

SCIENTIFIC AMERICAN

**BLACK HOLE
BLOWBACK:**

Building
Galactic Clusters

MARCH 2007
WWW.SCIAM.COM

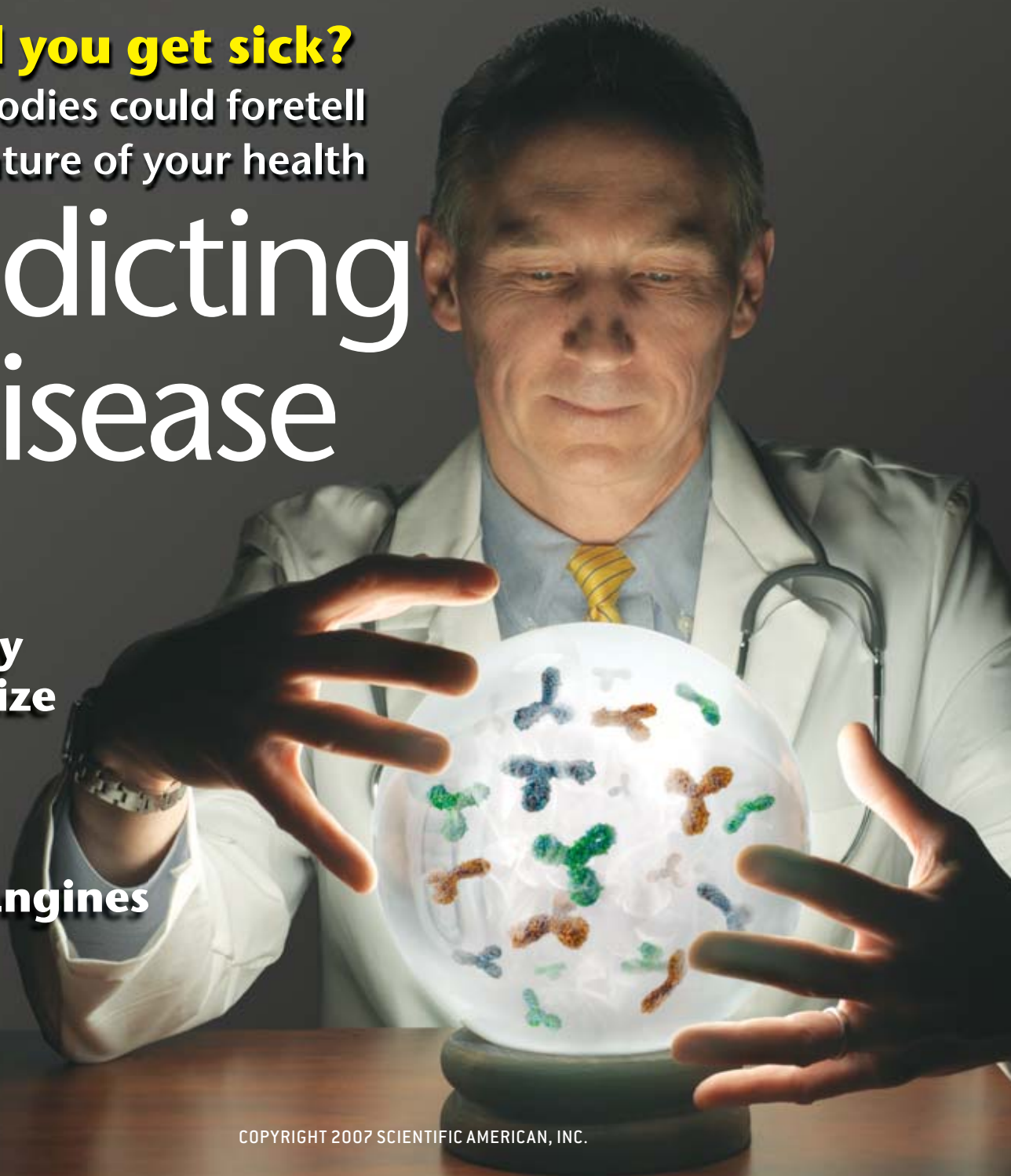
Will you get sick?

Antibodies could foretell
the future of your health

Predicting Disease

**Digitally
Memorize
Your Life**

**Cleaner
Diesel Engines**



march 2007
contents

SCIENTIFIC AMERICAN Volume 296 Number 3

features

ASTROPHYSICS

42 Black Hole Blowback

BY WALLACE TUCKER, HARVEY TANANBAUM
AND ANDREW FABIAN

A single black hole, smaller than the solar system,
can control the destiny of an entire cluster of galaxies.

BIOTECHNOLOGY

50 Mapping the Cancer Genome

BY FRANCIS S. COLLINS AND ANNA D. BARKER

The Cancer Genome Atlas will help chart a new course
across the complex landscape of human malignancies.

INFORMATION TECHNOLOGY

58 A Digital Life

BY GORDON BELL AND JIM GEMMELL

New systems may allow people to record everything that has touched their lives
and to store all these data in a personal digital archive.

ENVIRONMENT

66 Down Go the Dams

BY JANE C. MARKS

Many dams are being torn down these days, allowing rivers and the ecosystems
they support to rebound. But ecological risks abound as well. Can they be averted?

MEDICINE

72 New Predictors of Disease

BY ABNER LOUIS NOTKINS

Predictive autoantibodies appear in the blood years before people show symptoms of various
disorders. Tests that detected these molecules could warn of the need to take preventive action.

ENERGY

80 Diesels Come Clean

BY STEVEN ASHLEY

Improved engines and exhaust scrubbers, combined with a new fuel,
will make energy-efficient diesels nearly as green as hybrids.

NEUROSCIENCE

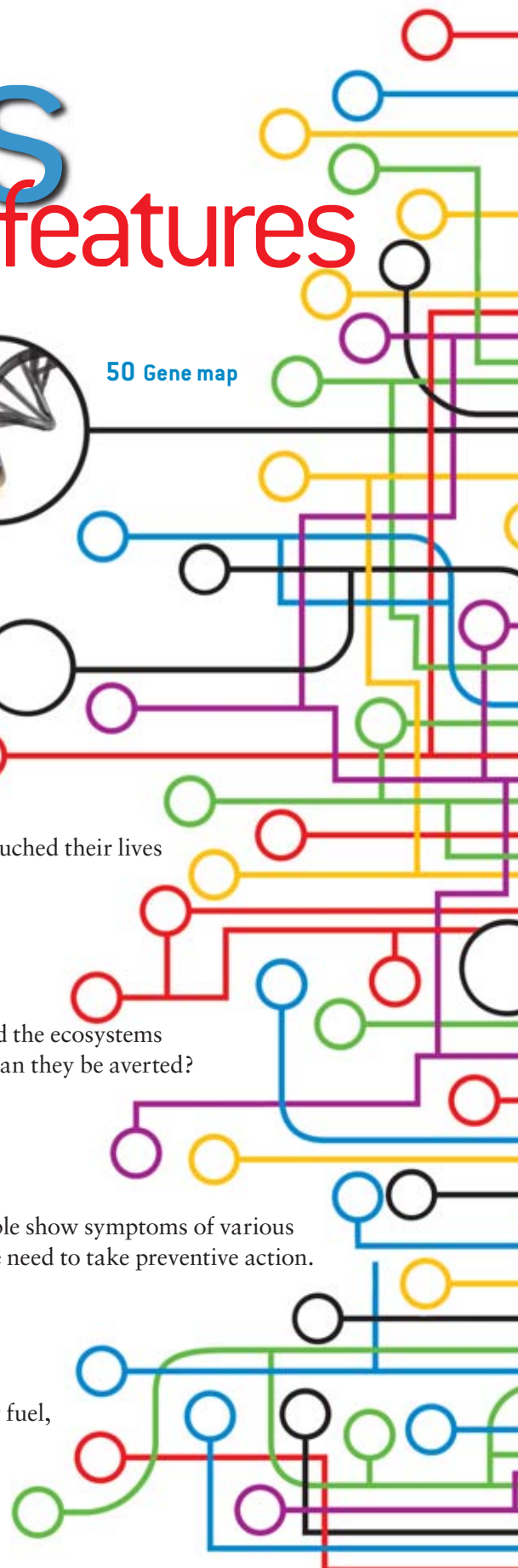
90 Illusory Color and the Brain

BY JOHN S. WERNER, BAINGIO PINNA AND LOTHAR SPILLMANN

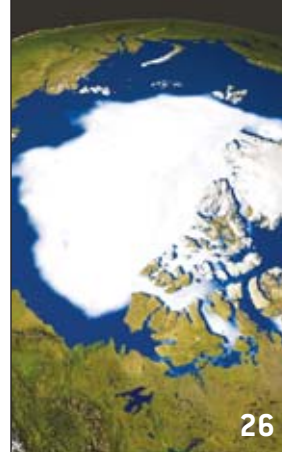
The brain may not separate perception of color from perception of form and depth.



50 Gene map



departments



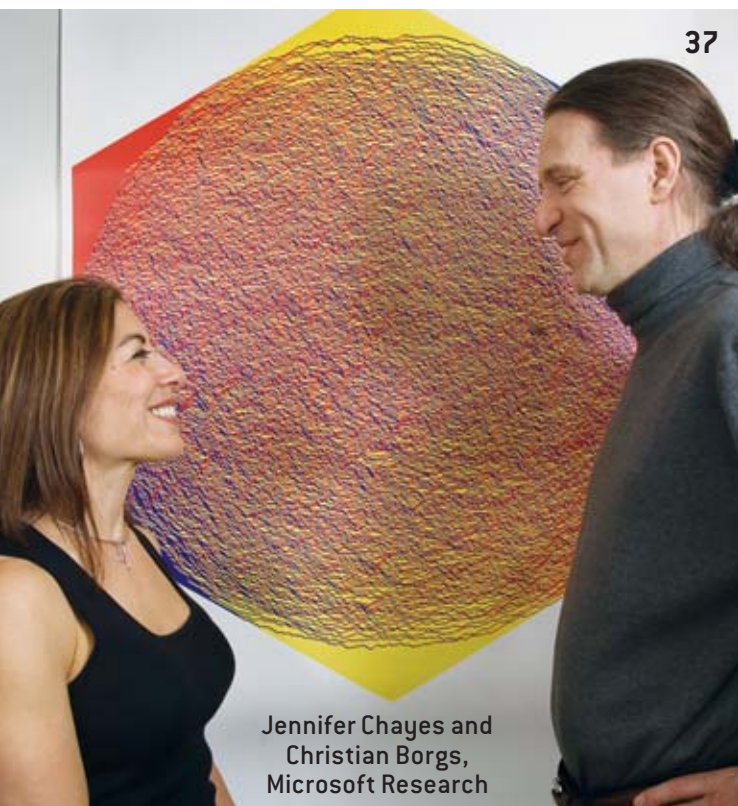
26



104

- 8 SA Perspectives**
The beef with cloned meat.
- 10 How to Contact Us**
- 10 On the Web**
- 12 Letters**
- 16 50, 100 & 150 Years Ago**
- 18 News Scan**
 - Stem cells to restore eyesight.
 - The Tevatron and the single top quark.
 - How tumors prepare “hot spots” for invasion.
 - International Polar Year.
 - What bedevils the beast from Down Under.
 - “Garlik” to block the identity vampires.
 - Data Points: Shrinkage and life span.
 - By the Numbers: Employed mothers.

- 37 Insights**
Christian Borgs and Jennifer Chayes head a theory group that pursues some of the most formidable challenges in pure math and theoretical computer science.
- 96 Working Knowledge**
Left ventricular assist devices.
- 98 Reviews**
In *I Am a Strange Loop*, Douglas R. Hofstadter explores the Gödelian vortex of the conscious mind.



37

Jennifer Chayes and Christian Borgs, Microsoft Research

columns

- 34 Skeptic** BY MICHAEL SHERMER
The new science of happiness needs some historical perspective.
- 34A Sustainable Developments**
BY JEFFREY D. SACHS
Preventing the spread of war will depend on strategies that recognize the shared interests of adversaries.
- 102 Anti Gravity** BY STEVE MIRSKY
Look out stomach, here comes the sword.
- 104 Ask the Experts**
Do the virtual particles in quantum mechanics really exist?
When you lose weight, where does it go?

Cover photography by Cary Wolinsky, with photoillustration by Jen Christiansen; photograph at left by Dan Lamont

Scientific American (ISSN 0036-8733), published monthly by Scientific American, Inc., 415 Madison Avenue, New York, N.Y. 10017-1111. Copyright © 2007 by Scientific American, Inc. All rights reserved. No part of this issue may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying and recording for public or private use, or by any information storage or retrieval system, without the prior written permission of the publisher. Periodicals postage paid at New York, N.Y., and at additional mailing offices. Canada Post International Publications Mail (Canadian Distribution) Sales Agreement No. 40012504. Canadian BN No. 127387652RT; QST No. Q1015332537. Publication Mail Agreement #40012504. Return undeliverable mail to Scientific American, P.O. Box 819, Stn Main, Markham, ON L3P 8A2. Subscription rates: one year \$34.97, Canada \$49 USD, International \$55 USD. Postmaster: Send address changes to Scientific American, Box 3187, Harlan, Iowa 51537. Reprints available: write Reprint Department, Scientific American, Inc., 415 Madison Avenue, New York, N.Y. 10017-1111; (212) 451-8877; fax: (212) 355-0408. Subscription inquiries: U.S. and Canada (800) 333-1199; other (515) 248-7684. Send e-mail to sacust@sciam.com Printed in U.S.A.



SA Perspectives

The Beef with Cloned Meat

The waiter places a perfectly grilled, prime-grade beefsteak before you and then reveals that it came from a cloned steer. Do you eat it? For most Americans, the answer is no. A survey conducted by the Pew Initiative on Food and Biotechnology found that the thought of dining on meat from animals copied via manual transfer of cell nuclei just does not sit well with six in 10 of us. Blame ethical or religious

concerns or mistrust of the meat industry, but the idea of cloned meat elicits distaste even in many confirmed carnivores.

Is that gut reaction justified? From a food-safety standpoint, probably not. In January, after reviewing available scientific reports about animal cloning, the U.S. Food and Drug Administration released a risk assessment asserting that food (including milk) from cloned cattle,

swine and goats is “likely to be as safe as” that from noncloned animals.

The FDA’s expected approval is unlikely to bring cloned meat products to store shelves soon, because cloning is still difficult and, hence, too expensive for routine food production. But ranchers and dairy producers may be willing to pay more than \$15,000 for a “rock star” breeding animal with superior genetics. Proponents claim that cloning these individuals will yield elite animals with desirable traits, such as general good health, disease resistance, greater productivity or leaner, tastier meat—without growth stimulants. The offspring of those clones will probably be the first to arrive at the dinner table. The distinction between this procedure and conventional animal husbandry would be the use of a genetic copy as breeding stock.

Detractors claim that this rosy scenario overlooks unresolved issues. Clone-based pregnancies, for example, result in more miscarriages, deformities and premature deaths than other methods do, but the FDA argues that these animal-welfare problems are not unique to cloning and that none are linked to human health risks. Many critics also fear overreliance on vulnerable monocultures of genetically identical animals that could be wiped out by a single disease. Even some members of the farm industry oppose animal cloning because cloned meat and dairy products could be shunned overseas, where food from genetically modified crops is often banned.

Perhaps the real issue here is one of full disclosure regarding our foodstuffs. Many meat eaters may be surprised to learn that the cattle industry has long employed a process called budding, in which technicians manually separate the undifferentiated cells in a fertilized cow egg that has undergone several divisions. Each of these cells is then grown into an identical individual, in some cases yielding hundreds of artificially induced twins, or “natural” clones.

Considering that the public has already been eating meat manipulated by high-tech means, an open debate might help inform or overcome skepticism about animal cloning. Such a discussion would require that Americans have ready access to detailed information about the food they eat. Industry marketing concerns forbid such full disclosure. We, however, believe that consumers should have the right to know whether their food was raised in a way they deem acceptable. Only clear and complete labeling of all food products, beyond today’s incomplete and sometimes misleading tags, can bring this about—and not just for cloned products, which might otherwise suffer unjustly in a system where food producers routinely game the meanings of “organic” and “natural.”



WHO'S THE CLONE? Only our rancher knows for sure.

THE EDITORS editors@sciam.com

JIM RICHARDSON Corbis

I How to Contact Us

EDITORIAL

FOR LETTERS TO THE EDITORS:

Letters to the Editors
Scientific American
415 Madison Ave.
New York, NY 10017-1111

or

editors@sciam.com

Please include your name
and mailing address,
and cite the article
and the issue in
which it appeared.

Letters may be edited
for length and clarity.

We regret that we cannot
answer all correspondence.

For general inquiries:

Scientific American
415 Madison Ave.
New York, NY 10017-1111
212-451-8200
fax: 212-755-1976

or

editors@sciam.com

SUBSCRIPTIONS

FOR NEW SUBSCRIPTIONS,
RENEWALS, GIFTS, PAYMENTS,
AND CHANGES OF ADDRESS:

U.S. and Canada
800-333-1199

Outside North America
515-248-7684

or

www.sciam.com

or

Scientific American
Box 3187
Harlan, IA 51537

REPRINTS

TO ORDER REPRINTS OF ARTICLES:

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

PERMISSIONS

For permission to copy or reuse
material from SA:

www.sciam.com/permissions
or
212-451-8546 for procedures

or

Permissions Department
Scientific American
415 Madison Ave.
New York, NY 10017-1111

Please allow three to six weeks
for processing.

