY LINCOLN F. PRATSON AND WILLIAM F. HAXBY; EXCEPT AS NOTED



FOLDED CARPET of sediments covers the seafloor offshore of Oregon. The undulations result from the head-on collision between the North American and Juan de Fuca plates. Like a colossal bulldozer, the North American Plate scrapes sediments off the down-going Juan de Fuca Plate and piles them into folds. To the north (*lower left*), the folds of sediment form distinct ridges. To the south (*upper right*), where part of the Juan de Fuca Plate breaks through its sedimentary cover, the folds are stacked so closely that they form terraces.

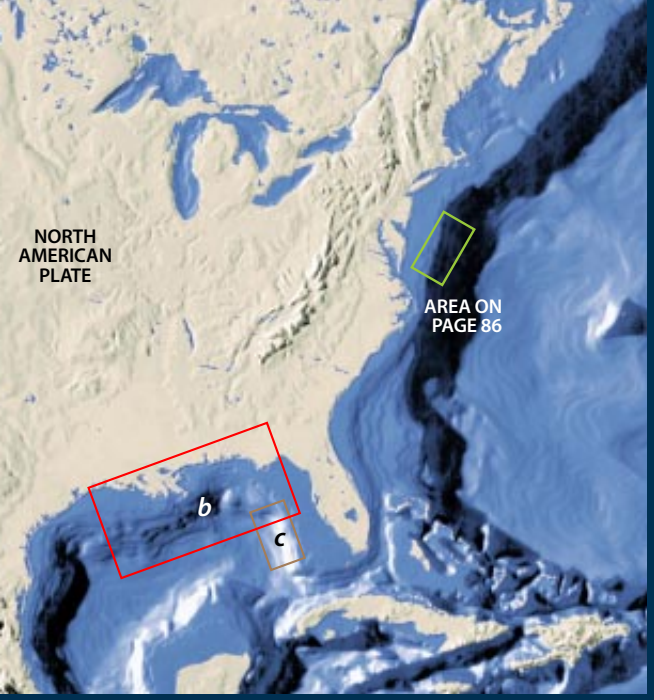


EDWARD BELL

THREE TYPES OF MARGINS exist along the borders of North America. The sides facing the Atlantic Ocean and the Gulf of Mexico are termed passive margins, because they are embedded within the North American Plate and simply ride along with this great slab as it moves. The western margin of the U.S., on the other hand, lies along the leading edge of the North American Plate, where it bumps and grinds its way past oceanic crust underlying

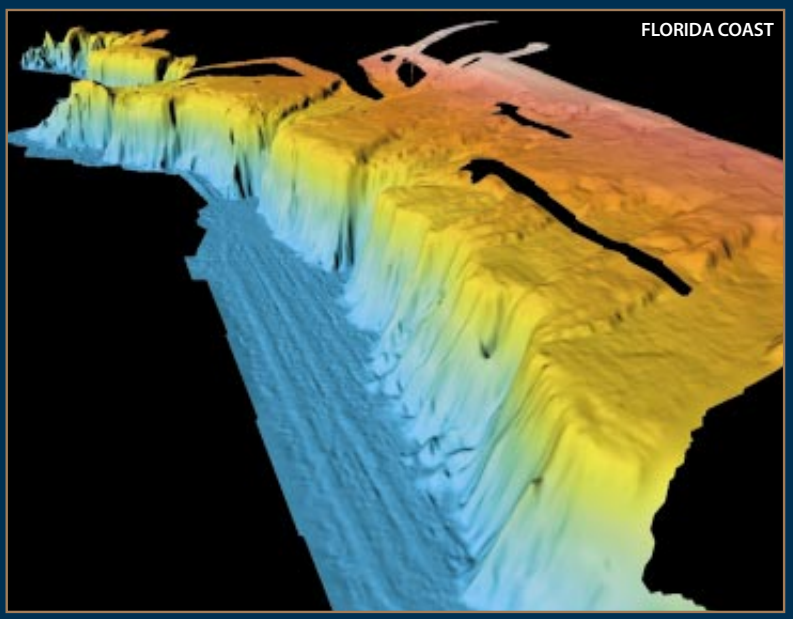
b

CRATERLIKE FEATURES pockmark the seafloor offshore from Mississippi to eastern Texas. These small basins are filled with thick accumulations of sediment—and, in some spots, billions of barrels of oil and gas. The ridges and domes between the basins hide shallowly buried salt bodies of various sizes and shapes. This salt formed initially by the evaporation of Gulf of Mexico waters (some 180 million years ago). The salt layer was then buried by a massive load of sediment eroded from the Rocky Mountains and carried into the gulf by the Mississippi River. Because salt resists compression, it will flow rather than compact. Hence, under the weight of the overlying sediment, giant blebs of the salt have bubbled upward and spread out toward the open sea.

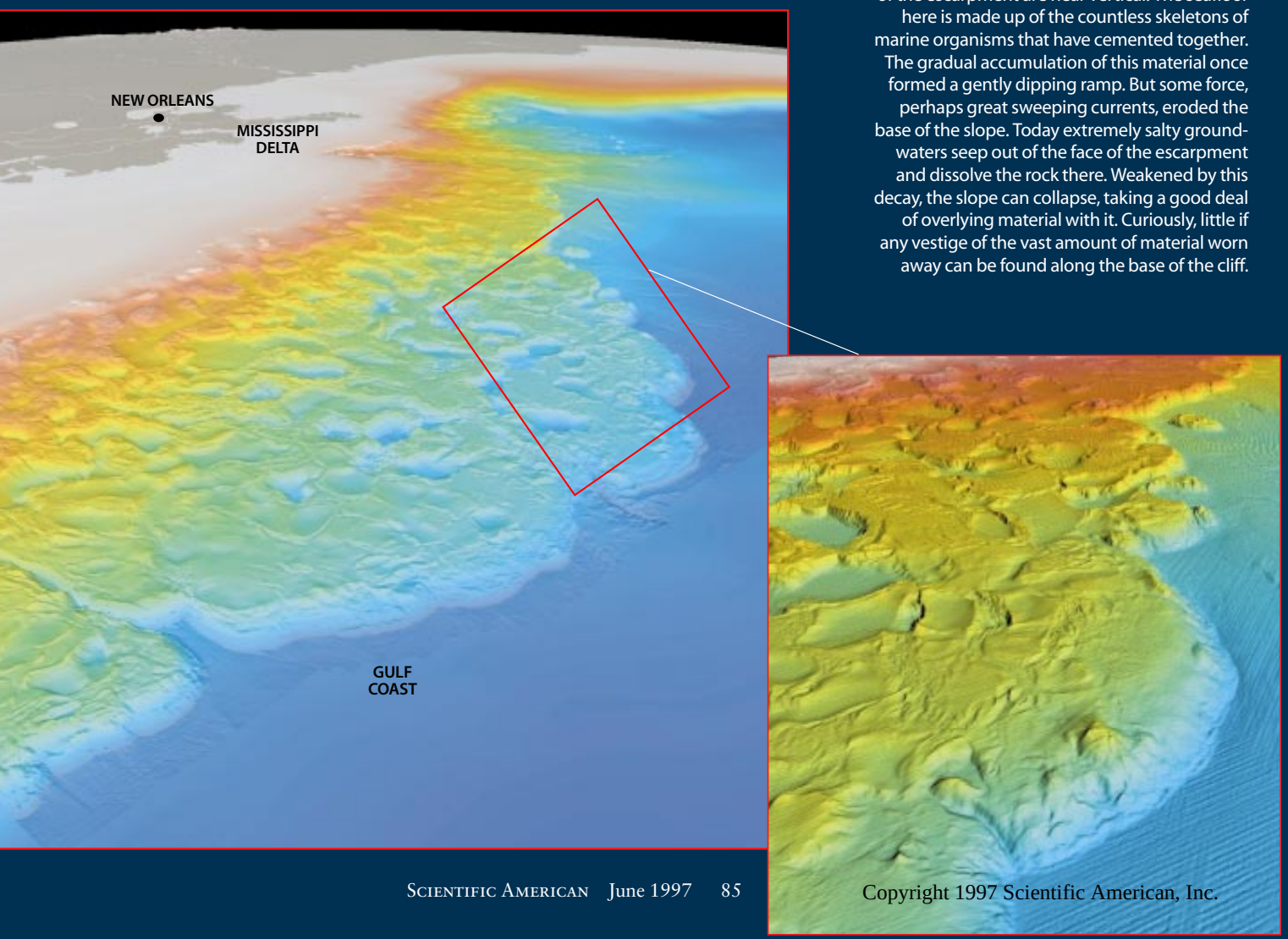


the Pacific Ocean. This collision takes two forms. Through most of the length of California (a strike-slip margin), the North American Plate slips sideways past the Pacific Plate along a system of near-vertical fractures in the earth collectively known as the San Andreas Fault. Farther up the coast (a convergent margin), the North American Plate is bulldozing its way over a sliver of oceanic crust named the Juan de Fuca Plate.

c



MILE-HIGH CLIFF marks the edge of the continental slope west of Florida. This undersea precipice, known as the Florida Escarpment, stands more than four times as high as the Empire State Building. Whereas the tilt of the continental slope elsewhere is typically just a few degrees, the face of the escarpment is, on average, slanted at 35 degrees. In many places the walls of the escarpment are near vertical. The seafloor here is made up of the countless skeletons of marine organisms that have cemented together. The gradual accumulation of this material once formed a gently dipping ramp. But some force, perhaps great sweeping currents, eroded the base of the slope. Today extremely salty groundwaters seep out of the face of the escarpment and dissolve the rock there. Weakened by this decay, the slope can collapse, taking a good deal of overlying material with it. Curiously, little if any vestige of the vast amount of material worn away can be found along the base of the cliff.



Seafloor Mapping Tools

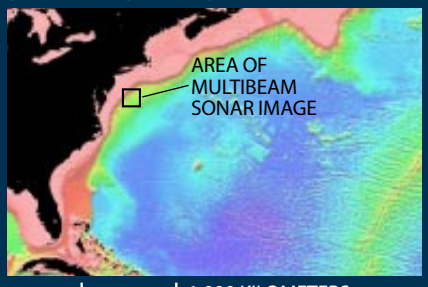
Our multibeam sonar images represent just one way scientists can visualize the seafloor. Other approaches are also used, and each has its peculiar advantages and shortcomings.

Satellites (*a*) cannot measure seafloor depth directly, but they can sense variations in the elevation of the water at the surface of the ocean. The U.S. Navy's Geosat satellite, for example, can measure the distance to the ocean surface to within five centimeters by bouncing radar pulses off the water below it. Because the precise position of the satellite is known, such determinations provide a measure of sea-surface height.

The ocean surface can vary in relief by as much as 200 meters. These undulations reflect minute differences in the earth's gravity from place to place that cause water to distribute itself unevenly. Most commonly, these gravitationally induced variations in the ocean surface are caused by rugged seafloor topography. For instance, a massive, submerged volcano that is 2,000 meters tall and 40 kilometers wide will pull water toward it, producing a bulge about two meters high in the ocean surface above it. But undersea features smaller than 10 kilometers across do not generally possess sufficient mass to affect the ocean surface and thus go undetected by satellite radars. What is more, gravity variations (particularly near continental margins) can reflect differences in the density of the underlying rock rather than topography. Still, satellites provide broad, if less than perfect, maps of regions not yet surveyed with ships.

Multibeam sonar (*b*) bounces sound off the seafloor to gauge ocean depth. In contrast to simple echo sounders, this modern technique employs an array of sound sources and listening devices mounted on the hull of the survey vessel. Every few seconds the sources emit a burst that reaches only a slim strip of seafloor aligned perpendicularly to the direction the ship is moving. At the same time, the listening devices begin recording sounds reflected from the bottom. This equipment is arranged to detect sounds com-

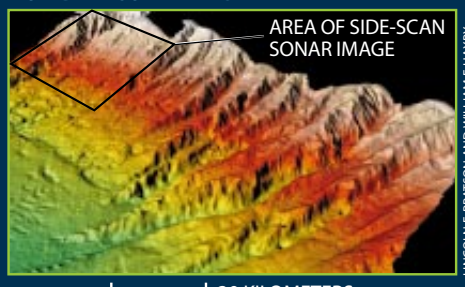
SATELLITE RADAR IMAGE



1,000 KILOMETERS

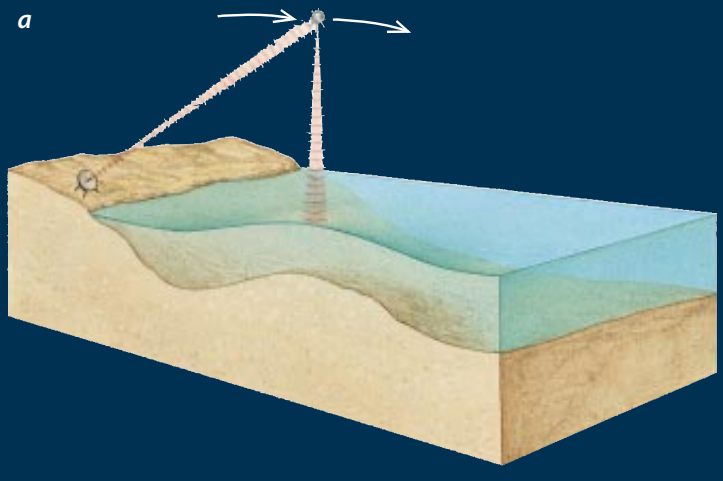
WALTER H. F. SMITH *National Oceanic and Atmospheric Administration*

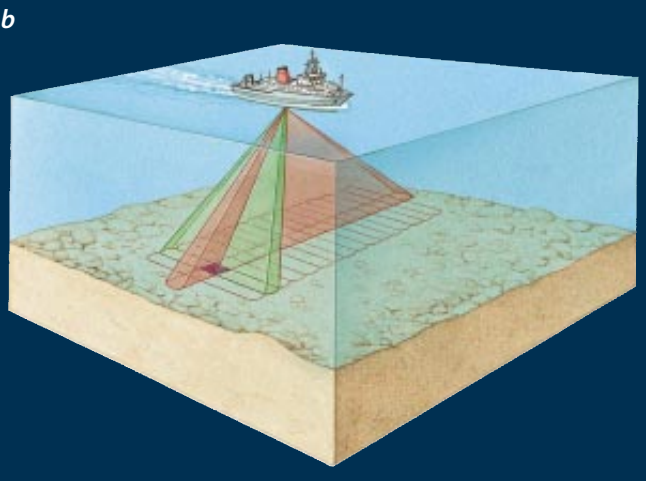
MULTIBEAM SONAR IMAGE



30 KILOMETERS

LINCOLN F. PRATSON AND WILLIAM F. HAXBY





ROBERTO OSTI

Continued from page 83

veys to study parts of the seafloor near the coasts of New Jersey and western Florida. All the vessels involved used multibeam sonars, the most modern form of instrumentation available for measuring the topography of the ocean bottom.

These surveys provide unprecedented views of the country's continental slope. Although no sunlight actually penetrates to these great depths, computers can render images of seafloor vistas as they would appear with the oceans drained. Such a perspective is particularly valuable in planning industrial activities offshore. For example, submarine cables increasingly carry international communications, and petroleum producers are moving drilling platforms into ever greater depths of water. These enterprises require maps of where the seafloor appears to be stable—not prone to subsea avalanches or violent currents. Disposal of waste at sea also demands this information, because currents running along the bottom can disturb the sites where waste settles. Bottom surveys further help geologists to locate offshore fault systems and to assess their risk of triggering earthquakes.

On a broader scientific level, undersea mapping is providing fundamental knowledge about the geologic forces that shape the ocean floor. Images such as those we have created offer scientists a way to take in vast stretches of undersea terrain in a glance—an ability they have long enjoyed while studying the surface of distant moons and planets. That perspective now offers some fascinating new insights into the marvelously complex evolution of the earth. SA

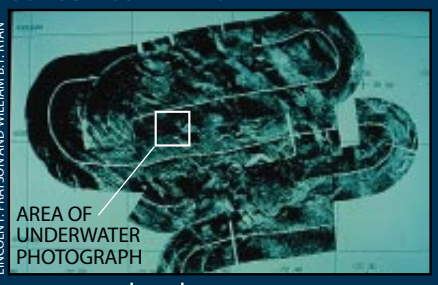
ing only from within a series of narrow seafloor corridors that are aligned parallel to the ship's direction. Thus, the sound reflections received at the ship emanate from the regions where the slim strip of sound and the listening corridors intersect. The timing of these reflections provides a profile of seafloor depth. Such profiles are recorded every few seconds while the survey ship moves over the seafloor, and so successive observations build up a continuous swath of coverage along the ship's track. By running the ship in the same pattern one mows a lawn, scientists can produce a complete map of an area. With less than 200 vessels outfitted with the necessary equipment, however, charting the entire seafloor in this way would require hundreds of years.

Side-scan sonar (*c*) provides yet a different perspective on what the seafloor looks like. The equipment is usually attached to a "sled" that is towed behind a ship. Two sonar units, affixed to either side of the sled, act as both sound sources and listening devices. These units emit bursts of sound outward, to either side. If the seafloor is flat and smooth, none of the energy emitted will be reflected back (as with a beam of light directed obliquely onto a mirror). But if the seafloor is rough, the sound hitting the bottom will be scattered in all directions, and some will return to the sonar sled (just as a beam of light illuminating ground glass will reflect in all directions). By equating the amplitude of the recorded echoes to different shades of gray and displaying the results to show the distance from the sled, scientists can obtain an image of the texture of the seafloor that looks similar to a black-and-white photograph. But like a single aerial photograph, a side-scan sonar image does not indicate the heights of the surface below.

The most accurate and detailed view of the seafloor is provided by underwater photography (*d*), using either cameras towed along the bottom, piloted submersibles or remotely operated vehicles. Such camera-carrying equipment gives researchers the opportunity to explore the seafloor up close. Yet because even the most intense illumination does not penetrate seawater effectively, photographic views obtained in this way are limited to the short distances that artificial beams of light can penetrate.

—L.F.P. and W.F.H.

SIDE-SCAN SONAR IMAGE

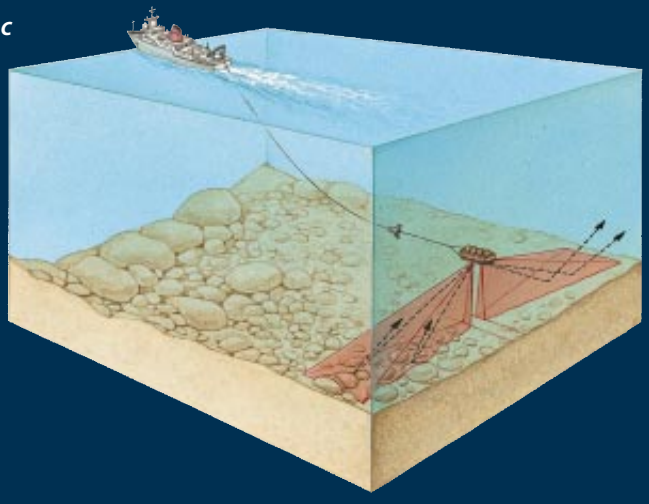


LINCOLN F. PRATSON AND WILLIAM B. F. RYAN

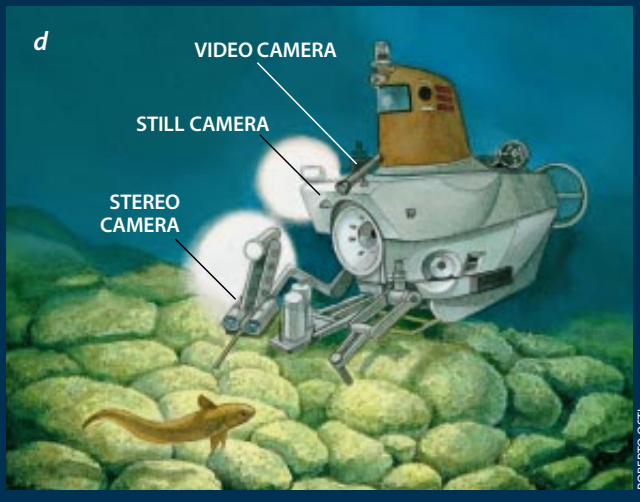
UNDERWATER PHOTOGRAPH

LINCOLN F. PRATSON

c



d



ROBERTO OSTI

The Authors

LINCOLN F. PRATSON and WILLIAM F. HAXBY have worked together for two years probing the continental margins of the U.S. Pratson completed his Ph.D. in geological sciences at Columbia University in 1993. He then studied the topography of the seafloor for the Office of Naval Research at Columbia's Lamont-Doherty Earth Observatory. In 1996 he joined the Institute of Arctic and Alpine Research at the University of Colorado. Haxby earned his doctorate from Cornell University in 1978. Since then, he has conducted investigations of the ocean basins as a research scientist at Lamont-Doherty Earth Observatory.

Further Reading

SWATH BATHYMETRIC MAPPING. Special issue of *Journal of Geophysical Research*, Vol. 19, No. B3; March 10, 1986.

IMAGING THE OCEAN FLOOR: HISTORY AND STATE OF THE ART. Peter R. Vogt and Brian E. Tucholke in *Geology of North America*, Vol. M: *The Western North Atlantic Region*. Edited by P. R. Vogt and B. E. Tucholke. Geological Society of America, 1986.

WHAT IS THE SLOPE OF THE U.S. CONTINENTAL SLOPE? Lincoln F. Pratson and William F. Haxby in *Geology*, Vol. 24, No. 1, pages 3–6; January 1996.

National Geophysical Data Center World Wide Web site available at http://www.ngdc.noaa.gov/mgg/mggd.html



Searching for Digital Pictures

Computers that can reason about images may be able to pick out distinct features of a person, place or object from photograph archives

by David Forsyth, Jitendra Malik and Robert Wilensky

The Internet and the digital libraries being connected to it can provide access to enormous amounts of information. Retrieving that information efficiently is another matter. Consider archives of still images or film clips: if a user wants to find something as simple as a picture of a horse near water, the only hope today would be that someone had properly captioned all such pictures in advance.

Experience from major archives shows that it is almost impossible to anticipate all likely queries. Furthermore, support for a new class of query—for example, "Show me all pictures that contain a horse and grass"—might require reexamining the entire collection.

People retrieve images from large collections in a variety of different contexts and with a wide range of purposes. One person might ask a museum's art collection to find out how often a particular shade of color was used in depicting skin in portraiture or query a database of satellite images to estimate what next year's corn crop would be. Another might browse a picture agency's collection to find the right photograph of a sunset for inclusion in a travel brochure cover or search a set of images of California waterways to compare the impact of recent floods with those of previous years.

Ultimately, fully automated image analysis seems the most desirable way to handle retrieval tasks. A seemingly straightforward means to support such diverse queries would be to employ a program that could discern whether a given image stored in a collection contained a specific object—such as a horse or a body of water. The program would have to identify the desired object correctly, independent of its color, size, appearance or angle of view. Unfortunately, current understanding of how to recognize objects in images falls far short of this task. Researchers are nonetheless addressing the problem from several directions, and a synthesis of approaches from computer vision and automated reasoning may eventually improve the ability to analyze image features.

Computer programs can analyze the content of an image to be retrieved in many ways. Some look for images that nearly match a given sample image. Others rely on general appearance: an image containing large numbers of small




yellow dots (fields of wildflowers) or one with a bright red central region (fires, sunsets or certain sports cars). Yet another alternative is to look for an object with a defined identity, such as water or a horse. These alternatives represent three fundamentally different notions of image retrieval: to find images that are iconically similar to a known exemplar; to analyze images in terms of "stuff"—regions of near-constant color or texture; or to identify "things" the way people do. (See also the description of pattern matching in "Configurable Computing," by John Villasenor and William H. Mangione-Smith, on page 66.)

Easy but Unhelpful

The ease of implementing each of these alternatives seems to be inversely related to its usefulness. Direct matching of one image to another is fairly straightforward, but its utility is limited to finding superficially related images. It is not very useful for finding objects, because changes in the position, composition or configuration of pictured objects will defeat most comparisons: a front view of a horse does not match a side view (for this reason, we believe direct matching against sample images is an algorithmic dead end).

Retrieving images based on "stuff" has greater potential, and many stuff analyses are not computationally daunting—it takes only a fraction of a second to calculate the percentages of red, blue or green pixels in a given image. Because "things" are made of "stuff," a natural first step is to try to retrieve images based on stuff queries alone, and most current image retrieval systems are oriented toward stuff-based queries. For example, perhaps the best-known image database system, Query by Image Content (QBIC), developed by a team of researchers at IBM, allows an operator to specify a desired image in terms of such properties as color, spatial distribution of color and texture (texture ranges from simple alternation of two colors, like a zebra's stripes, to more complicated multihued patterns). The system then displays a ranked selection of potential matches to those criteria. Another well-known system, Photobook, built by Alex Pentland and his colleagues at the Massachusetts Institute of Technology, largely shares QBIC's model of an image as a collage of flat, homogeneous regions but incorporates more sophisticated representations of texture and can automatically divide the picture into segments.

Although such systems are impressive, our belief is that stuff-based queries alone will ultimately be of limited utility. For example, a query about the proportions of colors in an image could not distinguish between the English and the French flags. Broadly speaking, human users are interested in things, and neither stuff nor iconic similarity alone will provide a sufficient foundation for content-based retrieval.

FINDING PICTURES accurately with image retrieval programs is difficult because the problem of object recognition is still poorly understood. To find pictures of purple flowers in a database, a user queried a prototype program for pictures containing small purple dots; some of the results of that search are shown across the top. Not all the pictures met the user's intention, but some do because flowers could be crudely described as purple dots. To find a more complex object, such as a tiger, the user's query must invoke a more sophisticated algorithm, one that searches for large-scale arrangements of color and texture. The results of one search for a tiger are shown across the bottom. The algorithm that retrieved them was much more accurate than others that ignored shape. Both programs were developed at the University of California at Berkeley.



To explore some of these issues, we have built a system for image retrieval as part of the Digital Library Project at the University of California at Berkeley. The system, constructed primarily by Ginger Ogle and Chad Carson, will eventually contain more than 600,000 photographs. About 50,000 images from a number of sources, including ground and aerial photographs, are currently available on-line at the project's World Wide Web site.

Our interface integrates a wide variety of queries, allowing users to search for both things and stuff. It is difficult to find things, however, and so in many cases it is better to ask for pictures that contain the stuff from which the desired things could be assembled. In our view, representations of (and therefore, queries for) objects are obtained by assembling stuff. This means that the most useful stuff queries are those that could be used to construct representations of things. In this view, a user interface should offer a variety of options: a user may query for things, or, if there is no satisfactory thing query, the user may specify various kinds of spatial assemblies of stuff that are known to be useful.

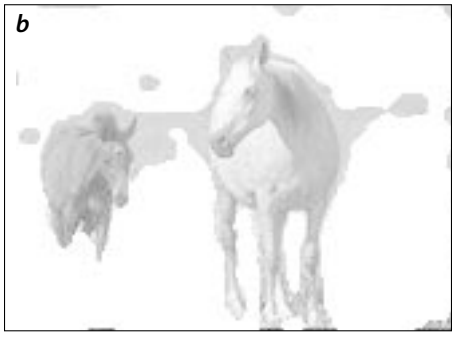
When one of the system's designers constructed a query for images of windsurfing, he asked for pictures with a blue-green content of at least 30 percent representing sky or sea and with at least one small yellow dot for the yellow stripe of a sail. A very high percentage of the images returned were relevant to the query. The system also displayed some irrelevant images, and it failed to bring up some pictures that did have to do with windsurfing. In information retrieval, there is typically a trade-off between accuracy and recall—retrieving only relevant items ensures that some will be missed, and retrieving all relevant items means retrieving some irrelevant ones as well. This example shows that a powerful stuff retrieval mechanism can sometimes provide a passable approximation to the goal of finding images containing objects, but it also illustrates the kind of mental gyrations that a user must perform to convert a search for a particular kind of thing into a query for stuff.

In our present implementation, users can request only a small number of things, such as horses, sailboats and unclad people. To specify stuff, users can give the percentages of different colors they expect in an image, along with the number and size of "color blobs" that should appear. They can also specify other features derived from image content, such as whether an image contains a horizon or not. Finally, they can search so-called metadata—extra information stored with the picture, including captions, the name of the photographer, the date of the picture, and so on.

Finding Horses Piece by Piece

Object-recognition algorithms work by grouping image elements into successively larger, more complex regions. They then generate hypotheses about what those composite regions might represent. A recognizer for four-legged animals such as horses might start by scrutinizing an image (*a*) for approximately horse-colored patches (*b*). It could define contours for those regions (*c*), then look for regions that might be the right shape for a horse's body (the body and legs are approximately cylinders) (*d*). The program could next weed out those regions with impossible spatial relations (body and legs must

Original photograph



Horse-shaded stuff

Perils of Object Recognition

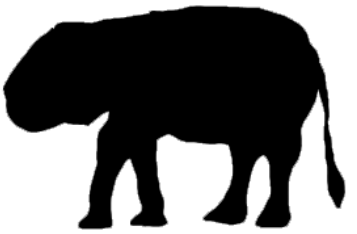
Given the right background knowledge, however, the computer may be able to deduce from the stuff in an image what things are present. The problem of identifying horses, people or anything else in an image is termed object recognition and has been a major focus of the field of computer vision for several decades. Object recognition has received a great deal of study in academia and elsewhere because of its industrial and military applications. Yet current techniques work only when images contain small numbers of objects whose shapes are precisely known and that can be viewed from a limited range of angles. They are not effective at recognizing even a single "ordinary" object, such as a person or a horse. Ordinary objects display a wide variation in height, weight or markings that confound computer vision programs while making little difference to the object's identity: a person is recognizably a person, for example, regardless of height, weight or dress.

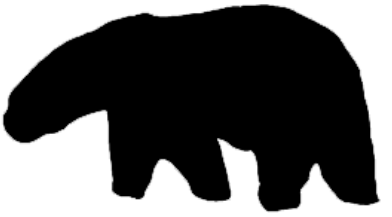
Furthermore, we have to separate objects from their backgrounds. This is a challenging problem even for human observers in some cases. Consider a leopard in the dappled light of a jungle. In order to attempt recognition, one needs to know which regions of the image belong together—the spots on a leopard, the leaves on a tree. In other words, the computer program must be able to group stuff together into things.

This so-called perceptual grouping problem has been much studied in the context of human vision by researchers in the Gestalt school of psychology. They have pointed out several factors that could be used to determine when parts of an image most likely arise from a single object in the scene. Similarity of color or texture is very powerful: humans readily form groups from parts of an image that are uniform in color, such as a red patch, or uniform in visual texture, such as a plaid region. On a more complex level, regions in a two-dimensional image that are symmetric about an axis can be grouped as projections of a symmetric three-dimensional object, such as a vase.

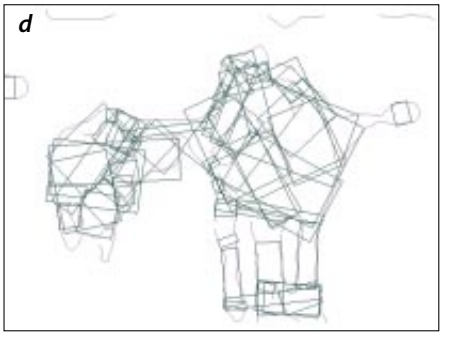
Computer vision researchers have been struggling for many years to turn such rules of thumb into working algorithms. Serge Belongie, Chad Carson, Hayit Greenspan and one of us (Malik) have developed a system that, though woefully inadequate compared with hu-

be perpendicular) (*e*). Final classifications can sometimes be made only with detailed knowledge of the color and surface texture of objects; even a person might be hard-pressed to distinguish on the basis of shape alone the silhouettes of an elephant with a coiled trunk (*left*) and a bear (*right*). —*D.F., J.M. and R.W.*





Boundaries of horse-shaded stuff



All possible body segments

DAVID FORSYTH AND MARGARET FLECK

Segments configured like a horse's body

man perceptual grouping capabilities, provides a useful decomposition of the image into a small set of regions, coherent in color and texture. Each coherent "blob" is described by attributes that represent its position, shape, color and texture. We can think of the blobs as a representation that captures the basic compositional features of the image.

We can now find images containing airplanes against the sky by looking for a gray blob inside a blue blob. It is possible to find pictures of tigers by looking for a blob with color and texture similar to what we expect in tiger skin and a blob with color and texture similar to the grass background of typical tiger habitat. The user can choose how specific he or she wants to be about the specification of the color, texture, shape or position attributes for each blob.