ADVERTISING

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

New York

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

West/Northwest

310-234-2699
fax: 310-234-2670

Detroit

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

Midwest

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

Southeast and Southwest

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

Direct Response

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

Australia

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

Belgium

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

Canada

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

France and Switzerland

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

Germany

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

Hong Kong

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

India

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

Japan

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

Korea

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

Middle East

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

The Netherlands

Insight Publicitas BV
+31-35-539-5111
fax: +31-35-531-0572

Scandinavia and Finland

M&M International Media AB
+46-8-24-5401
fax: +46-8-24-5402

U.K.

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

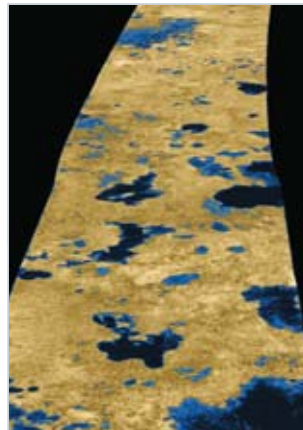
I On the Web

WWW.SCIAM.COM

UPDATED EVERY WEEKDAY

Visit www.sciam.com/ontheweb

to find these recent additions to the site:



Moon River: Titan's Polar Surface Dotted with Lakes of Methane

A missing methane sea

on Saturn's moon Titan
has been replaced by a
variety of lakes, according
to new radar images.

Fish Fin? How Climate Change Is Hurting Cold-Water Fish

A **bottom-loving fish** in the North Sea shows how
climate change can directly impact aquatic species—
and presage their local doom.

How Much AIDS Vaccine Do Poor Countries Really Want?

Public health groups take a page from business
to understand the developing world's weak demand
for drugs and vaccines.

FACT OR FICTION?

Urinating on a Jellyfish Sting Is an Effective Treatment

It worked for Monica on *Friends*, but how does the
alleged remedy hold up under scientific scrutiny?

PUZZLE

Sci-Doku

Try your hand at Sci-Doku, a Sudoku puzzle that uses
letters instead of numbers, with an added twist: a science-
related clue accompanies each puzzle, and the answer is
spelled out in one row or column of the solution.

Only at www.sciam.com

► New Podcasts

Every weekday: *60-Second Science*

Every Wednesday: *Science Talk*

Listen at www.sciam.com/podcast

EDITOR IN CHIEF: John Rennie
EXECUTIVE EDITOR: Mariette DiChristina
MANAGING EDITOR: Ricki L. Rusting
NEWS EDITOR: Philip M. Yam
SPECIAL PROJECTS EDITOR: Gary Stix
SENIOR EDITOR: Michelle Press
EDITORS: Mark Alpert, Steven Ashley, Graham P. Collins, Mark Fischetti, Steve Mirsky, George Musser, Christine Soares
CONTRIBUTING EDITORS: W. Wayt Gibbs, Marguerite Holloway, Michael Shermer, Sarah Simpson

EDITORIAL DIRECTOR, ONLINE: Kate Wong
NEWS EDITOR, ONLINE: Lisa Stein
ASSOCIATE EDITOR, ONLINE: David Biello
NEWS REPORTER, ONLINE: JR Minkel

ART DIRECTOR: Edward Bell
SENIOR ASSOCIATE ART DIRECTOR: Jana Brenning
ASSOCIATE ART DIRECTOR: Mark Clemens
ASSISTANT ART DIRECTOR: Johnny Johnson
PHOTOGRAPHY EDITOR: Emily Harrison
PRODUCTION EDITOR: Richard Hunt

COPY DIRECTOR: Maria-Christina Keller
COPY CHIEF: Molly K. Frances
COPY AND RESEARCH: Daniel C. Schlenoff, Michael Battaglia, Smitha Alampur, Michelle Wright, John Matson

EDITORIAL ADMINISTRATOR: Jacob Lasky
SENIOR SECRETARY: Maya Hartly

ASSOCIATE PUBLISHER, PRODUCTION: William Sherman
MANUFACTURING MANAGER: Janet Cermak
ADVERTISING PRODUCTION MANAGER: Carl Cherebin
PREPRESS AND QUALITY MANAGER: Silvia De Santis
PRODUCTION MANAGER: Christina Hippeli
CUSTOM PUBLISHING MANAGER: Madelyn Keyes-Milch

ASSOCIATE PUBLISHER, CIRCULATION: Simon Aronin
CIRCULATION DIRECTOR: Christian Dorbandt
RENEWALS MANAGER: Karen Singer
FULFILLMENT AND DISTRIBUTION MANAGER: Rosa Davis

VICE PRESIDENT AND PUBLISHER: Bruce Brandon
SALES DEVELOPMENT MANAGER: David Tirpack
SALES REPRESENTATIVES: Jeffrey Crennan, Stephen Dudley, Stan Schmidt

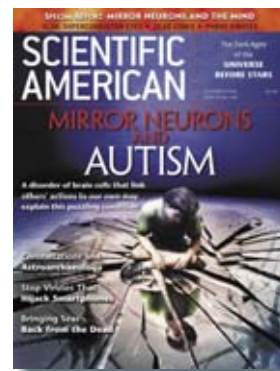
ASSOCIATE PUBLISHER, STRATEGIC PLANNING: Laura Salant
PROMOTION MANAGER: Diane Schube
RESEARCH MANAGER: Aida Dadurian
PROMOTION DESIGN MANAGER: Nancy Mongelli
GENERAL MANAGER: Michael Florek
BUSINESS MANAGER: Marie Maher
MANAGER, ADVERTISING ACCOUNTING AND COORDINATION: Constance Holmes

DIRECTOR, SPECIAL PROJECTS: Barth David Schwartz
MANAGING DIRECTOR AND VICE PRESIDENT, ONLINE: Mina C. Lux
DIRECTOR, WEB TECHNOLOGIES, ONLINE: Vincent Ma
SALES REPRESENTATIVE, ONLINE: Gary Bronson

DIRECTOR, ANCILLARY PRODUCTS: Diane McGarvey
PERMISSIONS MANAGER: Linda Hertz

CHAIRMAN EMERITUS: John J. Hanley
CHAIRMAN: Brian Napack
PRESIDENT AND CHIEF EXECUTIVE OFFICER: Gretchen G. Teichgraber
VICE PRESIDENT AND MANAGING DIRECTOR, INTERNATIONAL: Dean Sanderson
VICE PRESIDENT: Frances Newburg

“JEFFREY D. SACHS is a great man, doing great things for people around the world,” began a note from Trey Strawn of Winter Springs, Fla. **“His battle against poverty is inspiring. ‘Welfare States, beyond Ideology’ [Sustainable Developments], however, is unscientific, biased and wrong.”** Find out why readers such as Strawn responded with some heat to Sachs’s November 2006 column in the letters below. Correspondents were also drawn to fiery phenomena in other articles in that issue: from the mysterious sparkings in our brains’ inner space explored in a special section on mirror neurons to the first light of outer space described in an article about the Dark Ages of the universe.



NEURON SEE, NEURON DO

I am a clinician who often works with children and adolescents with Asperger’s syndrome, and I am an author of two books on the subject. The research and theory described in “Mirrors in the Mind,” by Giacomo Rizzolatti, Leonardo Fogassi and Vittorio Gallese, and in “Broken Mirrors: A Theory of Autism,” by Vilayanur S. Ramachandran and Lindsay M. Oberman, are consistent with, and help to explain, what these children experience and what others experience when they are with them. Those autistics who are higher functioning can learn about our language, such as the meaning of specific metaphors. They can even use these metaphors by translating, although not from their own intrinsic understanding.

Children and adults who think this way say they are often told that they are wrong, because they do not perceive something the same way others do. This correction is particularly disconcerting for autistics who cannot detect any overt reason for the “illogical” perceptions that most of us have and that we insist are the right answer. I was not familiar with the bouba/kiki effect, described by Ramachandran and Oberman, yet realized that I immediately “felt” that the rounded shape was a bouba and that it was likely that those on the autism spectrum would not. So it is undoubtedly from their perspective that I noted the statement “we discovered that children with autism perform poorly at the bou-

ba/kiki test, pairing the shapes and sounds incorrectly” would mean to the children that their perception was the wrong one. Even if they learn the “right answer,” it would be by memorizing something that was meaningless to them. Yet understanding what the authors describe can facilitate communication between the Asperger children and us, and it is important in education, especially with subjects who rely on shared understandings, so as to be meaningful and engaging for these students.

Here in Silicon Valley, we also seem to have many highly functional and successful individuals whose “Asperger thinking” is not only helpful but actually makes their work possible. Perhaps research will someday enable us to understand the exceptional gifts that some of these individuals have.

Paula Jacobsen
 Department of Psychiatry and Behavioral Sciences
 Stanford University School of Medicine

SOCIAL SERFDOM?

In “Welfare States, beyond Ideology” [Sustainable Developments], Jeffrey D. Sachs claims that Friedrich Hayek’s thesis that high taxation would be “the road to serfdom” was wrong, because the Nordic welfare states outperform the liberal Anglo-Saxon economies.

But Hayek did not say that social spending would lead to serfdom. He said that government control of production would, and one of the pillars of the Nor-

dic model has always been private businesses in open markets. In that environment our companies have been competitive. And that began before we had higher taxes than the U.S., so it is hardly an argument for greater taxation.

Since then, we have lagged behind, partly because social spending has pushed people onto the welfare rolls (hidden from the official unemployment statistics). For example, the McKinsey Global Institute estimates Sweden's real unemployment to be 15 percent. On a per capita income basis, only the oil kingdom of Norway can compete with the U.S. today. If the others were part of the U.S., Denmark would be the tenth poorest state, Sweden would be seventh and Finland fifth.

The Nordic countries have mixed economies, producing mixed results. Just what Hayek would have predicted.

Johan Norberg
Stockholm

Three cheers for Sachs, who wrote that a generous social-welfare state is a road to fairness, economic equality and international competitiveness. One of the things that most needs changing in the U.S. is health care. People put off going to the doctor because it is so expensive and can necessitate time-consuming referrals. If other countries have proved that fair health care can be done, we should be able to provide it here.

Judy Uhart
via e-mail

MOBILE MALADIES

"Malware Goes Mobile," by Mikko Hypponen, describes the dangers of smartphone viruses. Nevertheless, it somewhat misses the point about precautions.

Educating phone users is clearly not sufficient. It has not helped in the PC world, and the average phone user is much less technologically sophisticated than the average PC user. Antivirus software is also of limited use. For one, such software is always one step behind the most innovative malware. The real rea-

son smartphones are a promising malware target is structural. Present smartphones run software comprising on the order of seven million lines of code, which can be expected to contain literally tens of thousands of bugs.

In other words, the faults contained in the smartphone software present a huge target for the bad guys. Antivirus software will not help here, nor will education. Some faults (such as buffer over-



MALWARE can hijack smartphones, sending infected messages while running up the owner's bill.

runs) can be triggered by relatively simple means (for instance, sending a carefully crafted multimedia message that exploits a specific bug).

The only protection is a system structure that limits the damage that bugs can cause. Such a structure must be based on a minimal (to minimize the number of faults it contains) trusted computing base (the part of the system that can circumvent security). And it must utilize this trusted computing base to encapsulate system components and protect them from one another.

Gernot Heiser
Australia

BEFORE THE COSMOS LIGHTENED UP

"The Dark Ages of the Universe," by Abraham Loeb, states that "if, for example, both the electron and proton [in a hydrogen atom] point up, and the electron then flips so that it points down, the atomic state will decrease in energy and give off a photon with a wavelength of 21 centimeters." Could you help all of us

nonphysicists and explain what causes an electron to "flip"?

Don Kolehose
Belmont, Mich.

LOEB REPLIES: *In quantum mechanics an electron bound to an atom has different possible states. A transition, or decay, from a higher energy state to a lower level occurs spontaneously and is accompanied by the emission of a photon (a packet of light) that carries the energy difference between the two states. The only question is how much time it would take the electron to make any particular transition. The equations of quantum mechanics allow us to calculate the probability per unit time for different transitions. In particular, one finds that a hydrogen atom at the ground level, in which the proton and electron spins are aligned, will decay into a lower energy state in which they are antialigned over a typical period of 10 million years. Not all the atoms will have exactly the same decay time. The probability for a decay is exponential with this characteristic decay time.*

[Quantum mechanics is about calculating probabilities. An ensemble of atoms will behave statistically according to the calculated probability distribution. The behavior of each particular atom cannot be predicted deterministically, as in classical physics.]

ERRATUM "Keeping CO₂ Down," by Rebecca Renner [News Scan], stated that carbon dioxide was compressed into a supercritical fluid at 15 degrees Celsius. To reach that supercritical state, the CO₂ was warmed to a temperature above 31 degrees C.

CLARIFICATION "The Evolution of Future Wealth," by Stuart A. Kauffman [Forum], refers to the evolution of swim bladders to lungs in fish as an example of preadaptation. Most ichthyologists now believe the reverse to be true: swim bladders most likely developed from primordial lungs.

Send letters to editors@sciam.com or to Scientific American, 415 Madison Ave., New York, NY 10017. Letters become the property of SCIENTIFIC AMERICAN and may be edited for space and clarity.

150, 100 & 150 Years Ago

FROM SCIENTIFIC AMERICAN

Piaget's Physics ■ Babylon's Taxes ■ Mobile's Cotton

MARCH 1957

CHILDREN AND PHYSICS—"Does a child's first conception of velocity include comprehension of it as a function of distance and time, or is his notion more primitive and intuitive? Albert Einstein himself posed this question to me in 1928 when I was demonstrating some experiments on causality to him one day. I have since performed a very simple experiment which shows that a child does not think of velocity in terms of the distance-time relation. We place before the child two tunnels, one of which is obviously much longer than the other, and then we push a doll through each tunnel with a metal rod in such a way that the dolls arrive at the other end of both tunnels simultaneously. We ask the child:

'Is one tunnel longer than the other?'

'Yes, that one.'

'Did both dolls go through the tunnels at the same speed, or did one go faster than the other?'

'The same speed.'

'Why?'

'Because they arrived at the same time.'
—Jean Piaget"

MARCH 1907

ANCIENT BABYLON—"Before the hand of destiny consigned it to decay, Nippur must have been a highly desirable place of residence. The 'library' found there consists of about twenty-five thousand books and documents in the form of the clay tablets of the time. Prof. Albert T. Clay, of the University of Pennsylvania, has succeeded in translating many of the important tablets. It would seem that the saying regarding the futility of escape from death or the tax collector must first have been wailed forth from the overburdened soul of a resident of Nippur, 1400 B.C., for many of the documents found are records of receipts for rent or taxes. The taxes were paid, not in coin, but in natural products, such as corn, sesame, oil, dates, flour, and live stock."

MUSIC BY TELEPHONE—"Dr. Cahill's wonderful invention, the 'telharmonium,' consists in generating electrical oscillations corresponding with the acoustic vibrations of the various elemental tones desired, in synthesizing the different notes and chords required, in transmitting these oscillations by means of wires to any desired point, and in rendering the synthesized electrical vibrations audible by a translating device such as an ordinary telephone receiver or a speaking arc (which in our illustration is disguised as a hanging plant)."

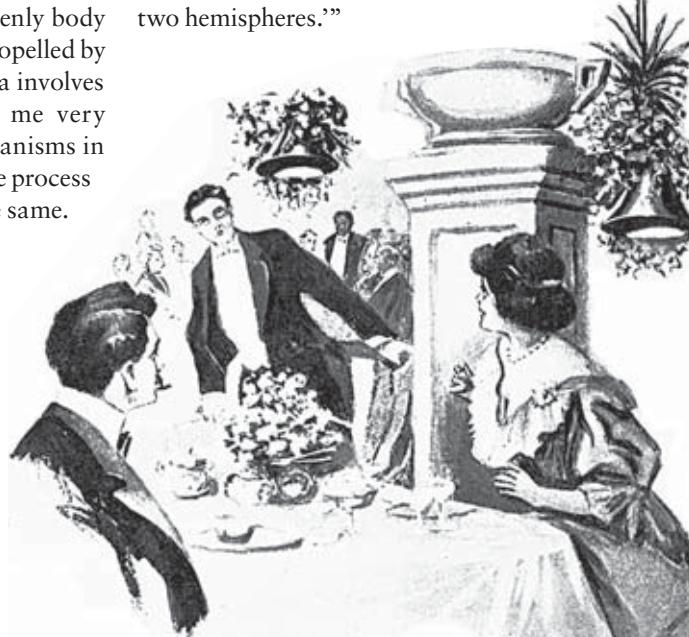
PANSPERMY—"It has been demonstrated that intense cold is not injurious to all germs. The rapidity of the photochemical changes induced by light and the rapidity of desiccation would be similarly diminished by cold. Hence we may perhaps conclude that the preservative effect of the low temperature of interstellar space assures the possibility of the conveyance of living germs from one solar system to another. Therefore spontaneous generation is unnecessary, as life can be transmitted from one heavenly body to another by minute germs propelled by the pressure of light. This idea involves another, which appeals to me very strongly, namely, that all organisms in the universe are related and the process of evolution is everywhere the same."
—Prof. Svante Arrhenius"

MARCH 1857

COTTON FIRES—"The recent great fire in Mobile, in which several thousands of bales of cotton were destroyed, has caused an extensive discussion in the Southern papers on the expediency of baling cotton with wire. The principal advantage is, that wire will not burn like rope, and, bursting, scatter the

cotton to the flames and the wind, causing the destruction of every other bale within its reach. Cotton bound with wire could scarcely be made to blaze, and if combustion be carried on at all it must be in a smouldering condition."

CABLE-LAYING ATTEMPT—"The Secretary of the Navy has ordered the U.S. steamers *Niagara* and *Mississippi* to England this summer, to assist in laying down the submarine telegraph cable between Newfoundland and Ireland. The *Niagara* is the largest steam vessel of war in the world, and the *Mississippi* is the most powerful paddle-wheel steamer in our Navy. It is not yet known what ships the English government will furnish to perform its part. The *Niagara* will receive on board at London or Liverpool one half of the cable, and the other half will be put on board the English ship. As justly remarked by the daily papers, 'It is a sign of advancing civilization when the ships of war of these two great nations thus meet in mid-ocean, not for a naval battle, but in a peaceful effort to join the two hemispheres.'"



VIRTUAL ORCHESTRA: Electrically synthesized music transmitted to speakers disguised as hanging plants, in a concept from 1907

Page Intentionally Blank

SCIENTIFIC AMERICAN Digital

Sight for Sore Eyes

PROGRESS IN CELL TRANSPLANTS TO HEAL DAMAGED RETINAS BY ALISON SNYDER

Millions of photoreceptor cells residing in the human retina gather light and transmit signals to the brain. When these light-collecting cells die, they take the person's sight with them. Medical researchers hoping to reverse blindness have turned their gaze toward stem cells, and recent experiments have shown that these cells could replace photoreceptors lost in macular degeneration.

As the most common form of blindness, macular degeneration affects 10 percent of Americans older than 65 years. It first targets a protective lining called the retinal pigment epithelium (RPE), which shuttles nutrients to the photoreceptor cells and is vital for their survival. A transplant of fresh RPE tissue could rescue dying photoreceptors. But the approach is not feasible considering the large amounts of tissue needed to treat the millions of Americans who show signs of early macular degeneration.

Scientists at the biotechnology firm Advanced Cell Technology in Worcester, Mass., have generated a more abundant source of RPE cells. In 2004 they devised a way to coax embryonic stem cells to turn into transplantable RPE tissue. In a follow-up

experiment, they injected the transformed cells into the eyes of rats that had a photoreceptor-killing genetic defect in their RPE cells. As the researchers reported in the September 2006 *Cloning and Stem Cells*, weeks later, when the effects of the disease would have normally set in, the rats receiving the treatment were able to track stripes on a rotating cylinder twice as well as those that did not. Their vision, though improved, was still far below normal.

But treating patients who have advanced degrees of macular degeneration or other photoreceptor diseases will ultimately require repairing the photoreceptor cells themselves. Last November researchers at University College London and other institutions announced that they had extracted cells from mouse retinas that were at different developmental stages and successfully transplanted them into blind mice. They found that immature photoreceptor cells from healthy newborn mice, rather than embryonic or adult mouse cells, migrated to the correct region of the retina and continued to develop into mature photoreceptor cells. The pupils that received these cells were also more sensitive to light than those that did not receive the transplant.

These findings have suggested the development stages at which to transplant cells—for instance, photoreceptor cells need to be



RETINA is a target for stem cell therapies. Visible in this retina cross section are the cells that form the optic nerve (pale red) and the rod (white) and cone (yellow) photoreceptors.

NEED TO KNOW:
EYEING RISKS

If scientists can generate and precisely identify photoreceptor cells derived from retinal stem cells, they will still have to address a host of other possible problems associated with stem cell transplantation.

Chief among them:

- Rejection of the transplanted cells
- Tumor development associated with stem cells
- Survival of the injected cells in a dying retina

relatively more mature than stem cells, according to Thomas Reh, who studies retinal development at the University of Washington. The human equivalent to the mouse cells, however, would have to be isolated from fetal retinas, posing the familiar problem of finding a source for the immature cells. Adult stem cells and cornea stem cells are two other possible sources for generating immature photoreceptor cells.

In his lab, Reh coaxes human embryonic stem cells into retinal stem cells, and currently about 6 percent of them subsequently turn into photoreceptor cells. That yield may sound small, but a low percentage is not necessarily discouraging, according to Evan Snyder, a stem cell researcher at the Burnham Institute for Medical Research in La Jolla, Calif. By studying what pushes those 6 percent into their fate as photoreceptor cells, researchers might figure out how to generate a larger number of transplantable cells. They might also come up with a way to select the right cells out of a mixed population; Anand Swaroop, an

ophthalmology researcher at the University of Michigan at Ann Arbor, is working on a way to identify and weed out the photoreceptor cells by focusing on proteins present on cell surfaces.

Having generated a cell source and overcome the safety concerns associated with transplanting stem cells, researchers still face possibly their biggest challenge: showing that the transplanted photoreceptors wire up to other neurons that eventually connect to the optic nerves. Each photoreceptor must make hundreds of critical connections. "Just because you have the right cell type doesn't mean you have the right circuitry," Snyder says. The immature photoreceptors transplanted from mouse retinas show activity, but Swaroop cautions that behavioral tests must determine that the photoreceptor cells are being repaired. A partial connection could generate the activity seen in the mice's pupils, but true vision improvement depends on the animals' ability to react to color and other visual cues. Seeing, after all, is believing.

PHYSICS

Alone at the Top

CLOSER TO GOD: FERMILAB MAKES SOLO TOP QUARKS BY ALEXANDER HELLEMANS

The world's biggest accelerator, the Large Hadron Collider (LHC) at CERN, the European laboratory for particle physics near Geneva, will come on line in a few months. Even so, for the next few years it may have a hard time upstaging the Tevatron collider at Fermi National Accelerator Laboratory (Fermilab) in Batavia, Ill., which appears to have generated "single" top quarks. The finding, reported last December, helps to narrow the search for the long sought after Higgs particle and raises the possibility that Fermilab will find it before the LHC does.

In 1995 Fermilab first produced top quarks, the heaviest and most elusive of the six quark types, in collisions between protons and antiprotons that generated both the top quark and its antimatter twin. These top-antitop pairs form via the so-called strong force, which binds quarks together. Very rarely, according to the Standard Model of particle

physics, top quarks may emerge in collisions via the weak force, which causes radioactive decay and can convert one flavor of quark to another. Such weakly made tops, however, would come without their antitop companions (instead a different antiquark, an anti-bottom quark, forms with the top quark).

Not only is single top production extremely rare, but its signature is not all that distinctive. "The backgrounds of events that look just like the single top quark events are very high," says Ann Heinson, a physicist at the University of California, Riverside, and co-leader of a group in DZero, one of the two research collaborations using the Tevatron.

From the trillions of collisions recorded since 2002, DZero has now identified 62 events indicative of single top quarks. That is not a lot of evidence, but still the data have strongly encouraged Tevatron researchers' hopes of beating CERN's LHC



ON TOP: Fermilab's main injector ring (*foreground*) feeds accelerated particles to the Tevatron ring (*background*), which is two kilometers wide and sits 10 meters below ground. Particle collisions have yielded top quarks without their antimatter twins.

in hunting down the Higgs. This “God particle,” predicted by the Standard Model, would explain why protons, neutrons and other types of matter have mass.

The detection of single top quarks serves as a kind of Tevatron dress rehearsal for finding the Higgs. “If the Higgs is of relatively low mass, it will have the same decay signature as that of the single top quark: a W particle, a bottom quark and an antibottom quark,” says Gregorio Bernardi of the University of Paris, who is co-leader of DZero’s Higgs physics group. This similarity would enable the group to utilize the advanced analysis techniques developed for the search of the single top. “We have improved our understanding and modeling of the background of the single top, and both improvements can also be transferred directly to the Higgs search,” Heinson adds. What sets the Tevatron ahead of the LHC is background “noise.” The Tevatron smashes protons with antiprotons so that their constituent quarks and antiquarks collide directly. The LHC, however, smashes protons with protons. Quarks end up colliding with antiquarks within a virtual “sea” of quark-antiquark pairs that continuously appear and disappear inside antiprotons. That complicates data analysis.

And for the Tevatron, things now look good. In January the Collider Detector at Fermilab (CDF) collaboration announced a mass determination of the W boson to within 0.06 percent—the best ever obtained. The new weighing of the W mass pushes the upper limit for the Higgs mass down from 166 billion electron volts (GeV) to 153 GeV, improving the odds that the Higgs mass is near its lower limit of about 114 GeV.

CDF member Tommaso Dorigo of the University of Padua argues that if the Higgs particle mass is around 114 GeV, the LHC will have a harder time finding a light Higgs particle than the Tevatron will. The LHC will have to detect the creation of two gamma-ray photons by a decaying Higgs particle, which are, in the strong background noise, much harder to sense than the bottom and antibottom quarks that would be produced by a disintegrating Higgs in the Tevatron.

A heavier Higgs, above 130 GeV, may escape Fermilab, because the Tevatron delivers just one seventh the wallop of the LHC. In fact, the LHC could find such a heavy Higgs rather quickly, explains David Plane, an experimental physicist at CERN. In the meantime, Plane remarks, “the Tevatron has the field pretty well to itself, until 2010 or maybe beyond.”

Alexander Hellemans, based in Paris, profiled Alain Connes and noncommutative geometry, an alternative to string theory, in the August 2006 issue.

NEED TO KNOW: MAKING SURE

Out of trillions of collisions since 2002, the DZero collaboration at Fermilab's Tevatron has found 62 events that signal the presence of “single” top quarks. “Data have to be at least doubled, even tripled, before you have a really clear scientific picture,” says Thomas Müller of Karlsruhe University in Germany, who is a member of the CDF collaboration, which is also looking for single top quarks at the Tevatron. DZero member Ann Heinson expects that within a year DZero will have analyzed a sufficient amount of data to claim an unequivocal identification of the particle.

Deadly Dialogue

HEALTHY TISSUE MAY INADVERTENTLY CALL IN TUMORS BY CHRISTINE SOARES

“What is it that decides what organs shall suffer a case of disseminated cancer?” asked London surgeon John Paget, exploring the deadly phenomenon of tumor metastasis in an 1889 edition of the journal *Lancet*. Cancer cells could spread with equal ease to any part of the body, he speculated, yet metastatic colonies seemed to favor certain organs, such as the lungs and liver. Paget imagined that the malignant cells might be like plant seeds that are carried by the winds in all directions but “can only live and grow if they fall on congenial soil.”

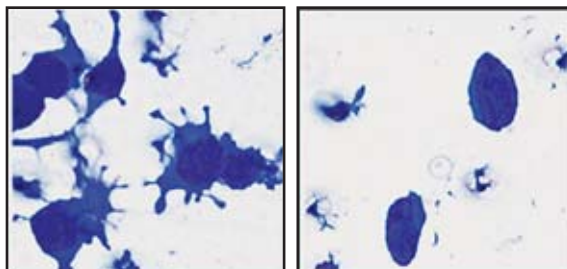
To this day, scientists are trying to understand whether properties of the “seeds” or the “soil” determine where spreading tu-

of healthy mice and found that inflammatory chemicals secreted by the cancer cells provoked a response in the animals’ lungs well before any malignant cells arrived there. That response, generated from local immune cells called macrophages and endothelial cells in the lung lining, was to begin producing compounds belonging to the so-called S100 family of proteins (specifically, S100A8 and S100A9). According to Maru, those proteins signaled more macrophages to rush to the scene, after which they, too, began emitting the S100 proteins. Finally the tumor cells themselves showed up to form colonies in the lung locations “bookmarked” by the S100 proteins.

Maru’s group found that, in addition to flagging these sites in the lungs, the S100 proteins—or some secondary chemical signal elicited by them—appeared to beckon to the malignant cells. Exposing cultured tumor cells to lung serum containing the S100 proteins even activated the cells’ formation of tiny feet, or invadopodia, in preparation for migration. And neutralizing the S100 proteins in live mice dramatically reduced the number of metastases formed by injected tumor cells.

The fact that S100 proteins generated in the lungs summoned both macrophages and tumor cells from afar indicates to Maru that normal immune responses may inadvertently invite metastases. “We don’t really understand why [else] macrophages would be migrating to the lung when you inject tumor cells into the animal’s back,” he says.

Studies have already demonstrated that cancer cells have an affinity for wounded or inflamed tissue and are perhaps drawn to sites of injury by the same signals that call for tissue repair crews. David Lyden and Shahin Rafii of Weill Cornell Medical College in New York City have shown, for example, that blood vessel progenitor cells from bone marrow homed in on future metastasis sites in the lungs and other organs, as if to prepare a cozy niche for tumor cells. Rendering those progenitors deaf to



HEEDING THE CALL: Responding to proteins generated in the lungs of mice, tumor cells project feet (*left*) in preparation to migrate. Cells cultured without the proteins (*right*) stay put.

mors take root, and growing evidence suggests that both play important roles. Researchers in Japan have recently added intriguing details to that theory in a study that reveals how distant tumor cells interact with future sites of metastasis in the lungs through a signaling mechanism involved in immune responses to pathogens. The participation of this pathway may also offer a clue as to why certain organs seem more susceptible to metastases than others, according to senior author Yoshiro Maru of Tokyo Women’s Medical University. “The lung is sensitive to microbes; it is a first defense,” Maru explains. “So any stress, including the presence of cancer, might be recognized by organs responsible for host defense mechanisms. That is our guess.”

Maru and his colleagues injected either melanoma or lung tumor cells into the backs

TAILORED TRANSMISSIONS

Researchers have observed bone marrow cells in mice migrating to premetastatic sites, but their locations varied depending on the type of cancer injected into the animals’ bodies. Marrow cells went mainly to the lungs and liver in response to lung carcinoma but fanned out to multiple organs when melanoma was present, mirroring typical patterns of metastasis for those cancers in humans.

Coupled with the recent observation of S100 proteins produced in the lungs, but not in other organs, in response to both cancer types, some investigators think that signals emanating from different kinds of cancer may provoke tailored responses from different tissues. These reciprocal signals could explain why most metastatic cancers seem to have preferences for certain organs. When colon cancer spreads, for instance, 95 percent of the time it heads to the liver.

signaling chemicals, or removing them from the bone marrow entirely, prevented tumor metastasis.

In a commentary on the Maru group's work in *Nature Cell Biology*, Lyden and Rafii note that collectively such results "set forth the novel concept that tumor metastatic potential may not only be dependent

on the [tumor-initiating ability] of the cancer cells, but also on the existence of 'hot spots' within each specific organ that are receptive to metastatic cells." If more of the precise signals emanating from such hot spots in various organs can be identified and blocked, Rafii believes, the soil can be made far less congenial to tumors.

CLIMATE

Pole Positions

THE INTERNATIONAL POLAR YEAR KICKS OFF BY CHARLES Q. CHOI

Arctic sea ice has shrunk to record low levels, and an ice shelf larger than Manhattan, which abruptly broke away from Canada's northernmost shore, could endanger ships and oil platforms this spring. To investigate these and other unprecedented changes occurring around the poles, more than 30 nations are initiating a global campaign to study the Arctic and Antarctic: an International Polar Year (IPY).

"Climate change is predicted to occur earliest and most dramatically at the polar regions, and that's what the observations seem to be showing," says Karl A. Erb, director of the National Science Foundation's office of polar programs. "Some places are warming faster than others, and some are cooling. The IPY research will help tease out what is driving those changes and differences."

In the first IPY, 125 years ago, a dozen countries collaborated on 15 dangerous missions. Some ended in disaster, some even generated sensational accusations of cannibalism, but altogether they yielded rare glimpses of the Arctic environment as it existed then. "The first IPY

had countries coming together to see the world as a system. Now we can see how the world is rapidly changing as a system," remarks oceanographer Stephanie Pfirman of Columbia University.

Just as the International Geophysical Year 50 years ago brought many firsts, such as ar-

tificial satellites, the latest IPY plans to continue this tradition of innovation. For example, researchers hope to deploy remote and autonomously operated vehicles to explore beneath the ice. "It'll be a challenge to push them into such harsh conditions where they can't routinely surface and transmit," says David Carlson, director of the IPY international program. Airborne sensors will communicate wirelessly and pack more instruments than before, he adds; besides temperature and pressure, they will measure cloud particles, radiation, humidity and even their own positions. Automated observatories at the South Pole will also help scientists operate telescopes from universities and get data back in real time.

Scientists will also explore polar biology in great depth. The studies include the genetic mechanisms behind adaptations to extreme cold and dark and the impact of polar changes both on wildlife and on the four million people living in northern communities.

The IPY officially starts this month and will actually run two years. It may include more than 200 projects. The U.S. has granted \$10 million so far. The program will leave behind observational networks for ongoing science, including sensor arrays to measure temperature and movement of glaciers or carbon emissions from the soil. Perhaps the most important legacy of the IPY will be human. "We'll be building scientific expertise in nations and collaborations between countries that will hopefully last for many generations," Erb states.

Charles Q. Choi is a frequent contributor.



ON THIN ICE: Researchers will probe the ends of the earth in great detail to explore, for instance, why Arctic sea ice (here, in the summer of 2005) has retreated so much.