Interestingly, this representation of objects as spatial configurations of regions of particular color and texture lends itself well to machine-learning techniques. Instead of relying on a human to think explicitly of the most salient aspects of a scene containing tigers or airplanes, statistical learning can be used. We have presented software known as a statistical classifier with example pictures of specific visual categories (such as airplane or tiger scenes) and trained it to recognize previously unseen pictures as being instances of these categories. At the moment, our classifiers can categorize scenes by looking only at information about color and texture; with appropriate information about region shapes, they should be able to distinguish between collections of regions that form objects and those that do not. Learning provides a natural framework for coping with irrelevant variations within a class of objects, because a classifier can adjust itself so that it does not pay attention to unimportant deviations (sky can be many different shades of blue, clouds can have endlessly different shapes, and so on).

Where's Waldo?

Takeo Kanade and his colleagues at Carnegie Mellon University have developed a face-recognition software module that provides a good example of the application of learning techniques. Other researchers have found it very difficult to detect individual eyes, noses and mouths accurately; Kanade and his co-workers have instead trained a neural network to detect simultaneously the presence of all these features in an appropriate configuration. They have also been studying the combined use of video and audio data for analyzing video and film clips. A computer might perform speech recognition on the audio track in a television news story, for example, and associate the mention of a celebrity's name with the face seen in the video data at the same time.

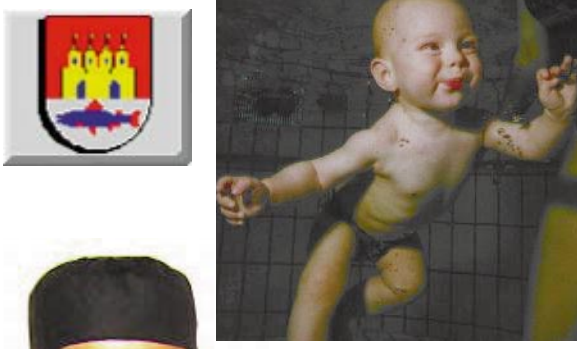
Color and texture can help find image regions. Several additional cues help to solve the difficulties in assembling the image regions that correspond to things. First, many objects are made of parts that have simple three-dimensional shapes, and the relations between these shapes are often also quite simple. Furthermore, such basic three-dimensional shapes usually appear as simply shaped image regions—for example, a cylinder almost always appears as a region with straight, nearly parallel sides. As a result, a program can identify such parts fairly easily.

Similarly, if there are constraints on the geometric relations between parts—as in the way that the range of motion of joints in people and animals limits the possible relations among body segments—it is often easy to tell whether image regions that appear to correspond to two parts of an object could in fact do so. In short, each image region generates hypotheses about its identity; these hypotheses then suggest new grouping strategies to identify a larger and more distinctive group of image regions. Statistical learning theory yields mechanisms by which one can determine which hypotheses should be accepted and which rejected. Once a sufficiently large group has been assembled in this way, the object is considered identified.

Margaret Fleck of the University of

In Your Face Algorithms

Rather than try to define explicitly the features that make up a face, researchers at Carnegie Mellon University trained a neural-network program with a large, labeled collection of faces and nonface images. Eventually the network learned to pick out eyes, nose and mouth. A World Wide Web interface allows all comers to submit images for face detection. A small selection of the several hundred pictures that were thought by the program to contain faces is shown at the right. *—D.F., J.M. and R.W.*

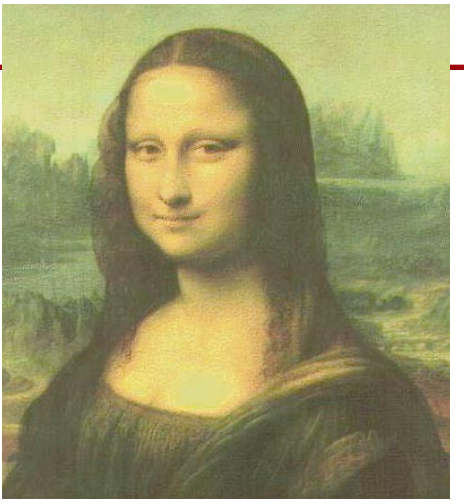


Iowa and one of us (Forsyth) have used these observations to build two systems that can recognize particular objects in a large collection of images. The first system finds images that contain people. Our present version works only for images that contain people wearing little or no clothing; the colors and textures of skin are surprisingly limited compared with those of garments. The program first segments the image into regions corresponding to skin (which can be done fairly accurately) and discards images that contain little or no skin. Once it has assembled these regions into groups that look approximately like cylinders (representing body segments), it tests the geometric relations between these groups to find collections that could have come from a limb. Finally, it tests relations between possible limbs to come up with regions that could represent more complex assemblies.

Because the model of the person that we use emphasizes relations between body segments, it is possible to modify it to find other bodylike objects. Our second system finds pictures of horses by testing relations among segments of hidelike pixels. We have tested the person finder on a total of nearly 5,000 images from extremely diverse sources; it finds people in 40 percent of those images that contain them and in only 4 percent of those that do not. The horse finder recognizes only 10 percent of the horse pictures it is given, but it marks only 0.4 percent of nonhorse pictures. We do not yet

DISTINGUISHING the same or different objects viewed from various angles can challenge object-recognition programs. Software developed by one of the authors can sometimes tell whether horses are in a picture regardless of background or pose.






Clockwise from top left: HARRI PULLI; KARINA MOELLER; KLAUS-PETER ZAUNER; PAUL QUALTROUGH; DA VINCI; KLAUS-PETER ZAUNER; SARITA J. BROWN; MICHAEL LOCKE; CARNEGIE MELLON UNIVERSITY; HIROYUKI KONISHI KONISHI

know how to perform such reasoning efficiently in scenes that contain many different kinds of objects to be recognized.

This limited performance, as far as we know, represents the state of the art of thing-based image retrieval from large collections. The performance of such an automated system should not be compared with the results that a perfect search process would produce, because there is no perfect search process. Even the best automated searches of thoroughly indexed material seldom retrieve more than 50 percent of relevant items without including a large number of irrelevant ones as well. And in cases where humans go through a collection of images manually, we have found that searchers may also miss a significant portion of relevant images—not to mention the prohibitive cost.

Although we have focused on the problem of image retrieval from large collections, it should be evident that the same issues pertain to other tasks involving images, such as filtering information from the Web. Digitized images and films are becoming common, both as large collections and as the content of more transient media, such as electronic mail messages or as Web pages.

We have made significant progress by simply assembling regions of coherent stuff according to hypotheses based on texture and spatial arrangement, but new techniques will be needed to bring retrieval rates up to the levels needed for useful archives. Ultimately, it should be possible to obtain a picture of a pouncing leopard by simply filling in a form ("spotty," "feline body plan," "leaping"), submitting it to a search engine and waiting for a selection of appropriate pictures to come back. This goal is achievable. It is worthwhile both because many people need to find things in collections of images and because in achieving it we will solve deep problems in understanding vision. SA

The Authors

DAVID FORSYTH, JITENDRA MALIK and ROBERT WILENSKY are faculty members in the computer science division and collaborators in the Digital Library Project at the University of California, Berkeley. Forsyth, a computer vision specialist, earned bachelor's and master's degrees in electrical engineering at the University of the Witwatersrand and a doctorate from the University of Oxford. He has been at Berkeley since 1994. Malik works on computer vision and computational modeling of human vision. He has been at Berkeley since 1986 and earned a bachelor's degree from the Indian Institute of Technology in Kanpur and a doctorate from Stanford University. Wilensky is director of the Digital Library Project and chair of the computer science division at Berkeley; he has worked in many areas of artificial intelligence and received both his B.A. and Ph.D. from Yale University.

Further Reading

DESCRIPTION AND RECOGNITION OF CURVED OBJECTS. R. Nevatia and T. O. Binford in *Artificial Intelligence,* Vol. 8, pages 77–98; 1977.

COLOR: WHY THE WORLD ISN'T GREY. Hazel Rossotti. Princeton University Press, 1983.

THE LEGACY OF GESTALT PSYCHOLOGY. I. Rock and S. Palmer in *Scientific American,* Vol. 263, No. 6, pages 84–90; December 1990.

QUERY BY IMAGE AND VIDEO CONTENT—THE QBIC SYSTEM. M. Flickner et al. in *Computer*, Vol. 28, No. 9, pages 23–32; September 1995.

OBJECT REPRESENTATION IN COMPUTER VISION—II. J. Ponce, A. Zisserman and M. Hebert. Springer LNCS, No. 1144, 1996.

PHOTOBOOK—CONTENT BASED MANIPULATION OF IMAGE DATABASES. A. Pentland, R. W. Picard and S. Sclaroff in *International Journal of Computer Vision,* Vol. 18, No. 3, pages 233–254; June 1996.

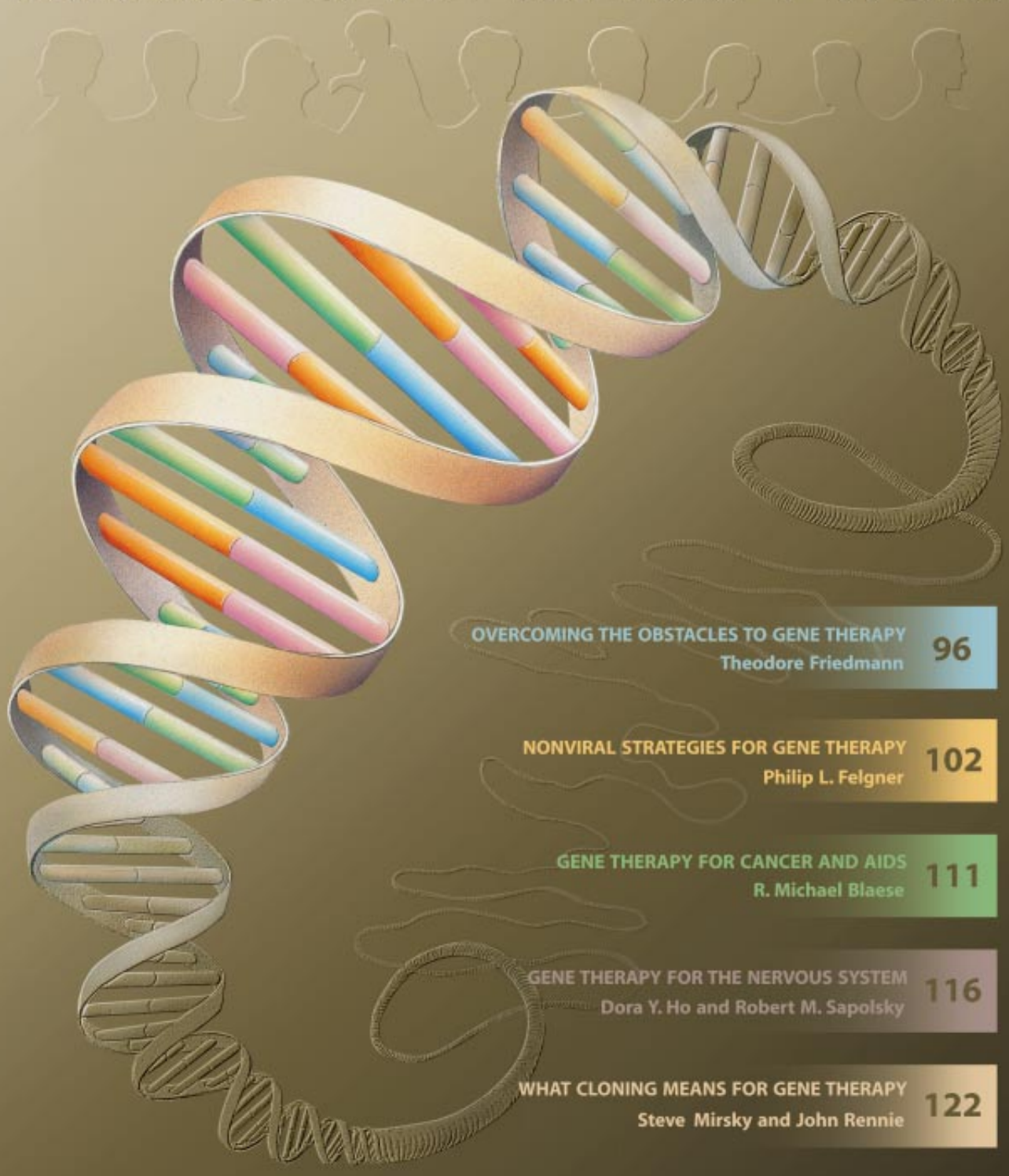
OBJECT-BASED NAVIGATION: AN INTUITIVE NAVIGATION STYLE FOR CONTENT-ORIENTED INTEGRATION ENVIRONMENT. K. Hirata et al. Proceedings of the ACM Hypertext Conference, Southampton, U.K., April 1997.

University of California at Berkeley Digital Library Project World Wide Web site available at http://elib.cs.berkeley.edu

SCIENTIFIC AMERICAN

— SPECIAL REPORT —

MAKING GENE THERAPY WORK



TOMO NARASHIMA AND LAURIE GRACE

Overcoming the Obstacles

Treating disease by providing needed genes remains a compelling idea, but clinical and basic researchers still have much to do before gene therapy can live up to its promise

by Theodore Friedmann

In the late 19th century, when the pioneering architect Daniel H. Burnham was planning some of the first modern skyscrapers, his associates were skeptical about erecting buildings that soared into the clouds. Burnham reportedly warned the skeptics against making "little plans," having "no magic to stir men's blood." He urged them to reach beyond traditional architectural boundaries, to think once inconceivable thoughts and to perform previously unimagined deeds—the hallmarks of revolutions.

Revolutionary changes have also occurred in medicine over the past few centuries. Witness the new understandings and practices that issued from the introduction of microscopy, anesthesia, vaccination, antibiotics and transplantation. Medicine is now preparing to undergo another epochal shift: to an era in which genes will be delivered routinely to cure or alleviate an array of inherited and acquired diseases.

Preparing for a radical change, yes, but not yet in the midst of it. By emphasizing hopes and downplaying uncertainties, some overzealous researchers, representatives of industry and members of the lay and scientific media have implied that gene therapy is already advanced enough for widespread application. It is not.

Arguably, the conceptual part of the gene therapy revolution has indeed occurred. Whenever a new gene is discovered, researchers and nonscientists immediately ask whether it can be used to treat some disorder, even when more traditional approaches might be applied. But the technical part of the revolution—the ability to correct disease—is another story. Investigators have accomplished the requisite first steps: they have shown that transferred genes can be induced to function in the human body, at times for several years. So far, however, no approach has definitively improved the health of a single one of the more than 2,000 patients who have enrolled in gene therapy trials worldwide.

This lack of a convincing therapeutic benefit is sobering. Yet it would be a mistake to doubt gene therapy's powerful future. Remember, the field is young; in the U.S., trials in patients have been carried out for fewer than 10 years. A more realistic interpretation of the unspectacular clinical results thus far is that they reflect researchers' imperfect initial gropings toward a difficult new technology and that the obstacles are more formidable than many of us had expected.

A central challenge, as a federally commissioned critique of the gene-therapy research effort noted in 1995, is perfecting methods for delivering therapeutic genes to cells. Often genes introduced into patients do not reach enough of the appropriate cells or, for reasons that are not always clear, function poorly or shut off after a time. Under those conditions, a gene that could potentially be helpful would have little chance of affecting a disease process.

In this article I will outline some of the most pressing technological stumbling blocks to successful gene transfer and the strategies being considered to cope with those difficulties. I will deal only with therapy affecting somatic cells, the kinds that are neither sperm nor egg. To date, research aimed at human gene therapy has avoided manipulations that would deliberately affect descendants of the treated individuals, perhaps in unintended ways. The need for enlightened public debate over the merits and risks of germ-line therapy has, however, been made more urgent by the recent cloning of an adult sheep [see "What Cloning Means for Gene Therapy," by Steve Mirsky and John Rennie, on page 122].

How Genes and Gene Therapy Work

Anyone who wants to understand the obstacles to gene therapy should first know a bit about what genes do and about how attempts at gene therapy are currently carried out. An individual gene in the human cell is a stretch of DNA that, in most cases, acts as a blueprint for making a specific protein; it spells out the sequence of amino acids composing that protein. All cells in a body carry the same genes in the chromosomes of the nucleus. But neurons, say, behave unlike liver cells because different cells use, or express, distinct subsets of genes and hence make separate sets of proteins (the main functionaries of cells). Put more precisely, each cell copies only selected genes into individual molecules of messenger RNA, which then serve as the templates from which proteins are constructed.

If a particular gene is mutated, its protein product may not be made at all or may work poorly or even too aggressively. In any case, the flaw may disturb vital functions of cells and tissues that use the normal gene product and can thereby cause symptoms of disease.

to Gene Therapy

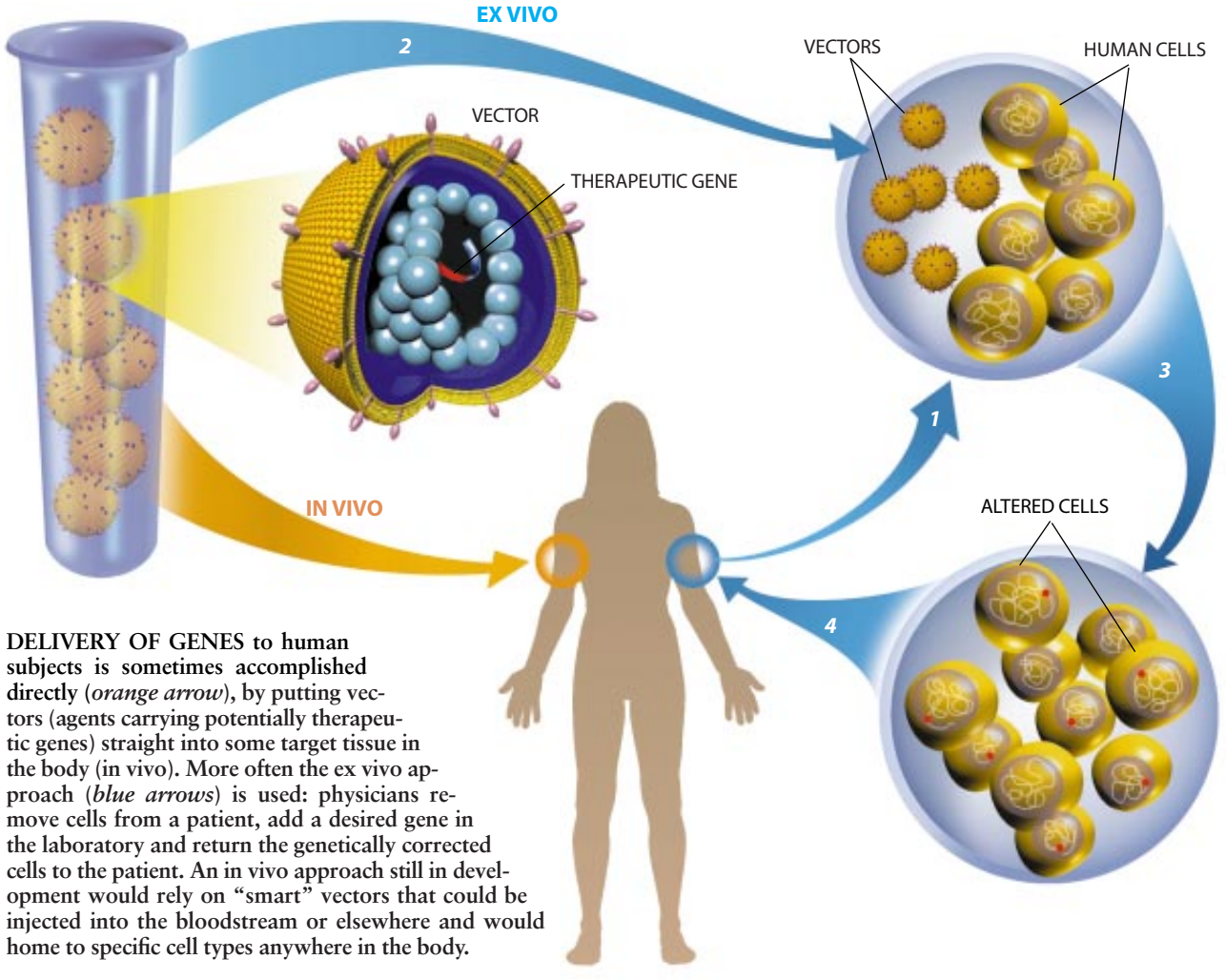
Historically, physicians have treated disorders stemming from inherited genetic mutations not by altering genes but by intervening in the biological events resulting from a mutation. For example, dietary restriction has long been prescribed for phenylketonuria, in which loss of a gene leads to the toxic buildup of the metabolic products of the amino acid phenylalanine. Unfortunately, nongenetic manipulations are usually only partly effective against inherited ills.

In the early 1970s this fact—combined with growing understanding of how genes function and with discovery of the genes underlying many inherited ills—led to the suggestion that better results might be achieved by attacking inborn diseases at their source. Among the genetic diseases that have been studied are cystic fibrosis (which mainly affects the lungs), muscular dystrophy, adenosine deaminase deficiency (which severely impairs immunity), and familial hypercholesterolemia (which leads to the early onset of severe atherosclerosis).

Surprisingly, as time went on, it became clear that even acquired maladies often have a genetic component that can theoretically be a target of a genetic correction strategy. Indeed, quite unexpectedly, more than half of all clinical trials for gene therapy these days aim at cancer, which in most cases is not inherited but results from genetic damage accumulated after birth [see "Gene Therapy for Cancer," by R. Michael Blaese, on page 111]. A number of trials also focus on AIDS, which is caused by the human immunodeficiency virus (HIV).

In principle, a normal gene can be delivered so that it physically takes the place of a flawed version on a chromosome. In practice, such targeted insertion of a gene into a chromosome is not yet achievable in people; fortunately, it often is not required. Most attempts at gene therapy simply add a useful gene into a selected cell type to compensate for a missing or ineffective version or to instill some entirely new property. Many proposed anticancer gene therapies under study take this last tack: they aim to induce cancer cells to make substances that will kill those cells directly, elicit a potent attack by the immune system or eliminate the blood supply that tumors require for growth.

Some gene therapy groups are also devising strategies to compensate for genetic mutations that result in destructive proteins. In one approach, called antisense therapy, short stretches of syn-



DELIVERY OF GENES to human subjects is sometimes accomplished directly (*orange arrow*), by putting vectors (agents carrying potentially therapeutic genes) straight into some target tissue in the body (in vivo). More often the ex vivo approach (*blue arrows*) is used: physicians remove cells from a patient, add a desired gene in the laboratory and return the genetically corrected cells to the patient. An in vivo approach still in development would rely on "smart" vectors that could be injected into the bloodstream or elsewhere and would home to specific cell types anywhere in the body.

thetic DNA act on messenger RNA transcripts of mutant genes, preventing the transcripts from being translated into abnormal proteins. Related tactics deploy small RNA molecules called ribozymes to degrade messenger RNA copied from aberrant genes. A rather different plan provides a gene for a protein, called an intracellular antibody, that can block the activity of the mutant protein itself. Some therapeutic strategies rely on the design of hybrids of DNA and RNA that might direct the repair of mutant genes.

Genes are currently provided to patients in two basic ways. In both cases, the genes are usually first put into transporters, or vectors, able to deposit foreign genes into cells. In the more common method, scientists remove cells from a selected tissue in a patient, expose them to gene-transfer vectors in the laboratory (ex vivo) and then return the genetically corrected cells to the individual. Other times researchers introduce the vectors directly into the body (in vivo), generally into the tissue to be treated. Our ultimate goal, of course, is to deliver vectors into the bloodstream or other sites and to have them act like homing pigeons, finding their own way to the desired cells—say, to organs that are hard to reach or to hidden cancer deposits. No such targeted carriers are yet ready for testing in patients, but work toward that end is advancing quickly.

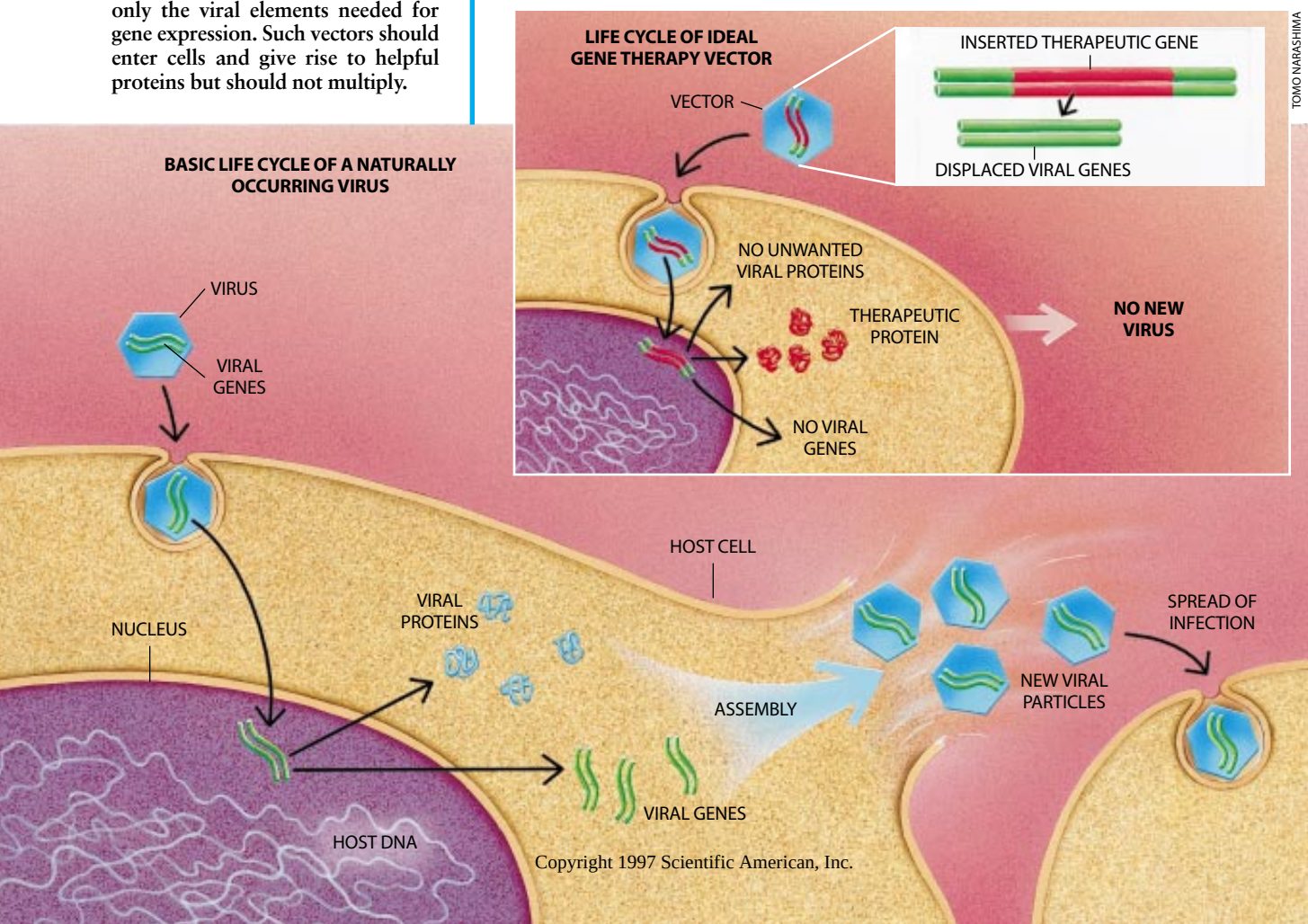
In the body, certain genes will be helpful only if their expression is regulated tightly: in other words, they must give rise to just the right amount of protein at the right times. Biologists have yet to achieve such precise control over foreign genes put into the body. For many gene therapy applications, however, exquisite regulation will not be essential. Nor will it always be necessary to put genes into the cells that are in need of fixing. Sometimes more accessible cell types (say, muscle or skin) might be turned into protein factories; these factories would release proteins needed by nearby cells or might secrete proteins into the bloodstream for transport to distant sites.

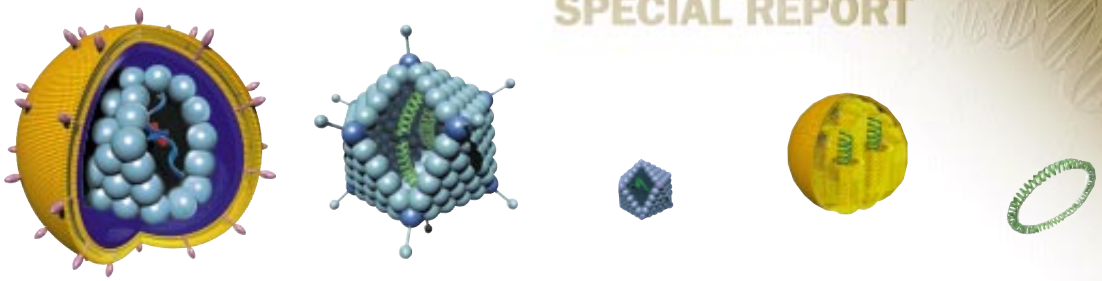
Retrovirus Vectors: Flaws and Fixes

The key to success for any gene therapy strategy is having a vector able to serve as a safe and efficient gene delivery vehicle. From the start, viruses—which are little more than self-replicating genes wrapped in protein coats—have drawn the most attention as potential vectors. They are attractive because evolution has designed them specifically to enter cells and express their genes there. Further, scientists can substitute one or more potentially therapeutic genes for genes involved in viral replication and virulence. In theory, then, an altered, tamed virus should transfer helpful genes to cells but should not multiply or produce disease.

The viruses that have been examined most extensively are retroviruses, which splice copies of their genes permanently into the chromosomes of the cells they invade. Such an integrated gene is cop-

NATURALLY OCCURRING virus (*bottom panel*) releases its genetic material into cells. Whether or not the genes become integrated into the DNA of the infected cell, they soon direct the synthesis of new viral particles that can injure the cell and infect others. To convert a wild-type virus into a safe gene therapy vector, scientists replace viral genes with ones specifying therapeutic proteins (*top panel*), while ideally leaving only the viral elements needed for gene expression. Such vectors should enter cells and give rise to helpful proteins but should not multiply.





SLIM FILMS

	Retroviruses	Adenoviruses	Adeno-Associated Viruses	Liposomes	"Naked" DNA
Some Potential Advantages	Integrate genes into host chromosomes, offering chance for long-term stability	Most do not cause serious disease; large capacity for foreign genes	Integrate genes into host chromosomes; cause no known human diseases	Have no viral genes, so do not cause disease	Same as for liposomes; expected to be useful for vaccination
Some Drawbacks of Existing Vectors	Genes integrate randomly, so might disrupt host genes; many infect only dividing cells	Genes may function transiently, owing to lack of integration or to attack by the immune system	Small capacity for foreign genes	Less efficient than viruses at transferring genes to cells	Inefficient at gene transfer; unstable in most tissues of the body

VECTORS UNDER STUDY as gene delivery vehicles include viral and nonviral carriers, only some of which are listed. Each vector type has its own set of advantages and disadvantages, and all are being modified rapidly to improve their effectiveness in patients.

ied and passed to all future generations of those cells. In contrast, many other kinds of viruses do not integrate their genetic material into a host's chromosomes. Their genes generally function in the body more transiently—in part because the genes do not replicate when recipient cells divide.

One group of ideal target cells for retrovirus vectors consists of so-called stem cells, which persist indefinitely and also produce more specialized descendant cells. Blood-forming stem cells, for example, give rise to every other type of blood cell (red cells, white cells of the immune system, and so on) and reconstitute the blood as needed; they also make more copies of themselves. At the moment, however, it is extremely difficult to identify human stem cells and modify them in safe, predictable ways.

Despite the appeal of retroviruses, which were first introduced as vectors in the early 1980s, they pose several challenges. They are promiscuous, depositing their genes into the chromosomes of a variety of cell types. Such lack of fine specificity for host cells can militate against direct delivery of the vectors into the body; uptake by cells that were not intended to receive the foreign gene could reduce transfer to the targeted population and might have unwanted physiological effects. Conversely, the retroviruses now receiving the most study fail to transfer genes to cell types that cannot divide or that do so only rarely (such as mature neurons and skeletal muscle cells). Current retrovirus vectors reach chromosomes only when the membrane surrounding the nucleus of the host cell dissolves, an event that occurs solely during cell division.

Also problematic is the fact that retroviruses splice their DNA into host chromosomes randomly, instead of into predictable sites. Depending on where inserted genes land, they might disrupt an essential gene or alter genes in ways that favor cancer development. Tumors would probably result only rarely, but even the remote chance of increasing cancer risk must be taken seriously.