Sympathy for the Devil

IDEAS EMERGE TO SAVE THE DYING TASMANIAN DEVIL BY WENDEE HOLT CAMP

During the past 10 years, a contagious and fatal cancer has decimated the world's Tasmanian devils. Pustulant tumors that become infested with maggots deform their faces, forcing teeth from their jaws. The devils eventually starve, but not before passing on the virulent cancer. Concerned that the disease could wipe out the devils, conservationists have already started planning how they might reintroduce the species if it goes extinct.

Resembling a small black dog with white splotches, these marsupial carnivores once lived in mainland Australia but today remain only on its island state of Tasmania. Devils have a ravenous appetite, an unearthly growl and a bone-crunching jaw strength that they use to devour carrion—skin, bones and all.

Around 1996, devils with the tumors started appearing in northeast Tasmania. Devil facial tumor disease (DFTD), as it is formally called, spread rapidly and now covers at least 56 percent of Tasmania—an island slightly larger than West Virginia. “In the areas where the disease has been for some years, we’ve seen an 80 percent population decline,” says biologist Menna Jones of the Tasmanian Department of Primary Industries and Water (DPIW). “There’s been no population recovery. They’re incurring no resistance. There’s very few animals over the age of two.” Devils normally live six to eight years.

The animals’ own demeanor contributes to the spread of DFTD: when the whirling dervishes fight and bite over a carrion meal or a mate, a diseased devil transfers cancer cells that infect the other individual.

Scientists initially suspected a virus but were unable to isolate one. Then Anne-Maree Pearse of the Tasmanian DPIW made a serendipitous discovery: devil DNA has 14 paired chromosomes, but devil tumor cells had only 13—and all had identical chromosomal rearrangements. Cancer tumors typically show genetic corruption, but having identical rearrangements would be nearly impossible. The best explanation: a rogue cell line emerged in a single devil that has taken on an infectious, cancerous existence.

With the population plummeting and sci-

entific answers potentially years away, conservation biologists are preparing for the worst. In 2006 Australian officials designated the once abundant species “vulnerable to extinction” and shipped 47 disease-free devils to mainland wildlife parks in “Project Ark”—a last-ditch effort to preserve the genetic diversity of devils across Tasmania for captive breeding.

An isolated 250-square-mile peninsula off southeast Tasmania may be a key for possible reintroduction. The Tasman peninsula once housed the world’s worst criminals; trained dogs, fences and armed guards kept escaped prisoners from reaching the mainland. That siege mentality may help keep sick devils out and healthy ones in. Researchers plan to build a bridge over the peninsula’s single narrow access point that will include water jets, spotlights and a cattle grid that devils cannot cross.

Research is suggesting other DFTD-beating strategies as well. Jones currently supervises an experiment to see whether removing infected devils can slow the cancer down. The first year’s results offer hope. “We have shown that disease suppression can be used to limit spread of the disease and can be used to prevent population decline,” she explains.

Jones acknowledges that, despite imperfect information, biologists must move rapidly. The demise of the devil could cause cascading effects in Tasmania’s ecosystem—especially since someone recently introduced red foxes to the area, a carnivore that has driven several local species to extinction. Devils can competitively keep fox populations down, because they fill a similar niche.

At this point, there is no single miraculous solution for saving the Tasmanian devil. Biologists still cannot even detect DFTD before tumors appear. But with millions of dollars being pumped into research, “Taz” may just be able to whirl and fight his way into the future.

Wendee Holtcamp is based in Houston.



FACIAL TUMOR DISEASE badly deforms Tasmanian devils and threatens to wipe them out.

CONTAGIOUS CANCER

Just how the transmissible tumor that is killing Tasmanian devils arose remains a mystery. A genetic mutation is the most likely answer; researchers initially suspected a cancer virus but thus far “we’ve found no viruses,” says veterinary pathologist Richmond Loh of the Tasmanian Department of Primary Industries and Water. The facial tumor disease, which is only one of two contagious cancers (the other being canine venereal tumor disease), somehow evades the devil’s immune system. The failure of the immune system to recognize the tumor as “nonself” may be a consequence of the low genetic diversity of the devil.

Devils in Tasmania’s isolated northwest corner have distinct genetic differences from the general population. Because the disease has encroached on that region only slowly, biologists are studying whether these genetic differences may confer disease resistance.

Scoring Your Identity

NEW TACTICS ROOT OUT THE FALSE USE OF PERSONAL DATA BY MARK FISCHETTI

As identity theft—and worry about it—burgeons around the world, online services are springing up to fight it. Their advanced tactics, collectively known as identity scoring, go beyond the usual credit monitoring to include online data mining, pattern recognition, even semantic analysis of information

about a subscriber that appears on Web pages. The approach is enjoying increasing success in unveiling suspicious activity. public databases and Web sites for information about customers and presents them with a detailed profile. The portrait allows them to see whether criminals may be trying to use their personal facts to apply for credit cards, take out a loan, or register a fake driver's license or marriage certificate. Four days after Garlik went live at www.garlik.com, more than 10,000 Brits had registered. An annual subscription costs £30 (\$59).

In the U.S., MyPublicInfo.com in Arlington, Va., will piece together a customer's "public identity profile" for \$79.95 and will alert him or her to dubious changes for \$4.95 a month. The firm has already served 100,000 consumers.

Since the initial rush at Garlik, demand for data patrol has continued at five times the anticipated rate, says chief executive officer Tom Ilube. Fraud is rising rapidly because people are posting personal facts on



SHREDDING will not thwart thieves who swipe information online.

Among the latest firms to join the battle is Garlik, a start-up in Richmond, England. In October 2006 it began offering its "data patrol" service to British residents, 100,000 of whom fell prey to identity theft last year. Garlik combs credit reports,

public databases and Web sites for information about customers and presents them with a detailed profile. The portrait allows them to see whether criminals may be trying to use their personal facts to apply for credit cards, take out a loan, or register a fake driver's license or marriage certificate. Four days after Garlik went live at www.garlik.com, more than 10,000 Brits had registered. An annual subscription costs £30 (\$59).

Since the initial rush at Garlik, demand for data patrol has continued at five times the anticipated rate, says chief executive officer Tom Ilube. Fraud is rising rapidly because people are posting personal facts on

essential₂th

NEED TO KNOW:
OPT-OUT CONTROL

the Web and because government agencies are steadily making databases available online. These databases include birth, marriage and death certificates, credit histories, voter registrations and property deeds.

Criminals interviewed by Garlik say it now takes them two or three hours instead of two or three weeks to build a convincing stolen identity. They are also cooperating with one another much more, Ilube says: "Identity theft is becoming organized crime."

Although identity scoring monitors the same financial details that credit bureaus do, its real strength comes from scrutinizing online sources for crucial data such as Social Security numbers and birth dates. "Identity scoring will tell you if someone is doing criminal activity in your name, which credit bureaus can't tell you," says analyst Avivah Litan of Gartner Group in Stamford, Conn. She predicts that by 2009, more U.S. consumers will sign up for identity scoring than for credit monitoring, which 24 million U.S. consumers paid for in 2006.

Nevertheless, Litan cautions that the services cannot tell if a crook tries to take over

someone's online brokerage accounts, because these vendors typically do not make account information public. And they cannot tell if a thief is selling your credentials to illegal immigrants, who are unlikely to be represented in databases. They also do not offer restitution services.

Garlik may be furthest along in spotting subtle vagaries. In addition to data-harvesting and pattern-recognition software, Garlik employs semantic techniques to analyze Web activity. The company has built so-called reference ontology software that scans data about Web pages (known as RDF, for resource description framework) that might refer to a Garlik subscriber's information. The software also rates "provenance"—the strength of relations among data, which can indicate a high likelihood that the real Joe Smith actually works for Widget Corp. and not Sprocket Corp. and lives in Portland, Ore., not Portland, Me.

Given the sheer weight of information out there, Ilube says, "semantic techniques will become increasingly important" to understanding the relevance of all the data.

"Social network sites such as MySpace and search engines such as Google are aware of identity theft," says Simon Davies, director of Privacy International in London, "but they're not giving users tools to protect information users provide." One rare exception, he says, is the "Shield & Scramble" tool offered by www.outeverywhere.com, a gay lifestyle site.

Davies states that Web companies should agree on a common mechanism that allows users to keep their information from being searched, as a means to thwart identity thieves. He doubts Web providers will do so, however, and says that legislation requiring opt-out controls may be needed.

egoodguys

It is what makes bullet-resistant vests and saving
police officers possible. It is chemistry.
americanchemistry.com
essential



DATA POINTS: SHRINK TO DEATH

Height loss occurs with age, but a drop of more than three centimeters correlates with an increase in illness and death, according to a 20-year survey of British men. The increase may result from mechanisms underlying bone loss, which reduces stature. Researchers have long known that osteoporosis affects mortality by compromising breathing and digestion, although the shrinkage in such cases tends to be at least six centimeters.

Ages of subjects, years: **60 to 79**

Average height lost, centimeters: **1.67**

Average weight change, kilograms: **+3.14**

Percent who lost at least 3 centimeters: **14.4**

Weight change, kilograms: **-0.2**

Percent increase in death compared with those who shrank less than 1 centimeter: **64**

When adjusted for health, smoking, weight loss and other factors: **45**

Adjusted percent risk increase:

From cardiovascular disease: **39**

From respiratory disease: **75**

From other, noncancer causes: **227**

SOURCE: Archives of Internal Medicine, December 11/25, 2006

NANOTECH

The Password Is G-L-O

Chemists in Israel have designed a molecule that can process a password in a manner similar to electronic keypads found in home security systems. The molecular lock, which resembles an iron-binding compound secreted by bacteria, “opens” when a combination of ultraviolet light and two chemical signals activates its fluorescent molecules. Information carried via the glowing molecules could thereby authorize a user or trigger another process. Because the molecular keypad relies on fluorescence, it can operate

on the level of a single molecule. Moreover, molecules would build up from more than one password attempt, causing the lock to jam and block further tries. The researchers suggest the device could be used, along with current molecule-based cryptography methods, to protect high-security information. The data would be safe even if an unauthorized individual knew the lock’s location and the decryption keys. They published their work in the January 17 *Journal of the American Chemical Society*. —Alison Snyder

BEHAVIOR

Cutting Back, Not Cutting Down

Smokers who cut back on cigarettes may not necessarily be curtailing the amount of toxic substances they inhale. Heavy smokers who light up less often still breathe in at least two times more toxic substances per cigarette than light smokers, according to a study by Dorothy Hatsukami and her colleagues at the University of Minnesota. Evidently, when heavy



FEWER CIGARETTES often translates to longer drags.

smokers puff on fewer cigarettes, they compensate for lower levels of nicotine by inhaling more deeply or longer on each one. The study supports earlier epidemiological work that revealed that individuals who cut the number of cigarettes they smoke do not have decreased health risks. “The bottom line is there may not be any health benefit to reducing the number of cigarettes,” Hatsukami says. If smokers want to reduce their risk of cancer and disease, she adds, they “need to stop smoking.” The study appears in the December 2006 *Cancer Epidemiology, Biomarkers and Prevention*. —Alison Snyder

ASTROPHYSICS

The Long and Short of It

Current understanding divides gamma-ray bursts into those that last less than two seconds and those of longer duration. The latter occur when a massive, young star goes supernova and leaves behind a black hole. The former happen when an old neutron star spirals into a preexisting black hole. The two kinds also differ in the type and intensity of energy released. But GRB060614, detected last June, fits neither category. Its duration of 102 seconds supports the longer-lasting variety, but the intensity of its emis-

sions more closely matches the shorter-lasting one. “All the data seem to point to a new but perhaps not so uncommon kind of cosmic explosion,” says Neil Gehrels of the NASA Goddard Space Flight Center, lead author of one of four papers in the December 21 *Nature* describing the phenomenon. Indeed, other bursts also hint at similar “hybrids.” A possible explanation is the creation of black holes so powerful that the dying stars get no chance to supernova.

—David Biello

BRIEF
POINTS

■ **Bovine biotech:** researchers have created eight Holstein cattle that lack the gene for the prion protein, rendering them resistant to mad cow disease. The animals appear healthy without the gene.

Nature Biotechnology,
January 2007

■ **Scientists have pinpointed one of the genes manipulated by Gregor Mendel—specifically, the gene that controls the color of pea seeds. It marks the third of the monk's seven pea genes to be precisely identified.**

Science, January 5, 2007

■ **A ceramic called zircon, envisioned as a container for nuclear waste, appears less stable than thought. A study finds that zircon holding 10 percent by weight of plutonium 239 would last only 1,400 years, rather than 250,000 years, as was hoped.**

Nature, January 11, 2007

■ **Not as chimp:** Factoring in the number of gene duplicates (the gene copy number) and not just the difference in equivalent genes, humans and chimps only have a 94 percent genetic similarity.

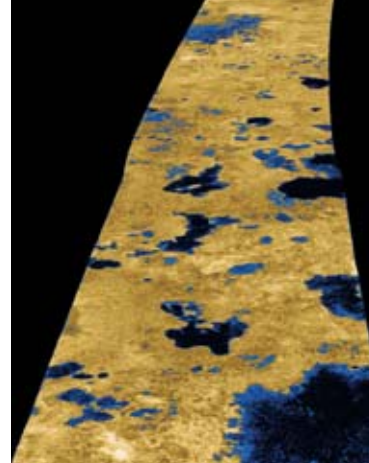
PLoS ONE, December 20, 2006

PLANETARY SCIENCE 

Titanic Lakes of Methane

Methane smog enshrouds Titan, Saturn's mysterious moon. Because of that smog, researchers have surmised that liquid methane should exist on the surface, but probes had failed to find any. Recent radar imaging by the Cassini space probe, however, has now uncovered 75 lake-like areas, some spanning 70 kilometers, near the northern pole. Scientists believe these to be liquid-filled depressions, because the temperature (−179 degrees Celsius) and pressure (1.5 times that at Earth's surface) there are ripe for liquid methane and its breakdown product, ethane, to persist. The lakes may fill up either from liquid stored underground or through evaporation and subsequent hydrocarbon rain. Future flybys should reveal how the lakes vary seasonally, as well as whether lakes exist elsewhere on Titan's surface. The report appears in the January 4 *Nature*.

—David Biello



METHANE LAKES dapple Titan's surface in this false-color radar image taken by the Cassini probe.

NEUROSCIENCE 

Back to the Future

The human mind taps into the same parts of the brain while imagining the future as it does when recollecting the past. Neuroscientists at Washington University in St. Louis put 21 volunteers in a functional magnetic resonance imaging machine and asked them to recall or imagine events, such as seeing themselves at a party with Bill Clinton. Eight different regions displayed extra activity—that is, increased blood flow—when

dealing with imagining the future, including Brodmann's area, the medial posterior parietal cortex and the posterior cerebellum. An additional 15 regions played a role in either remembering the past or imagining the future, including those previously identified as important for remembering locations already visited. The *Proceedings of the National Academy of Sciences USA* published the study online January 1. —David Biello

LONGEVITY 

A Gene for Aging Smartly

If you live to 100, as roughly one in every 10,000 people do, you will want both your mind and body intact. A certain gene seems to help accomplish just that. Nir Barzilai of the Albert Einstein College of Medicine and his colleagues examined 158 elderly people of Ashkenazi



LONG LIFE and mental acuity may have a genetic connection.

Jewish descent. Centenarians who passed a 30-question test were two to three times as likely to have a common variant of the so-called CETP gene as those who did not. Those between 75 and 85 who passed the test were five times as likely. The CETP gene variant leads to larger than normal cholesterol particles in the blood, their size perhaps making them less likely to lodge in the lining of blood vessels, a process that boosts the risk of heart attack and stroke. The findings appear in the December 26 *Neurology*.

—Charles Q. Choi



News Scan briefs with this icon have extended coverage on www.sciam.com/ontheweb

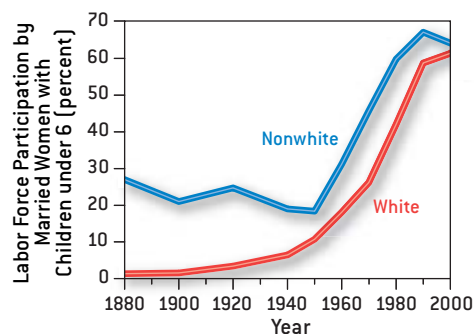
Breaking the Mold

A REAL FAMILY VALUE: MOTHERS WHO WORK OUTSIDE THE HOME BY RODGER DOYLE

In the 19th-century U.S., finding a mother of young children who worked was rare. The big exception was black mothers, many of whom were employed as domestics. That situation has changed, most likely because the economy since the early 20th century began creating an increasing number of white-collar jobs that could help support a more affluent way of life. The movement of mothers out of the home and into offices and shops happened all over the Western world, with some countries such as Austria, Belgium and the Netherlands substantially exceeding the U.S. in the proportion of mothers with very young children in the workforce.

Folk wisdom has long held that working mothers do severe harm to their children, and indeed one academic, University of Alberta chemist Gordon Freeman, claimed in a 1990 article in the *Canadian Journal of Physics* that they inflict “serious psychological damage” on their children, leading to teenage sex, drug use and other problems.

Extensive research supports neither the folk wisdom nor Freeman’s claims. In a 2006 overview, two well-known researchers in the field, Adele E. Gottfried of California State University at Northridge and Allen W. Gottfried of C.S.U.-Fullerton, documented that no link exists between the employment of mothers—even mothers of very young children—and criminal activity perpetrated by their offspring. Whether the mother’s employment was part-time, full-time, sporadic or consistent did not seem to matter. According to one study, children whose mothers had recent or long-term employment were in better health than those whose mothers had been unemployed for more than two years. Another positive result of the employment of moth-



ers is that the families’ fathers became more involved with their children, particularly on weekdays when their wives worked. Paternal involvement, once established, continued at least through their children’s late teens.

The research has revealed some negatives, too. One 2002 study found evidence among white, non-Hispanic families of a fairly small cognitive deficit in first-grade schoolchildren attributable to maternal employment, a deficit that persisted only to age seven or eight. Another study, published in 1996, suggested that as the number of children increases, full-time employment of their mothers may lead to lower happiness and increased conflict in marriages. (The opposite, however, appears to be true for mother-stepfather families.) Another downside is that parents in dual-earner families feel that they do not spend enough time with their children, with their spouses or by themselves. On net, however, the preponderance of evidence shows that the employment of mothers is a positive influence on families.

Most studies of working motherhood have focused on middle-class people and not on the poor, who to a large extent are non-white and on welfare. For these women, juggling work and child care, often without a husband, can be extremely stressful. Still, they often find that work is rewarding, as it increases self-esteem and self-reliance, mitigates depression, provides opportunities to meet new people and allows them to buy things for their children.

Rodger Doyle can be reached at rodgerpdoyle@verizon.net

FURTHER READING

A Long-Term Investigation of the Role of Maternal and Dual-Earner Employment in Children’s Development.

Adele E. Gottfried and Allen W. Gottfried in *American Behavioral Scientist*, Vol. 49, No. 10, pages 1310–1327; June 2006.

Working Mothers and the Welfare State: Religion and the Politics of Work-Family Policies in Western Europe and the United States.

Kimberly J. Morgan. Stanford University Press, 2006.

Percent of Women with Children Who Were Employed in 2005					
AGE OF CHILDREN	TOTAL	WHITE	BLACK	HISPANIC	ASIAN
Under 3	55	55	56	41	51
3–5	63	63	66	53	60
6–13	72	72	72	64	69
14–17	78	79	75	66	72



(Can't Get No) Satisfaction

The new science of happiness needs some historical perspective By MICHAEL SHERMER

Imagine you have a choice between earning \$50,000 a year while other people make \$25,000 or earning \$100,000 a year while other people get \$250,000. Prices of goods and services are the same. Which would you prefer? Surprisingly, studies show that the majority of people select the first option. As H. L. Mencken is said to have quipped, “A wealthy man is one who earns \$100 a year more than his wife’s sister’s husband.”

This seemingly illogical preference is just one of the puzzles that science is trying to solve about why happiness can be so elusive in today’s world. Several recent books by researchers address the topic, but my skeptic’s eye found a historian’s long-view analysis to be ultimately the most enlightening.

Consider a paradox outlined by London School of Economics economist Richard Layard in *Happiness* (Penguin, 2005), in which he shows that we are no happier even though average incomes have more than doubled since 1950 and “we have more food, more clothes, more cars, bigger houses, more central heating, more foreign holidays, a shorter working week, nicer work and, above all, better health.” Once average annual income is above \$20,000 a head, higher pay brings no greater happiness. Why? One, our genes account for roughly half of our predisposition to be happy or unhappy, and two, our wants are relative to what other people have, not to some absolute measure.

Happiness is better equated with satisfaction than pleasure, says Emory University psychiatrist Gregory Berns in *Satisfaction* (Henry Holt, 2005), because the pursuit of pleasure lands us on a never-ending hedonic treadmill that paradoxically leads to misery. “Satisfaction is an emotion that captures the uniquely human need to impart meaning to one’s activities,” Berns concludes. “While you might find pleasure by happenstance—winning the lottery, possessing the genes for a sunny temperament, or having the luck not to live in poverty—satisfaction can arise only by the conscious decision to do something. And this makes all the difference in the world, because it is only your own actions for which you may take responsibility and credit.”

Harvard psychologist Daniel Gilbert goes deeper into our

psyches in *Stumbling on Happiness* (Knopf, 2006), in which he claims, “The human being is the only animal that thinks about the future.” Much of our happiness depends on projecting what *will* make us happy (instead of what actually does), and Gilbert shows that we are not very good at this forethought. Most of us imagine that variety is the spice of life, for example. But in an experiment in which subjects anticipated that they would prefer an assortment of snacks, when it actually came to eating the snacks week after week, subjects in the no-variety group said that they were more satisfied than the subjects in the variety group. “Wonderful things are especially wonderful the first time they happen,” Gilbert explains, “but their wonderfulness wanes with repetition.”

Incomes have more than doubled since 1950, yet we are no happier.

This habituation to even a multiplicity of wonderfulness is what economists call “declining marginal utility” and married couples call life. But if you think that an array of sexual partners adds to the spice of life, you are mistaken: according to an exhaustive study published in *The Social Organization of Sexuality* (University of Chicago Press, 1994), married people have more sex than singles—and more orgasms. Historian Jennifer Michael Hecht emphasized this point in *The Happiness Myth* (Harper, 2007). Her deep and thoughtful historical perspective demonstrates just how time- and culture-dependent is all this happiness research. As she writes, “The basic modern assumptions about how to be happy are nonsense.” Take sex. “A century ago, an average man who had not had sex in three years might have felt proud of his health and forbearance, and a woman might have praised herself for the health and happiness benefits of ten years of abstinence.”

Most happiness research is based on self-reported data, and Hecht’s point is that people a century ago would most likely have answered questions on a happiness survey very differently than they do today.

To understand happiness, we need both history and science. ■

Michael Shermer is publisher of Skeptic (www.skeptic.com). His new book is Why Darwin Matters.



Threats of War, Chances for Peace

Preventing the spread of war will depend on strategies that recognize the shared interests of adversaries By JEFFREY D. SACHS

Although climate change, deforestation and depletion of groundwater are all serious threats to sustainable development, the biggest threat to future well-being remains the specter of war. The world was at the brink of a nuclear conflict during the 1962 Cuban missile crisis and could quickly find itself there again in South Asia, the Middle East, the Korean peninsula or some other hot spot. The Cuban crisis was transformed, through President John F. Kennedy's political vision and dexterity, into the beginning of arms control in the 1963 Nuclear Test Ban Treaty. That historic breakthrough offers timely lessons for today.

The events of late 1962 through mid-1963 are well known. Soviet leader Nikita Khrushchev gambled by trying to position offensive surface-to-ground nuclear missiles in Cuba, cheating on promises to limit the Soviets' Cuban arsenal to defensive weapons. The U.S. caught the Soviets in midcourse of installing the missiles and imposed a naval quarantine. The Soviets agreed to withdraw the offensive missiles in return for a U.S. commitment not to invade Cuba and a secret pledge to remove nuclear missiles from Turkey at a later date. After coming within hours of war, the U.S. and the Soviet Union went on a few months later to sign a test ban agreement.

How does one go from the brink of war to a breakthrough peace treaty in under a year? Kennedy's methodological starting point was to avoid vilifying the Soviet Union or declaring the adversary to be evil. At every step, Kennedy assumed that Soviet counterparts were rational, though not necessarily beyond mistakes in their chosen actions. He assumed that the Soviet Union would seek tactical advantages where it could get them but would pull back from self-annihilation.

Today's game theorists would describe Kennedy's strategy as "generous tit-for-tat" (GTFT). A player adopts a position of cooperation as long as the other side does, too. If the second player begins to cheat, the first player stops cooperating as well, to show the cheater that there are adverse consequences to the collapse of this arrangement. The door remains forgivingly open to future cooperation, however, if the cheater re-

verts to form. And generously, the first player might initiate renewed cooperation, with a view to enticing the former cheater to reciprocate. GTFT is so successful and robust that many evolutionary biologists suppose that the basic strategy is somewhat hardwired in human attitudes.

As Kennedy later explained his thinking: "In short, both the United States and its allies, and the Soviet Union and its allies, have a mutually deep interest in a just and genuine peace and in halting the arms race. Agreements to this end are in the interests of the Soviet Union as well as ours.... And even the most hostile nations can be relied upon to accept and keep those treaty obligations, and only those treaty

obligations, which are in their own interest." He stressed the need to avoid humiliating one's adversary. "And above all, while defending our own vital interests, nuclear powers must avert those confrontations which bring an adversary to a choice of either a humiliating retreat or a nuclear war. To adopt that kind of course in the nuclear age would be evidence only of the bankruptcy of our policy—or of a collective death-wish for the world."

Kennedy's sentiments were radical at the time, but he believed that the potential for cooperation was grounded in our common humanity. "For in the final analysis, our most basic common link is that we all inhabit this small planet. We all breathe the same air. We all cherish our children's futures. And we are all mortal." As we face today's challenges and threats, we will do well to grasp the insight that our counterparts and adversaries, like us, are searching for survival and for a future for their children. As occurred 45 years ago, that critical insight might prove to be the key to keeping us alive and secure. SA

Jeffrey D. Sachs is director of the Earth Institute at Columbia University.



An expanded version of this essay is available at www.sciam.com/ontheweb

Page Intentionally Blank

SCIENTIFIC AMERICAN Digital

Graph Theory and Teatime

Deep in the heart of Microsoft, Jennifer Chayes and Christian Borgs lead a who's who of mathematics and computer science. The goal? To explore anything they please By GARY STIX

Every weekday afternoon some 20 mathematicians and theoretical computer scientists gather in the Seattle suburbs to share tea. The conversation runs from the latest on number theory to the fairest way to decide a closely contested election. The gathering spot is not the faculty lounge of an elite university but rather a meeting area in Building 113, the nondescript glass and steel structure that houses the Theory Group of Microsoft Research.

A decade ago two mathematical physicists—Jennifer Chayes and Christian Borgs—gave up permanent academic positions for the allure of being able to go out and hire the best minds in discrete mathematics, statistical physics and theoretical computer science. By most measures, the pair have succeeded in re-creating the rarefied world of a top university department, right down to the tea ritual. In essence, the group resembles a smaller version of the Mathematical Sciences Research Center in its heyday at the old Bell Labs, home to Claude E. Shannon, Richard Hamming, Narendra Karmarkar and other quantitative luminaries, before corporate upheavals ultimately forced a scaling back. “It would be very hard, if not just impossible, for a university to assemble such a group within a 10-year time frame,” remarks Bart Selman, professor of computer science at Cornell University and also a former Bell Labs researcher. “Clearly, Microsoft resources play a role here.”

Microsoft Research was established in 1991 to emphasize basic research in computer science at a time when other industrial labs were revamping to focus on more applied endeavors. The Theory Group, whose members routinely publish papers with titles such as “The D_4 Root System Is Not Universally Optimal,” probably has the least relevance to product development of any Microsoft department.

The disconnect is intentional. In 1996 Nathan Myhrvold, a former classmate of Chayes at Princeton University who was then Microsoft's chief technology

officer, suggested that Chayes and Borgs come to work at Microsoft. “Are you crazy?” Chayes asked Myhrvold. “You can't make money from what we do.”

Myhrvold promised that they would not be enlisted to write code for a new version of Microsoft Office. “He wanted us to do the most way-out stuff,” Chayes remembers. “He said, ‘Look, I'm not hiring two engineers,’” Borgs chimes in a moment later. The Microsoft offer solved a fundamental problem related to



CHRISTIAN BORGS AND JENNIFER CHAYES: THEORIZING AT MICROSOFT

- Co-managers of the Theory Group of Microsoft Research, which explores mathematics, computational theory and physics that may provide the conceptual framework for critical advances in software and networking.
- Chayes on the nature of the group's work: “We do research for the day after tomorrow.”

time and space. The two had married four years earlier. Chayes was a tenured professor of mathematics at the University of California, Los Angeles. Borgs had a chaired professorship in statistical physics at the University of Leipzig in Germany.

“We went from living on the other side of the world to doing everything together,” Chayes says. Every paper they write bears joint authorship; every intern candidate interviewed receives questions from both. The compatible trajectories stretch back to their youth. Neither followed the rectilinear path set out for them by their parents. Borgs, 49, came from a traditional family in Düsseldorf, Germany, and was expected to take over their 120-year-old chemical business. Chayes, 50, a rebellious “child of the sixties” and the daughter of a Jewish father and a Muslim mother who had immigrated to the U.S. from Iran, was supposed to become a physician. (Her brother, James Tour, also paid no heed to his parents’ plans, going on to become a chemist at Rice University and a major figure in nanotechnology.)

The collaborating spouses held Myhrvold to his word and went on to hire some of the best and brightest. There are nine full-time researchers, eight postdoctoral students, five academics on sabbatical from other institutions—and 150 to 200 visitors annually who arrive for stays that range from a day to a month. “Their list of visitors reads like a veritable who’s who of theoretical computer science,” observes Lenore Blum, a computer scientist at Carnegie Mellon University.

The quasi-academic environment has enabled notable researchers to continue their work undistracted—or, if they so choose, to branch out in new directions. Oded Schramm devised a mathematical proof that shows how certain random two-dimensional objects, when distorted, retain the same statistical properties—a characteristic called conformal invariance. One of Schramm’s colleagues, Wendelin Werner, received the Fields Medal for this work. (Schramm was a few weeks too old to qualify for the medal, bestowed only on those younger than 40.) “Oded basically invented a new branch of mathematics, which I predict will be studied 100 years from now,” Chayes says.

Another notable was Michael Freedman, who won the Fields Medal while at the University of California, San Diego, for his work on the Poincaré conjecture. He moved to the Theory Group in 1997 and began to explore in earnest how topological quantum field theory could be applied to create a quantum computer with very low error rates, taking advantage of the fact that topological properties resist perturbations (errors). Freedman has since formed his own group within Microsoft that focuses on quantum computation.

A younger researcher at 32, Henry Cohn has, along with postdoc Abhinav Kumar, published seminal work on how

densely spheres can be packed together within eight and 24 dimensions. Mathematicians are fascinated by what Cohn calls these “miracle dimensions” because of packing efficiencies generally not found elsewhere. Such calculations may ultimately enable better error-correction codes for transmitting digital bits on noisy channels.

Chayes and Borgs have also been able to build on their original university work on the mathematics of phase transitions: sudden discontinuities in a physical state, such as when water turns to ice. Similarly, whenever increasing loads are placed on two parallel microprocessors, a phase transition occurs in which balancing work among the processing elements becomes much more difficult. In their papers, Chayes and Borgs have shown that once the transition has occurred, it may be virtually impossible to improve on a near-optimal solution to partitioning a workload—the programmer of a parallel processor cannot just shift some of the load from one processor to another to achieve the best balance. “You may as well start over,” Chayes says. “That’s a disaster for computation.”

The couple's overlapping orbitals have been good for their careers and for Microsoft.

Besides computer science, this type of optimization problem has implications for modeling the precise networks of chemical bonds, genes and synapses that are found

in investigations of protein folding, gene activation in microarray chips, and the changes in neural connections that occur during learning. Chayes and Borgs have undertaken a collaborative initiative with Riccardo Zecchina of the International Center for Theoretical Physics in Trieste, Italy, and other European researchers to explore a technique, called survey propagation, that might find better solutions for the hard optimization problems found after a phase transition occurs.

Chayes and Borgs’s prior university labors on graph theory and phase transitions have been of some use to the enterprise. Since they joined Microsoft, the World Wide Web has come into its own. “All of a sudden the stuff we were doing has become relevant,” Chayes notes. Graph theory serves as a powerful tool for modeling the complexity of the Web. Chayes and Borgs have shown how the patterns formed by links fanning out from spam sites differ in appearance from connections to normal sites, a tool that is being incorporated into search engines by Microsoft product developers.

For the pair, the fusion of work and personal life has proved essential for building both the Theory Group and continuing their own research. Certainly Borgs understands Chayes when she gets angry at her husband and shouts, “You’re perturbing around the wrong ground state.” The couple’s overlapping orbitals have been good for their own careers, for Microsoft and for the larger community of mathematicians and computer scientists. ■

DO NOT ATTEMPT THIS AT HOME: Looking down the barrel of a jet from a supermassive black hole is not something you would be advised to do. The jet packs enough punch to create bubbles of hot gas with an energy equivalent to millions or billions of supernova explosions.

Black Hole **BLOWBACK**

*A single black hole, smaller than the solar system,
can control the destiny of an entire
cluster of galaxies*

By Wallace Tucker, Harvey Tananbaum and Andrew Fabian



If you drew a large-scale map of the universe, it would look rather like a map of the U.S. Interstate Highway System. Galaxies line up in filaments that criss-cross intergalactic space like freeways. In between the roads are regions of relatively low density: the cosmic countryside. And at the crossroads, where multiple filaments converge, are clusters of galaxies: the cosmic megacities.

The size of these clusters is daunting. It takes light a little more than a second to reach Earth from the moon and eight minutes to reach Earth from the sun. Light from the center of our Milky Way galaxy must make a journey of 25,000 years to reach us. Even that is fairly quick compared with the time required for light to cross a galaxy cluster—about 10 million years. In fact, clusters are the largest gravitationally bound bodies in the universe. The roadlike filaments may be larger in sheer size, but they are not coherent bodies held together by gravity.

The gravitational binding means that the galaxies and other material within a mature cluster have settled into an overall dynamic equilibrium. Galaxies buzz around within it and are kept from flying apart by dark matter, a mysterious form of matter that has eluded detection except through its gravitational effects. The interactions of these components produce a rich array of phenomena that astronomers are only beginning to grasp.