Researchers have made good progress recently in confronting the shortcomings of retroviruses as gene delivery vehicles. For instance, to increase specificity and thus enable retrovirus vectors to direct themselves to particular cells in the body, researchers are altering the viral envelope (the outermost surface). Like other viruses, retroviruses deposit their genetic cargo into a cell only if proteins projecting from their surface find specific mates, or receptors, on the cell. Binding of the viral proteins to the cellular receptors enables a retrovirus to fuse its envelope with the cell membrane and to release viral genes and proteins into the cell's interior. To make retroviruses more selective about the cells they invade, investigators are learning how to replace or modify natural envelope proteins or to add new proteins or parts of proteins to existing envelopes.

In an experiment showing that the replacement strategy is feasible, Jiing-Kuan Yee of the University of California at San Diego, with my laboratory at that university, substituted the envelope protein of the mouse leukemia virus with that of the human vesicular stomatitis virus. (The mouse virus, which causes no known disease in people, is the retrovirus that has been evaluated most extensively as a gene therapy vector.) The altered mouse retrovirus then infected cells bearing receptors for the human vesicular stomatitis virus instead of cells with receptors for the mouse virus.

Work on modifying existing envelope proteins is also proceeding well. Yuet Wai Kan and his colleagues at the University of California at San Francisco have recently linked a protein hormone to the envelope protein of the mouse leukemia virus. This hormone enabled the virus to infect human cells that displayed the receptor for that hormone.

Prospects for generating retrovirus vectors able to insert therapeutic genes into the chromosomes of nondividing cells are looking up as well. Inder M. Verma, Didier Trono and their colleagues at the Salk Institute for Biological Studies in San Diego have capitalized on the ability of HIV, a retrovirus, to deposit its genes into the nucleus of nondividing brain cells without waiting for the nuclear wrapping to dissolve during cell division.

The team removed genes that would allow HIV to reproduce and substituted a gene coding for a protein that was easy to trace. This vector then brought the traceable gene into nonreplicating cells, not only when the vector was mixed with cells in culture but also when it was injected directly into the brains of rats. HIV itself might one day prove to be a

useful vector if worry that the disabled vectors might somehow become pathogenic can be allayed. Another tactic would transfer certain of HIV's useful genes—particularly those coding for the proteins that transport genes to the nucleus—into retroviruses that do not cause human disease.

Finally, efforts are under way to ensure that retrovirus vectors will place genes less randomly into human chromosomes. Workers toiling in this taxing realm have recently been assisted by new understanding of how genes integrate into predictable sites in the DNA of other organisms, such as yeast.

Pros and Cons of Other Virus Vectors

Vectors derived from viruses other than retroviruses present their own sets of advantages and disadvantages. Those based on the ubiquitous human adenoviruses have gained the most popularity as alternatives to retroviruses in part because they are quite safe; the naturally occurring forms typically cause nothing more serious than chest colds in otherwise healthy people. Moreover, they infect human cells readily and, initially at least, tend to yield high levels of the therapeutic protein.

Adenovirus vectors dispatch genes to the nucleus but apparently do not insert them into chromosomes. This feature avoids the possibility of disturbing vital cellular genes or abetting cancer formation, but, regrettably for some applications, the genes are often effective only temporarily. Because the DNA eventually disappears, treatments for chronic conditions, such as cystic fibrosis, would have to be repeated periodically (every so many months or years). In some situations, though—say, when a protein is needed only temporarily to induce an immune response to cancer or to a pathogen—short-term expression of a foreign gene may be preferable. Another drawback, shared with retroviruses, is lack of specificity for target cells. As is true for retroviruses, however, scientists are rapidly devising ways to target adenovirus vectors to tissues of the researchers' choosing.

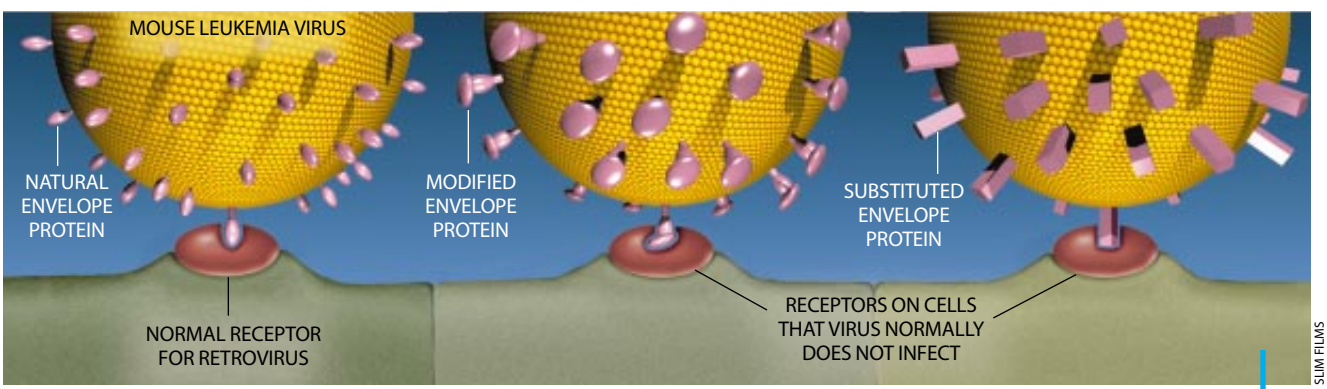
At the moment the more serious stumbling block to use of adenovirus vectors in patients is the body's strong immune response against them. During an initial round of treatment, such vectors might infect the appropriate cells and generate high amounts of the desired proteins. But soon host defenses come into play, killing the altered cells and inactivating their new genes. Further, once the immune system is alerted to the viruses, it eliminates them quickly if they are delivered a second time. Such responses probably have contributed to a shutdown of gene expression in a number of adenovirus gene-transfer studies in patients. Advancing understanding of the shortcomings of adenoviruses is now leading to a new generation of vectors that should reduce defensive interference. These enhancements have been achieved in part by removing or mutating the adenovirus genes most responsible for eliciting immune attacks.

Several other viruses are being explored as vectors as well—among them, adeno-associated viruses, herpesviruses, alphaviruses and poxviruses. None is perfected yet, but each is likely to have its own therapeutic niche. For example, adeno-associated viruses appeal because they cause no known diseases in people. What is more, in their natural form, they integrate their genes into human chromosomes. They are likely to be useful for some applications that now depend on retroviruses, but they are smaller and so may not be able to accommodate large genes. Herpesviruses, in contrast, do not integrate their genes into the DNA of their hosts. But they are attracted to neurons, some of which retain the viruses in a more or less innocuous state for the lifetime of the affected person. Herpesviruses therefore have potential as vectors for therapy aimed at neurological disorders [see "Gene Therapy for the Nervous System," by Dora Y. Ho and Robert M. Sapolsky, on page 116].

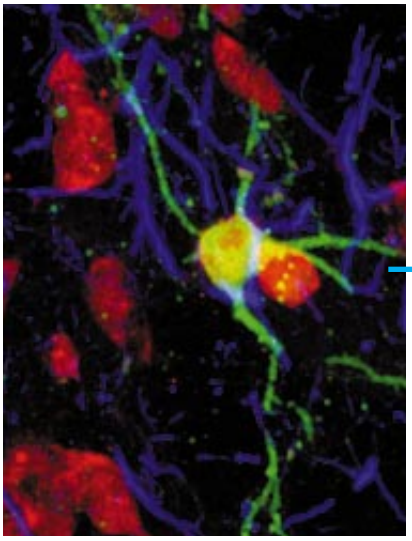
Perfecting Nonviral Delivery Systems

As a group, vectors produced from viruses continue to show great promise, although researchers must always work to ensure that the viruses will not change in ways that will enable them to cause disease. This consideration and others have encouraged development of various nonviral methods for therapeutic gene transfer. In common with viruses, these synthetic agents generally consist of DNA combined with molecules that can condense the DNA, deliver it to cells and protect it from degradation inside cells. And, like virus vectors, they will almost certainly be used in medical practice eventually but are still in need of refinement. The genes transferred by nonviral vectors become integrated into the chromosomes of recipient cells in the laboratory but have done so only rarely after delivery into the body. Whether lack of integration will be an advantage or disadvantage depends, as I have noted, on the particular goal of therapy.

Liposomes, which are small fatty (lip-



VIRUSES MUST BIND to particular surface molecules, or receptors, to gain entry into cells. The naturally occurring mouse leukemia virus normally binds through its envelope protein to a receptor found on many cell types (*left*). By altering the envelope protein to include new components (*center*) or replacing it with other proteins (*right*), scientists have directed the virus to cells it would otherwise ignore. Similar tactics can target other vectors to selected cell types.



U. BLÖMER, F. H. GAGE AND INDER M. VERMA

HUMAN BRAIN CELL took up an HIV-based vector containing a gene for a traceable protein (*shown in yellow*). This success implies that disarmed forms of HIV, a retrovirus, can potentially be used to deliver therapeutic genes into neurons, which, being unable to divide, are resistant to traditional retrovirus vectors.

id) spheres, have been studied almost as long as retrovirus vectors. These synthetic bubbles can be designed to harbor a plasmid—a stable loop of DNA derived from bacterial viruses known as phages—in which original genes have been replaced by those intended to be therapeutic. Gene transfer by liposomes (or "lipoplexes," as current versions are increasingly called) is much less efficient than virus-mediated transfer but has advanced enough for these vectors to enter clinical trials for such diseases as cancer and cystic fibrosis. Meanwhile alterations in the chemical composition of liposomes are addressing the efficiency problem and are beginning to produce vectors that mimic viruses in their targetability and prowess at gene transfer [see "Nonviral Strategies for Gene Therapy," by Philip L. Felgner, on page 102].

Newer kinds of vectors sheathe DNA in nonlipid coats. These coats include amino acid polymers and other substances intended to target therapeutic genes to the proper cells in the body and to protect the genes from being broken down by cellular enzymes. These complexes—studied intensively by Max Birnstiel and Matt Cotten of the Institute of Molecular Pathology in Vienna and by David T. Curiel of the University of Alabama at Birmingham—have performed well in cell culture experiments. They are now being further modified and are undergoing testing in animal studies and in patients.

Some scientists are also exploring injecting so-called naked DNA—without a lipid wrapping—into patients. Initial results suggest that the naked-DNA strategy has exciting potential for immunization against infectious diseases, and even against certain kinds of cancer.

Alternatives to plasmids are being pursued as well. Notably, workers are learning to construct miniature chromosomes, or artificial human chromosomes, into which therapeutic genes can be spliced. These constructs will contain the minimum amount of genetic material needed to avoid degradation in the nucleus and loss during cell division. They will also incorporate elements that enable the artificial chromosomes to copy themselves accurately (and only once) each time a cell divides, just as ordinary chromosomes do.

Looking Ahead

In the future, as now, investigators will choose one or another gene delivery method on the basis of their therapeutic goal. If a patient inherited a genetic defect and needs a continuing supply of the normal gene product throughout life, a vector that can integrate the therapeutic gene into the patient's chromosomes, where it will stay in perpetuity, might be best. Then a retrovirus or adeno-associated virus may be selected. If only short-term activity of a gene is needed, such as to arouse the immune system against cancer cells or an infectious agent, nonintegrating delivery vehicles, such as adenovirus vectors, liposomes or even naked DNA may be more suitable.

But the tools that finally come into common use almost certainly will not be the prototypes being tested today. And because no single technique will be perfect for every disorder, there will be many choices. The ideal gene transfer systems of the future will combine the best features of different vectors. Each system will be tailored to the specific tissue or cell type requiring modification, to the needed duration of gene action and to the desired physiological effect of the gene product. Scientists will also want to develop ways to alter the level of gene expression at will and to shut off or completely remove introduced genes if therapy goes awry.

Even when these gene delivery vectors are perfected, the challenges will not end. For instance, cells often modify foreign genes in ways that ultimately cause the genes to stop working. This activity is being addressed vigorously but is not yet solved. In addition, we still have few clues as to how the defensive systems of patients will respond when they encounter a seemingly foreign protein from a therapeutic gene. To prevent an inactivating immune reaction, physicians might have to treat some patients with antirejection drugs or try to induce immune tolerance to the encoded protein by carrying out gene therapy very early in a patient's life (before the immune system is fully competent).

Although I have dwelled on certain technical challenges to gene therapy, I am nonetheless highly optimistic that it will soon begin to prove helpful for some diseases. Our tools are improving rapidly, and some of the burgeoning clinical trials clearly are on the verge of demonstrating real merit in ameliorating disease, even with today's imperfect techniques. Notably, it seems likely that gene-based immunotherapies for some malignancies, such as neuroblastoma and melanoma, will be shown convincingly in the next few years to slow the development of further disease and to force existing tumors to regress; they should then become helpful additions to existing therapies. But I must emphasize that it is only through insistence on rigorous science, carefully designed clinical studies and less exaggerated reporting of results that researchers can ensure the timely, ethical and effective flowering of this exciting new field of medicine. SA

The Author

THEODORE FRIEDMANN, who holds an M.D. from the University of Pennsylvania, is director of the gene therapy program, professor of pediatrics and Muriel Whitehull Chair of Biomedical Ethics at the University of California, San Diego. He joined the faculty of U.C.S.D. in 1969, after studying genetics and biochemistry at the National Institutes of Health and at the Salk Institute for Biological Studies in San Diego. In 1994 he served as Newton Abraham Professor at the University of Oxford (and earned a master of arts degree there). Twenty-five years ago he co-authored (with Richard Roblin) the first major scientific paper outlining the need for gene therapy, the ethical considerations surrounding it and the techniques that could produce virus vectors.

Nonviral Strategies for

Many drawbacks of viral gene delivery agents might be overcome by nonviral systems. Studies in patients suggest these systems have potential as therapies and as vaccines

by Philip L. Felgner

As Theodore Friedmann notes in "Overcoming the Obstacles to Gene Therapy" on page 96 of this issue, many efforts at developing gene therapy employ modified viruses to shuttle into human cells genes coding for potentially therapeutic proteins. The aim is to induce cells that are invaded by a virus to transfer the gene to the cell nucleus. The cells should then "express," or manufacture, the needed protein specified by the gene.

Viruses are effective at transferring genes into cells because they have evolved specialized mechanisms that allow them to bind to specific types of cells and to deliver their cargo efficiently into the cellular interior. Yet the therapeutic use of viruses as gene delivery vehicles, or vectors, entails problems. Some viruses can disrupt the DNA of the cells they infect, with potentially harmful results. Furthermore, weakened viruses can conceivably change inside the body and regain their pathogenic activity. An additional serious limitation is that a patient may generate an immune response to the microbe. Such responses can quickly make a gene therapy useless, because they may either destroy the virus itself or possibly kill the infected cells before the therapeutic gene has a chance to help a patient.

For these reasons, researchers have long wanted to deliver therapeutic genes to cells without using infectious agents. Further, physicians have come to realize over the past five years or so that for many of the conditions that might ultimately be treated by gene therapy, repeated treatments will probably be needed, rather than a one-time procedure aimed at producing a permanent cure. Nonviral techniques could be especially suitable for repeated use, because they do not elicit the immune responses that can damage viral vectors.

One approach I have studied, which is now being tested in humans, employs complexes consisting of DNA and nonimmunogenic lipids. Recent years have also seen the surprising discovery that injecting "naked" DNA into experimental animals and patients can provoke expression of encoded proteins. This second approach, as will be seen, holds particular promise for new vaccines.

Electrical Problems

Scientists have long been enchanted by the possibility that they could alter cells in selected ways by putting foreign DNA into them. John Holland of the University of California at San Diego and several of his contemporaries showed as long ago as the mid-1950s that cells could take up nucleic acids (RNA or DNA) extracted from viruses and express them as proteins. This discovery provided an incentive for scientists to improve the efficiency of gene "transfection," the delivery of functional genes to cells.

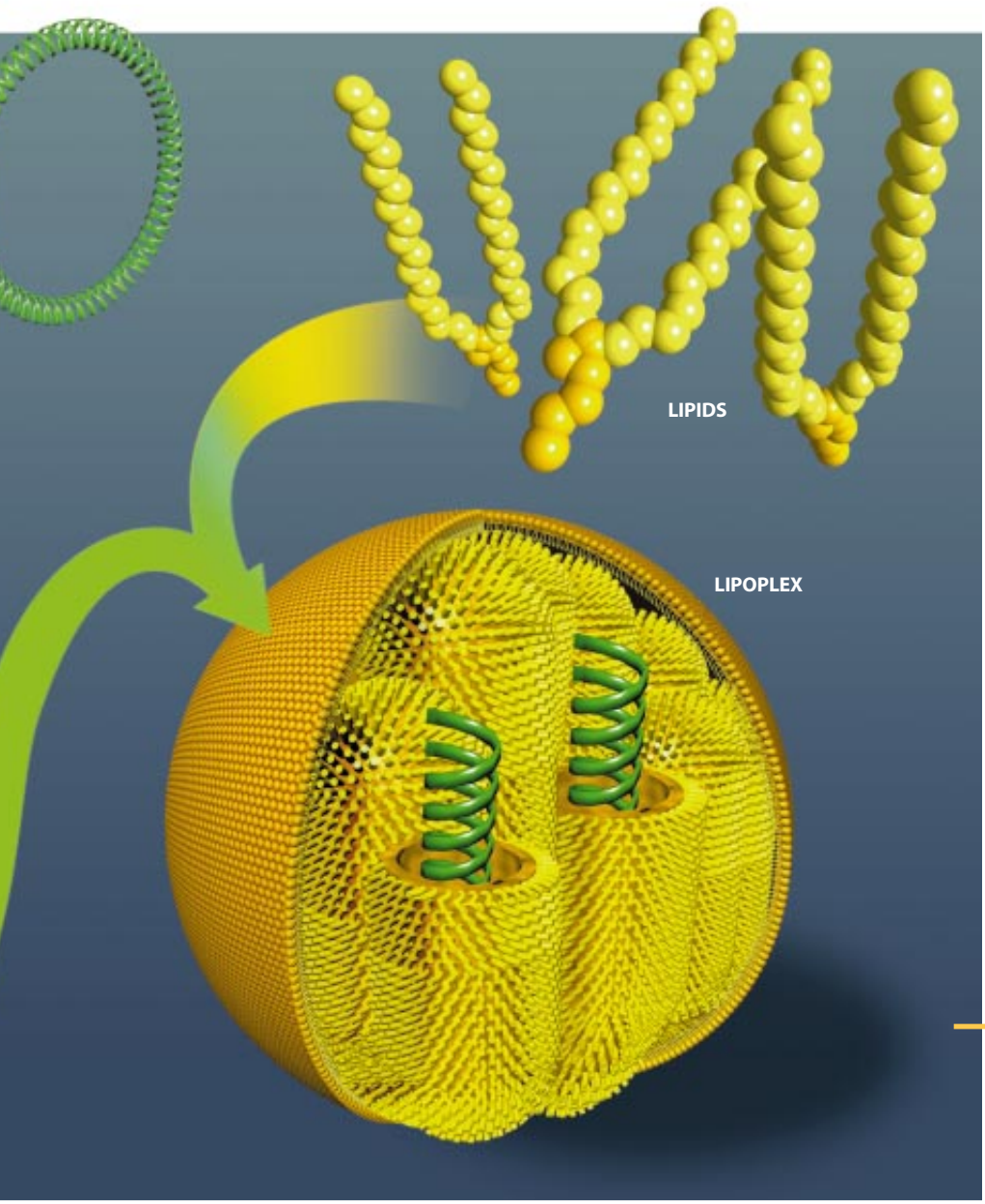
During the 1960s, investigators established that a principal obstacle to the uptake of DNA by cells was that in water solution, such as the milieu that bathes cells in the body, the molecule has a negative electrical charge.

Gene Therapy

This means it tends to be repelled from the membranes of cells, because they are also negatively charged. Researchers therefore developed techniques that combined DNA with chemicals that electrically neutralized it and so allowed it to be absorbed more easily. One technique made use of a positively charged organic polymer called DEAE-dextran. Another procedure employed the mineral calcium phosphate.

Researchers established by these means that human cell cultures could take up genes and permanently express them. One important demonstration employed the gene for the enzyme thymidine kinase, isolated from the herpes simplex virus. Some cells can incorporate the gene when it is complexed with calcium phosphate and produce the enzyme in a stable fashion.

In the late 1970s the modern biotechnology industry got its start with the discovery of techniques for taking individual genes from cells and splicing them into plasmids—flexible loops of DNA that multiply naturally within bacteria. Recombinant DNA techniques therefore allowed investigators to produce large numbers of copies of specific genes. Paul Berg and his colleagues at Stanford University melded recombinant DNA and chemical cell transfection technology by delivering recombinant plasmids derived from bacteria into cultured mammalian cells. Expression of genes spliced into such plasmids suggested the genes were being taken to the nucleus, because genes usually cannot be expressed unless they are acted on by nuclear proteins. These nonviral transfection methods were then applied to produce many of the mammalian cell lines that industry uses today to produce medicinal recombinant proteins, such as the Factor VIII administered to hemophiliacs to correct potentially life-threatening blood-clotting problems.



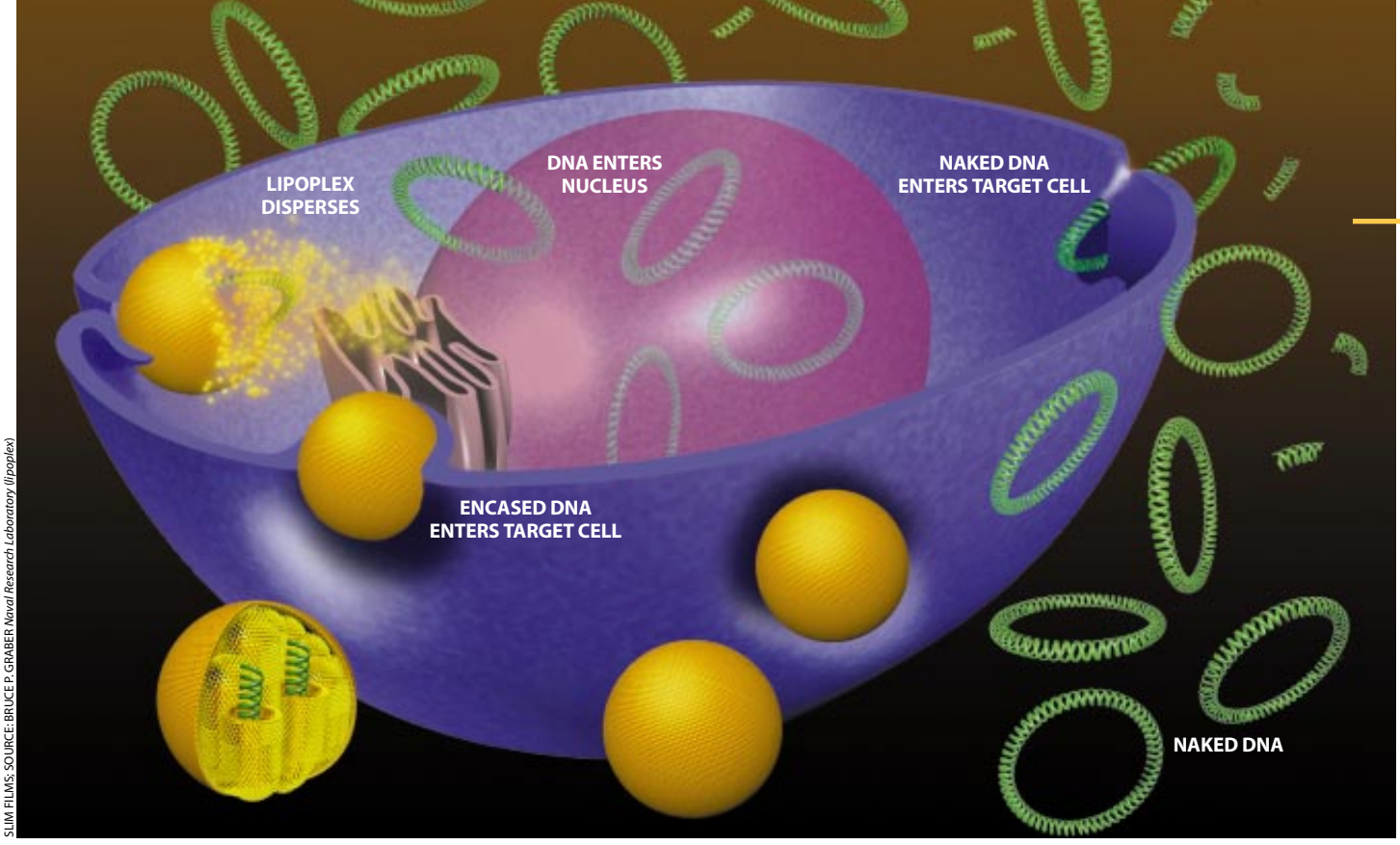
SLIM FILMS; SOURCE: BRUCE P. GRABER *Naval Research Laboratory (lipoplex)*

LIPOPLEXES for experimental gene therapy are made in part by copying in bacteria plasmids that contain therapeutic genes. When the extracted plasmids are mixed with lipids, lipoplexes form spontaneously.

Lipids to the Fore

Despite the commercial value of chemical techniques, most investigators considered them too inefficient for gene therapy. Yet Berg and Demetrios Papahadjopoulos of the University of California at San Francisco also successfully transfected cells with genes by exposing the cells to gene-containing liposomes, which are minute hollow spheres composed of a lipid (fatty) membrane on the outside and an aqueous solution within. Claude Nicoloau of Harvard Medical School achieved the same result. This work also suggested that the genes of the plasmids were getting to the nucleus, where the cell's machinery was acting on the introduced genes and causing them to be expressed as proteins.

First described in the 1960s by Alec D. Bangham of the Babraham Institute in Cambridge, England, liposomes can form spontaneously



SLIM FILMS; SOURCE: BRUCE P. GRABER *Naval Research Laboratory (lipoplex)*

when certain types of lipids are suspended in an aqueous solution. Liposomes resemble animal cells in that the outer membrane consists of a double layer of lipid molecules. This feature arises because the lipid molecules that are employed have a water-loving and a water-hating end. In water solutions they form double-layered membranes in which the water-loving heads face out toward the aqueous external environment and also in toward the water-filled center of the liposome. This cell-like aspect had long suggested to investigators that liposomes "loaded" with some medicinal substance might fuse with cells and deliver the liposomes' contents into the cellular interior.

The results with liposomes were encouraging, but a technical problem slowed their practical exploitation for delivering plasmids into cells. The internal diameter of a liposome—between about 0.025 and 0.1 micron—is typically much less than the longest dimension of a DNA plasmid, which may be as much as two microns. The mismatch meant that when liposomes were synthesized in the presence of plasmids, only a few tightly wound plasmids were encapsulated.

Nevertheless, optimistic scientists believed ways could be found to improve the encapsulation rate. Working at Syntex Research in Palo Alto, Calif., my colleagues and I theorized that we might be able to modify liposomes to capture plasmids more effectively and to deliver more readily their contents into cells.

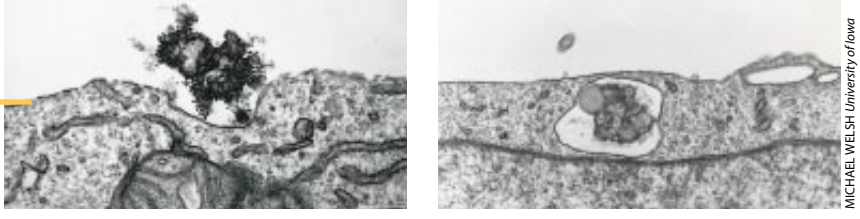
Our idea was to make liposomes in which some of the standard lipids were replaced by ones carrying a positive charge at the water-loving end. This, we thought, should make the liposomes interact more easily with DNA and RNA, as well as with cell surfaces. At the time, however—1983—there were few examples of positively charged (cationic) lipids that had the right shape to organize into liposomes.

We therefore synthesized variant cationic lipids that we predicted would have the right properties. The molecules behaved as expected, and the resulting cationic liposomes bound firmly to the surface of cells in tissue culture. Furthermore, simply by mixing plasmids with about eight times their mass of cationic lipids, we could effectively capture all the DNA. I was pleased to find that I could readily create conditions that produced physically stable complexes.

Exploiting Discovery

The structures formed from mixtures of plasmids and cationic lipids are more variable and complicated than simple liposomes. We often find plasmids enclosed in tubelike lipid structures, for example. And under the right conditions, a lipid tube containing a plasmid can fold up to form a dense particle with a lipid wall. This structure is somewhat like that of some viruses. Because structures formed from cationic lipids are so different from simple liposomes, researchers have recently agreed to give them a new name: lipoplexes.

I had expected to have to modify lipoplexes further before they would



MICHAEL WELSH *University of Iowa*

ASSIMILATION of a lipoplex by a cultured cell is seen in these micrographs. The lipid in the lipoplex protects its DNA cargo.

LIPIDS in lipoplexes help DNA enter a cell (*left* and *front*), before cellular organelles take up the lipids. Naked-DNA plasmids (*right*) can occasionally enter a cell through a tear.

deliver their payload into cells. But, remarkably, an undergraduate intern in my lab showed that these rudimentary structures transfected cells at a significant rate. Mixing cationic lipids with DNA has since become a standard technique for inserting genes into cultured cells, and many cationic lipid preparations are now available commercially for this purpose.

Researchers have lately initiated studies of lipoplexes in humans. The first candidate therapy consists of lipoplexes containing genes for an immune system protein known as HLA-B7, a so-called major histocompatibility antigen. When cancer cells express HLA-B7, they stimulate a patient's immune system to recognize them as foreign and selectively destroy them. In clinical protocols sponsored by Vical, more than 90 patients who had failed to respond to standard cancer treatments were given injections into their tumors of lipoplexes containing DNA encoding HLA-B7.

In most cases, investigators could demonstrate that the treatment stimulated production of HLA-B7 protein. Sixty of the patients suffered from malignant melanoma. In about a third of these cases the lipoplex-injected tumor either shrank or disappeared. Advanced melanoma often spreads widely throughout the body, so injecting drugs into visible tumors may not cure most cases. But our encouraging preliminary findings indicate that lipoplexes might help treat melanoma. Sometimes treatment with lipoplexes makes even uninjected tumors shrink. Other studies with lipoplexes containing the gene for HLA-B7 aim to establish the safety and efficacy of similar treatments for inoperable colon, kidney and breast cancer.

Vical is also sponsoring a clinical trial of a lipoplex formulation containing the gene for the immune system protein interleukin-2 (IL-2). IL-2 is infused into patients to treat kidney cancer, but it has serious toxic side effects. Injected into a tumor, lipoplexes might stimulate the immune system by producing locally high concentrations of IL-2 while avoiding most of the toxic side effects.

Additionally, Genzyme General Division is studying lipoplexes in patients with cystic fibrosis. For treating that disease, the structures carry a gene for the protein that is defective in the illness—the cystic fibrosis transmembrane conductance regulator. The lipoplexes are delivered with an aerosol spray into a patient's lungs, where expression of the protein should mitigate some of the most serious symptoms.

The levels of gene expression obtainable in animals with lipoplex formulations are in some instances comparable to the levels obtainable with viral gene delivery systems. Still, the efficiencies of the two systems differ substantially. Some viruses are almost 100 percent efficient at delivering their genome into cells, so that 1,000 viruses can infect almost 1,000 cells of the right type. To transfect 1,000 cells with lipoplexes would require about 10 million copies of the gene in a comparable number of lipoplexes, making this approach some 10,000 times less efficient.

One strategy we and others are pursuing to improve the efficiency of lipoplexes is to incorporate into their outer surface specialized proteins or protein fragments resembling those that target viruses to particular cell types. We might also include other molecules to facilitate gene survival and functioning in transfected cells. For example, so-called membrane fusion proteins of viruses, which normally help viruses escape from the cell's internal waste disposal system, could be incorporated. In addition, researchers are experimenting with ways of attaching to genes certain viral proteins, known as nuclear targeting signals, that help to direct viral genes to the cell nucleus.

Meanwhile we and others are also pursuing applications for naked DNA. In the late 1980s, long before the lipoplexes had entered clinical trials, my colleagues Jon A. Wolff of the University of Wisconsin and Robert W. Malone and I, along with some others, made a surprising discovery. Hoping to determine which lipid formulations were most effective at gene delivery to cells, we had decided to measure the expression of a particular, easily monitored gene after we injected various mouse tissues with lipoplexes carrying that gene.

It was a memorable day when we analyzed the data from our first experiment and saw we had obtained levels of gene expression in skeletal muscle comparable to the best results we had obtained by transfecting cells in culture. But the next observation was unprecedented. We found that DNA by itself, which we had employed in some tests as a control experiment, gave expression levels that were similar to or greater than the lipid formulations.

Naked DNA

We repeated these experiments several times in our different laboratories, but the results always came out the same way: naked DNA, injected into the muscle of an animal, was expressed as protein. Moreover, we could achieve quite a high local concentration of the protein, up to 100 nanograms of the gene product per gram of muscle tissue.

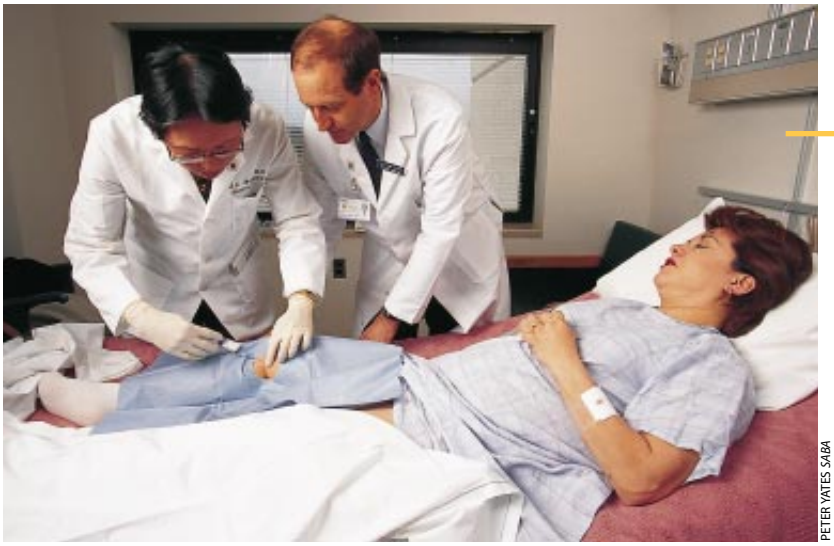
Despite the electrical repulsion between cell surfaces and DNA, it appeared that a few cells were nonetheless able to assimilate the molecule. To this day, it is unclear how. Perhaps a small amount of tissue damage or increased pressure at the injection site plays a role.

In principle, it seemed that it might be possible to inject patients' muscles with naked DNA that would then produce therapeutic amounts of a selected protein. There is a great need, for example, for better ways to deliver to patients insulin for treating diabetes, or the blood-clotting protein Factor VIII or IX for treatment of hemophilias. Yet in these early studies the locally high concentrations of protein produced inside the muscle would not have been sufficient to be effective against these diseases when the proteins were diluted into the three liters of plasma in the bloodstream.

Thanks to improvements in plasmid design, my colleagues and I have recent-

PHILIP L. FELGNER

MOUSE MUSCLE injected with naked-DNA plasmids takes up the plasmids and produces the protein they encode. In this experiment the encoded protein produced a marked blue coloration.



PETER YATES *SABA*

LIPOPLEXES to fight cancer are administered to a melanoma patient by Alfred E. Chang and Gary J. Nabel of the University of Michigan Medical Center.

ly made progress. We find we can stimulate the production of red blood cells in mice by injecting the animals with naked plasmids encoding erythropoietin. This hormone is used to help patients grow new red blood cells after chemotherapy and radiation therapy. Perhaps in the future, injections of similar recombinant plasmids into muscle will be a less expensive alternative to delivering erythropoietin itself.

In the nearer term, naked DNA has promise for use in vaccines, because even minute amounts of protein can stimulate a protective immune response. Immunologic studies have indicated that proteins can elicit two different kinds of immunity. The first, humoral immunity, develops after a pathogen is destroyed by the immune system. The damaged microbes are then taken up by specialized cells that display the foreign proteins to antibody-secreting cells called *B* lymphocytes. The *B* lymphocytes respond by producing antibodies that recognize the specific foreign proteins. These antibodies will rapidly bind to the original pathogen whenever it is encountered again, neutralizing it or marking it for destruction by other components of the immune system.

The second kind of response, known as cellular immunity, occurs when invading pathogens colonize cells and force them to make more of the pathogen. Short pieces of the pathogen's proteins are then displayed on the surface of the infected cell. The immune system responds by producing other lymphocytes: activated *T* lymphocytes that recognize these fragments. These lymphocytes do not produce antibodies but rather destroy directly cells that are displaying the critical fragments. If the pathogen invades the body again, infected cells display the fragments and so are destroyed.

Dennis A. Carson of U.C.S.D. and I reasoned that naked plasmid DNA, because it invades cells, might be able to stimulate cellular as well as humoral immunity. Our very first experiment, conducted with Gary H. Rhodes, showed that injecting into mice a plasmid encoding a coat protein from the human immunodeficiency virus (HIV) stimulated the mice to generate antibodies that bound to the HIV protein. Rhodes subsequently demonstrated that the plasmids elicited cellular immune responses, too. In laboratory tests, *T* lymphocytes from these animals attacked cells displaying fragments of HIV proteins.

Looking Ahead

The crowning achievement came when Suezanne E. Parker from my laboratory showed that plasmids carrying a gene from influenza virus could be used to immunize mice and so protect them from what would otherwise be a lethal viral dose. Margaret Liu and her colleagues at Merck have confirmed these findings in animals and expanded them into the development of a variety of potential DNA vaccines that produce long-lasting cellular and humoral responses. Merck has a naked-DNA influenza vaccine candidate in human clinical trials.

We can expect to see clinical testing of DNA vaccines against herpes, malaria and HIV in the not too distant future. In the longer term, tuberculosis, papilloma, chlamydia and hepatitis could be targets. As well as preventing infection, such vaccines might stimulate the immune systems of those already sick. A clinical trial of a naked-DNA therapy for lymphoma, a cancer of the white blood cells, is being planned.

Lipoplexes and naked DNA are not the only nonviral approaches to gene therapy. Investigators are also studying various nonlipid cationic polymers that form complexes with DNA. These structures, known as polyplexes, are showing promise in the laboratory and in clinical trials.

One important goal for nonviral gene therapy will be to develop delivery systems that can be injected into the bloodstream and will deliver their DNA sequences to the appropriate tissues, such as lung, liver, spleen or bone marrow. Gene delivery systems that can be swallowed as a pill could make gene therapy even more convenient. And if delivery systems could be made to target tumor cells specifically, they could in various ways improve the treatment of cancer. Ultimately, gene therapy could be employed to correct mutated genes in the cells of people with genetic diseases or cancer. A technique known as gene targeting offers a possible approach; it has been used successfully with lipoplexes in cultured cells.

Nonviral gene delivery to cells via lipoplexes, polyplexes and naked DNA is now an important and expanding field of pharmaceutical research and development. If the current pace of progress is maintained, coming decades should see many products based on this technology being administered on a routine basis for the treatment and prevention of common diseases. SA

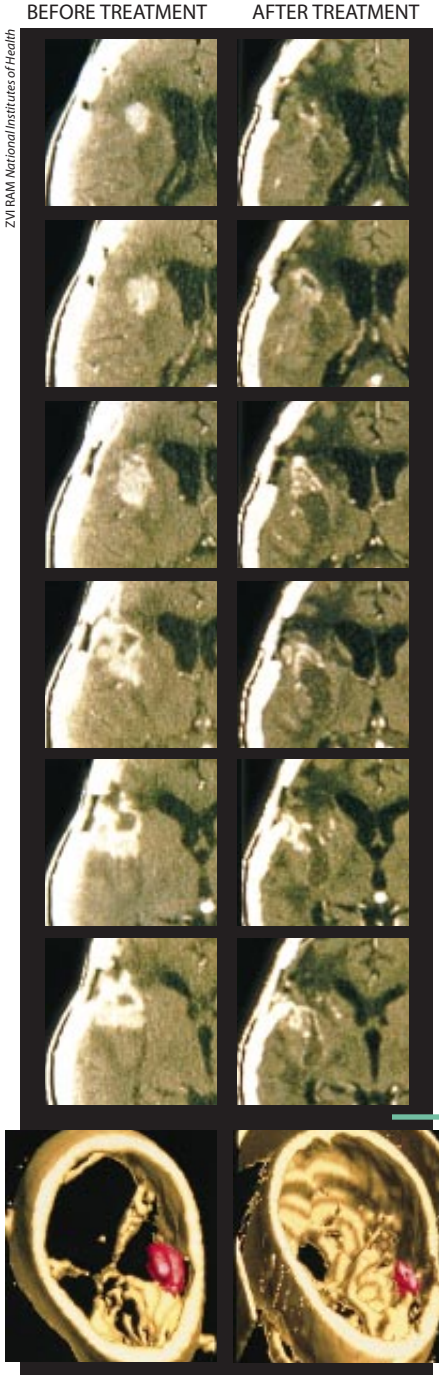
The Author

PHILIP L. FELGNER is chief scientist at Vical, Inc., in San Diego. After earning his Ph.D. in biochemistry at Michigan State University, Felgner studied lipid biophysics at the University of Virginia. In 1982 he joined Syntex Research in Palo Alto, Calif., and six years later he co-founded Vical. His wife, Jiin H. Felgner, is formulations scientist at the company. When not researching lipoplexes, Philip Felgner relaxes by playing classical guitar.

Gene Therapy for Cancer

Inserted genes could in theory arrest tumor growth or even AIDS

by R. Michael Blaese



In 1997 an estimated 1.38 million Americans will be newly diagnosed with cancer. Sadly, the main treatments currently available—surgery, radiation therapy and chemotherapy—cannot cure about half of them. This sobering fact has spurred serious efforts to develop additional strategies for treating the disease—ones based on the biology behind it. To that end, scientists are turning toward gene therapies, which involve introducing into the body genes that can potentially combat tumors.

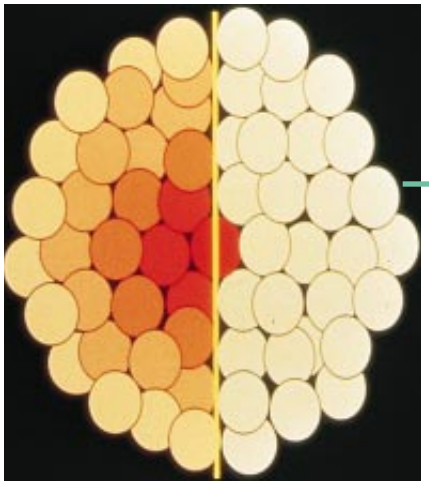
Researchers initially explored gene therapies for remedying conditions caused by defective genetic instructions, or mutations, passed on from one generation to the next. Most cancers are not inherited in this way but instead result from acquired mutations, produced by external factors such as tobacco smoke or high doses of radiation—or just pure bad luck. These mutations accumulate in cells over time, ultimately rendering the cells unable to control their own growth—an inability that leads to cancer [see "What You Need to Know about Cancer," special issue of Scientific American; September 1996].

Gene therapies in general deliver instructions—in the form of DNA sequences—to diseased cells so that they will produce a therapeutic protein of some kind. This type of therapy is possible because viruses, bacteria, plants and people all share the same genetic code. Researchers have learned a great deal in little time about how certain genes govern the fundamental processes of life and how they contribute to disease. And because genes from one species can be read and understood by another, experimenters can transfer genes between cells and species in their efforts to devise treatments.

For treating cancer, experimental gene therapies take varied forms: some involve imparting cancer cells with genes that give rise to toxic molecules. When these genes are expressed (that is, used by cells to make proteins), the resulting proteins then kill the cancer cells. Other designs aim to correct or compensate for acquired genetic mutations. Still others attempt to activate the processes by which such defects are normally repaired. And a host of ideas are coming from insights into how tumors evade recognition and destruction by the immune system, how they spread away from their sites of origin, how they gain a new blood supply and how they accomplish other feats that allow them to endure and spread.

Most of these approaches have yet to pass even the most preliminary clinical tests demonstrating their overall safety and efficacy, but these ideas may lead to

BRAIN TUMOR (*white*) appears in six MRI images of sequential slices taken before (*left*) and two weeks after (*right*) the patient received an experimental "suicide" gene therapy. The author and his colleagues inserted a gene for an enzyme called thymidine kinase (tk) from a herpesvirus into the patient's affected brain cells. They then gave the patient ganciclovir. The tk enzyme converts this otherwise harmless drug into a toxic metabolite, capable of killing not only the cancer cells containing the tk gene but even some surrounding cancer cells as well. The computer models (*bottom two images*) reveal how quickly this same tumor (*red*) responded: a single treatment brought about the dramatic reduction shown.



R. MICHAEL BLAESE

BYSTANDER EFFECT, shown schematically, is critically important to the success of some gene therapies. On the left, gap junctions connect the tumor cells. As a result, toxin generated in one gene-modified cell (*orange*) can spread to neighboring "bystander" tumor cells (*yellow*), which have not themselves been altered. The tumor on the right lacks gap junctions and cannot be cured unless the therapeutic gene reaches each individual cancer cell—an unlikely bet for now.

better cancer treatments in the future.

Aside from promising actual treatments, gene therapy techniques have thus far helped physicians evaluate existing interventions. For example, in recent years doctors have increasingly relied on bone marrow transplants for treating cancers that fail to respond to traditional therapies. Frequently, the procedure is used for battling advanced stages of leukemia, a cancer that affects white blood cells, which are made by bone marrow. Before performing transplantation, oncologists take bone marrow from leukemia patients in remission. This apparently healthy bone marrow is stored away, and the patients are given superhigh doses of chemotherapy or radiation to kill off any residual cancer. Because these high doses destroy normal bone marrow, such an aggressive treatment would ordinarily kill the patient as well as the cancer. In a transplantation, though, the patient is "rescued" by receiving a transfusion of his or her own saved bone marrow.

This method should, in theory, cure leukemia. But sometimes the disease recurs anyway. Clinicians have often wondered what goes wrong. Do the high-dose radiation treatments sometimes fail to kill off all residual cancer cells, or are there sometimes undetected leukemia cells in the presumed disease-free marrow stored away during remission? To distinguish between these possibilities, scientists needed to devise a nontoxic and permanent tag so that they could mark the extracted bone marrow cells and find them later in the body. Malcolm K. Brenner of St. Jude Children's Research Hospital in Memphis, Tenn., did just that starting in late 1991, by inserting a unique sequence of bacterial DNA not found in humans into the patient's saved bone marrow. Brenner knew that if he detected this bacterial DNA in the recovering patient's blood and bone marrow after transplantation, it would prove that the "rescue marrow" was restoring the blood system. In addition, detection of this tag in recurrent cancer cells would prove that the rescue marrow was a source of leukemic cells.

This is in fact what Brenner and others have found in some cases, forcing a critical reappraisal of the use of bone marrow transplantation. For instance, it is now recognized that for certain types of cancer, it may be necessary to give additional treatments to the marrow itself to rid it of any contaminating cancer cells before transplantation. To that end, marrow "gene-marking" studies, identical to those described above, are helping to find the best solution. These studies allow researchers to compare sundry methods for purging the marrow of residual cancer. Different marker genes are used to tag marrow samples purged in different ways. As a result, physicians can determine how well marrow that has been purged in some way helps the patient following therapy. And, if the cancer recurs, they can also determine whether it has come from bone marrow that was purged in a particular way.

Gene Vaccinations

In terms of treatment, scientists have for more than three decades tried to find ways to sic the immune system on cancer—a tactic termed immunotherapy or vaccine therapy. And with good reason. Because immunity is a systemic reaction, it could potentially eliminate all cancer cells in a patient's body—even when they migrate away from the original tumor site or reappear after years of clinical remission. The problem with this strategy has been that the immune system does not always recognize cancer cells and single them out for attack. Indeed, many tumors manage to hide themselves from immune detection.

Recently, however, research in basic immunology has revealed means for unmasking such cancers. In particular, it now seems possible to tag cancer cells with certain genes that make them more visible to the immune system. And once awakened, the immune system can frequently detect even those cancer cells that have not been tagged.

The immune response involves many different cells and chemicals that work together to destroy in several ways invading microbes or damaged cells. In general, abnormal cells sport surface proteins, called antigens, that differ from those found on healthy cells. When the immune system is activated, cells called *B* lymphocytes produce molecules known as antibodies. These compounds patrol the body and bind to foreign antigens, thereby marking the antigen bearers for destruction by other components of the immune system. Other cells, called *T* lymphocytes, recognize foreign antigens as well; they destroy cells displaying specific antigens or rouse other killer *T* cells to do so. *B* and *T* cells communicate with one another by way of proteins they secrete, called cytokines. Other important accessory cells—antigen-presenting cells and dendritic cells—further help *T* and *B* lymphocytes detect and respond to antigens on cancerous or infected cells.

One gene therapy strategy being widely tested at the moment involves modifying a patient's cancer cells with genes encoding cytokines. First the patient's tumor cells are removed. Into these tumor cells, scientists insert genes for making cytokines, such as the *T* cell growth factor interleukin-2 (IL-2) or the dendritic cell activator called granulocyte-macrophage colony-stimulating factor (GM-CSF). Next, these altered tumor cells are returned to the patient's skin or muscle, where they secrete cytokines and thereby catch the immune system's attention. In theory, the altered cells should solicit vigorous immune cell activity at the site of the reinjected tumor. Moreover, the activated cells, now alerted to the cancer, could circulate through the body and attack other tumors.

In certain instances, these gene-modified tumor vaccines do seem to awaken the immune system to the presence of the cancer, and some striking clinical responses have been observed. All these

clinical studies, however, are preliminary. In most cases, patient responses to these treatments have not been carefully compared with responses to conventional treatments alone. Also, the response patterns are not predictable, and they are not consistent from one tumor type to another or among patients who have the same type of cancer.

Another problem with these studies is that nearly every person tested so far has had widely disseminated terminal cancer. Usually these patients have previously received intensive anticancer therapy, which has weakened their immune systems. Thus, even if gene vaccines did activate immunity in these individuals, the responses might not be easily noticeable. Gene-modified tumor vaccines are most likely to prove beneficial in patients with minimal tumor burdens and robust immunity. Testing patients in this category, though, must wait until researchers are finished testing more seriously ill patient groups and have established the risks associated with the treatment. As this research so well illustrates, the development of new cancer therapies is a very complex and lengthy process.

A related gene therapy involves antigens that are found predominantly on cancer cells. During the past three to four years, scientists have made remarkable progress in identifying antigens produced by tumor cells. In addition, they have uncovered the genes that encode these tumor-associated antigens, particularly those on the most serious form of skin cancer, malignant melanoma. Now that at least some of these antigens have been described, it might be possible to develop a vaccine to prevent cancer, much like the vaccines for preventing tetanus or polio. The approach might also help treat existing tumors.

Preventive Immunizations

As with the cytokine vaccines, these antigen-based cancer vaccines require gene transfer. They work best when administered to cells that are readily accessed by the immune system. For example, Philip L. Felgner of Vical in San Diego and Jon A. Wolff of the University of Wisconsin and their colleagues observed that injecting a DNA fragment coding for a foreign antigen directly into muscle triggered a potent immune response to the antigen in mice [see "Nonviral Strategies for Gene Therapy," by Philip L. Felgner, on page 102]. The explanation for this reaction is simple: a bit of the foreign DNA enters the cells of the muscle or other nearby cells and directs them to produce a small amount of its protein product. Cells containing this newly synthesized foreign protein then present it to roving antibody-producing *B* cells and *T* cells. As a result, these sensitized immune components travel the body, prepared to attack tumor cells bearing the activating antigen.

The same basic strategy is revolutionizing the development of vaccines for preventing many infectious diseases. When these DNA immunizations are tested against cancer, the genes for newly identified tumor antigens are delivered directly into the body by way of vaccinia or adenovirus particles that have been rendered harmless or by such nonviral gene delivery systems as naked DNA. At present, the tests involve patients with widely spread cancer. It is clearly too late in these cases for DNA vaccines to prevent disease, but the studies should demonstrate whether the antigens can meet the essential requirement of eliciting a defensive response in the human body. Further, the studies offer a sense of whether DNA vaccines might have any merit for treating existing cancers. Given how sick many of these patients are, though, the results so far are difficult to interpret.

Yet another gene immunotherapy for cancer currently being tested in patients and in the laboratory involves antibodies. Thanks to highly variable regions on individual antibodies, these molecules are exquisitely specific. They can distinguish the slightest differences between foreign or mutated and very similar self-antigens. As it turns out, specific antibody molecules exist naturally in the outer membranes of some cancer cells—such as lymphomas that develop from *B* cells, which are committed to producing antibody molecules. Because a single lineage or clone of cells produces one specific antibody, all cancers of these cells will contain the same specific membrane molecule. This antibody then provides a unique molecular marker by which the cancer cells might be differentiated from similar but noncancerous antibody-producing cells.

Occasionally scientists have managed to produce antibodies to the antibodies found on cancer cell membranes. And some patients treated with these so-

Gene Therapies Being Studied in Cancer Patients

Approach	Number of U.S. Trials Approved since 1988 or Awaiting Federal Approval*
Antisense therapy (to block synthesis of proteins encoded by deleterious genes)	4
Chemoprotection (to add proteins to normal cells to protect them from chemotherapies)	7
Immunotherapy (to enhance the body's immune defenses against cancer)	58
Pro-drug, or suicide gene, therapy (to render cancer cells highly sensitive to selected drugs)	21
Tumor suppressor genes (to replace a lost or damaged cancer-blocking gene)	6
Antibody genes (to interfere with the activity of cancer-related proteins in tumor cells)	2
Oncogene down-regulation (to shut off genes that favor uncontrolled growth and spread of tumor cells)	2

SOURCE: OFFICE OF RECOMBINANT DNA ACTIVITIES, NATIONAL INSTITUTES OF HEALTH

*This table was up-to-date as of April 1997. It includes only those trials requiring approval by the federal government.

called anti-idiotype antibodies have responded exceedingly well. Unfortunately, producing anti-idiotype antibodies is laborious. Thus, even though the approach can sometimes provide an effective treatment, it has seen only limited use. More recently, gene transfer techniques have offered other options. Because antibodies are gene products, scientists have been able to prepare anti-idiotype DNA vaccines that include the DNA encoding the critical cancer marker (the idiotype). This DNA sequence has then been linked with a gene encoding the cytokine GM-CSF. So far this double-whammy cancer vaccine has been tested only in laboratory animals, but it shows exciting promise.

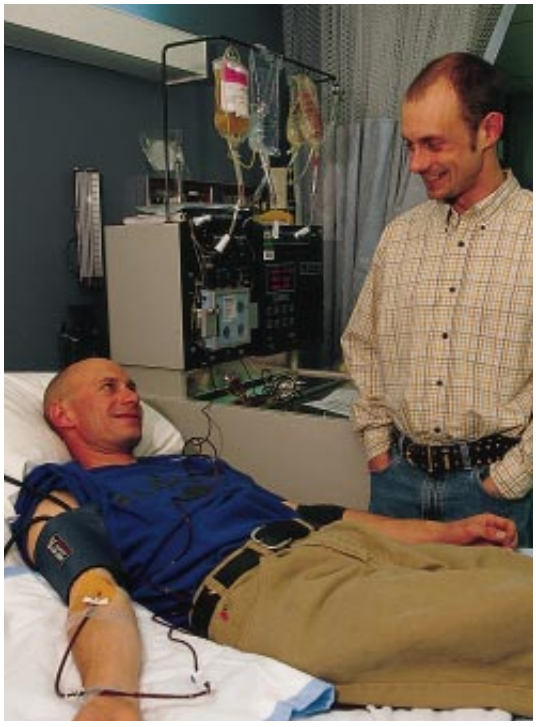
Another double-whammy therapy in the works couples antibodies and *T* lymphocytes. Some rare patients have cancers that their *T* cells do recognize. But the *T* cells from these patients usually attack only their own tumor cells or those from a small fraction of cases with the same type of cancer and tissue type. Also, people rarely produce antibodies to tumors. In contrast, mice immunized with human cancers do make antibodies that react strongly to those same human cancer cells. In some cases, the mouse antibodies bind to nearly all the tumor cells of one cancer in a test tube—even if they have been taken from many different individuals with the same kind of cancer. The mouse antibodies, though, are usually not effective in killing the cancer cells in humans. Even if the murine antibodies do have cancer-killing activity in a patient, the response is usually very short-lived because the patient soon produces inactivating antibodies against the mouse antibodies.

Therefore, oncologists have long hoped to find some way to combine the targeting ability of the murine antitumor antibodies with the killing ability of human *T* cells. Recombinant DNA technology offers the necessary tools. Researchers have successfully isolated the antitumor antibody genes from mouse cells and recombined parts of them with gene segments encoding the receptor that killer *T* cells use to recognize their targets. The modified receptor gene redirects killer *T* cells, which often do not recognize cancers, to see what the less discriminating mouse antibodies see. Indeed, killer *T* cells rearmed with chimeric *T* cell receptors kill cancer cells in a test tube quite efficiently. Early clinical experiments using this strategy are now under way in cancer patients, as well as in those infected with HIV, the AIDS-causing agent [*see box on these two pages*], and other pathogens.

Gene Therapy for AIDS

Although about half of all clinical gene-therapy research today focuses on cancer, the next largest group of studies—about 10 percent—is devoted to combating infection by the AIDS-causing human immunodeficiency virus (HIV). Experimental gene therapies for HIV aim to do one of two things: stop HIV from replicating inside infected cells, or prevent the virus from spreading to healthy ones.

The ultimate target for all these efforts would be stem cells, which develop into immune and blood cells. These cells might then be rendered resistant to HIV long before they had matured. For now, though, researchers can at best test various strategies on blood cells called monocytes and on so-called helper, or CD4, *T* cells, the immune cells most heavily ravaged by HIV. It is not difficult to isolate *T* cells from someone's blood, give them a therapeutic gene and then return them to that same person or another recipient.



MARTIN H. SIMON *SABA*

TWIN of an AIDS patient donates cells that will be genetically altered and given to his brother.

Some researchers hope one day to administer therapeutic genes in vivo through delivery vehicles, or vectors, that can seek out infected or otherwise susceptible cells. The favored vectors at the moment are viruses. These viruses—which so far include adeno-associated virus and such retroviruses as HIV itself—are modified so that they are no longer pathogenic but can still pass genetic information on to human cells. HIV particles altered to serve as vectors, for example, are attracted to the same CD4 *T* cells as wild-type HIV particles are but do not multiply dangerously in cells. Thus, these vectors

Other Gene Therapies

Immunotherapy aside, cancers can be battled on other genetic fronts. There has been intense interest in identifying the precise DNA defects that cause cancer. Some mutations, scientists have learned, are associated with specific types of cancer. Other mutations occur in many varieties. Furthermore, there are different kinds of mutations. Some activate genes, called oncogenes, that drive uncontrolled growth in cells. Other mutations—those in so-called tumor suppressor genes—result in the loss of a normal brake on uncontrolled cell growth.

One of the most commonly mutated tumor suppressor genes in human cancer is *p53,* a gene whose protein product normally monitors the DNA in a cell as it divides. If the DNA is flawed, the p53 protein may halt cell division until the damage can be fixed or may induce cell suicide (apoptosis). When a normal copy of *p53* is reintroduced to cancer cells in tissue culture, those cells return to a more regular growth pattern or self-destruct. Either outcome would be useful in cancer treatment, and so a great deal of effort has gone into developing methods for inserting normal *p53* genes into cancers growing in the body.

There are still major roadblocks: as Theodore Friedmann notes in the first article of this section on page 96, current technologies for delivering genes to specific organs or cell populations are inefficient. In addition, there are no perfected means for extending the effects of such locally delivered genes to other ar-

are capable of providing therapeutic genes to precisely those cells that most need them.

A range of genes are being tested for their ability to disable HIV. One gene type, containing what is called a dominant negative mutation, generates inactive versions of proteins that HIV normally makes in order to replicate. When an infected, treated cell produces these useless look-alikes, the altered proteins trip up their ordinary cousins—either by binding to them or by taking their place in molecular reactions. Clinical testing of a dominant negative mutation of the HIV gene *rev* began in 1995.

Scientists are also evaluating the merit of delivering genes that would be transcribed into short RNA strands that mimic essential viral control RNAs. The hope is that these RNA decoys might bind to HIV regulatory proteins and block them from functioning. Genes transcribed into ribozymes (catalytic RNAs) capable of degrading viral RNA might similarly interrupt HIV replication. A related idea involves delivering genes encoding proteins that are made by the host cell and that interact with HIV particles. For instance, soluble forms of the protein CD4 might bind to HIV particles extracellularly, thereby keeping them from infecting *T* cells that display CD4 molecules on their outer surface.

Scientists are also exploring for HIV treatment so-called suicide genes, which are not unlike those being tested as cancer gene therapies. Because the gene would presumably get into any cell normally invaded by the selected vector, researchers want to be sure the suicide gene will be expressed only in the subset of recipient cells that harbor HIV infection. So they plan to attach the gene to control elements that become active, and switch on the gene, only in cells that are infected by HIV.

Another design borrowed from gene therapies for cancer relies on enhancing the ability of functioning helper *T* cells to recognize infected cells and orchestrate an immune response against them. For instance, a gene coding for part of the antibody molecule that recognizes and binds to the gp120 protein on HIV's surface can be integrated with genes encoding the molecule, or receptor, on killer *T* cells that is normally responsible for recognizing diseased cells. The chimeric receptor that results from this mix takes better notice of HIV and thus redirects the *T* cells to destroy the infected HIV cell. A Phase 1 clinical trial (looking at safety) is currently under way.

Other therapeutic genes being scrutinized give rise to antibody fragments that act within infected cells. By binding to some newly made viral protein, these intracellular antibodies, or intrabodies, prevent virus particles from being assembled. Finally, clinical trials of gene vaccinations for preventing AIDS have been approved. As in other gene vaccinations, these immunizations supply to cells patrolled by the immune system genes coding for HIV molecules that distinguish the virus as a foreign invader. The immune system then reacts to these antigens by producing antibodies that wander through the body, ready to attack any cells presenting the antigens, should they ever appear. —*R.M.B.*

eas in the body. Until physicians can do so, these gene therapies will help tackle only tumors at isolated sites.

Even so, animals have shown significant improvements when the *p53* gene is delivered either through the bloodstream (in complexes with lipids that allow cells to take up the gene) or to tumors directly (using modified viruses to shuttle the gene into cells). An early clinical trial has reported some tumor regressions at local sites. In theory, though, there is one major limitation to using gene transfers to activate tumor suppressor genes or to neutralize oncogenes—the corrective gene must be delivered to every tumor cell. Otherwise, the unaccessed cells will continue growing uncontrollably. It is impossible to correct the genes in every tumor cell—even those in a single site—using current technology. And although additional treatments might help correct more tumor cells, repeated gene transfers using modified viruses often are not feasible; the immune system frequently recognizes the virus the second time and destroys it before it can deliver genes to tumors.

Fortunately, though, the beneficial effects of an initial injection sometimes appear to reach cells that have not been gene corrected. Indeed, several different experimental gene therapies for cancer report the appearance of a "bystander effect." This phenomenon is invoked to explain why a treatment sometimes kills a higher proportion of tumor cells than can be accounted for by the number of cells actually expressing some new gene. Researchers have reported this kind of discrepancy in some *p53* gene therapy trials but cannot yet explain it: presumably, if normal *p53* genes did generate a bystander effect, cancer would not develop in the first place. But the bystander effect has been seriously studied in conjunction with other treatments, such as "suicide" gene therapy, in which a gene inserted into a cancer cell renders it supersensitive to some drug that ordinarily has no anticancer effect.

In the original application of suicide gene therapy, my colleagues Edward H. Oldfield, Zvi Ram and Ken Culver and I inserted the gene for an enzyme called thymidine kinase (tk) from a herpesvirus into cancerous brain cells of patients. In cells infected with the herpes simplex virus, this enzyme can convert the otherwise nontoxic drug ganciclovir into a metabolite, or by-product, that acts as a potent viral killer. We found that this same toxic metabolite could kill dividing cancer cells; in some tumors, it killed neighboring cancer cells as well. To create this bystander effect, the toxic metabolite spread from the cell in which it was produced to its neighbors via gap junctions—channels that allow small compounds to move between cells. In the original clinical trial testing this treatment for brain tumors, about one quarter of the patients responded. And clinicians are testing other suicide gene therapies involving different anticancer compounds, some of which are also expected to produce a bystander effect.

In various early explorations of gene therapy technology, researchers are just beginning to learn about its potential—and its limitations. As with so many other new and unexplored areas of science, some ideas will probably prove useful; many more will fall by the wayside. Ideas that are unworkable now may eventually become highly successful, when our technological capabilities increase. Even though in the future current methods for using genes for treatment will be looked back on as crude and inefficient, these methods have already offered important lessons. And they have indicated many new paths in the quest for cancer control. SA

The Author

R. MICHAEL BLAESE is chief at the Clinical Gene Therapy Branch of the National Human Genome Research Institute of the National Institutes of Health. Work in his laboratory there has focused on developing gene therapies for a range of inherited diseases, as well as for cancer and AIDS. Blaese received his M.D. from the University of Minnesota in 1964.