Like metropolises on Earth, clusters are more than the sum of their inhabitants. Processes occurring at the scale of a cluster can dictate events on much smaller scales, such as the growth of galaxies and the fueling of the supermassive black holes at the hubs of those galaxies. In turn, the black holes blow out huge amounts of high-speed material that can drive the evolution of the entire cluster. At first glance, these inter-

the more populated region: it contains more mass than all the stars in all the galaxies in the cluster.

The gas, which is heated primarily by the slow gravitational collapse of the cluster, gives off x-rays. Optical telescopes cannot see the gas, and x-rays cannot penetrate Earth's atmosphere, so the discovery and study of this gas has depended on orbiting observatories. Two decades ago astronomers peering with NASA's Einstein X-ray Observatory and other instruments noticed that the x-rays carry away so much energy that the gas should steadily cool off and settle into the center of the cluster—thus the term “cooling flow.” One of us (Fabian) led the way in investigating these flows using Einstein and later Germany's ROSAT x-ray satellite. He and his colleagues calculated the flows would have quite dramatic effects. If they

For a black hole to affect an entire galaxy cluster would be like a blueberry affecting the entire EARTH.

connections between large and small are enigmatic. The diameters of the black holes in question are smaller than the solar system. For them to affect an entire galaxy cluster would be like a blueberry affecting the entire Earth.

The Case of the Disappearing Gas

THESE INTERACTIONS explain some long-standing paradoxes in the urban life of the cosmos. One is the so-called cooling flow problem, which has to do with multimillion-degree gas that fills the space between the galaxies in a cluster. If the galaxies within a cluster are the urban cores of the megacity, this gas is the suburban sprawl. Like the suburbs that surround most American downtowns, the gas is actually

persisted for a billion years, the gas deposited in the central regions of the cluster could form trillions of new stars.

The only trouble was, no one could find them. Observers looked in vain for large amounts of cool gas and hordes of newly formed stars. If a black hole had swallowed them all, it would weigh as much as a trillion stars, and not even the biggest black hole is that massive. Another one of us (Tucker) maintained that large-scale, long-term cooling flows do not exist. A possible explanation was that long-lasting outbursts of energy from the central galaxy of the cluster heated the gas enough to offset the radiative cooling. Radio astronomers had for years been accumulating evidence for such activity. But it was questionable whether the outbursts provided enough energy distributed over a large enough volume to halt the cooling flows, so the paradox remained: the hot cluster gas must cool, but the end product of the cooling mysteriously escaped detection.

Resolving this paradox was a major goal of two powerful x-ray telescopes launched in 1999: NASA's Chandra X-ray Observatory and the European Space Agency's XMM-Newton. Because the gas in clusters radiates away its energy fairly slowly, it preserves a record of activity in the clusters over the past few billion years. For instance, it retains the elements and energy injected into it by supernova explosions in the cluster galaxies. Like archaeologists unearthing the past, astronomers have used the new telescopes to excavate the relics present in galaxy clusters and piece together their history.

Bubble, Bubble

THE BRIGHTEST CLUSTER observed by x-ray instruments is the Perseus cluster because of its high intrinsic luminosity and relative cosmic proximity to Earth (about 300 million light-years). In the 1990s ROSAT discovered two vast holes in the x-ray gas in the central 50,000 light-years of the cluster. They look like an hourglass centered on the giant galaxy

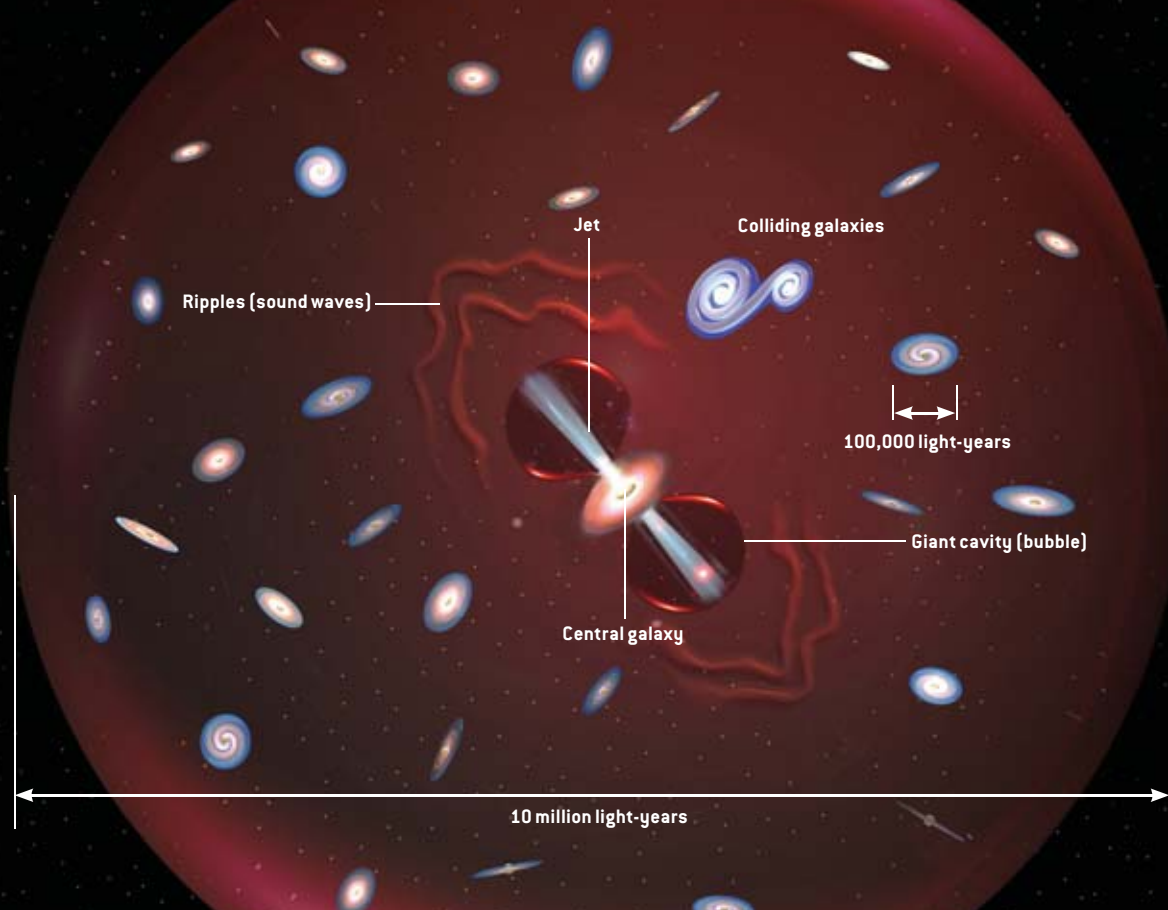
Overview/Monster Bubbles

- Using a combination of radio and x-ray telescopes, astronomers have discovered vast bubbles of high-energy particles hundreds of thousands of light-years across. The energy required to create these structures beggars belief—it is as if 100 million stars went supernova at once.
- The only thing capable of creating such a monster is a giant black hole. Not all matter that comes close to a black hole is doomed. As magnetized hot gas swirls in an accretion disk toward a hole, strong electromagnetic forces build up and spit out some of the gas in a narrow jet.
- Not only can the jet create bubbles, it adds heat and magnetism to the intergalactic gas in a galaxy cluster, explaining long-standing puzzles in astronomy. The process appears to be a part of a multimillion-year cycle that regulates the growth of supergiant galaxies in the centers of clusters.

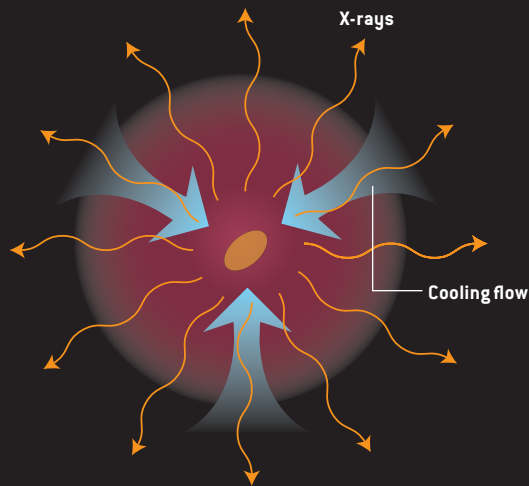
ANATOMY OF A CLUSTER

The largest things in the universe worthy of being called "things" are galaxy clusters. They consist of 1,000 or so galaxies, buzzing through a ball of hot gas (red) like bees in a

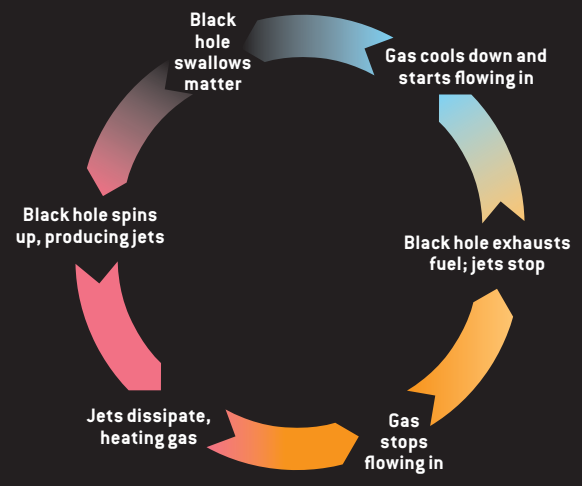
hive and prevented from dispersing by gravity. At the core of the cluster is an especially large galaxy—the site of the most violent processes in the modern universe.



As the x-rays carry off energy, the cluster gas should cool and flow inward. Over billions of years, it should form trillions of new stars. Yet few such stars are seen.



Cycle of heating and cooling explains why those stars are not seen. Black hole jets return energy to the gas and choke off the inward flow.



ALFRED T. KAMAJIAN (top); JANA BRENNING (bottom)

NGC 1275. With Chandra, Fabian and his colleagues went back for a closer look. Their Chandra data showed the cavities in exquisite detail, revealing that they were aligned with previously observed radio jets emanating from the center of the giant galaxy [see illustration on page 48]. The x-ray cavities are not empty but are filled with magnetic fields and energetic particles such as protons and electrons. These energetic, low-density bubbles are rising buoyantly and pushing aside the x-ray-emitting hot gas.

Other clusters have bubbles, too. Chandra observations revealed x-ray cavities with associated radio emission in the Hydra A, Hercules A and Abell 2597 clusters. The observatories also uncovered bubbles that were faint in both radio waves and x-rays, indicating that the energetic particles within have dissipated most of their energy. These “ghost cavities” have detached from the central galaxy and may be relics of past bubbles.

The most spectacular activity seen by Chandra to date was discovered in the cluster MS 0735.6+7421 (MS 0735 for short) by Brian R. McNamara, now at the University of Waterloo in Ontario, and his colleagues. Although the image of this cluster is not as detailed as the one for Perseus, it tells an amazing story. Each of the cluster’s two x-ray cavities is 600,000 light-years across—more than six times larger than the disk of our

bubbles—a strong indication that bubbles keep the gas from cooling down. But there is a missing link in this argument: How does the energy get transferred from bubble to gas?

Sub-Subwoofer

AN OBVIOUS ANSWER is that the bubbles generate strong shock waves in a manner similar to atmospheric explosions on Earth. As the energetic material from an explosion pushes into the atmosphere at supersonic speeds, it sweeps up surrounding air into a thin shell. Collisions between the swept-up particles convert the kinetic energy of the explosion into heat. Strong shock waves are also observed in a wide range of cosmic phenomena, such as the remnants of supernova explosions.

H. L. Mencken is said to have observed: “For every complex problem there is an answer that is clear, simple, and wrong.” Unfortunately, it appears that the heating of cluster gas by strong shock waves is a case in point. Telescopes do not see the hot, thin shells that would be produced by such heating. Also, it is likely that heating from strong shock waves would be too concentrated in the central regions of a cluster to offset widespread cooling of the cluster gas.

A more plausible energy-transfer mechanism is heating by sound waves. The intergalactic gas inside clusters may be ten-

The pitch of the sound waves translates to a note of B-flat, 57 OCTAVES below middle C.

Milky Way galaxy. The size of the cavities and the observed density and temperature of the gas around them indicate they are 100 million years old and contain as much kinetic energy as 10 billion supernovae. Even astronomers, who regularly deal in billions and trillions, are impressed by the immensity of the bubbles and the titanic amount of energy they represent.

This energy is enough to resolve the paradox of cooling flows. In fact, John R. Peterson, now at Purdue University, and others used spectra measured by the XMM telescope to show that cooling flows do not occur in clusters that have such

uous by human standards (the equivalent of just a few thousand hydrogen atoms per cubic meter), but sound waves can still propagate through it. They evolve into weak, barely supersonic shock waves that heat gas fairly gently.

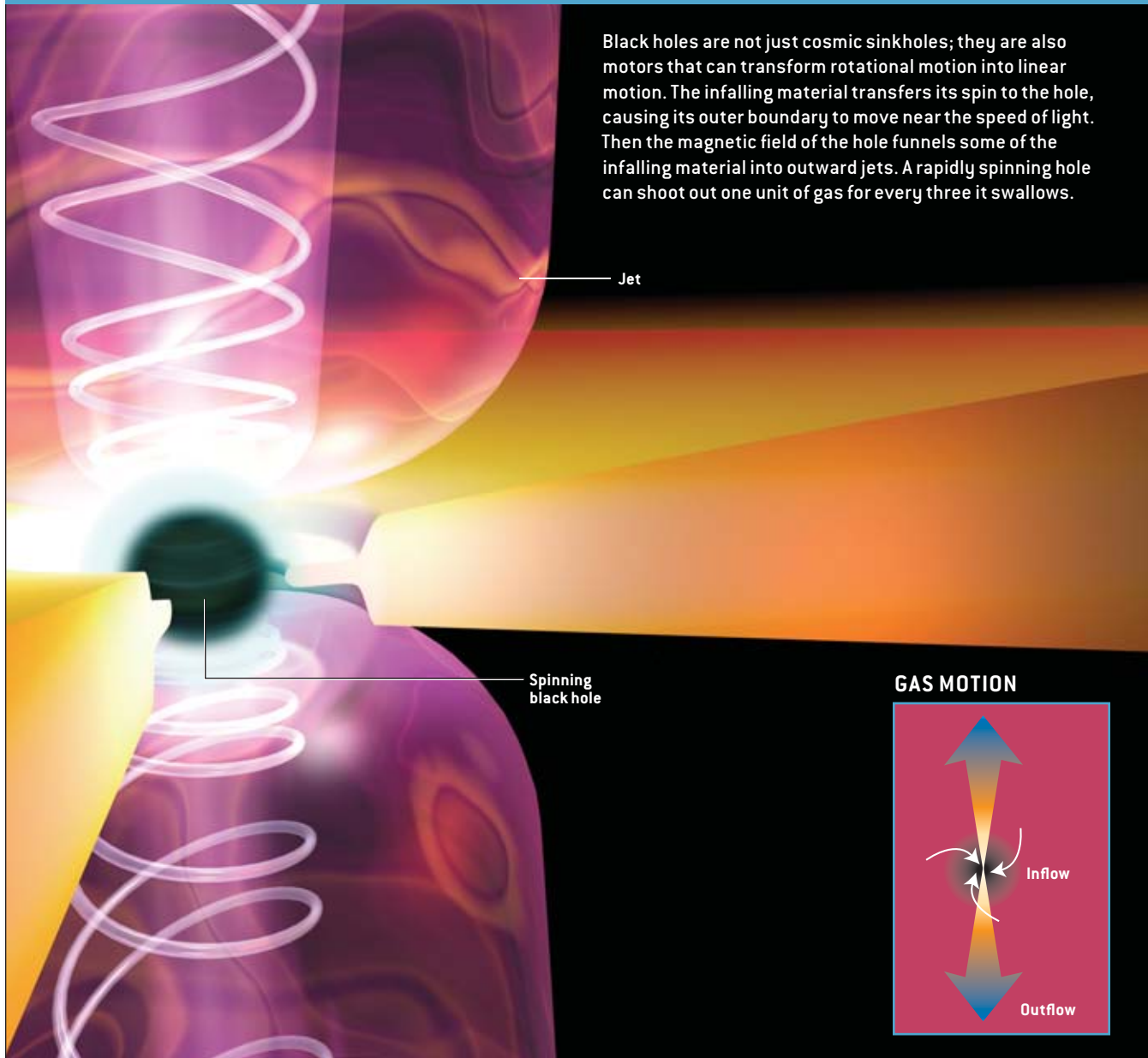
By special processing of the Perseus cluster images, Fabian’s team uncovered the smoking gun for this idea: a series of nearly concentric ripples. The density and pressure of the gas, though not its temperature, change abruptly at the innermost ripple, indicating that this feature is a weak shock wave. At the outer ripples, the density and pressure vary gradually, indicating that these ripples are sound waves. The spacing of the ripples (35,000 light-years) and the calculated speed of sound in the gas (1,170 kilometers a second) imply that 10 million years passed between the events producing the ripples. The pitch of the sound waves translates to a note of B-flat, 57 octaves below middle C. What they may lack in musicality, they make up for in sheer power.

Similar features show up in the Virgo cluster, the nearest cluster to us, about 50 million light-years away. William Forman of the Harvard-Smithsonian Center for Astrophysics and his colleagues observed the central, dominant galaxy in this cluster, M87, using Chandra. They found a web of filamentary structures, each about 1,000 light-years across and 50,000 light-years long. The filaments, like the ripples in the Perseus cluster, may be the product of sound waves generated by a series of buoyant bubbles that in turn arise from outbursts—in this case, spaced about six million years apart.