Gene Therapy for the

Inserting genes into brain cells may one day offer doctors a way to slow, or even reverse, the damage from degenerative neurological disease

by Dora Y. Ho and Robert M. Sapolsky

The prospect of acquiring any chronic illness is disturbing. But for most people, the threat of neurological impairment evokes a special dread. Afflictions such as Parkinson's disease or amyotrophic lateral sclerosis (Lou Gehrig's disease) progressively rob control of the body. Damage to the spinal cord can create equal misery in just an instant. And Alzheimer's disease attacks the very essence of one's personality as it destroys the mind.

Unfortunately, physicians and medical researchers have made only limited progress in the battle against such diseases, in large part because the brain and spine are so vulnerable. Unlike many types of cells, neurons (nerve cells) in the central nervous system of adults are typically unable to divide. That fact of life creates the central tragedy of neurological illness or injury: under normal circumstances, neurons that are lost are gone for good, and injured nerve tissue of the brain and spinal cord cannot be expected to repair itself.

But scientific advances might yet change that grim situation. Some of the most ambitious research in neurology aims to replace lost cells in damaged tissue by transplanting neurons or by delivering growth factors—chemicals that can stimulate surviving neurons to extend their reach or that can awaken the cells' dormant ability to regenerate. Such therapies would be immensely beneficial, but it will probably take many years before they become routine. Preventing neuron loss in the first place is a more modest goal—one that is perhaps not so distant.

In the past few years, researchers have learned a great deal about how neurons die after a sudden medical insult such as a stroke, seizure or head injury, as well as during progressive diseases such as Parkinson's or Alzheimer's. Some attempts to take advantage of these recent discoveries suggest that administering certain drugs may protect threatened neurons or even that lowering the temperature of the brain can avert the death of fragile cells during a neurological crisis. What is more, new knowledge about how neurons succumb to various diseases has raised the exciting possibility of protecting these cells by modifying their genes.

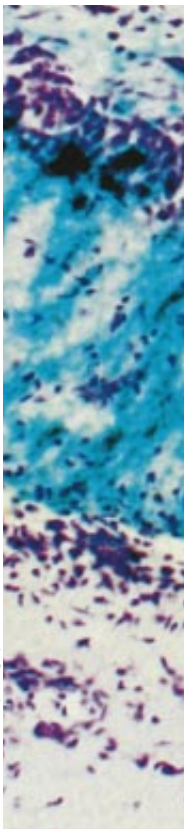
Reprogramming for Survival

Genes instruct cells to make specific proteins, such as the enzymes that catalyze various chemical reactions. Nerve cells, for example, produce enzymes that synthesize neurotransmitters—substances that carry chemical signals across the tiny gaps (synaptic spaces) between one neuron and another. Gene therapy targeted to failing neurons could potentially provide them with a gene specifying a protein that is able to shield these cells from whatever threat may loom.

To create such remedies, researchers must first decide what kinds of proteins would be most helpful. In some cases, the goal would be to augment the production of a particular brain protein when the naturally occurring version is dysfunctional or made in inadequate amounts. Or, in theory, one might want to add a novel protein, found in a different type of tissue—or even in a different organism entirely.

Another strategy, called the antisense approach, also constitutes a form of gene therapy [see "The New Genetic Medicines," by Jack S. Cohen and Michael E. Hogan; Scientific American, December 1994]. Antisense tactics aim to limit the manufacture of proteins that are doing damage. Some types of amyotrophic lateral sclerosis and certain other neurological diseases result from the destructively intense activity of a normal protein or from the action of an abnormal protein that works in an injurious manner. Antisense therapies might also help when neurons synthesize proteins that (for reasons that are still inexplicable) exacerbate a neurological crisis. To that end, many researchers are now trying to find ways to block the production of so-called death proteins, which induce endangered neurons to commit cellular suicide.

Once the basic understanding of a particular neurological disease is in place, it does not take great imagination to come up with a list of genes that might save neurons from destruction. The challenge is in figuring out how to deliver those genes. In principle, one can insert a gene into brain tissue by directly injecting appropriately coded segments of pure DNA. Unfortunately, this brute-force method is rarely successful, because neurons are not particularly efficient at picking up such "naked" DNA. A better technique is to encase



ROBERT M. SAPOLSKY (*photograph*); SLIM FILMS (*virus*); TOMO NARASHIMA (*cat brain*)

Nervous System

the gene in a fatty bubble called a liposome. Because of the chemical nature of these tiny containers, liposomes easily transport DNA into target neurons by fusing with the cell membrane and releasing their contents into the interior. The cell, for reasons not entirely understood, will then incorporate some of this material into its nucleus, where its own DNA resides, and will use the gene as a blueprint for making the therapeutic protein [see "Nonviral Strategies for Gene Therapy," by Philip L. Felgner, on page 102].

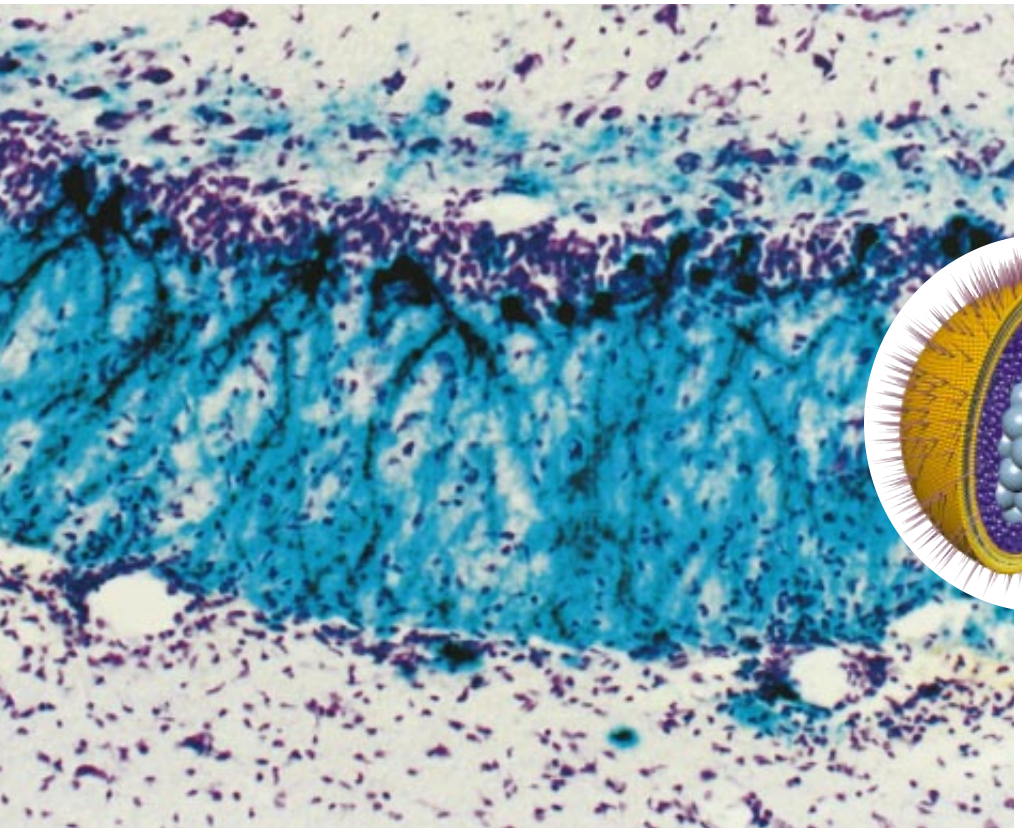
An even better tool for putting genes into cells is a virus. In the course of a typical infection, viruses insert their genetic material into cells of the victim, where this added genetic code directs the synthesis of various molecules needed to make new viral particles. Although natural viruses can be immensely destructive, scientists can tame and convert some of them into microscopic Trojan horses, which then can carry a therapeutic gene and quietly deposit it inside a cell without causing damage. For gene therapy in the central nervous system, investigators are focusing much of their effort on just a few viral types, including adenoviruses and herpesviruses.

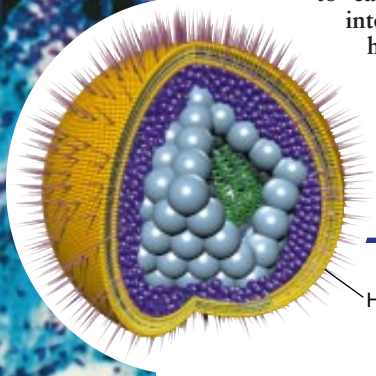
Combating Parkinson's

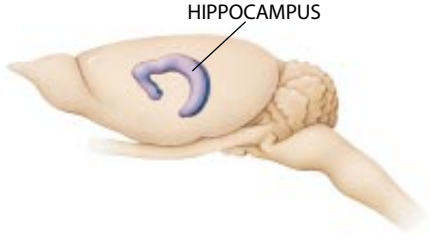
Experiments with these viral vectors, as such delivery vehicles are called, have provided the first hints that gene therapy in the nervous system can work. One promising area of research is directed against Parkinson's disease. This devastating disorder arises because a part of the brain, known as the substantia nigra, degenerates over time. This region helps to regulate motor control, and its destruction makes it hard for a person to initiate movements or execute complex coordinated motion. The loss also brings on the classic parkinsonian tremor [see "Understanding Parkinson's Disease," by Moussa B. H. Youdim and Peter Riederer; SCIENTIFIC AMERICAN, January].

Parkinson's disease ensues after the death of nigral neurons that secrete the neurotransmitter dopamine. For complex reasons, these neurons also generate oxygen radicals, rogue chemical groups that cause damaging reactions within the cell. Hence, there is a fair amount of ongoing destruction in the substantia nigra as a normal part of aging. (This process contributes to the mild tremor typical of senescence.) Sometimes Parkinson's disease appears to strike people who are predisposed to having an excess of oxygen radicals in their brain tissue or who have been exposed to environmental toxins that cause these oxygen radicals to form. Other cases seem to involve people who have normal amounts of these chemicals but who have impaired antioxidant defenses.

Whatever the underlying cause, it is clear that the symptoms of Parkinson's disease result primarily from the absence of dopamine after too many neurons in the substantia nigra die. Thus, a straightforward way to correct this deficit, at least temporarily, would be to boost the amount of dopamine where it is in short supply. Dopamine is not itself a protein,



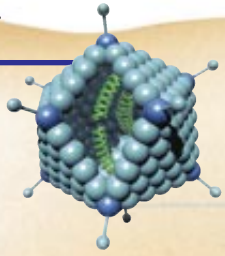




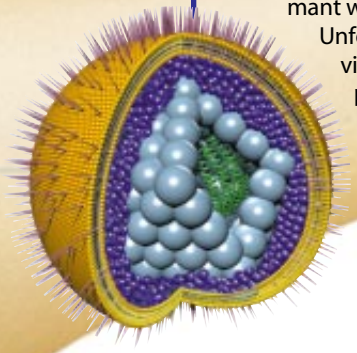
NEURONS (*bright blue in photomicrograph*) **from the hippocampal region of a rat brain have taken up a herpesvirus that was engineered to carry a gene specifying a protein that can turn cells blue when suitably prepared for microscopic analysis. This ability to carry foreign genes into neurons raises hope that modified viruses can be engineered to provide therapeutic genes to the human brain.**

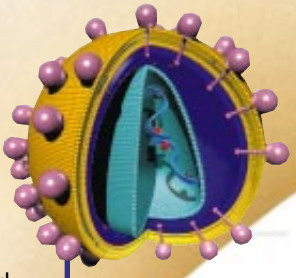
A Toolbox of Gene Delivery Agents for Nerve Cells

ADENOVIRUS, a pathogen that commonly causes respiratory ailments, is capable of infecting neurons. But in its naturally occurring form, it can damage cells and evoke a strong immune response, so care must be taken to inactivate it and minimize its immunogenicity.



HERPES SIMPLEX VIRUS type 1, the agent that causes common cold sores, is especially able to carry genes into nerve cells. (Between outbreaks, the cold sore virus often remains dormant within sensory neurons.) Unfortunately, like adenoviruses, unmodified herpesviruses damage cells and cause an immune response.





RETROVIRUS incorporates its genes into the DNA of the host cell. Many retroviruses infect only cells that divide regularly and thus cannot be used to treat neurons. Others (in the so-called lentivirus family, which includes HIV, the AIDS virus) can infect cells that do not divide, and so these retroviruses may someday serve in gene therapy for the nervous system.

ADENO-ASSOCIATED VIRUS does not damage infected nerve cells or induce an immune response. It is also much more compact than other viruses under study for gene therapy. Consequently, it may be more successful at traversing the small pores that allow few substances to cross the blood-brain barrier. But the small size may limit the amount of therapeutic genetic information that can be loaded into the virus.

TOMO NARASHIMA (*neuron*); SLIM FILMS (*viruses*)

but the enzymes that synthesize this neurotransmitter are. So increasing the manufacture of one enzyme critical to that process (tyrosine hydroxylase) should enhance the synthesis of this much needed brain chemical for as long as the dopamine-producing cells of the substantia nigra survive.

Although administering a chemical precursor to dopamine—a substance called L-dopa—also works to augment levels of this neurotransmitter, the drug reacts throughout the brain, causing serious side effects. The lure of gene therapy in this context is that corrective changes would take effect just within the substantia nigra.

Several scientists have been working hard to exploit this possibility. In a pair of recent collaborative studies, five research teams reported success using herpesviruses as gene vectors to correct symptoms in rats that were surgically treated in a way that caused them to exhibit some of the manifestations of Parkinson's disease. Application of gene therapy increased the production of the corrective enzyme, raised the level of dopamine near the cells that had been deprived of this neurotransmitter and partially eliminated the movement disorders in these animals.

Dale E. Bredesen and his colleagues at the Burnham Institute in La Jolla, Calif., recently explored an even more sophisticated scheme. Investigators had shown previously that transplanting neurons from the substantia nigra of fetal rats corrected some of the parkinsonian defects that were surgically induced in adult rats. This strategy worked because the robust young neurons were able to grow and produce dopamine for the nearby cells in need. A problem emerged, however. For some reason, the grafted neurons tended to activate an internal suicide program (a process termed apoptosis) and died after a while. So Bredesen and his co-workers carried out gene therapy on the fetal neurons before transplanting them; the researchers hoped to coax these cells to produce large quantities of a protein called bcl-2, which suppresses cell suicide.

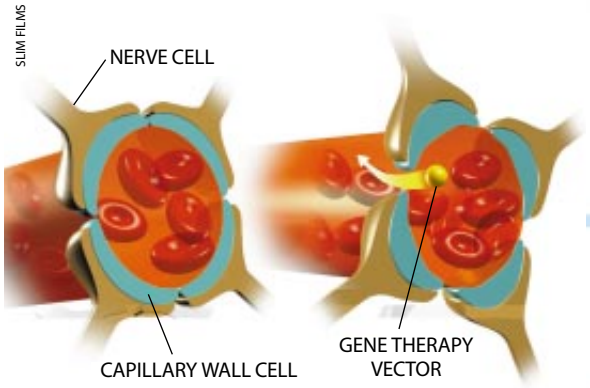
The result was dramatic: four weeks later the rats that had received standard grafts were only marginally better, whereas the creatures that obtained the added gene in their grafts were substantially improved. Treatment for Parkinson's disease would require a longer period of effectiveness still—one would want the grafts to survive for years. Physicians have already carried out human fetal cell transplants to help patients with severe Parkinson's disease, but these attempts have met with mixed results. Perhaps one or two clever gene modifications to the human fetal cells before they are transplanted would make that procedure work much better.

Battling Stroke

The success of current research with animals indeed sparks hope that new treatments will eventually emerge for Parkinson's and other progressive degenerative diseases of the brain. Gene therapy also offers the prospect of stemming tissue damage during such acute

neurological crises as the overstimulation of a seizure or the loss of oxygen and nutrients that occurs during a stroke.

Under these conditions, the most vulnerable cells in the brain are the many neurons that respond to an extremely powerful neurotransmitter called glutamate. Glutamate normally induces recipient neurons to take up calcium, which causes long-lasting changes in the excitability of synapses stimulated by this neurotransmitter. This process may, in fact, be the cellular basis of memory.



BLOOD-BRAIN BARRIER consists of tightly packed cells that line the capillaries within the brain (*left*). The tiny spaces between these cells allow only small molecules to reach this organ's many neurons (*tan*). It is possible that certain treatments could be devised that would widen these gaps temporarily (*right*), permitting larger substances (such as therapeutic viral vectors) to pass into brain neurons.

But during seizure or stroke, neurons are unable to mop up glutamate from synapses or clear the tidal wave of calcium that floods into many brain cells. Instead of fostering mild changes in the synapses, the glutamate and calcium do serious damage: the cellular architecture of the affected neurons crumbles, and newly generated oxygen radicals create further havoc. This destruction then kills cells directly or signals the initiation of internal suicide programs that will cause the swift demise of the flagging neurons.

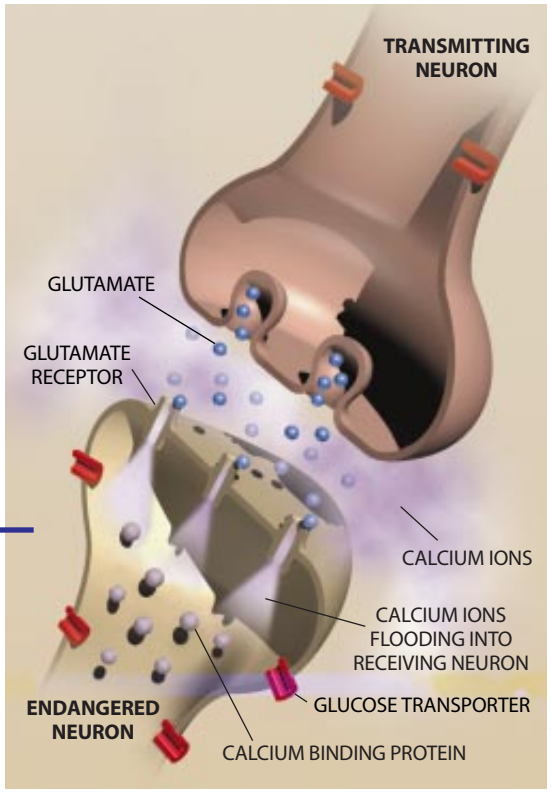
Our group has examined the possibility that gene therapy could interrupt this calamitous sequence of events. For our first experiments, we extracted some brain cells from a rat and cultured them in a petri dish. We then subjected these neurons to a modified herpesvirus engineered to carry a gene for a protein that transports energy-rich glucose molecules across the cell membrane. In a patient suffering a neurological crisis, a similar type of therapy might increase the influx of glucose just when the beleaguered neurons would benefit most from extra energy (which is needed, among other tasks, to pump the excess calcium out of these cells).

Early experiments showed that our treatment enhanced the uptake of glucose and helped to maintain proper metabolism in neurons subjected to the test-tube equivalent of seizure or stroke. We later found that we could lessen the damage from stroke in rats by injecting the viral vector into the vulnerable region of the brain before an injurious event occurred. It is obviously not possible for a person to forecast when a seizure or a stroke will happen. But, as Matthew S. Lawrence and Rajesh Dash discovered when they worked in our laboratory, there is a window of a few hours after a seizure when the gene treatment to these rats still helps to protect neurons from additional damage—which suggests that humans, too, might one day benefit from a similar kind of therapy.

Another form of potential gene therapy for stroke and trauma targets the activation of suicide programs. Howard J. Federoff and his colleagues at the University of Rochester, and Lawrence, working with our group, independently constructed herpesvirus vectors that included the suicide-suppressing *bcl-2* gene. Application of this vector tends to shield the brain cells of rats from damage when crisis conditions occur, even if the treatment begins only after the insult.

Investigators also report advances that might one day prevent the so-called lipid storage diseases—genetic disorders that cause defects in certain enzymes and lead to a fatal accumulation of fatty molecules in the brain. To explore such therapies, researchers have produced mice that carry a mutation in the gene encoding the enzyme beta-glucoronidase. (People suffering from a rare malady called Sly syndrome have the same gene mutation.) John H. Wolfe and his colleagues at the University of Pennsylvania transplanted into an afflicted mouse fetal neurons engineered to produce beta-glucoronidase. These researchers found that the implants were able to dispose of damaging lipids throughout the animal's brain.

Despite the many encouraging first



CRISIS CONDITIONS ensue during a seizure or stroke when too much of the neurotransmitter glutamate (*blue*) accumulates in the synaptic gap between neurons. Specialized molecules, called glutamate receptors, on the receiving cell (*tan*) then allow calcium ions (*purple*), which are normally present in high concentration only on the outside, to flood into the cell's interior and cause permanent damage. Future genetic therapies could augment the amount of calcium binding protein (which safely sequesters these ions) or increase the number of glucose transporters (which would bring more energy-rich molecules to fuel the endangered neuron).

THOMAS HOEPKER *Magnum Photos*

WILLEM DE KOONING (1904–1997), the noted abstract painter, suffered late in his life from Alzheimer's disease, a crippling affliction that strikes at least one in five Americans older than 85. Investigators are optimistic that Alzheimer's disease, Parkinson's disease and other common neurodegenerative disorders will eventually be ameliorated by genetic therapies of the kind being pioneered by the authors and others today.

steps in applying gene therapy to the nervous system, sundry hurdles remain. For example, significant problems persist in engineering viral vectors. Cripple the virus too much, and it becomes difficult to maintain sufficient potency to infect cells. Strip away too little, and the virus will damage the host neurons. Because the viral vectors now available each suffer from one or the other of these shortfalls, a great deal of refinement will be needed before scientists can safely begin testing gene therapies in people with neurological disease.

Another difficulty emerges simply because the brain—a vital but delicate organ—is encased in a relatively impregnable skull. Thus, injecting a therapeutic drug directly into the affected tissue is rather difficult. Most researchers conducting animal studies resort to neurosurgery: drilling a hole in the skull and injecting the vector directly into the endangered part of the brain. But, clearly, human patients would require something less invasive for routine treatment. Although one could give a vector intravenously (if it could be designed to enter only nerve tissue), the virus would be unlikely to get past the blood-brain barrier, a specialized network of capillaries that lets only small molecules pass into brain tissue. So, without further special measures, virtually all of the viral vector would wastefully end up in places other than the brain.

Even if these obstacles could be overcome, some final stumbling blocks would still stand in the way. After a viral vector reaches a patch of neurons, it does not go far. (Viruses that are able to replicate can spread readily in brain tissue, but these agents cannot be used for gene therapy, because they invariably provoke a damaging immune response.) A safe viral vector traverses a limited area, where it infects only a small percentage of neurons. Hence, these viruses are not particularly effective in reaching diseased tissue. Furthermore, for most of the vectors tried so far, activity persists for a few weeks at most—too brief a period to combat slow but relentless degenerative illnesses. So researchers will need to find ways to improve the spread, efficiency and duration of these engineered infections.

Future Shocks?

Despite the many challenges, studies of gene therapy for diseases of the nervous system—like most ambitious efforts that have shown some initial successes—have generated an aura of optimism. Perhaps with adequate effort, gene therapy for the brain will eventually be commonplace.

A glimpse of the possibilities ahead comes from the work of Anders Björklund and his colleagues at the University of Lund in Sweden. These researchers, who pioneered methods to transplant fetal neurons, have engineered grafts to produce large quantities of a nerve growth factor. They implanted some of these engineered cells into mature rats, targeting a region of the brain that is critical for learning and memory—an area that, not surprisingly, slowly degenerates during normal aging. Remarkably, this maneuver reversed cognitive decline in aged rats.

That success suggests that gene therapy might serve not just to blunt the edges of disease but to improve memory, sensation and coordination in older people. At present, scientists have enormous strides to make before they can hope to aid geriatric patients in this way. But ultimately, gene therapists may be able to offer powerful medicines for rejuvenating aging brains.

Such treatments might also be able to make younger people's minds work "better than well," to borrow a now popular phrase describing the effects of Prozac. Few areas of medical research pry at such a Pandora's box as does work on improving normal brain function. But this prospect—and its possible abuses—will be difficult to sidestep if scientists are to continue to pursue genetic treatments directed at specific neurological diseases. So further research on applying gene therapy to the nervous system, like some other swiftly moving currents in the flow of biomedical inquiry, will surely force vexing ethical questions to float to the surface. SA

The Authors

DORA Y. HO and ROBERT M. SAPOLSKY have worked together for almost seven years. Ho, a researcher in the department of biological sciences at Stanford University, investigated the genetics of the herpes simplex virus during her Ph.D. studies, which she completed in 1990. Since that time, she has focused her work on developing herpesviruses for gene therapy and understanding the molecular biology of diseases of the central nervous system. Sapolsky, a professor of biological sciences and neuroscience at Stanford, initially concentrated his research on the physiology of stress, an enterprise that at times takes him to Kenya to study baboons in the wild. In 1990 he shifted part of his research effort from stress hormones in the brain toward gene therapy for disorders of the nervous system.

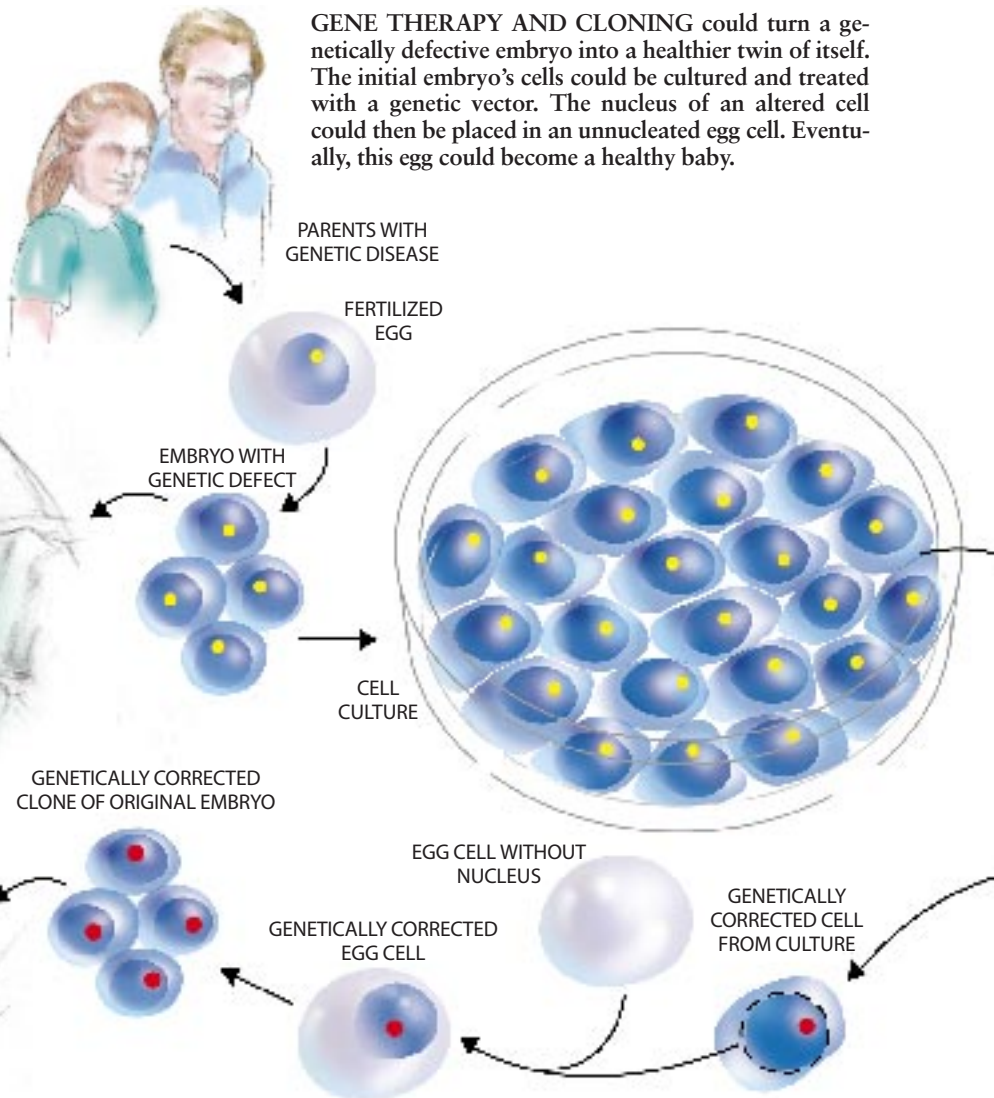
What Cloning Means

The recently debuted technology for cloning is usually discussed as a means of creating genetic copies of whole adult individuals. This is far from its only use, however. Cloning could be combined with other biotechnologies, either to achieve more novel goals or to improve on previous methods. Although the technique is still in its infancy, and needs to be studied and developed much further, educated musings about cloning's ability to inform gene therapy are already being brought to the table. An area that might particularly benefit is germ-line gene therapy—genetic modifications that could correct a problem for future generations. "I think cloning is going to be used as a tool that will make gene therapy work," comments Lee Silver, a molecular biologist at Princeton University and an expert on reproductive technologies. "For the first time, germ-line gene therapy becomes realistic."

Germ-line therapy, which is not yet being studied in humans, could ideally prevent deadly or debilitating disorders such as sickle cell anemia or cystic fibrosis. Such diseases are typically transmitted silently from generation to generation by people carrying one copy of a defective gene; the disease becomes manifest when two carriers have a child who inherits two copies.

Today prenatal genetic testing can reveal whether a fetus or embryo is affected with many of these conditions. The parents then have the option of aborting and rolling the genetic dice again with another pregnancy. In some cases, however, the dice are guaranteed to come up snake eyes. "If both parents are sickle cell diseased," Silver says, "then all of their embryos will also carry the disease. You can't select, because there are no good embryos." But gene therapy, aided and abetted by cloning, could theoretically correct the condition for their children, and all subsequent progeny as well.

The recipe would begin with a fertilized egg growing, in the laboratory, into



GENE THERAPY AND CLONING could turn a genetically defective embryo into a healthier twin of itself. The initial embryo's cells could be cultured and treated with a genetic vector. The nucleus of an altered cell could then be placed in an unnucleated egg cell. Eventually, this egg could become a healthy baby.

DANA BURNS-PIZER

for Gene Therapy

a mass of early embryonic tissue. A functioning gene—say, for the blood's oxygen-carrying protein, beta globin, which is mutated in sickle cell anemia—would then be inserted into the embryonic cells by tailored viruses or other vectors. (A marker sequence inserted along with the gene might help identify which cells took up the gene correctly.) The DNA of one of those cells could then be implanted into a new egg cell from the mother, beginning the pregnancy afresh. In effect, this last step replaces the original embryo with a healthier clone of itself.

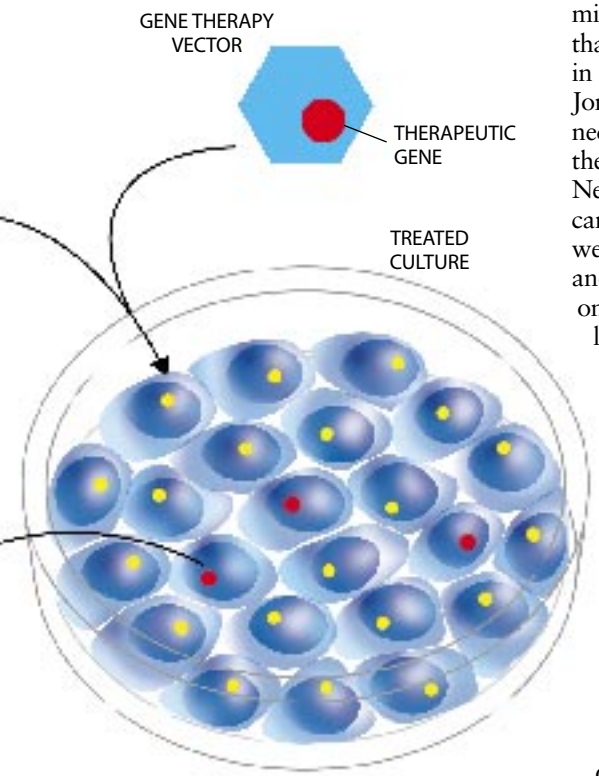
Germ-line therapy does not require a cloning step, but cloning might make it far easier. Very early stage embryonic cells, if separated, retain the ability to regenerate into whole embryos (indeed, that is how identical twins, triplets and quadruplets arise). Gene therapists could therefore alter the DNA of the embryonic cells and return one to the mother for gestation. The problem is that embryonic cells lose their "pluripotent" capacity after a few cell divisions, so the gene therapists would be forced to work on relatively few cells. The inefficiency of current gene manipulation techniques would consequently undermine many therapeutic attempts. With cloning, however, the age and number of cells eligible for manipulation is unlimited.

In theory, cloning would allow therapy on cells from a more advanced pregnancy (although this would raise more troubling ethical issues for many parents). In a variation on this theme, the gene therapy might also be conducted on cells from one of the parents. A child cloned from those altered cells would be free of the genetic defect but in other ways a genetic duplicate of its donor parent.