THE AUTHORS

WALLACE TUCKER, HARVEY TANANBAUM and ANDREW FABIAN have been leaders of x-ray astronomy almost as long as there has been such a thing. Tucker, science spokesperson for the Chandra X-ray Center, studies dark matter, clusters of galaxies, and supernova remnants. On top of research papers, magazine articles (including three past *Scientific American* pieces) and half a dozen books, he has written three award-winning plays on Native Americans. Tananbaum is director of the Chandra X-ray Center, a member of the National Academy of Sciences, and recipient of the 2004 Rossi Prize in astronomy. His research delves into x-ray binaries, quasars and active galaxies, and optically quiet x-ray bright galaxies. Fabian is a professor at the University of Cambridge, a Fellow of the Royal Society, and recipient of the 2001 Rossi Prize. He has co-authored more than 500 research papers on clusters of galaxies and accreting black holes of all sizes.

THE MOST POWERFUL ENGINE KNOWN TO SCIENCE



Black holes are not just cosmic sinkholes; they are also motors that can transform rotational motion into linear motion. The infalling material transfers its spin to the hole, causing its outer boundary to move near the speed of light. Then the magnetic field of the hole funnels some of the infalling material into outward jets. A rapidly spinning hole can shoot out one unit of gas for every three it swallows.

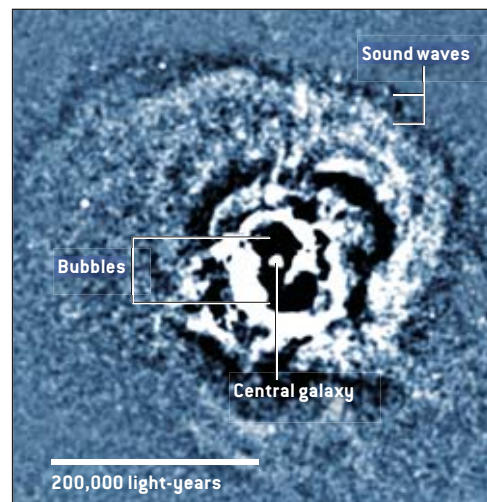
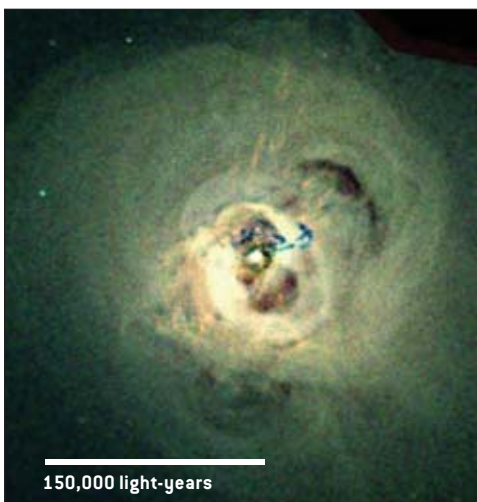
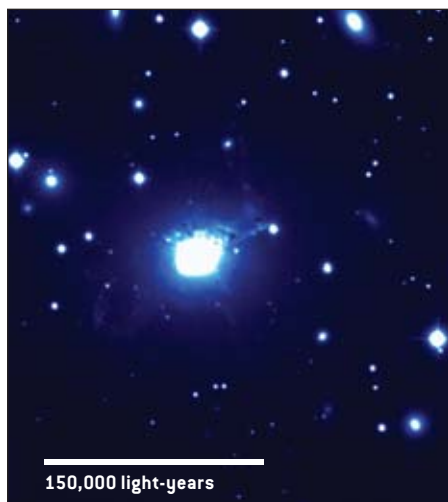
Thus, the sound waves are about an octave higher in pitch than those in Perseus. Forman's team also detected a ring of hotter emission with a radius of about 40,000 light-years, which is probably a weak shock front, as well as a large x-ray cavity about 70,000 light-years from the galaxy's center.

The question then becomes one of how energy in the sound waves heats the gas. The observation from Perseus that the temperature of the inner ripples does not increase across the shock front may be the key. Heat conduction could be rapidly carrying away the energy of gas particles heated by the shock waves, or high-energy electrons escaping from the bubbles or from behind the waves could speed away and heat the gas. Either process would keep down the temperature at the shock front itself.

Electromagnetic Tornadoes

THE BIGGEST QUESTION of all, though, is what blew the bubbles to begin with. Only one type of object known to science could generate that much energy: a supermassive black hole. Although most people think of black holes as the ultimate sinkholes, they can also whip up and shoot out matter at high speeds. Exactly how they do so has been the subject of intense study in recent years.

Simulations show that a black hole can act as a giant motor. Gas falling into the hole revs it up. Magnetic fields then convert this rotational energy to linear motion, catapulting out a portion of the gas. This process was first proposed in the late 1970s by Roger D. Blandford, now at Stanford University, and Roman Znajek, who was at the University of Cam-



PERSEUS CLUSTER looks sedate when viewed in visible light (*left*) but comes alive when viewed in x-rays (*center*). The space between the galaxies is filled with hot gas threaded by bright loops, filaments and streaks. Two bubbles, which appear vacant in these images but actually

contain high-energy particles, straddle the central galaxy, NGC 1275. Increasing the image contrast (*right*) brings out ripples, thought to be sound waves that transport energy to the intergalactic gas. For animations, visit chandra.harvard.edu/photo/2003/perseus/animations.html

bridge at the time but has since left academia. The rotating black hole twists up the fabric of space around it, forcing the magnetic field in the infalling gas into a funnellike shape—an electromagnetic tornado that flings fields and charged particles outward in two opposing jets. Slowly spinning black holes produce feeble jets; most of the inflowing gas continues down into the hole and is lost forever. Rapidly spinning black holes, though, expel roughly a quarter of the inflowing gas.

Supermassive black holes in the centers of galaxies are expected to spin up over time as they accrete gas. By the time the black hole has swallowed enough gas to double its mass, its horizon, or outer boundary, should be whipping around at nearly the speed of light. According to Einstein's theory of relativity, the hole can never reach the speed of light no matter how much gas it devours; each additional clump of gas produces diminishing returns. A variety of observational methods for estimating the spin of black holes confirm that many are whirling rapidly enough to produce powerful jets. A similar phenomenon happens on a smaller scale. Stellar-mass black holes, with masses of a dozen suns (rather than a billion), can pump out powerful jets of particles at near the speed of light, heating and pushing aside ambient gas.

Calculations show that black hole jets have two major components: a matter-dominated outflow that moves at about a third of the speed of light, forming the outer sheath of the funnel, and an inner region along the axis of the funnel that contains a rarefied gas of extremely high energy particles. It is the inner region that carries much of the energy and creates the dramatic structures observed by radio and x-ray astronomers.

One of the most astounding features of jets is the pencil-thin shape that they can maintain despite traveling hundreds of thousands of light-years, far beyond the confines of their parent galaxies. Moreover, they manage to do this while radiating hardly any of their energy. The pressure of the gas near a black hole can get a jet started as a narrow beam, and it may be that inertia keeps the jet narrow, much the way a blast of water emerges from a hose or steam erupts from a high-pressure tea kettle. The tightly coiled magnetic field that is spun out with the jet may also play a role.

Regardless of the confining mechanism, the pressure of the gas through which jets move gradually takes its toll. The jets slow down and billow out, creating enormous magnetized clouds of high-energy particles. These clouds continue to expand, pushing out the surrounding gas to create the dark x-ray cavities observed by Chandra.

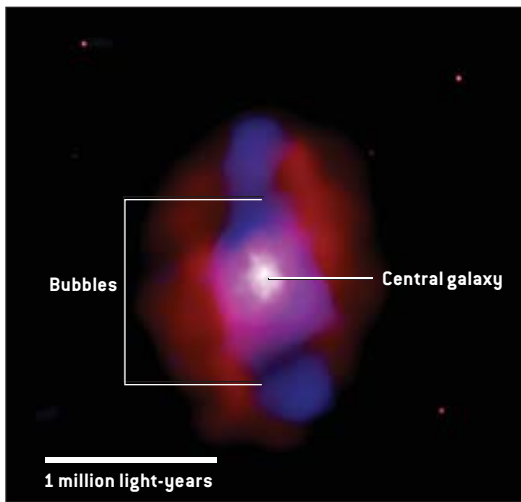
The Cosmic Hydrologic Cycle

THIS SEQUENCE OF EVENTS—gas falls into a rapidly spinning black hole to form outward-moving megajets that carve out gigantic bubbles of high-energy particles and heat vast volumes of space—is a blowback of truly cosmic proportions. The black hole is both responding to and influencing events on the scale of the entire galaxy cluster.

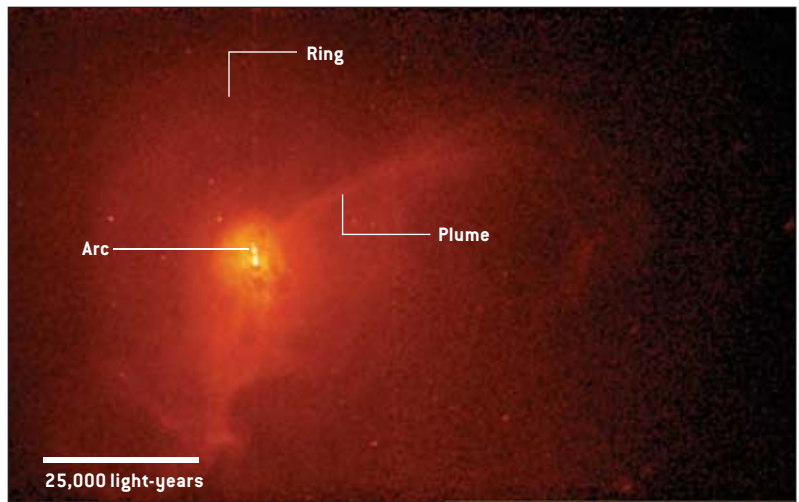
A likely scenario goes as follows. Initially the gas in the cluster is very hot, and the supermassive black hole in a centrally located large galaxy is quiet. Over about 100 million years, gas in the central region of the cluster cools and drifts toward the central galaxy in a cooling flow. Some of the gas in this cooling flow condenses into stars that become part of the central galaxy, and some sinks all the way down to feed the supermassive black hole. In so doing, it creates an accretion disk and activates high-power jets.

The jets blast through the galaxy and out into the cluster gas, where their energy converts to heat. The heat greatly diminishes the cooling flow, if not shutting it off altogether. It is a case of biting the hand that feeds: by shutting down the cooling flow, the supermassive black hole chokes off its own supply of gas and gradually goes dormant. The jets fade away, leaving the cluster gas without a heat source. Millions of years later the hot gas in the central region of the cluster finally cools sufficiently to initiate a new season of growth for the galaxy and its supermassive black hole, and thus the cycle continues.

This scenario is supported by high-resolution x-ray and radio images of the Virgo, Perseus, Hydra and other clusters, which show evidence of repetitive outbursts from the vicinity of the central galaxies' supermassive black holes. Magnetized rings, bubbles, plumes and jets ranging in size from a few thousand to a few hundred thousand light-years strongly sug-



MOST POWERFUL ERUPTION EVER SEEN has been under way for 100 million years in the cluster MS 0735. The bubbles (blue) seen in this composite radio and x-ray image are 250 times more powerful than those in the Perseus cluster.



JETS POKING OUT OF GALAXY M87 in the Virgo cluster are relative weaklings—just 0.01 percent of the power of those in MS 0735—but make up for it with their fine details, including curved plumes (which are possibly the remnants of previous outbursts), arcs (possibly shock waves) and faint rings (possibly sound waves).

gest that intermittent violent activity has been going on in these clusters for hundreds of millions of years.

One startling implication is that supermassive black holes are still growing at a rapid rate even in the present day. Astronomers had thought their growth had tapered off. In the case of cluster MS 0735, the activity indicates that the supermassive black hole has gulped down the equivalent of 300 million suns in the past 100 million years—nearly doubling in size and mass over that relatively brief interval. Yet the central black hole shows no other signs of activity, such as bright x-rays or visible light, which are usually emitted by active holes. It is only through the x-ray cavities that we can discern the properties of this extraordinary system.


Cosmic Consequences

THE SCENARIO IS ENRICHED by galaxy collisions, an ever present hazard in the central regions of galaxy clusters. A smaller galaxy passing too close to the giant central galaxy is torn asunder—its stars assimilated, some of its gas lost down the black hole drain, its own central black hole merged with the one in the giant galaxy. The enormous cavities observed in MS 0735 were probably the end result of a sequence of events initiated when a merging galaxy caused a huge influx of gas into a supermassive black hole.

The role of collisions in clusters may help scientists understand the evolution of galaxies in the early universe. In a sense, clusters are living fossils, the only places in the universe that retain the conditions that prevailed billions of years ago, when galaxies were closer together and mergers were common. A growing body of research indicates that many aspects of galaxy formation and evolution—the size and shape of galaxies, the rate of star formation—can be understood in terms of a cosmic cycle involving mergers of galaxies. Large-scale computer simulations by Philip F. Hopkins of the Harvard-Smithsonian Center for Astrophysics and his colleagues show that the mergers of gas-rich galaxies trigger bursts of star formation and inflow of gas into the central region. The inflowing gas fuels rapid growth of the supermassive black hole and intense radiation from its vicinity. Blowback ejects

much of the gas from the galaxy, star formation abruptly slows, and accretion onto the black hole declines—until another merger occurs.

Most of the black hole feedback that shaped the evolution of galaxies occurred about eight billion to 10 billion years ago. Since then, the universe has thinned out too much—except in clusters. The blowback processes in clusters are similar (though not identical) to those that occurred in the ancient universe, allowing astronomers to study the jets, bubbles and waves that shaped our galaxy and others.

It may seem strange that supermassive black holes, objects with masses that range from a few million to hundreds of millions of solar masses, can have such an impact on galaxies whose masses range from a few billion to a few hundred billion solar masses, let alone galaxy clusters with masses measured in the hundreds of trillions of solar masses. The reason is the concentrated nature of supermassive black holes and their gravitational fields. Supermassive black holes are by far the largest supply of gravitational potential energy in an entire galaxy. By tapping this energy through accretion disks and the launching of megajets, blowback vastly increases the reach of these black holes—making it one of the most important processes at work in the universe. 

MORE TO EXPLORE

Black Holes and Time Warps. Kip Thorne. W. W. Norton, 1994.

Cooling Flows in Clusters of Galaxies. A. C. Fabian in *Annual Review of Astronomy and Astrophysics*, Vol. 32, pages 277–318; 1994.

A Deep Chandra Observation of the Perseus Cluster: Shocks and Ripples. A. C. Fabian et al. in *Monthly Notices of the Royal Astronomical Society*, Vol. 344, No. 3, pages L43–L47; September 2003. Available at <http://arxiv.org/abs/astro-ph/0306036>

Energy Input from Quasars Regulates the Growth and Activity of Black Holes and Their Host Galaxies. Tiziana Di Matteo, Volker Springel and Lars Hernquist in *Nature*, Vol. 433, pages 604–607; February 10, 2005. arxiv.org/abs/astro-ph/0502199

Magnetically Driven Jets in the Kerr Metric. J. Hawley and J. Krolik in *Astrophysical Journal*, Vol. 641, No. 1, Part 1, pages 103–116; April 10, 2006. arxiv.org/abs/astro-ph/0512227

For the latest from the Chandra and XMM-Newton orbiting observatories, see <http://chandra.harvard.edu> and <http://xmm.esac.esa.int/>

MAPPING THE CANCER GENOME

Pinpointing the genes involved in cancer will help chart a new course across the complex landscape of human malignancies