Cloning may remove some of the practical barriers to germ-line gene therapy, but it does not alter the ethical ones. Many researchers, not to mention the general public, are deeply concerned that germ-line techniques could be misapplied toward eugenic goals with authoritarian or even genocidal overtones. So even if cloning does enable the technology, there may not be a sudden rush to perform germ-line gene therapy.

Cloning may also benefit somatic gene therapy as a tool for basic research. By making it easy to obtain large numbers of genetically identical cells for study, cloning should help elucidate how embryonic cells commit to become a particular cell type. "That process of commitment involves shutting off genes that would otherwise have played a role in becoming a liver or a brain," reflects Jon Gordon, professor of obstetrics, gynecology and reproductive science at the Mount Sinai School of Medicine in New York City. "I think the fact that we can now reverse that gives us hope that we can understand that process better and understand diseases that are based on or manifest as errors in this process, like cancer." Cloning might therefore help therapists determine which genes they should be aiming to correct in various illnesses. If so, cloning's greatest utility may not be for making more people but for making more people healthy.

—Steve Mirsky and John Rennie



STEVE MIRSKY is Reuter Fellow in Medical Journalism at Columbia University. JOHN RENNIE is editor in chief of Scientific American.

Further Readings on Gene Therapy

Obstacles to Therapy

Gene Therapy for Human Genetic Disease. Theodore Friedmann and Richard Roblin in *Science*, Vol. 175, pages 949–955; March 3, 1972.

Human Somatic Gene Therapy: Progress and Problems. M. K. Brenner in *Journal of Internal Medicine*, Vol. 237, No. 3, pages 229–239; March 1995.

Recombinant DNA Research. Agency: National Institutes of Health. *Federal Register*, Vol. 61, No. 131, pages 35774–35777; Monday, July 8, 1996.

Nonviral Gene Delivery

Direct Gene Transfer into Mouse Muscle in Vivo. Jon A. Wolff, Robert W. Malone, Phillip Williams, Wang Chong, Gyula Acsadi, Agnes Jani and Philip L. Felgner in *Science*, Vol. 247, pages 1465–1468; March 23, 1990.

Direct Gene Transfer for Immunotherapy. G. J. Nabel and P. L. Felgner in *Trends in Biotechnology*, Vol. 11, No. 5, pages 211–215; May 1993.

DNA Vaccines. J. J. Donnelly, J. B. Ulmer, J. W. Shiver and M. A. Liu in *Annual Review of Immunology*, Vol. 15, pages 617–648; 1997.

Lipidic Vector Systems for Gene Transfer. R. J. Lee and L. Huang in *Critical Reviews in Therapeutic Drug Carrier Systems*, Vol. 14, No. 2, pages 173–206; 1997.

Cancer and AIDS

Gene Transfer as Cancer Therapy. Glenn Dranoff and Richard C. Mulligan in *Advances in Immunology*, Vol. 58, pages 417–454; 1995.

Steps toward Gene Therapy, 2: Cancer and AIDS. R. M. Blaese in *Hospital Practice*, Vol. 30, No. 12, pages 37–45; December 15, 1995.

Gene Therapy Strategies for Novel Cancer Therapeutics. Maryland E. Rosenfeld and David T. Curiel in *Current Opinion in Oncology*, Vol. 8, No. 1, pages 72–77; January 1996.

Nervous System Disorders

Long-Term Behavioral Recovery in Parkinsonian Rats by an HSV Vector Expressing Tyrosine Hydroxylase. M. J. During et al. in *Science*, Vol. 266, pages 1399–1403; November 25, 1994.

Use of Herpes Simplex Virus Vectors for Protection from Necrotic Neuron Death. D. Y. Ho et al. in *Viral Vectors: Gene Therapy and Neuroscience Applications*. Edited by M. Kaplitt and A. Loewy. Academic Press, 1995.

Gene Transfer to Neurons Using Herpes Simplex Virus-Based Vectors. D. J. Fink, N. A. DeLuca, W. F. Goins and J. C. Glorioso in *Annual Review of Neuroscience*, Vol. 19, pages 245–287; 1996.

Bringing Schrödinger's Cat to Life

by Philip Yam, *staff writer*

I am sorry that I ever had anything to do with quantum theory," Erwin Schrödinger reportedly complained to a colleague. The Austrian physicist was not lamenting the fate of his now famous cat, which he figuratively placed in a box with a vial of poison in 1935. Rather he was commenting on the strange implications of quantum mechanics, the science behind electrons, atoms, photons and other things submicroscopic. With his feline, Schrödinger attempted to illustrate the problem: according to quantum mechanics, particles jump from point to point, occupy several places at once and seem to communicate faster than the speed of light. So why don't cats—or baseballs or planets or people, for that matter—do the same things? After all, they are made of atoms. Instead they obey the predictable, classical laws quantified by Isaac Newton. When does the quantum world give way to the physics of everyday life? "That's one of the $64,000 questions," chuckles David Pritchard of the Massachusetts Institute of Technology.

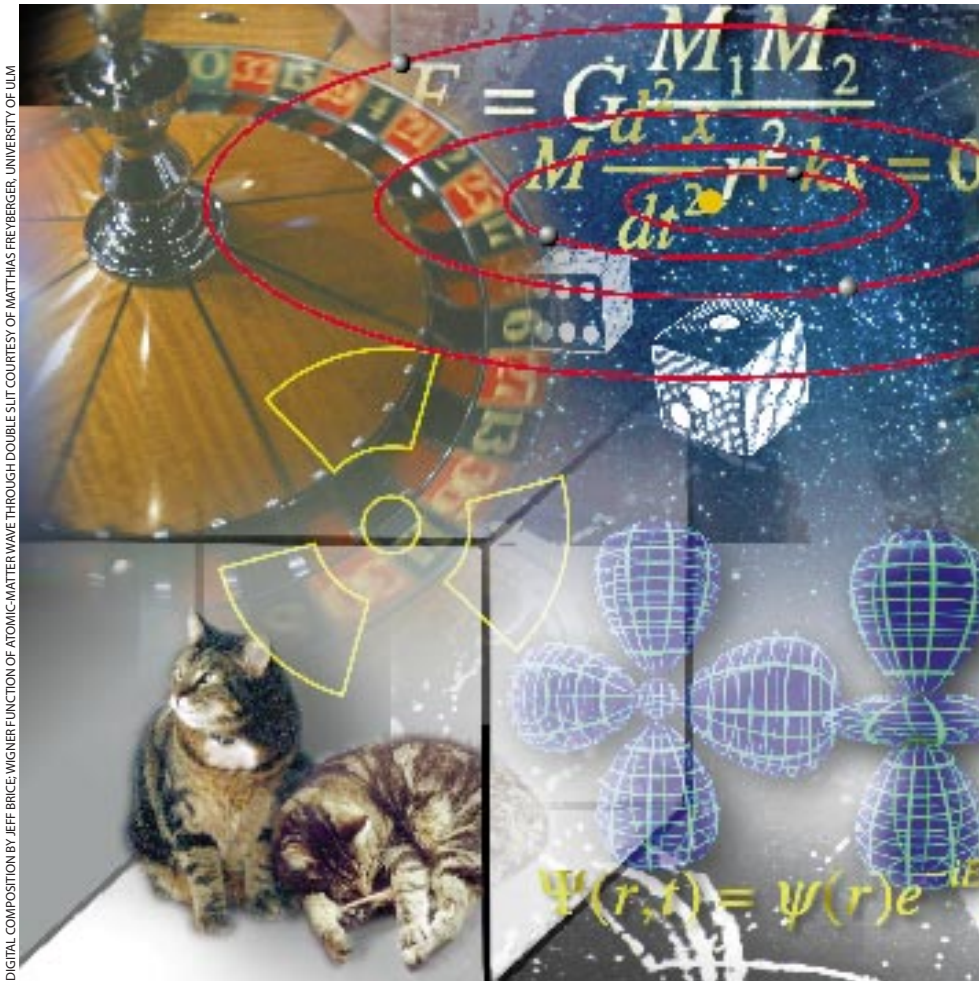
Pritchard and other experimentalists have begun to peek at the boundary between quantum and classical realms. By cooling particles with laser beams or by moving them through special cavities, physicists have in the past year created small-scale Schrödinger's cats. These "cats" were individual electrons and atoms made to reside in two places simultaneously, and electromagnetic fields excited to vibrate in two different ways at once. Not only do they show how readily the weird gives way to the familiar, but in dramatic fashion they illustrate a barrier to quantum computing—a technology, still largely speculative, that some researchers hope could solve problems that are now impossibly difficult.

The mystery about the quantum-classical transition stems from a crucial quality of quantum particles—they can undulate and travel like waves (and vice versa: light can bounce around as a particle called a photon). As such, they can be described by a wave function, which Schrödinger devised in 1926. A sort of quantum Social Security number, the wave function incorporates everything there is to know about a particle, summing up its range of all possible positions and movements.

Taken at face value, a wave function indicates that a particle resides in all those possibilities at once. Invariably, however, an observation reveals only one of those states. How or even why a particular result emerges after a measurement is the point of Schrödinger's thought experiment: in addition to the cat and the poison, a radioactive atom goes into the box. Within an hour, the atom has an even chance of decaying; the decay would trigger a hammer that smashes open the vial of antifeline serum.

The Measurement Problem

According to quantum mechanics, the unobserved radioactive atom remains in a funny state of being decayed and not decayed. This state, called a superposition, is something quantum objects enter quite readily. Electrons can occupy several energy levels, or orbitals, simultaneously; a single photon, after passing through a beam splitter, appears



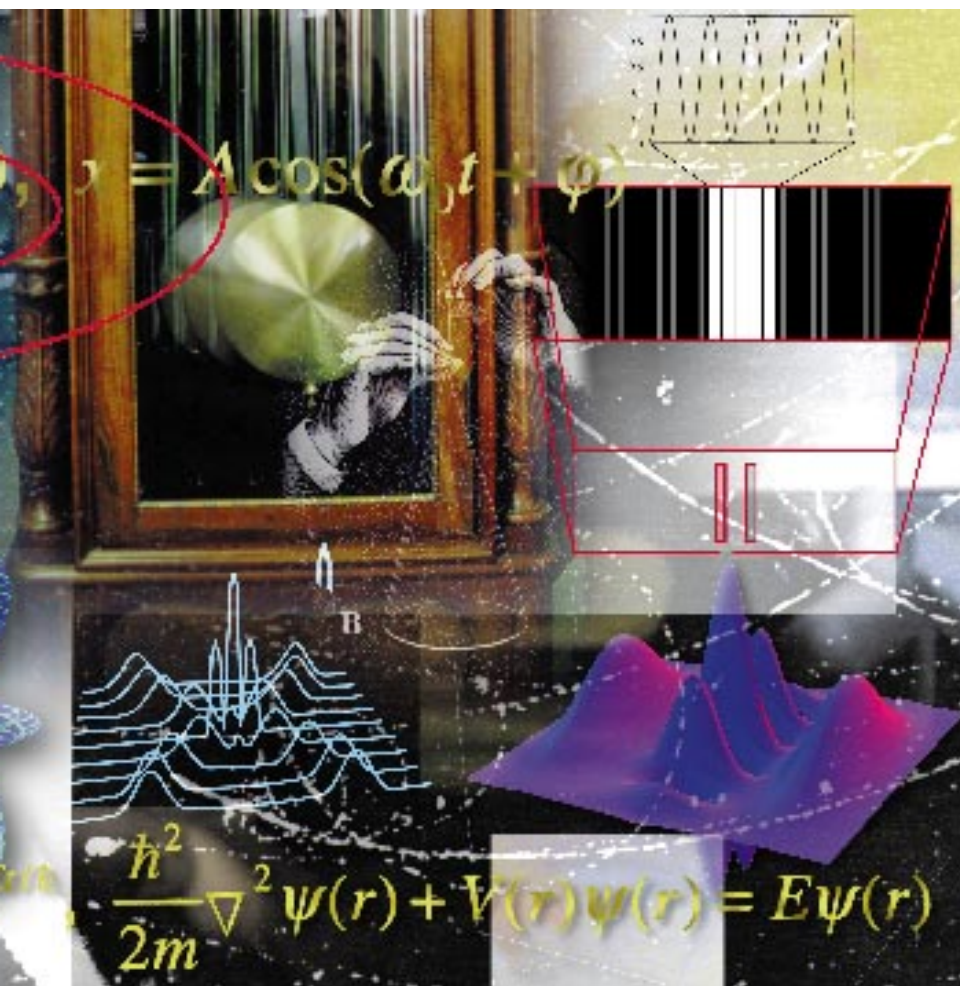
DIGITAL COMPOSITION BY JEFF BRICE; WIGNER FUNCTION OF ATOMIC MATTER WAVE THROUGH DOUBLE SLIT COURTESY OF MATTHIAS FREYBERGER, UNIVERSITY OF ULM

Recent experiments have begun to demonstrate how the weird world of quantum mechanics gives way to the familiarity of everyday experience

to traverse two paths at the same time. Particles in a well-defined superposition are said to be coherent.

But what happens when quantum objects are coupled to a macroscopic one, like a cat? Extending quantum logic, the cat should also remain in a coherent superposition of states and be dead and alive simultaneously. Obviously, this is patently absurd: our senses tell us that cats are either dead or alive, not both or neither. In prosaic terms, the cat is really a measuring device, like a Geiger counter or a voltmeter. The question is, then, Shouldn't measuring devices enter the same indefinite state that the quantum particles they are designed to detect do?

For the Danish physicist Niels Bohr, a founder of quantum theory (and to whom Schrödinger's regretful comment was directed), the answer was that measurements must be made with a classical apparatus. In what has come to be called the standard, or Copenhagen, interpretation of quantum mechanics, Bohr postulated that macroscopic detectors never achieve any fuzzy superposition, but he did not explain exactly why not. "He wanted to mandate 'classical' by hand," says Wojciech Zurek of Los Alamos National Laboratory. "Measurements simply became." Bohr also recognized that the boundary between the classical and the quantum can shift depending on how the experiment is arranged. Furthermore, size doesn't necessarily matter: superpositions can persist on scales much larger than the atomic.

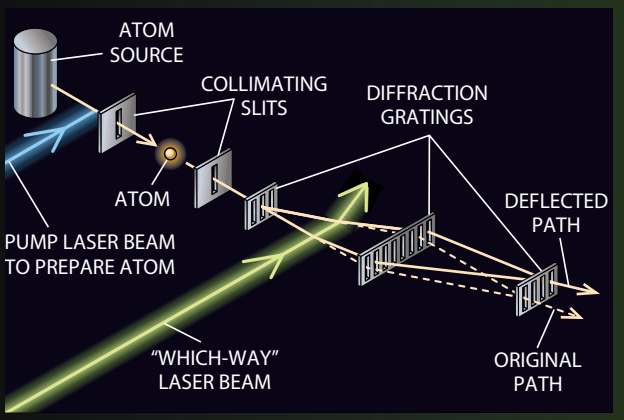


In November 1995 Pritchard and his M.I.T. colleagues crystallized the fuzziness of measurement. The team sent a narrow stream of sodium atoms through an interferometer, a device that gives a particle two paths to travel. The paths recombined, and each atom, acting as a wave, "interfered" with itself, producing a pattern of light and dark fringes on an observing screen (identical to what is seen when a laser shines through two slits). The standard formulation of quantum mechanics states that the atom took both paths simultaneously, so that the atom's entire movement from source to screen was a superposition of an atom moving through two paths.

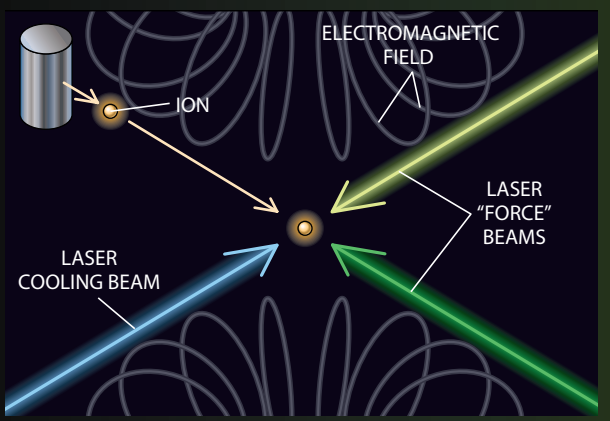
The team then directed a laser at one of the paths. This process destroyed the interference fringes, because a laser photon scattering off the atom would indicate which path the atom took. (Quantum rules forbid "which-way" information and interference from coexisting.)

On the surface, this scattering would seem to constitute a measurement that destroys the coherence. Yet the team showed that the coherence could be "recovered"—that is, the interference pattern restored—by changing the separation between the paths to some quarter multiple of the laser photon's wavelength. At those fractions, it was not possible to tell from which path the photon scattered. "Coherence is not really lost," Pritchard elucidates. "The atom

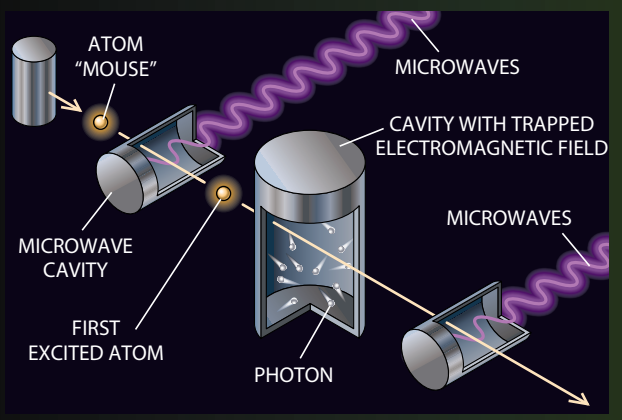
FRAMEWORK OF PHYSICS must somehow connect the exotica of quantum mechanics—its dead-and-alive cats, orbitals, oscillating ions and matter waves—with the more intuitive counterparts from classical physics: probabilities, planetary motions, pendulum swinging and double-slit, light-wave interference.



FUZZINESS OF QUANTUM MEASUREMENT is demonstrated with sodium atoms split and recombined to produce an interference pattern (*not shown*). A laser deflecting off an atom would reveal which path the atom took and thereby eliminate interference. But the pattern reemerges if the path lengths are varied, showing how deeply quantum systems can become "entangled" with classical apparatus.



SCHRÖDINGER'S CAT made from a beryllium ion is first trapped by an electromagnetic field and then cooled with a laser. Laser "force" beams prepare the ion in a superposition of two spin states. These states are then eased apart so that the ion resides in two places at once.



CAT-AND-MOUSE EXPERIMENT is done with a trapped electromagnetic field (confined photons). A rubidium atom is excited by microwaves into a superposition of two states. As it passes through the center cavity, it relays its superposed state to the electromagnetic field. A second atom serves as the "mouse" that probes the resulting state of the field. (The second microwave cavity, identical to the first, provides a way to create quantum interference and is essential to measurements.)

JARED SCHNEIDMAN DESIGN

became entangled with a larger system." That is, the quantum state of the atom became coupled with the measuring device, which in this case was the photon.

Like many previous experiments, Pritchard's work, which is a realization of a proposal made by the late Richard Feynman many years ago, deepens the mysteries underlying quantum physics rather than resolving them. It demonstrates that the measuring apparatus can have an ambiguous definition. In the case of Schrödinger's cat, then, is the measurement the lifting of the lid? Or when light reaches the eye and is processed by the mind? Or a discharge of static from the cat's fur?

A recent spate of Schrödinger's cat experiments have begun to address these questions. Not all physicists concur that they are looking at bona fide quantum cats—"kitten" is the term often used, depending on the desired level of cuteness. In any event, the attempts do indicate that the quantum-classical changeover—sometimes called the collapse of the wave function or the state-vector reduction—has finally begun to move out of the realm of thought experiments and into real-world study.

Here, Kitty, Kitty

In 1991 Carlos Stroud and John Yeazell of the University of Rochester were experimenting with what are called Rydberg atoms, after the Swedish spectroscopist Johannes Rydberg, discoverer of the binding-energy relation between an electron and a nucleus. Ordinarily, electrons orbit the nucleus at a distance of less than a nanometer; in Rydberg atoms the outer electron's orbit has swollen several 1,000-fold. This bloating can be accomplished with brief bursts of laser light, which effectively put the electron in many outer orbitals simultaneously. Physically, the superposition of energy levels manifests itself as a "wave packet" that circles the nucleus at an atomically huge distance of about half a micron. The packet represents the probability of the excited electron's location.

While swelling potassium atoms, the Rochester workers noticed that after a few orbits, the wave packet would disperse, only to come back to life again as two smaller packets on opposite ends of its large orbit. With his colleague Michael W. Noel, Stroud showed last September that the two packets constituted a Schrödinger's cat state—a single electron in two locations.

An electron, though, is essentially a mere point. Closer to the macroscopic realm is an ion (a charged atom), which consists of many elementary particles. In May 1996 Chris Monroe, David J. Wineland and their colleagues at the National Institute of Standards and Technology (NIST) in Boulder, Colo., created a Schrödinger's cat out of a beryllium ion. They first trapped the ion with electromagnetic fields, then hit it with a laser beam that stifled the ion's thermal jitters and thereby cooled it to within a millikelvin of absolute zero. Then the researchers fired two laser beams, each of a slightly different frequency, at the ion to manipulate its spin, an intrinsic, quantum feature that points either up or down. With the lasers, the researchers made the ion take on a superposition of spin-up and spin-down states.

So much for the preparations; next came the more macroscopic part. By manipulating the tuning of the two lasers, the NIST team could swing the spin-up state to and fro in space, and the spin-down state fro and to. A snapshot would show the ion in the spin-up state at one physical location and simultaneously in the spin-down state at a second position. The states were 80 nanometers apart—large on the atomic scale. "We made one ion occupy two places that are very far separated compared with the size of the original ion," Monroe says.

Last December, Michel Brune, Serge Haroche, Jean-Michel Raimond and their colleagues at the Ecole Normale Supérieure (ENS) in Paris took matters a step further. "We were able to monitor the washing-out of quantum features," Haroche explains. To see how the superposition collapsed to one state or another, they in effect dangled a quantum mouse in front of their Schrödinger's cat to check whether it was alive or dead.

The cat was a trapped electromagnetic field (a bunch of microwave photons in a cavity). The researchers sent into the cavity a Rydberg atom that had been excited into a superposition of two different energy states. The Rydberg atom transferred its superposed state to the resident electromagnetic field, putting it into a superposition of two different phase, or vibrational, states. With its two phases, the field thus resembled the Schrödinger's cat in its odd superposition between life and death.

For the mouse, the ENS team fired another Rydberg atom into the cavity. The electromagnetic field then transferred information about its superposed phases to the atom. The physicists compared the second atom with the first to glean superposition information about the electromagnetic field.

MICHAEL NOEL AND CARLOS STROUD *University of Rochester*

WRITING ON AN ATOM is theoretically plausible. An electron in a superposition of 2,500 energy levels has a wave function sufficiently complex to encode a message. Words are written by assigning color and saturation to the amplitude and phase of the wave function.

More interesting, however, was the team's ability to control crucial variables and to determine how coherent states become classical ones. By varying the interval between the two atoms sent into the cavity (from 30 to 250 microseconds), they could see how the collapse of the superposition varied as a function of time, and by enlarging the electromagnetic field (by putting more photons in the cavity), they could see how the collapse changed with size. "This is the first time we can observe the progressive evolution of quantum to classical behavior," Haroche says.

"This is a breathtaking experiment," Zurek enthuses. "Seeing a Schrödinger's cat is always surprising, but being able to see the cat forced to make a choice between 'dead' and 'alive,' to observe for the first time quantum weirdness going away, is the real coup." Moreover, the ENS results jibed with most theorists' technical expectations. "What it tells me," Zurek remarks, "is that the simple equations we've been writing down seem to be a good approximation."

Losing Coherence

Zurek is the leading advocate of a theory called decoherence, which is based on the idea that the environment destroys quantum coherence. He formulated it in the 1980s (although some of it harkens back to Bohr and other quantum founders) and with various collaborators has been investigating its consequences ever since.

The destabilizing environment essentially refers to anything that could be affected by—and hence inadvertently "measure"—the state of the quantum system: a single photon, a vibration of a molecule, particles of air. The environment is not simply "noise" in this theory; it acts as an apparatus that constantly monitors the system.

The ENS experiment makes that effect clear. "The system decoheres because the system leaks information," Zurek notes. Some photons can escape the cavity and hence betray the state of the remaining ones to the rest of the universe. "So in a sense, Schrödinger's cat is having kittens crawling out," Zurek says.

Having the environment define the quantum-classical boundary has the advantage of removing some of the mystical aspects of quantum theory that certain authors have promulgated. It does away with any special need for a consciousness or new physical forces to effect a classical outcome. It also explains why size per se is not the cause of decoherence: large systems, like real-life cats, would never enter a superposition, because all the particles that make up a feline influence a vast number of environmental parameters that make coherence impossible. Given a one-gram bob on a pendulum and a few reasonable assumptions, the interference terms in the system's wave function drop to about $2.7^{-1,000}$ of their original value in a nanosecond—a virtually instantaneous disappearance of quantum weirdness. "The old intuition going back to Bohr is on the money," although now there is a physical mechanism to substantiate his mandate, Zurek concludes.

Still, Zurek's decoherence model is flawed in some eyes. "In my view, decoherence doesn't select a particular outcome," opines Anthony J. Leggett of the University of Illinois. "In real life, you get definite macroscopic outcomes."

Zurek argues that the environment does indeed dictate the quantum possibilities that end up in the real world. The process, which he refers to as environment-induced superselection, or einselection, tosses out the unrealistic, quantum states and retains only those states that can withstand the scrutiny of the environment and thus might become classical. "The selection is done by the environment, so you will not be able to predict which of the allowed possibili-

ties will become real," Zurek observes.

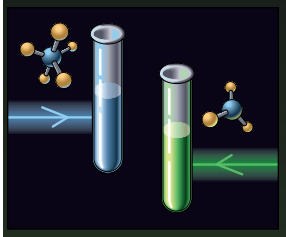
The explanation feels less than satisfying. Zurek's approach is "very appealing. It allows you to calculate things, to see how the interference fringes wash out as the superposition gets bigger," NIST's Monroe says. "But there's still something funny about it. He's sweeping things under the rug, but it's hard to say what rug." The problem is that decoherence—and in fact any theory about the quantum-classical transition—is necessarily ad hoc. Quantum superpositions must somehow yield outcomes that conform to our everyday sense of reality. That leads to circuitous logic: the results seen in the macroscopic world arise out of the quantum world because those results are the ones we see. A solution of sorts, advocated by a few prominent cosmologists, is the unwieldy "many worlds" interpretation, which holds that all possibilities stipulated by the wave function do in fact happen. They go on to exist in parallel universes. The idea, however, is untestable, for the parallel universes remain forever inaccessible to one another.

Jobs for Quantum Cats

Researchers have proposed and demonstrated several technologies exploiting entangled and superposed quantum states, such as quantum computing. A few other schemes include the following:

Quantum Chemistry

Using lasers, researchers can place molecules in a superposition of reaction pathways; then they can control the chemical process by adjusting the degree of interference. Last December workers separated isotopes with a similar technique. Obstacles include less than practical efficiency levels and difficulty in controlling phase characteristics of the laser.



Quantum Key Cryptography

A much better prospect than quantum computing is quantum key cryptography. Legitimate communicators create shared keys using the polarization of photons. Eavesdropping on these keys would immediately be noticed, because it would disrupt the key photons' states. Quantum cryptography has been shown to function over several kilometers in optical fibers.

Radical Reworkings

The problems with decoherence and the many-worlds idea have led a sizable minority to support a view called GRW theory, according to Leggett. The concept was put forward in 1986 by GianCarlo Ghirardi and Tullio Weber of the University of Trieste and Alberto Rimini of the University of Pavia.

In the GRW scheme, the wave function of a particle spreads out over time. But there is a small probability that the spreading wave "hits" a mysterious "something" in the background. The wave function suddenly becomes localized. Individual particles have only a small chance of a hit, about once every 100 million years. But for a macroscopic cat, the chance that at least one of its roughly 10^{27} particles makes a hit is high, at least once every 100 picoseconds. The cat never really has a chance to enter any kind of superposition. Hence, there is no need for decoherence: the macroscopic state of the cat results from spontaneous microscopic collapses.

A few problems plague this model. One is that the timing factor that triggers the hit is entirely arbitrary; proponents simply choose one that produces reasonable results. More important, though, is the source of the trigger. "Basically, [there is] a sort of universal background noise that cannot itself be described by quantum mechanics," Leggett explains. The noise is not simply random processes in the environment; it has a distinct mathematical flavor. Roger Penrose of the University of Oxford argues in his book *Shadows of the Mind* that the trigger may be gravity, which would neatly sidestep certain technical objections.

Other, more radical proposals abound. The most well known was put forth by the late David Bohm, who postulated that "hidden variables" underpin quantum mechanics. These variables—describing properties that in a way render wave functions as real forces—would eliminate the notion of superpositions and restore a deterministic reality. Like the many-worlds idea, Bohm's theory cannot be verified: the hidden variables by definition remain, well, hidden.

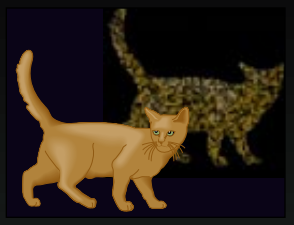
Given such choices, many working physicists are subscribing to decoherence, which makes the fewest leaps of faith even if it arguably fails to resolve the measurement problem fully. "Decoherence does answer the physical aspects of the questions," Zurek says, but does not get to the metaphysical ones, such as how a conscious mind perceives an outcome. "It's not clear if you have the right to expect the answer to all questions, at least until we develop a better understanding of how brain and mind are related," he muses.

Bigger superpositions may enable researchers to start ruling out some theories—GRW and decoherence predict them on different scales, for instance. "What we would like to do is to go to more complex systems and entangle more and more particles" than just the mere 10 trapped before, Haroche of the ENS says. Future NIST experiments are particularly suited to serve as "decoherence monitors," Monroe contends. "We can simulate noise to deliberately cause the superposition to decay." Leggett has proposed using sensors made from superconducting rings (called SQUIDs): it should be possible to set up large currents flowing in opposite directions around the ring simultaneously.

Still, there's a long way to go. "Even in the most spectacular experiments, at most you've shown a superposition for maybe 5,000 particles. That's a long way from the 10^{23} characteristic of the macroscopic world," says Leggett, who nonetheless remains supportive. "My own attitude is that one should just try to do experiments to see if quantum mechanics is still working."

Shrinking transistors, now with features less than a micron in size, may also lead to insights about the quantum-classical changeover. In a few years they may reach dimensions of tens of nanometers, a realm sometimes called the mesoscopic scale. Da Hsuan Feng of Drexel University speculates that quantum mechanics perhaps really doesn't lead to classical

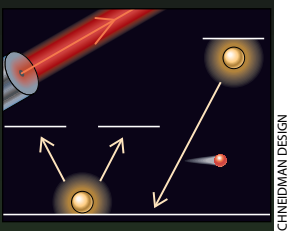
Quantum Teleportation



The idea has less to do with *Star Trek* than with reconstructing destroyed information. The crux is the Einstein-Podolsky-Rosen effect, which shows that two photons can remain entangled, no matter how far apart they are, until a measurement is made (which instantaneously puts both in a definite state). Alice takes one EPR photon, Bob the other. Later, Alice measures her EPR photon with respect to a third photon. Bob can use the relational measurement to re-create Alice's non-EPR photon. Whether Bob truly rematerialized the photon or just created an indistinguishable clone is unclear. Researchers at the University of Innsbruck reportedly demonstrated the phenomenon, which might have use in quantum cryptography.

Quantum Laser Optics

Lasers ordinarily require a population inversion, a condition in which atoms in an excited state outnumber those in the ground state; the excited atoms emit laser photons as they drop to the ground state. In 1995 researchers sidestepped this requirement. In lasing without inversion, two coupling lasers give ground-state atoms two paths to one higher energy level. Interference between the paths renders the ground-state atoms invisible, and so fewer excited atoms are needed. Such lasers do not require as much power and in principle could emit light in the desirable x-ray region.



JARED SCHNEIDMAN DESIGN

mechanics; rather both descriptions spring from still undiscovered concepts in the physical realm between them.

Quantum Computing

Even if experiments cannot yet tackle the measurement problem fully, they have much to contribute to a very hot field: quantum computing. A classical computer is built of transistors that switch between 0 or 1. In a quantum computer, however, the "transistors" remain in a superposition of 0 *and* 1 (called a quantum bit, or qubit); calculations proceed via interactions between superposed states until a measurement is performed. Then the superpositions collapse, and the machine delivers a final result. In theory, because it could process many possible answers simultaneously, a quantum computer would accomplish in seconds tasks, such as factoring large numbers to break codes, that would take years for a classical machine.

In December 1995 researchers successfully created quantum two-bit systems. Monroe and his colleagues crafted a logic element called a controlled-NOT gate out of a beryllium ion. The ion is trapped and cooled to its lowest vibrational state. This state and the first excited vibrational state constitute one bit. The second bit is the spin of one of the ion's electrons. Laser pulses can force the bits into superpositions and flip the second bit depending on the state of the first bit. Other variations of gates couple two photons via an atom in a cavity or transmit an entangled pair of photons through a network of detectors.

Yet the creation of a useful quantum computer, relying on superpositions of thousands of ions performing billions of operations, remains dubious. The problem? Loss of superposition. The logic gates must be fast enough to work before the qubits lose coherence. Using data from the NIST gate experiment, Haroche and Raimond calculated in an August 1996 *Physics Today* article that given the gate speed of 0.1 millisecond, the bits would have to remain in a superposition for at least a year to complete a meaningful computation (in this case, factoring a 200-digit number).

Other physicists are less pessimistic, since error-correcting codes (which are indispensable in classical computing) might be the solution. "It gives you instructions on how to repair the damage," says David DiVincenzo of the IBM Thomas J. Watson Research Center in Yorktown Heights, N.Y.

Moreover, DiVincenzo points out that a new method of quantum computation, making use of nuclear magnetic resonance (NMR) techniques, could raise coherence times to a second or more. Say a liquid—a cup of coffee—is placed in a magnetic field; because of thermal vibration and other forces, only one out of every million nuclei in the caffeine molecules would line up with the magnetic field. These standouts can be manipulated with radio waves to put their spins in a superposition of up and down. Maintaining coherence is easier here than in the other techniques because the nuclear spins undergoing the superpositions are well protected from the environment by the surrounding turmoil of seething molecules, the mad scramble of which averages out to zero. The calculating caffeine sits effectively in the calm eye of a hurricane. Two groups have recently demonstrated quantum computing by NMR, using a four-qubit version to sum 1 and 1. More complex systems, using perhaps 10 qubits, could be had by the end of the year.

The drawback is readout. With no way to detect individual spins, researchers must measure all the molecules' spins—both qubit and nonqubit ones. Complex molecules capable of sustaining many spins are therefore "noisier" than simpler ones. "They'll be able to do some nice stuff," Monroe says, "but beyond about 10 bits, they'll run into fundamental problems." The output from 10 bits is only 0.001 as strong as that from a single bit; for 20, the output is down by one million. So the NMR technique may not enter a meaningful computational realm of at least 50 bits.

There might be other uses for quantum superpositions, though. Stroud proposes data storage on an atom, because an electron in a Rydberg atom could be made to inhabit a superposition of 2,500 different energy levels. "That means that the electron's wave function can be quite complex, encoding a great deal of information," Stroud expounds. He demonstrated the possibility theoretically by writing "OPTICS" on an atom. Other uses for quantum superposition, such as in cryptography, chemistry and even teleportation, have been demonstrated. Schrödinger's boxed cat may have outwitted the best philosophical minds so far, but it seems to have found plenty of technological reasons to stay put. SA

Further Reading

DECOHERENCE AND THE TRANSITION FROM QUANTUM TO CLASSICAL. Wojciech Zurek in *Physics Today,* Vol. 44, No. 10, pages 36–44; October 1991.

WHERE DOES THE WEIRDNESS GO? David Lindley. BasicBooks, 1996.

SCHRÖDINGER'S MACHINES. Gerard J. Milburn. W. H. Freeman and Co., 1997.

by Shawn Carlson

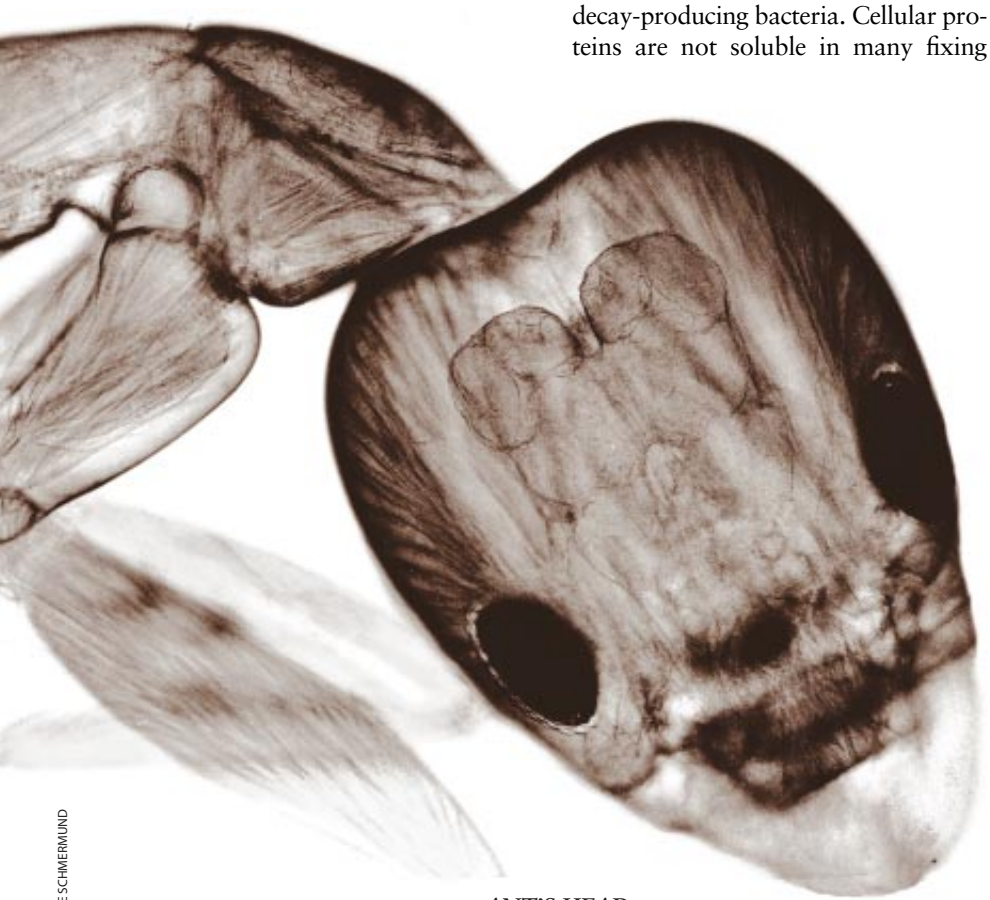
Getting Inside an Ant's Head

The toy microscope set I got one childhood Christmas almost put me off microscopy for good. Squinting through a cheap plastic eyepiece, manipulating the sloppy focus control with one hand and struggling with the other to position the factory-prepared specimen slides resulted in more neck aches and frustration than reward. So years later when a friend offered me his old professional-quality microscope for a mere fraction of its market value, I hesitated. Besides, I could scarcely see through the dirt-encrusted optics, and the gearing for the traveling stage was so gunked up that it scarcely traveled at all. But a little patience, along with cleaning fluid and grease for the gears, quickly restored this binocular beauty to mint condition. Today I'm convinced it was the best $100 I ever spent.

My vintage Spencer has been my constant companion for more than a decade. With it I have poked about inside plant cells, swum abreast of wriggling sperm and gotten delightfully lost inside floating forests of phytoplankton. Unfortunately, no biological specimen can be enjoyed under a microscope for long before natural processes begin to destroy it. Tissues quickly dry out and shrink, and bacteria begin breaking down the delicate structures of life almost immediately after a specimen dies.

Microscopists therefore routinely preserve their treasures before scrutinizing them carefully. The process, called fixation, replaces the water in the specimen's tissues with a chemical that is incompatible with life. The fluid pressure keeps the cells plump while the chemical kills decay-producing bacteria. Cellular proteins are not soluble in many fixing agents, however; they precipitate to form a sort of plaster that sticks to cell walls. This process stiffens the cell wall and can actually increase its index of refraction, making the cell stand out under the microscope, but it ultimately destroys potentially useful information.

When amateur scientist extraordinaire George Schmermund of Vista, Calif., recently brought together his interests in insects and microscopy, he uncovered a wonderful fact. Different tissues within an insect absorb the fixing agent he used at different stages in the fixing process. By observing the specimen throughout fixation, rather than waiting until the process is complete, as microscopists are often taught to do, he found that separate organ systems within the animal became highlighted sequentially. This let Schmermund take the stunning pictures of insect insides that you see here. His technique empowers any amateur microscopist to explore insect interiors in a detail unmatched by any other method I know.



GEORGE SCHMERMUND

ANT'S HEAD
photographed during fixation shows details that later disappear.

Begin your own fantastic voyage by securing some suitable specimens. Small ants and fleas make ideal subjects. Schmermund kills them by placing them in his freezer for a few hours. His fixing agent is ordinary isopropyl (rubbing) alcohol. Isopropyl alcohol is inexpensive, readily available and easy to handle.

My local drugstore stocks isopropyl alcohol in concentrations of 70 and 91 percent. Buy the highest concentration you can find. For reasons that will become clear in a moment, if you intend to create a permanent library of mounted specimens you'll need to purchase pure (anhydrous) isopropyl alcohol from a chemical supply company. (Note that denatured alcohol is a blend of mostly ethanol and methanol; it is not isopropyl alcohol. Both substances are also fixing agents, but methanol is highly poisonous. If you choose to experiment with denatured alcohol, take appropriate precautions to protect your family and pets.)

The box on page 132 gives a recipe for diluting your fixing solution to any desired concentration. Schmermund took these photographs using only two

dilutions: 35 and 70 percent. The time required at each stage depends on the specimen's size. Soak garden ants and fleas for at least one hour. Larger insects may take as long as six hours. The volume of solution should be at least 20 times that of your specimen's body, but for ant-size specimens that is still a tiny amount of alcohol. Schmermund soaks his specimens in bottle caps and transfers the chemicals with eyedroppers.

To get views like those you see here, you must regularly check the insect while fixing. Transfer the insect to a well slide (a slide with a polished depression to receive the sample) with a drop of fixing solution, rest a slide cover on top, then set sail into uncharted waters. Transfer and position the specimen with laboratory tweezers (available at a laboratory supply store and often at your local swap meet) and an eyedropper.

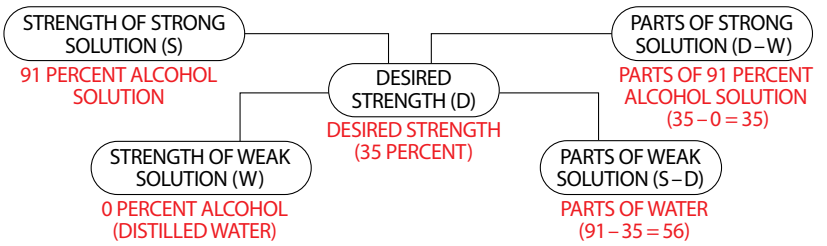
Unfortunately, neither water nor alcohol mixes with the compounds used to glue your specimen permanently to the slide. So to share your specimen with posterity, you must first replace every trace of both water and alcohol with a solvent that will mix with the glue. The process that removes the alcohol is called clearing. Clearing will make your specimen largely transparent and will destroy much of the contrast created during the fixing stages. You can lock in some permanent contrast by staining your specimen with various commercially available dyes, such as borax carmine or alum cochineal. But no method I know of produces the level of contrast in specific tissues that naturally develops during the fixing process.

Xylene is the clearing agent of choice for the amateur scientist. I recently purchased 32 ounces (a lifetime supply!) for $5 in the paint section of my local hardware store. But watch out: xylene is poisonous, it dissolves plastic on contact, and its fumes can make you quite sick. So be particularly careful to protect yourself and your family and pets. In addition, although water mixes with alcohol and alcohol mixes with xylene, water and xylene do not mix. The water must be completely removed with a final fixing step in 100 percent anhydrous isopropyl alcohol before clearing.

Once the specimen has been soaked in pure alcohol for two hours or more, place it in a solution of equal parts of xylene and anhydrous isopropyl alcohol.

A Solution for Dilutions

A general formula for diluting fixing solution to a desired concentration is shown in black. In the example in red, 35 parts of 91 percent alcohol are mixed with 56 parts of water to make a 35 percent solution.



Cover your container to keep the xylene from evaporating. Your specimen will probably float at first. Let it soak for an hour or two after it sinks to the bottom and then replace the mixture with pure xylene. Your specimen should become fairly translucent at this stage. Dark patches mark water that was not removed during fixation. These can often be cleared up by returning the specimen to the anhydrous alcohol for several hours and then re-clearing it.

GEORGE SCHMERMUND

FLEA LEG **(top)** ***has clearly visible muscle striations. The mouthparts of a wood tick are revealed in detail*** **(bottom).**

Scientific supply companies sell Canada Balsam and a variety of plastic resins for the permanent mounting of specimens. The resin dissolves in xylene, and so it infuses into the insect's tissues before hardening. To make your final slide, place your cleared specimen in the depression in the well slide, add a drop of resin and gently angle the slide cover down on top, being careful not to trap any air bubbles. Let the resin set before moving the slide.

One final note. Some microscopists have marveled at the striking depth of field visible in Schmermund's handiwork. Here's his secret. These photographs were taken with 35-millimeter film at low magnification and then enlarged in a darkroom. Low magnification means good depth of field. SA

As a service to amateur scientists everywhere, the Society for Amateur Scientists (SAS) has put together a kit that contains everything you'll need to fix, clear and mount small insects. The kit contains anhydrous isopropyl alcohol, xylene, laboratory tweezers, six well slides, a small killing jar for your freezer, small petri dishes for fixing and clearing, slide covers and mounting resin. Send $54.95 plus $5 shipping to the Society for Amateur Scientists, 4735 Clairemont Square, Suite 179, San Diego, CA 92117. For telephone orders, call (619) 239-8807. For information about this and other amateur science projects, check out the SAS's World Wide Web site at http://www.thesphere.com/SAS/ or call (800) 873-8767.

I gratefully acknowledge informative conversations that I had with George Schmermund.

by Ian Stewart

The Sifting Sands of Factorland

Factorizing a large number is among the most challenging problems in number theory. Last month I described efficient methods for testing a number with 200 or so digits for primeness. But although these methods may prove that a number is not prime, they do not yield the explicit factors. Can we fill that gap?

Recall that a factor of a number is any number that divides it exactly and that a number is prime if it has no factors other than itself and 1. Next, some "geography." Primeland, the world of prime numbers, is scattered increasingly thinly and rather randomly along the number line. But Factorland, the world of integers, is very different. Embedded in it are the primes, which have only one factor—themselves. Most numbers are not prime; obviously one number in two is a multiple of 2, one in three a multiple of 3, and two thirds are multiples of either 2 or 3. (Why not five sixths? Think about it.) But a significant proportion of numbers don't have any small factors at all. If you can find just one factor of a number, then the rest are factors of the quotient, which is smaller; so the main difficulty is to find the first one.

That is truly hard: in 1903 the American mathematician F. N. Cole spent three years of Sundays with pencil and paper to discover that $2^{67}-1 = 193{,}707{,}721 \times 761{,}838{,}257{,}287$. In the past few decades, progress has been more rapid. Carl Pomerance of the University of Georgia surveyed recent developments in the December 1996 issue of *Notices of the American Mathematical Society;* this column is based on his article.

The school method for factoring numbers is formalized trial and error: start with 2, then try 3, and so on up to the square root of the number. This method is hopelessly inefficient. In 1970 improved methods could just about handle arbitrary numbers with 20 digits. By 1980 the limit had risen to 50 digits, by 1990 to 116 digits, and by 1994 to 129 digits. In 1996 a new champion arose, an algorithm that factored a 130-digit number in a sixth of the time required by previous methods.

Part of this improvement has been in computer power, but making a computer run a million times faster can add only a few digits to the record. Most of the advances have been conceptual.

Early this century the mathematicians Allan J. C. Cunningham and Herbert J. Woodall initiated the Cunningham Project, to factor numbers of the form $r^k \pm 1$ for r between 2 and 12, and k large. These numbers include many recreational favorites: the Mersenne numbers $M_k = 2^k - 1$, the Fermat numbers $F_k = 2^{2^k} + 1$, and the so-called repunit numbers $R_k = 1111\ldots 1 = (10^k - 1)/9$. A popular means of attacking this problem is the "quadratic sieve" due to Pomerance.

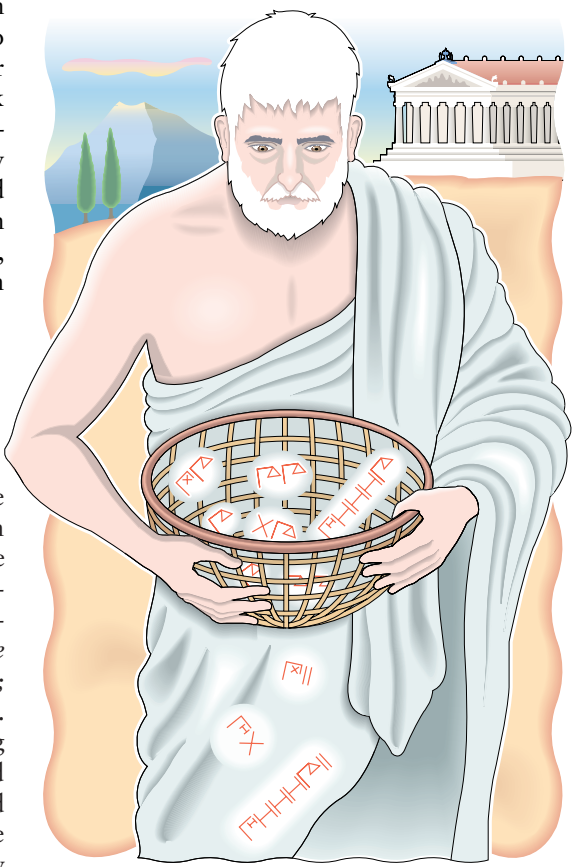
The method is called a sieve because it treats Factorland as a sandy desert, sifting out unwanted numbers one grain at a time. The great granddaddy of sieves, known as the sieve of Eratosthenes, goes back to the ancient Greeks. Wander through Factorland, sifting out all multiples of 2, the first prime. Find the first survivor, 3: this is the next prime. Wander through what remains of Factorland, sifting out all multiples of 3. Then sift out all multiples of 5, and so on. After the infinite desert has been sifted infinitely many times, only Primeland remains in the sieve.

Pomerance begins with a problem he was given in high school: factor the number 8,051, time limit five minutes, no calculators. He decided there had to be a shortcut, but only too late did he spot it—to write 8,051 as a difference of two squares: $8{,}051 = 8{,}100 - 49 = 90^2 - 7^2$. Now the algebraic identity $a^2 - b^2 = (a - b)(a + b)$ tells us that this may be written as $(90 - 7) \times (90 + 7) = 83 \times 97$. The trick actually goes back to 17th-century French mathematician Pierre de Fermat.

In the 1920s Maurice Kraitchik—author of *Mathematical Recreations*—improved on Fermat's method. He observed that instead of representing the number n as a difference of squares, it is often enough to represent some multiple of n in this way. Specifically, suppose we can write

$$kn = a^2 - b^2 = (a + b)(a - b)$$

There is a class of uninteresting solutions, in which $a + b$ and $a - b$ are multiples of n, and interesting ones, in which they are not. For an interesting solution, the highest common factor (hcf) of n and $(a - b)$, designated hcf$(n; a - b)$, must be a nontrivial factor of n—that is, not equal to n or 1. (One can also use hcf$(n; a + b)$, but it is



SIEVE OF ERATOSTHENES provides a method to select prime numbers.

JUAN VELASCO *Computer Graphics*

likely to be a larger number, posing a harder problem.) If you try to find the hcf using the normal school method—"Factor n and $a - b$ and see what primes they have in common"—then you won't get far, because it requires you to factor n. But there is a much more efficient way to calculate the hcf, known to Euclid more than 2,000 years ago.

Say we want to factor 415,813. We might observe that $15 \times 415{,}813 = 2{,}498^2 - 53^2$, so a factor is hcf(415,813; 2,498 – 53) = hcf(415,813; 2,445). Now we proceed as follows:

1. Divide 415,813 by 2,445 and find the quotient and remainder: $415{,}813 = 170 \times 2{,}445 + 163$.
2. Observe, in the above equation, that if any number divides 415,813 and 2,445 exactly, it must also divide 163 exactly. So hcf(415,813; 2,445) = hcf(2,445; 163).
3. Repeat the process, dividing 2,445 by 163 to find the quotient and remainder: $2{,}445 = 15 \times 163 + 0$.

Because the remainder is 0, 163 divides 2,445 exactly, so hcf(2,445; 163) = 163. This tells us that 163 is a factor of the original 415,813; dividing out we find the other one, 2,551.

This scheme relies on guessing a good multiplier, here 15. Kraitchik had an idea for avoiding guesswork. Start from the smallest number x whose square is bigger than n = 415,813, namely, x = 645. List the numbers $Q(x) = x^2 - n$ for increasing x and see if any of them have easy factorizations. Here we get

x	$Q(x)$	factors
645	212	$2^2 \times 53$
646	1,503	$3^2 \times 167$
647	2,796	$2^2 \times 3 \times 233$
648	4,091	4,091
649	5,388	$2^2 \times 3 \times 449$

and so on, up to

690 60,287 $19^2 \times 167$

At this stage we observe that the product $Q(646)Q(690) = (3^2 \times 167) \times (19^2 \times 167) = 3^2 \times 19^2 \times 167^2$ is a perfect square. Kraitchik realized that this lets us write some multiple of n in the form $(646 \times 690)^2 - (3 \times 19 \times 167)^2$, from which a factor can be derived as before. Here the method yields the factor 2,551.

As a challenge, you might try to factor the numbers 777,923; 603,181; and 21,720,551 in this way. Answers will be in a future Feedback section.

What is the essence of Kraitchik's method? It is to find values of $Q(x)$ that have simple factors and combine such values to get a perfect square. In 1975 John Brillhart and Michael A. Morrison formalized the procedure. Start by choosing a "factor base," which is a list of relatively small primes. In the above example, any list that includes just 3, 19 and 167 will do. Calculate the successive numbers $Q(x)$ as before but now retain only those that are products of primes in the factor base. The jargon for this property is that x is "smooth." To each smooth x, assign a vector of 0's and 1's: an entry 0 means that the corresponding prime divides $Q(x)$ to an even power, an entry 1 means an odd power. Here the vectors look like this:

x	$Q(x)$	vector
645	212	not smooth
646	1,503	(0,0,1)
647	2,796	not smooth
…		
…		
690	60,287	(0,0,1)

Now look for a set of vectors whose corresponding entries all add up to an even number: here (0,0,1) + (0,0,1) = (0,0,2). Then the corresponding product of $Q(x)$'s is a perfect square, because all primes occur to even powers.

The main obstacle to speed is recognizing the smooth values of x. For example, we would like to pick out x = 646 and 690 quickly and simply, without working out all the other $Q(x)$'s. In 1981 Pomerance realized that a variant of the sieve of Eratosthenes could be employed to sift out the smooth values of x "unbelievably fast." The improved algorithm is known as the quadratic sieve. It doubled the length in digits of numbers that could be factored.

As it happened, Pomerance had trouble getting anyone to try out his new algorithm, because a rival method, the "continued fraction algorithm," was believed to be faster. The first to use the quadratic sieve was Joseph L. Gerver of Rutgers University, who factored a 47-digit number (itself a factor of $3^{22} - 1$) from the Cunningham Project. In 1984 James A. Davis and Diane B. Holdridge of Sandia National Laboratories tried it on the more challenging number 111… 111, a repunit with 71 1's. Davis improved the method still further, factoring the number successfully.

Ironically, another Sandia team had just finished making a computer chip for cryptography, based on the belief that factoring 100-digit numbers was impossible. That size was too close to 71 to look safe, and the chip was scrapped. In 1994 the method inspired Arjen K. Lenstra, then at Bellcore, and his colleagues to organize an Internet-based project to factor a 129-digit number. Anyone who wanted to be involved ran a section of the code, generating a partial list of vectors. He or she then e-mailed the results to a central agency. As soon as a set of vectors with even sums turned up, everyone was told to stop. Factor bases with a million primes in them are now typical.

There are other new methods, too. In 1988 the mathematician John Pollard wondered whether algebraic number theory could aid factoring. In this subject the ordinary numbers are augmented by roots of polynomial equations, to create a "field." A field is a set of numbers that you can add, subtract, multiply and divide using the normal laws of algebra and without generating any numbers outside the set. If a field includes square roots, for example, one can write the prime number 13 as a product of irrational factors: $13 = (\sqrt{14} + 1)(\sqrt{14} - 1)$.

Pollard's idea, now called the number field sieve, was to use such "factorizations" to chisel away at large numbers in an attempt to reveal genuine factors. It proved its worth when Arjen Lenstra, with Hendrik W. Lenstra, Jr., of the University of California at Berkeley and Mark S. Manasse of Digital, used it to factor the ninth Fermat number. With over 150 digits, this number was well beyond the reach of the quadratic sieve.

The area of practical factoring continues to grow, with a steady stream of new ideas. Will future mathematicians discover dramatically better ways to sift the sands of Factorland in search of those elusive factors? Only time will tell. SA

Correction: In last month's column, n *refers to the number being tested, not the number of its digits.*

FEEDBACK

In the Feedback section of the October 1996 column I asked, "Any more mathematical sculptors or modelers out there?" There sure are. Paul Tucker of Dillsburg, Pa., told me about some fascinating mathematical carpentry. Tucker, a carpenter himself, was in the Amelia Givins Public Library in Mount Holly Springs, a small town in his locality. He noticed that the doors and partitions contained latticelike structures formed from interlocking spiral moldings. Architectural experts had stated that these spirals were formed by bending dowel rods, but Tucker saw at once, by looking at the grain, that they must have been milled.

INTERLOCKING SPIRALS *decorate this library door* **(detail at right).**

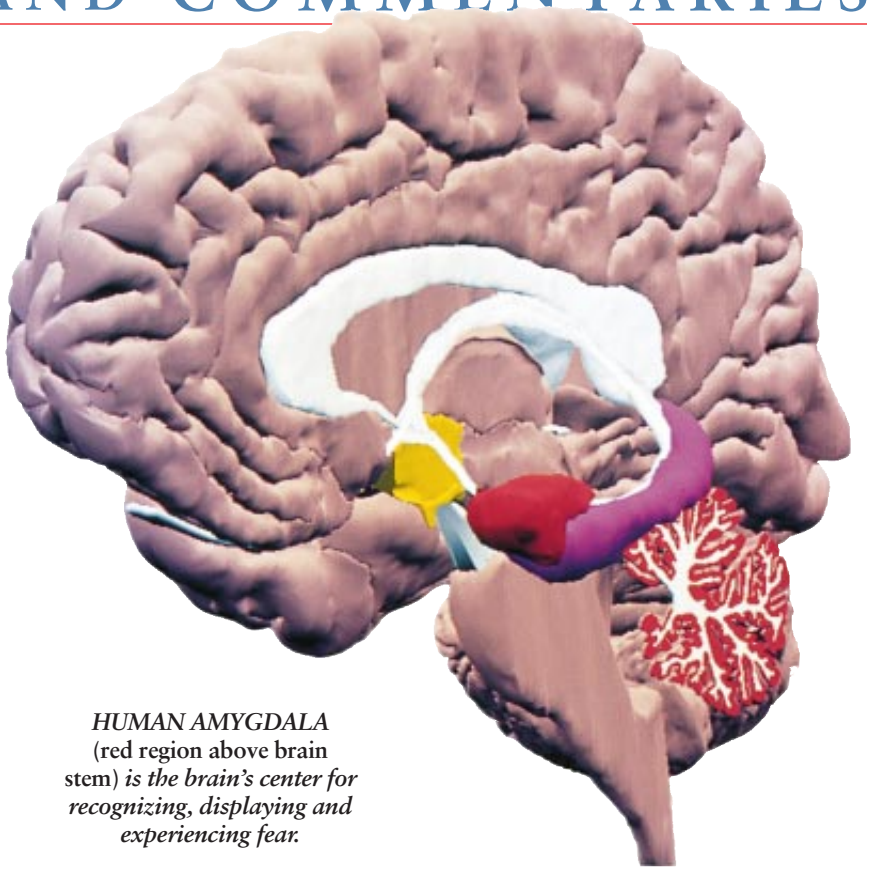
PAUL TUCKER

Armed with the sole useful evidence, a brass plate bearing the date "September 15, 1885," Tucker visited the patent office. With its help, and that of the local historical society, he traced the moldings to Moses Y. Ransom, who lived in Cleveland, Ohio, in the late 1800s. Ransom devised clever ways to mill helical spirals and thread them through one another. His patents, for those who want to track them down, are No. 307,332 (October 28, 1884) and No. 326,329 (September 15, 1885). —*I.S.*

THINKING AND FEELING

Review by Antonio R. Damasio

The Emotional Brain: The Mysterious Underpinnings of Emotional Life
BY JOSEPH LEDOUX
Simon & Schuster, New York, 1996
($25)



HANK MORGAN *Photo Researchers, Inc.*

HUMAN AMYGDALA **(red region above brain stem)** ***is the brain's center for recognizing, displaying and experiencing fear.***

We may never understand exactly why emotion was given the cold shoulder of science for almost 100 years. By the last quarter of the 19th century, Charles Darwin, William James and Sigmund Freud had thought and written brilliantly about the nature of emotion, about the possible biological mechanisms behind it and about the ways in which it could be disturbed. The British neurologist John Hughlings Jackson had even made a first stab at discerning the anatomy of human emotion by planting it in the right cerebral hemisphere. There would have been every reason to expect that the budding brain sciences would concern themselves with emotion in much the same way they had been taking on language or visual perception.

Curiously, it never came to pass. Emotion was consistently left out of the mainstream of what became neuroscience and cognitive science. A handful of psychologists, such as Stanley Schachter of Columbia University, carried on important studies on emotion; psychiatrists and pharmacologists concerned with mood disorders developed and applied drugs that gave indirect information on the mechanisms of emotion. By and large, however, neuroscience and cognitive science neglected emotion until quite recently. In what may have been the last gasp of post-Cartesian and post-Kantian intellectual dryness, emotion was not trusted, in real life or in the laboratory. Too subjective, it was said, too elusive, let's face it, too irrational for the likes of modern science.

Some of us have long thought this attitude was wasteful, not to say silly, and proceeded as if a field called "neurobiology of emotion" already existed. That missing field was created in the process, and Joseph LeDoux, author of *The Emotional Brain*, stands out among the creators. At his laboratory at New York University, LeDoux has made a rich contribution to the understanding of one of the star emotions, fear. Working in animals and concentrating on a pivotal brain region known as the amygdala, LeDoux has performed a number of ingenious experiments that throw light onto some of the neural mechanisms of the fear response. Much of what he has discovered is applicable to other emotions and to human beings. It also provides a valuable blueprint for further animal studies.

The Emotional Brain draws its strengths heavily from the author's own work. There is no substitute for the firsthand knowledge of how the process of discovery unfolds—sometimes exciting, sometimes painful. LeDoux incorporates his experience to produce an account that is informative (for those who may wish to learn about one approach to systems neuroscience), useful (for specialists) and pleasant (for all). The writing is direct, without subterfuge or pretension, and the author even acknowledges colleagues—a duty so rarely observed in science trade books that it should certainly be regarded as a virtue.

LeDoux frames his account with a well-articulated reflection on the past and future of emotion research. With remarkable courage, he takes on the long-standing controversy over the role of the body in the processing of emotions. This controversy is probably the central issue in the neurobiology of emotion. William James proposed that when the brain triggered emotions they were enacted in the body and that the changes resulting from such an enactment were subsequently represented in the brain and apprehended consciously. A number of critics countered that the essential components of the mechanism occurred within the brain and that body responses were not essential for the feeling of an emotion. Although its arguments were muddled and uninformed, the anti-James camp was the winner, perhaps because the incompleteness of James's framework (inevitable given the limited knowledge of the time) rendered it vulnerable.

That this state of affairs prevailed until recently can only be explained by sub-

sequent researchers' reluctance to examine the problem in any depth. LeDoux is not shy in his assessment: "It's hard to believe that after all these years we actually still don't have a clear and definitive understanding of the role of body states in emotions. However, I'm placing my bets in favor of feedback playing a role. Emotional systems evolved as ways of matching bodily responses with the demands being made by the environment, and I can't see many ways that a full-blooded emotional feeling could exist without a body attached to the brain that is trying to have the feeling."

He also replies eloquently to the oft-asked question of the emotional status of tetraplegics who have lesions in the cervical spinal cord (and who therefore have little feeling or motor control below the level of the neck). The continuing ability of tetraplegics to feel some emotion seems to speak against the role of body feedback in emotion. But few, if any, injuries ever destroy the cord completely and thus some body sensory input still goes through. Moreover, a considerable amount of visceral input bypasses the cord altogether and enters the central nervous system at the brain stem level via cranial nerves. Patients with spinal cord lesions do display some changes in emotional processing, a clear witness to the value of on-line body signaling since even partial damage can have an effect.

LeDoux's argument is completed by my own proposal for an "as-if-body-loop," as presented in my book *Descartes' Error.* I have suggested, and LeDoux agrees, that some emotional responses can change somatosensory representations in the brain directly, "as if" the latter were receiving signals from the body, although in effect the body is bypassed. People probably have both body-loop and as-if-body-loop mechanisms to suit diverse processing conditions. The critical point, however, is that both mechanisms are body-related.

I also enjoyed LeDoux's perspective on the closely related argument that the body would not provide enough variety of signals to signify the different emotions we experience. I have pointed out the weaknesses in this argument, and I am pleased LeDoux endorses the idea that body signals can easily provide the diversity of emotional patterns. He writes, "When all the interactions between the various systems are taken together, the possibilities for the generation of emotion-specific patterns of feedback are staggering. This is especially true when considered from the point of view of what would be necessary to scientifically document the existence of these patterns, or, even more difficult, to prove that feedback is not important."

I do not wish to disappoint readers bored by so much agreement between reviewer and author, so I will say I do not endorse LeDoux's general attitude toward feelings (though I understand well why he holds it). In essence, LeDoux believes the investigative effort in our field should focus on the biological responses that constitute an emotion because these responses can be identified and manipulated in animal experiments, whereas feelings, which are the perception of the emotional responses, can be studied only with the cooperation of the self that experiences them. His fear of feelings is really twofold. On the one hand, animals may have feelings, but we cannot study them effectively. On the other, the "soft" study of feelings was associated in the past with some of the confusion that arose from poorly conceived studies of emotion.

I disagree, because times and technologies have changed. The modern techniques of cognitive neuroscience allow us to study emotion in animals and both emotion and feeling in humans. The combination of animal and human studies will eventually reveal, in seamless fashion, the continuity of processes that begin with a triggering mind stimulus, proceed to emotional responses and to their sensory representation, and conclude with the conscious readout known as feelings.

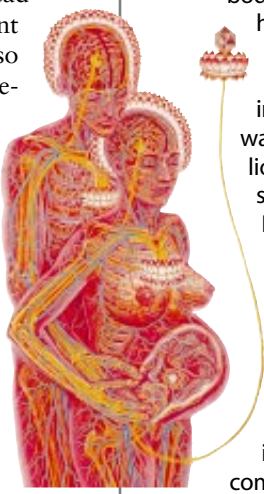
In the meantime, *The Emotional Brain* is an excellent introduction to the strange history of the neurobiology of emotion and a preview of what lies ahead.

ANTONIO R. DAMASIO is the author of Descartes' Error: Emotion, Reason and the Human Brain (G.P. Putnam's Sons, 1994). *He is Van Allen Professor and head of neurology at the University of Iowa College of Medicine.*

BRIEFLY NOTED

NAKED TO THE BONE: MEDICAL IMAGING IN THE TWENTIETH CENTURY, *by Bettyann Holtzmann Kevles. Rutgers University Press, New Brunswick, N.J., 1997 ($35.95).*

This has been the century of the x-ray. Devices for viewing the interiors of bodies are so ubiquitous they have become almost invisible (as the things they reveal once were). X-rays inspired artists with a new way of seeing and sold millions of pairs of unnecessary orthopedic shoes. Bettyann Kevles illuminates the evolution of x-ray and other imaging technologies, from both the technological and the cultural perspectives, showing how the metaphors and mechanics of imaging have become entwined.



COURTESY OF ALEX GREY

DEADLY FEASTS: TRACKING THE SECRETS OF A TERRIFYING NEW PLAGUE, *by Richard Rhodes. Simon & Schuster, New York, 1997 ($24).*

Is the outbreak of "mad cow" disease the harbinger of an epidemic that could dwarf AIDS? Richard Rhodes traces the unfinished history of "slow virus" diseases from Stone Age New Guinea to the "industrial cannibalism" of the modern food chain. His lucid writing tends toward the credulous here, and he does not have the benefit of hindsight that aided his marvelous histories of the atomic age, but he remains an engrossing reporter.

THE INFLATIONARY UNIVERSE: THE QUEST FOR A NEW THEORY OF COSMIC ORIGINS, *by Alan H. Guth. Helix Books/Addison-Wesley, Longman, Reading, Mass., 1997 ($25).*

One of the primary inventors of the new cosmology tells how he and a handful of other physicists crafted an understanding of the first 10^{-30} second of the history of the universe. Alan H. Guth's long-awaited account is remarkable both for its clear explanation of the nearly inconceivable and for the personal sense Guth conveys of what it felt like to do theoretical physics during the explosive heydays of the 1970s and 1980s.

Continued on page 143

ELICITING SCIENCE'S BEST

Review by Tim Beardsley

Frontiers of Illusion: Science, Technology, and the Politics of Progress
BY DANIEL SAREWITZ
Temple University Press, Philadelphia, 1996 ($19.95)

Political revolutions come and go, but it is science that has brought about the 20th century's most startling changes. Partisans will point variously to antibiotics, improved crops and the World Wide Web or to global warming, mass extinction and the Bomb, but nobody can deny that science made these developments possible.

The U.S. government spends some $75 billion annually on research to ensure military preparedness, improve health, provide new sources of energy and foster economic growth. Careful public assessment of what knowledge should be pursued would therefore seem to be in order. Many leaders of science have, however, resisted that innocent-sounding suggestion. For decades, science's spokesmen argued that the internal assessment procedure known as peer review should be the only guide, because the unpredictability of basic research makes attempting to target specific projects counterproductive. It was taken as axiomatic that more research would lead to greater societal benefits.

These "truths" of science policy—that basic research should be unfettered by consideration of its likely practical consequences and that more research is always better—are among the sacred cows that Daniel Sarewitz lines up and unceremoniously slaughters in his bracing critique of the research enterprise. An earth scientist affiliated with the Geological Society of America who spent four years in the office of the former head of the Committee on Science, Space and Technology in the House of Representatives, Sarewitz builds his case using the words of science's high priests. He has a keen eye for facts that contradict the standard mythology, and the contradictions are many and flagrant.

Although the U.S. research investment is huge compared with that of other nations, many other countries seem to equal or surpass the American quality of life, as physicist John M. Rowell noted in 1992. This nation spends 30 percent of its civilian research and development budget on medical research, as compared with about 5 percent in other industrial countries, yet the U.S. has far higher health care costs and lags behind its competitors in infant mortality and life expectancy at birth. Japan, famously, has made life comfortable for its citizens with a historically low investment in basic research.

It is, then, perhaps less than startling that since the collapse of the Evil Empire, scientists have failed to persuade Congress of the necessity of appropriating ever increasing sums to maintain technological dominance. (Sarewitz identifies the end of the age of physics as occurring on October 21, 1993, the day Congress canceled funding for the Superconducting Super Collider). Federal research and development expenditures have decreased by 3.3 percent in real terms since 1994, and the American Association for the Advancement of Science (AAAS) estimates that a further 14 percent cut is in store over the next five years. The establishment is not pleased. In March the leaders of more than 20 professional scientific societies called instead for a 7 percent budget increase for scientific research and development in 1998, an amount that would put science on track for a doubling of its budget in 10 years. But harrumphing may not be enough to win that support.

Perhaps, Sarewitz suggests, it is time for scientists to reassess their contract with society and tackle the connection—or lack thereof—between progress in science and technology and progress in society. He argues, contrary to popular wisdom, that basic research can be targeted to problems, citing the fundamental physics conducted during the Manhattan Project in the 1940s. Even the transistor, which emerged from the ivory tower of Bell Labs in the 1960s, sprang from a culture where "we never forgot the fact that we existed to serve the telephone business," Sarewitz quotes one researcher as saying. So if the direction of basic research can be steered in part, is it too much to ask scientists to aim toward

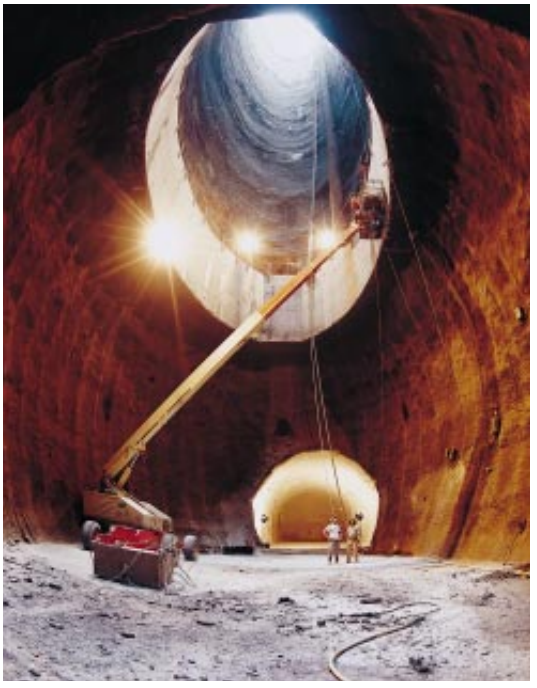
THE ILLUSTRATED PAGE

The Family of Cats
ILLUSTRATIONS
BY JOSEPH WOLF
Pomegranate Artbooks, Rohnert Park, Calif., 1997 ($16.95)

Joseph Wolf was one of the greatest natural illustrators of the 19th century, a status well documented in this small-format reproduction of 43 lithographs he created for the 1883 *A monograph of the Felidae, or family of cats.* Amazingly, Wolf never saw living examples of these animals (such as *Felis serval* at right), but he breathed vitality into the specimens and sketches he scrutinized. The majesty of the great cats—as well as the fascinating, stylized curatorial imaginations of a late-colonial naturalist—radiates from every page. —*Corey S. Powell*

AMERICAN MUSEUM OF NATURAL HISTORY

"Beautiful coats and mighty muscles bring into play the artist's power of modeling."



JOHN BIRD U.S. Department of Energy

END OF AN ERA?
After spending $2 billion on tunnel excavation and other work, the government canceled the Superconducting Super Collider in 1993.

responsible use of their discoveries?

The usual response is that investigators are not responsible for abuses of their work by others: pesticides don't kill scores of Third World farmworkers every year; people do. Yet within the U.S., companies are increasingly being forced to accept responsibility for their creations. Sarewitz wants researchers and technologists generally to hold themselves more responsible.

The author also takes aim at the notion of scientific accountability—that scientists discharge their responsibilities to society simply by conducting work of high intellectual integrity. The corollary is that any outside attempt to influence research agendas stems from ignorance of the power of reproducible experiments and of peer review.

Sarewitz holds that those principles are not enough: researchers must level with society in a way that some now seem unwilling to contemplate. They might, for example, have to confront possible dangers inherent in their work at an earlier stage. They might even consider putting aside a theoretically intriguing question in favor of a more urgent practical priority.

Sarewitz advances a new guiding myth for science: working toward sustainability, or developing technologies and systems that might allow humans to survive into the long-term future. It is an imprecise notion, but energizing just the same. Coincidentally (or not), Jane Lubchenco, an ecologist at Oregon State University who is this year's president of the AAAS, made a similar call for a "new social contract with science" in February. Science has made tremendous contributions to national well-being, she noted at the AAAS's annual meeting in Seattle, but a consensus of environmental scientists now holds that human activities have "unwittingly compromised" ecological systems essential to that well-being. In this "fundamentally different" situation, according to Lubchenco, scientists should rethink the direction of their work.

Sarewitz's book should make a lot of scientists rethink their work and the arguments they use to defend it. He is not looking, Pollyanna-like, for science to provide quick "rational" solutions to complex political problems; investigators are usually themselves deeply divided on the applications of technology. His arguments are generally cogent, although his distaste for what he calls reductionist science is puzzling. Reductionist is not the opposite of cross-disciplinary, as Sarewitz seems to think, and making progress on grave problems such as the global dying of coral reefs will need reductionist approaches.

But this quibble is terminological. Scientists horrified that one of their number should launch this fusillade should remember that Sarewitz is not advocating spending less government money on science, as some libertarian economists have been doing lately. Rather he wants researchers to strengthen their claim on the public purse, by wielding their enormous influence more proactively and by being more realistic about what they can deliver. Objections, anyone?

TIM BEARDSLEY is a staff writer for SCIENTIFIC AMERICAN.

BRIEFLY NOTED

Continued from page 141

BIOMIMICRY, *by Janine M. Benyus. William Morrow, New York, 1997 ($25).*
Materials scientists are learning from nature how to make tougher ceramics, more tenacious glues and stronger fibers. Perhaps less appreciated is the way that agricultural researchers are figuring out how to pattern farms after prairie ecosystems or that managers are learning to organize a business like an old-growth forest. The author ably brings together many disparate tracks of biomimetic work in a wide-ranging overview of this emerging and still speculative field.

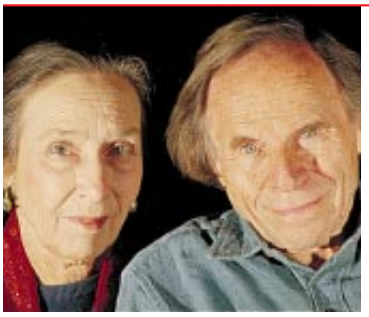
MINING THE SKY: UNTOLD RICHES FROM THE ASTEROIDS, COMETS, AND PLANETS, *by John S. Lewis. Helix Books/Addison-Wesley, Reading, Mass., 1996 ($26).*

IMAGINED WORLDS, *by Freeman Dyson. Harvard University Press, Cambridge, Mass., 1997 ($22).*
Resurrecting the decades-old dream of mining asteroids, John S. Lewis envisions a future of resources that are "for all practical purposes, infinite"; all we need to do is unleash economic forces, he argues, and the vast resources of the solar system could be ours for the taking. Short fictional interludes bring Lewis's zealous vision to life while underscoring how far it is from our current reality.

NASA OFFICE OF SPACE SCIENCE, JET PROPULSION LABORATORY

Like Lewis, the redoubtable Freeman Dyson draws on the spirit of science fiction in his broad-based look into the future. Dyson, too, envisions nearly unbounded possibilities, from space travel to "radiotelepathy" to genetic engineering, springing from the human imagination. But his essays (originally delivered as a series of lectures at the Hebrew University of Jerusalem) venture into more complex social and ethical territory, mulling over the fallacies of ideology-driven technology and asking how a world powered by the engine of capitalism can take care of its poor.

Despite some idiosyncrasies, these scientific speculations provide a thrilling jolt of perspective. They force us to step back from the world that is and ask, Where do we want to go, and why?



COMMENTARY

WONDERS

by Philip and Phylis Morrison

Wrapping up Science and Art

DUSAN PETRICIC

Technology, the aggregated tools of complex societies, sustains our everyday lives. Its benefits are utilitarian. We count on it to heal our own ills and indeed to work ill on our wartime enemies. Its flaws require a separate story, a long one.

Yet with similar vigor, high technology has entered the domain of play. Bold ocean-faring yachtsmen now rival the clipper masters of the tea trade. One person may be both captain and crew, using servos to replace dozens of able seamen in the rigging. Hardware and software alike take part; reliable radioed forecasts of favorable winds to come are only less important than the light-composite hulls. The lonely runner, too, benefits from well-designed shoes and track underfoot, possibly even more from the new physiology of pacing and diet (not to mention illicit steroids able to reconstruct an athlete's very musculature).

What of the ancient visual arts? We still consider painters and sculptors for the most part as individual artists, whose recognizable debts to pigments and welding gases are rather detached from higher values. In New York one gifted pair of artists has for a generation been assembling a new repository of the visual arts. Beautiful, fully individual, without any ancillary function or any trophy to win, their art is part and parcel of today's social and technological web. The two are Christo and Jeanne-Claude, husband and wife, whose original works have been viewed worldwide and enjoyed on the largest public scale.

That very scale demands imaginative use of current engineering—its expert powers of design and construction, its varied and novel materials in massive amounts—and daring teams of specialized artisans. These works of art come into existence only after full engagement with public agencies, with private owners of property, with uneasy neighbors, and after wide popular response.

What light catches and what our eyes see is mainly surface. Two-dimensional matter neither weighs nor costs as much as the bulk, and so surfaces of great area can be produced by linking long one-dimensional threads. That allows amazing scope for our world's swift-powered machines; building on ancestral handicrafts, they can fashion affordable surfaces by the square mile. Therefore, the primary medium of the ingenious works of Christo and Jeanne-Claude is cloth in plenty, mainly of synthetic fibers. These knowing artists also grasp that novelty means change; their displays are as fragile as they are unforgettable. Text, image and memory hold lasting records of their fascinating but transient work. Dazzling volumes and vivid tapes document the long creative process itself, its brief climax and equally its deft removal.

Their art is not merely conceptual, it is stunningly apparent to our senses and strongly rooted in our daily affairs. Creating such wonders is no matter of a lonely studio; it is our society that builds here, after the artists have opened the path to the engineers, craftsmen and suppliers, to knowing designs and scrupulous tests and fairly drawn contracts and, yes, to the persuasion of skeptics by presenting them, too, with a share in the beauty of the work.

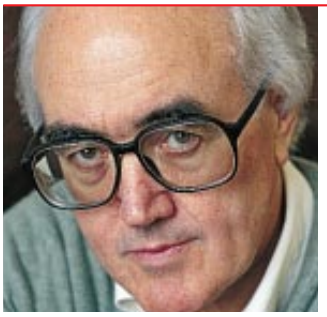
The scale of their high-tech art demands imaginative engineering.

We ourselves were first caught by the riveting film of their 1969 outdoor project, *Wrapped Coast*. Southward some miles from coastal Sydney, Australia, the seafront stretches out cliffed and bare before the winds that blow out of the Southern Ocean. What the film showed us was a landform quite new in color and texture, a surprise package cloth-wrapped by professional mountain climbers, laborers and plenty of eager students. That ocean scenery became a set for some powerful drama, abstracted beyond any natural scene, yet large enough so that rough and jointed nature seemed rather far off. Then the camera took us airborne; soon enough we could hardly locate the artifact in those wild wastes below. Two dozen acres of woven propylene makes a heroic stage, yet one that is all but unnoticeable seen on a shoreline that stretches out to far horizons. Art had disclosed the human scale in context.

In 1976 *Running Fence* emerged from the cold California waters in Bodega Bay, to extend as a shimmering, yielding curtain of fabric suspended 18 feet high for 25 miles across Sonoma County. A few country roads were crossed before the curtain ended in the small inland city of Petaluma. This is grazing land, and some 50 landowners, mostly resident ranchers, agreed to admit the passage of this gentle barrier for two weeks. Persuasive explanations by Jeanne-Claude, long formal hearings, volumes of statements, scrupulous environmental engineering, a wise and generous judge, and Christo's courage overcame all the litigious Californians.

Two thousand carefully stayed steel poles held the white nylon on steel cable; everything was tested, well documented, inspected and patrolled. Fragile, mobile, it was visually dominant as it crossed the skyline to disappear over every ridge. Looking at one long stretch all in ripple for the moment, Christo excitedly said, "See how it describes the wind!" In the event, it was as benign and as widely admired as it was beautiful, including the busy control post the

Continued on page 146



CONNECTIONS

by James Burke

Notice the Difference?

DUSAN PETRICIC

I was jogging along the Azorean pavements in predawn dark on New Year's Day and listening to the radio through my headphones when the incongruity struck me. Here I was, in the mid-Atlantic and plugged into the usual semiliterate stuff on the BBC World Service, all because back in the early 1940s somebody made a fuss about B-29s. Those planes' delicate onboard vacuum valves were vulnerable to the shakes and to temperature extremes—pretty much standard operating procedure on any bombing mission. No surprise then that the postwar Bell Labs work of John Bardeen, Walter H. Brattain and William B. Shockley got the 1956 Nobel for producing the solid-state, drop-it-on-the-floor-no-problem transistor made out of germanium and so brought enlightenment to the early-morning ear.

Solid-state amplifiers turned out to be just what people working on masers wanted. Researchers could get around coaxing microwaves out of excited ammonia gas molecules and graduate to the big time with doped crystals. Once the molecules of these crystals got excited enough to give off light, voilà a laser. The beam spread so little you could shine it at the moon and see where it lit. One reason for the excitement, in more sense than one, was a dopant called neodymium. If you pointed just a little illumination at it, at the right frequency, the molecule would shoot off a ton of laser light. So you got a lot more bang for your buck, as it were.

Ironically, the discovery of neodymium, in 1885, was associated with another form of blindingly bright light. While looking for materials that would emit an incandescent glow, the Austrian chemist Carl Auer von Welsbach came across it among the so-called rare earth elements. Welsbach was to get his name in lights that year by impregnating a cotton gas mantle with a mixture of these earths, thus making gaslight bright enough almost to beat Edison to the punch. Well, not quite. But he kept gas shares high for a decade or so longer than you might have expected after the first lightbulb. Even today his mantle is in your portable camping gaslight.

Welsbach was such a dazzler possibly because he had learned his stuff in the lab of the man who put the "Bunsen" in burner. And Robert Bunsen had got his gas, like everybody else, because of the chicanery surrounding the work of the seedy ninth earl of Dundonald, in Scotland. The earl was seeking a way to avoid bankruptcy when he roasted up some coal (he owned a tinpot mine and not much else), lit the fumes that came off, and made one of history's great discoveries without realizing it. Like an idiot, he mentioned the fumes to William Murdoch, James Watt's sidekick, who promptly snitched the idea. Dundonald eventually died penniless in a Paris garret, and Murdoch was recognized as the inventor of coal gas.

I think I'm getting this right, but it is *19th-century German psychophysics.*

From 1813 on, gaslight began to change life utterly, from factory night shifts to evenings out on the town, to increasing levels of literacy, to classes for the budding artisan at the new Mechanics Institutes. As well as making the plumber's candle famous. Weighing one sixth of a pound and burning 120 grains an hour, this humble illuminator became the standard by which the bright life was now to be quantified (that is, it became the official measure of candlepower).

A way of checking brightness was with a new gizmo called a photometer. This made sure that a gas company's clients were getting their money's worth. Some photometers were like small, double telescopes, in which a prism brought two images (of a plumber's candle and of a gas flame) side by side in the eyepiece. Then you moved a separate lens to magnify one image to the point where it appeared to be as bright as the other. The amount you had moved the magnifying lens to achieve this told you how much brighter the gas flame was.

Key word in this process: "appeared." Which is where some Germans entered the story, with a law of nature that I bet you've waited all your life to hear about: the "law of the just noticeable difference." In terms of light, this law came to matter most with instruments used to check stellar magnitudes (where perhaps humankind's awareness of just noticeable differences had begun, back when the ancients had classified by how much one star was brighter than another).

The modern law was first generally applied in October 1850, by a professor at the University of Leipzig named Gustav Theodor Fechner, father of psychophysics. Yet the idea was originally his teacher's, which is why it's Weber-Fechner's Law, and the point of it all was to measure by how much any sensation had to be increased before the increase was noticed. E. H. Weber tested it on touch by asking weight lifters at what point they noticed the extra kilos on the barbell. Fechner did similar things to the

Continued on page 146

Wonders, *continued from page 144*

Federal Aviation Administration had set up to shepherd the light planes that flocked to examine that long and luminous fence, now entirely gone.

Biscayne Bay is a shallow arm of the sea that enters some 10 miles into greater Miami. Our airliner banked a little to let us see the local sight of 1983, a set of eleven unused islets in the bay, mostly untended dumps for litter. Now every green islet was cleansed and fringed by a floating, anchored pad of pink propylene a couple of hundred feet wide. Were these *Surrounded Islands,* tropical lilies on the blue-green waters, Prospero's magic come lightheartedly to the New World? It goes without saying that the two-year preparation for the two-week display was marked by exceptionally demanding amphibious engineering. That earned many required permits, a few as thick as telephone books, capped by the provision of a 24-hour patrol, 100 monitors on inflatables.

It took 21 years of political maneuvering before fabric enwrapped the Reichstag in the summer of 1995, a fit way to make a new start, a century of European ruin and rebirth all in one for the people who crowded the space. How carefully it was wrapped! Fitted steel frames (200 tons of them) under the cloth protected every urn and muse of stone that bedeck the building. Sky by day, electrical floods by night, made the noble drapery glow blue, red and gold as it shifted in the summer winds. Some five million came to see their gift.

Many other projects have gone past, their materials recycled. Some lie ahead. One is *The Gates,* fluttering golden banners floating over two dozen miles of the walks of Central Park, a scheme pursued for nearly 20 years and not yet gained. Another project is *Over the River,* fabric stretched above a river's course for miles to echo its form and surface. Eighty-nine western rivers have been examined; six or so are possible choices. But before any achieve their transient reality, Christo will presage their coming with collages and drawings, whose worldwide sale pays for it all.

The ideas are not so important, Jeanne-Claude wrote: "What is important is TO DO IT." Both artists were born the same day of the same year: June 13, 1935. We wish a very happy birthday to Jeanne-Claude and to Christo. SA

Connections, *continued from page 145*

other senses. Between them, they showed that the just noticeable difference in any stimulus was a constant relating to the level of the basic stimulus.

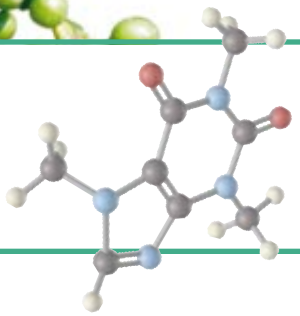
The whole idea of something *being* noticed originated from the work (alas, not noticed by many of his contemporaries) of a colleague of Weber's brother, Johann Herbart, the first to use the phrase "threshold of consciousness." Herbart's obsession was with how people learned anything. He had come up with the concept of a mass of experience one gradually accumulated: the "apperception mass." Any new experience that came along was referred to it, and if you'd had the experience before, the event was subconsciously noted as a ho-hum affair and nothing to bother your awareness with. But when something even *partly* new happened (I think I'm getting this right, but it *is* 19th-century German psychophysics), then—bingo!—it crossed the magic threshold, and you became conscious of it. You see the connection with weight lifters.

It was Herbart's friendship with a Swiss ex-farmer turned teacher-to-the-world, Johann Heinrich Pestalozzi, that got Herbart into making perception scientifically definable, because I suppose Pestalozzi didn't know (and didn't care) how to. By 1801 Pestalozzi had published *How Gertrude Teaches Her Children,* and by 1802 Herbart had written his *Pestalozzi's Idea of an ABC of Sense Perception.* Pestalozzi was by now famous for his school, where the kids learned from experience. No books, no formal classes. Development, not training. Showing the children a mountain before showing them the word for it.

By 1806 there was a Pestalozzi school in Philadelphia, run by one of his teachers, Joseph Neef. In 1825 Neef was recruited to start a school in the new utopian commune at New Harmony, Ind. This had been founded by the English libertarian Robert Owen, thanks to money from a businessman, William McClure. He was a geologist in his spare time and had published the first proper geologic map of the U.S. in 1809.

It was in the tristate area of Missouri, Oklahoma and Kansas that McClure's map identified the deposits from which, in 1952, zinc ores would provide the first major supply of germanium. I bet you're wondering where I dug this one up! SA

DECAFFEINATING COFFEE



CAFFEINE is a small, bitter-tasting alkaloid. High-quality Arabica coffee beans (the source of most specialty coffees) are typically 1 percent caffeine by weight, whereas cheaper and more bitter Robusta beans have twice that amount.

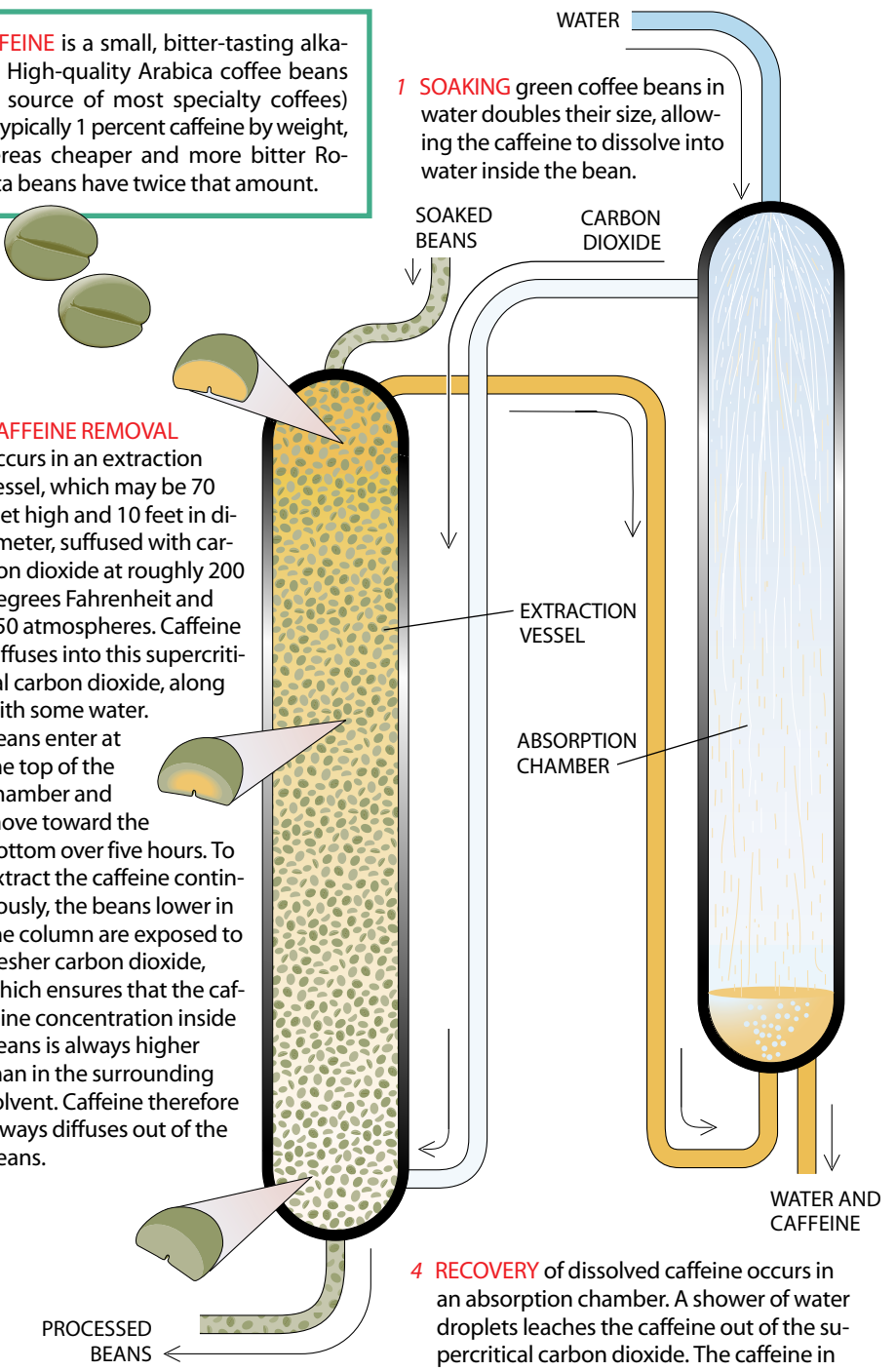
by Saul N. Katz

Spurred by the belief that excessive coffee drinking had poisoned his father, the German chemist Ludwig Roselius, in about 1900, found a number of compounds that dissolved the natural caffeine in coffee beans without ruining the drink's taste. Chloroform and benzene did the job but were toxic, so for 70 years methylene chloride became the solvent of choice.

When it was discovered in the 1980s to be a suspected carcinogen, the chemical was abandoned by all the big U.S. coffee labels. The Food and Drug Administration continues to permit the use of methylene chloride if the residues in the coffee are below 10 parts per million. Processing for specialty decafs still often uses it because it perturbs other flavorings so little.

Many other solvents can serve to debuzz coffee. An "all-natural" label may mean that ethyl acetate is the solvent in use, because that chemical occurs naturally in fruit. Water also works as a means of decaffeination. The so-called Swiss water process soaks green coffee beans in a solution that contains the chemical components of beans dissolved from a previous batch, except for the caffeine. Because the water is already saturated with sugars and peptides, only the caffeine passes from the beans into the water.

Another process, illustrated here, uses supercritical carbon dioxide as a solvent; in this state, the carbon dioxide is intermediate between a gas and a liquid. The variety of caffeine extraction methods demonstrates that a lot of sleepless nights have gone into helping the world get a good night's rest.



JEFFREY L. ROTMAN *Peter Arnold, Inc. (photograph);* MICHAEL GOODMAN *(illustrations)*

SAUL N. KATZ retired in 1989 as a principal scientist at the Maxwell House Division of General Foods. He holds several patents on the process for supercritical fluid extraction of caffeine.