By Francis S. Collins and Anna D. Barker

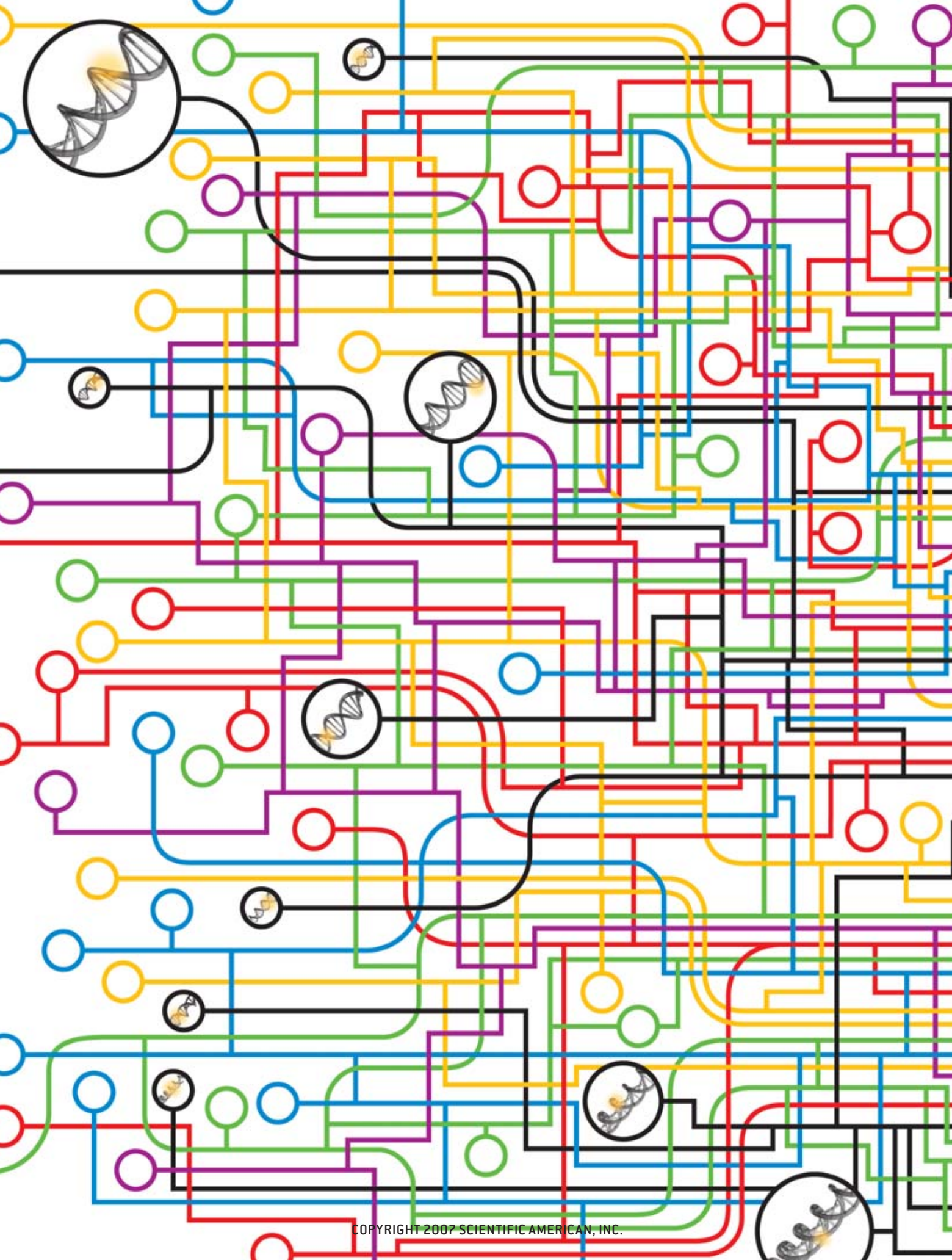
“If we wish to learn more about cancer, we must now concentrate on the cellular genome.” Nobel laureate Renato Dulbecco penned those words more than 20 years ago in one of the earliest public calls for what would become the Human Genome Project. “We are at a turning point,” Dulbecco, a pioneering cancer researcher, declared in 1986 in the journal *Science*. Discoveries in preceding years had made clear that much of the deranged behavior of cancer cells stemmed from damage to their genes and alterations in their functioning. “We have two options,” he wrote. “Either try to discover the genes important in malignancy by a piecemeal approach, or ... sequence the whole genome.”

Dulbecco and others in the scientific community grasped that sequencing the human genome, though a monumental achievement itself, would mark just the first step of the quest to fully understand the biology of cancer. With the complete sequence of nucleotide bases in normal human DNA in hand, scientists would then need to classify the wide array of human genes according to their function—which in turn could reveal their roles in cancer. Over the span of two decades Dulbecco’s vision has moved from pipe dream to reality. Less than three

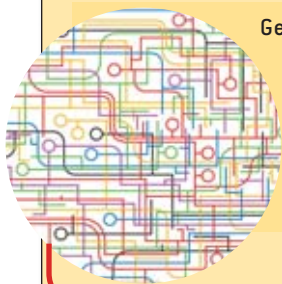
years after the Human Genome Project’s completion, the National Institutes of Health has officially launched the pilot stage of an effort to create a comprehensive catalogue of the genomic changes involved in cancer: The Cancer Genome Atlas (TCGA).

The main reason to pursue this next ambitious venture in large-scale biology with great urgency is cancer’s terrible toll on humankind. Every day more than 1,500 Americans die from cancer—about one person every minute. As the U.S. population ages, this rate is expected to rise significantly in the years ahead unless investigators find ways to accelerate the identification of new vulnerabilities within cancerous cells and develop novel strategies for attacking those targets.

Still, however noble the intent, it takes more than a desire to ease human suffering to justify a research enterprise of this magnitude. When applied to the 50 most common types of cancer, this effort could ultimately prove to be the equivalent of more than 10,000 Human Genome Projects in terms of the sheer volume of DNA to be sequenced. The dream must therefore be matched with an ambitious but realistic assessment of the emerging scientific opportunities for waging a smarter war against cancer.



MANY PATHWAYS TO MALIGNANCY

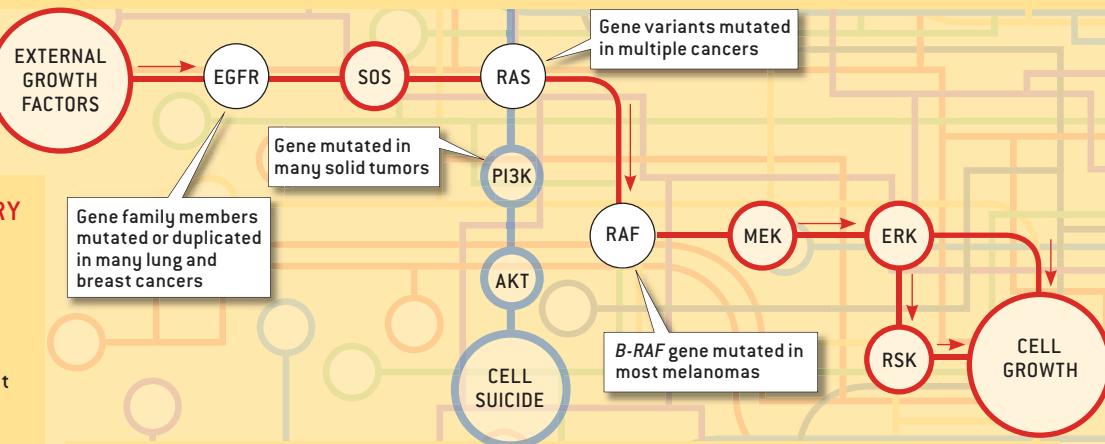


Gene malfunctions underlie the ability of cancer cells to escape normal constraints on a cell's behavior. Because genes give rise to proteins that serve as cellular building blocks, signals and regulators of other genes, a mutation that disables one gene, or causes it to be overactive, can have multiple deranging effects on the cell (*below*).

Nevertheless, cells usually need to accrete several cancer-promoting, or oncogenic, mutations in separate genes to acquire the hallmark properties of malignancy (*box at right*). Identifying all the genes whose alteration can produce those traits should one day reveal which mutations are the true drivers of specific types of cancer—even of a specific patient's malignancy—and therefore the most effective ways to intervene in the disease.

▲ COMPLEX CIRCUITRY

The extraordinarily complex molecular interactions in a human cell can be viewed as a network of parallel and intersecting pathways. A simplified depiction (*right*) of just one such pathway that promotes cell proliferation begins with a family of epidermal growth factor receptors (EGFR) in the cell membrane. Their stimulation by growth factors outside the cell transmits signals to additional proteins and genes, ultimately prompting the cell to “grow” by dividing.



▲ ONCOGENIC MUTATIONS

In a significant portion of lung and breast tumors, members of the *EGFR* gene family are mutated or duplicated, which boosts the number or function of the receptors they encode, overstimulating this growth pathway. Damage to downstream genes can have similar results. Changes in the *B-RAF* gene, seen in some 70 percent of melanomas, also promote hyperactive cell proliferation. Versions of the *RAS* gene are mutated in many cancer types, which can affect cell growth as well as intersecting pathways—for example, interfering with a suicide program that normally destroys damaged cells.

A Disease of Genes

THE IDEA THAT ALTERATIONS to the cellular genome lie at the heart of all forms of cancer is not new. Since the first identification in 1981 of a cancer-promoting version of a human gene, known as an oncogene, scientists have increasingly come to understand that cancer is caused primarily by mutations in specific genes. The damage can be incurred through exposure to toxins or radiation, by faulty DNA repair processes or by errors that occur when DNA is copied prior to cell division. In relatively rare cases, a cancer-predisposing mutation is carried within a gene variant inherited from one's ancestors.

Whatever their origin, these mutations disrupt biological

pathways in ways that result in the uncontrolled cell replication, or growth, that is characteristic of cancer as well as other hallmarks of malignancy, such as the ability to invade neighboring tissues and to spread to sites throughout the body. Some mutations may disable genes that normally protect against abnormal cell behavior, whereas others increase the activity of disruptive genes. Most cells must acquire at least several of these alterations before they become transformed into cancer cells—a process that can take years.

Over the past two decades many individual research groups have used groundbreaking molecular biology techniques to search for mutations in genes that are likely candidates for wreaking havoc on normal patterns of cell growth and behavior. This approach has identified about 350 cancer-related genes and yielded many significant insights into this diabolical disease. A database of these changes, called the catalogue of somatic mutations in cancer, or COSMIC, is maintained by Michael Stratton's group at the Wellcome Trust Sanger Institute in Cambridge, England. But no one imagines that it is the complete list.

So does it make sense to continue exploring the genomic basis of cancer at cottage-industry scale when we now possess the means to vastly increase the scope and speed of discovery? In recent years a number of ideas, tools and technologies have emerged and, more important, converged in a manner that

Overview/*Cancer Connections*

- Changes in the structure or activity of genes underlie the malignant behavior of cancer cells.
- Identification of genes involved in certain cancers is already advancing diagnosis and treatment.
- The Cancer Genome Atlas is a monumental initiative to eventually identify all the genetic alterations in different forms of cancer so that gene changes driving the disease can be targeted directly.

LUCY READING-IKKANDA; ACKNOWLEDGMENT: SPECIAL THANKS TO JEFFREY SETTLEMAN OF THE CENTER FOR CANCER RESEARCH AT MASSACHUSETTS GENERAL HOSPITAL AND DAPHNE W. BELL OF THE NHGRI CANCER GENETICS BRANCH FOR THEIR ADVICE ON THE PATHWAY ILLUSTRATION

Hallmarks of Cancer

The six abnormal capabilities listed below together give tumors their lethal power to overrun their native tissue and spread through the body.

Self-sufficiency in growth signaling

Cancer cells amplify external growth cues or generate their own.

Insensitivity to antigrowth signals

Cancer cells become deaf to quiescence cues from surrounding tissue.

Evasion of cell suicide

Mechanisms that should trigger or carry out a self-destruct program in damaged cells are disabled or overridden.

Limitless replicative potential

Cancer cells evade intrinsic limits on the number of times a normal cell can divide.

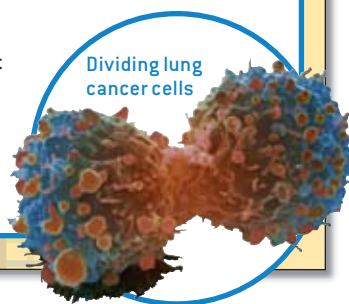
Sustained blood vessel growth

Tumors emit signals promoting the development of new blood vessels to deliver oxygen and nutrients.

Invasiveness and motility

Cancer cells defy multiple signals and forces that hold a cell in place and prevent it from traveling to—and thriving in—other tissues.

Adapted from "The Hallmarks of Cancer," by Douglas Hanahan and Robert A. Weinberg, in *Cell*, Vol. 100; January 7, 2000.



has convinced many leading minds in the cancer and molecular biology communities that it is time for a systematic, collaborative and comprehensive exploration of the genomics of cancer.

The Human Genome Project laid a solid foundation for TCGA by creating a standardized reference sequence of the three billion DNA base pairs in the genome of normal human tissues. Now another initiative is needed to compare the DNA sequences and other physical characteristics of the genomes of normal cells with those of cancerous cells, to identify the major genetic changes that drive the hallmark features of cancer [see box above]. The importance of international partnerships in large-scale biology to pool resources and speed scientific discovery was also demonstrated by the Human Genome Project, and TCGA is exploring similar collaborations.

Finally, the Human Genome Project spurred significant advances in the technologies used to sequence and analyze genomes. At the start of that project in 1990, for example, the cost of DNA sequencing was more than \$10 per "finished" nucleotide base. Today the cost is less than a penny per base and is expected to drop still further with the emergence of innovative sequencing methods [see "Genomes for All," by George M. Church; *SCIENTIFIC AMERICAN*, January 2006]. Because of these and other technological developments, the large-scale approach embodied in TCGA—unthinkable even

a few years ago—has emerged as perhaps the most efficient and cost-effective way to identify the wide array of genomic factors involved in cancer.

Proofs of Concept

PILES OF DATA are, of course, not worth much without evidence that comprehensive knowledge of cancer's molecular origins can actually make a difference in the care of people. Several recent developments have provided proofs of concept that identifying specific genetic changes in cancer cells can indeed point to better ways to diagnose, treat and prevent the disease. They offer encouraging glimpses of what is to come and also demonstrate why the steps toward those rewards are complex, time-consuming and expensive.

In 2001, when the Wellcome Trust Sanger Institute began its own effort to use genomic technologies to explore cancer, the project's immediate goal was to optimize robotics and information management systems in test runs that involved sequencing 20 genes in 378 cancer samples. But the group hit pay dirt a year later when they found that a gene called *B-RAF* was mutated in about 70 percent of the malignant melanoma cases they examined. A variety of researchers swiftly set their sights on this potential new therapeutic target in the most deadly form of skin cancer. They tested multiple approaches—from classic chemical drugs to small interfering ribonucleic acids—in cell lines and in mice, to see if these interventions could block or reduce the activity of *B-RAF* or inhibit a protein called MEK that is overproduced as a result of *B-RAF* mutations. Just five years later the most promising of these therapies are being tested in clinical trials.

Other research groups have already zeroed in on genetic mutations linked to certain types of breast cancer, colon cancer, leukemia, lymphoma and additional cancers to develop molecular diagnostics, as well as prognostic tests that can point to an agent in the current arsenal of chemotherapies to which a particular patient is most likely to respond. Cancer genomics has also helped to directly shape the development and use of some of the newest treatments.

The drug Gleevec, for example, was designed to inhibit an enzyme produced by a mutant fused version of two genes, called *BCR-ABL*, that causes chronic myelogenous leukemia. Gleevec is proving dramatically effective against that disease and showing value in the treatment of more genetically complex malignancies, such as gastrointestinal stromal tumor and several other relatively rare cancers that involve similar

THE AUTHORS

FRANCIS S. COLLINS and **ANNA D. BARKER** are leaders of The Cancer Genome Atlas initiative in their positions as, respectively, director of the National Human Genome Research Institute and deputy director for Advanced Technologies and Strategic Partnerships of the National Cancer Institute. Collins led the Human Genome Project to its completion of the human DNA sequence, and Barker has headed drug development and biotechnology research efforts in the public and private sectors, with a particular focus on fighting cancer.

Genes and Cancer

A connection between genetic abnormalities and the aberrant features of cancer cells was first suggested more than 100 years ago by German biologist Theodor Boveri and others. But over the past few decades evidence that gene alterations directly cause the deranged behavior of cancer cells began accumulating. Calls arose by 1986 to sequence the normal human genome to study malignant gene changes comprehensively. The Human Genome Project was completed in 2003. The Cancer Genome Atlas project will start cataloguing the gene mutations found in three types of human cancer this year.

1890–1914

Studies of abnormal chromosome distribution during cell division suggest a role in malignancy.

Theodor Boveri ▶

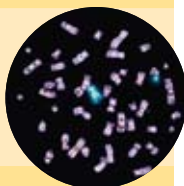


1950s–1960s

Multiple discoveries reveal that tumor viruses can cause cancer by injecting their genes into cells.

1960

First genetic defect associated with a specific cancer—an abnormality known as the Philadelphia chromosome—is discovered in chronic myelogenous leukemia (CML) cells.



1976

Scientists discover that *src*, a nonviral gene found in animal cells, can cause cancer.

1979

P53, later found to be the most frequently mutated gene in human cancer, is discovered.

1981

H-RAS is the first human oncogene (a gene whose alteration is cancer-promoting) to be discovered.

1983

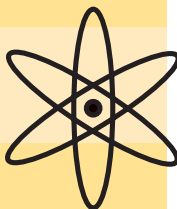
Altered methylation of DNA, suspected to affect gene activation, found in cancer cells.

1986

Renato Dulbecco, writing in *Science*, calls for sequencing the human genome to advance cancer research.

1986

U.S. Department of Energy considers sequencing the human genome to further study of radiation effects.



1986

First tumor suppressor gene, *RBI*, is identified.

1987

Fused gene *BCR-ABL* in the Philadelphia chromosome is found to cause CML.

1990

Model of multistep tumor genesis clarifies the role of accumulated gene changes in cellular transformation to malignancy.



1990

Human Genome Project begins.



enzymes. Herceptin, an agent that targets a cellular signal-receiving protein called *HER2*, is successful against breast cancers with an abnormal multiplication of the *HER2* gene that causes overproduction of the receptor protein.

Strategies for selecting treatments based on specific gene mutations in a patient's cancer are also being tested in studies of the drugs Iressa and Tarceva for lung cancer, as well as Avastin for lung, colon and other cancers. The performance of these new gene-based diagnostics, prognostics and therapeutics is certainly good news, although the list of such interventions remains far shorter than it would be if researchers in academia and the private sector had ready access to the entire atlas of genomic changes that occur in cancer.

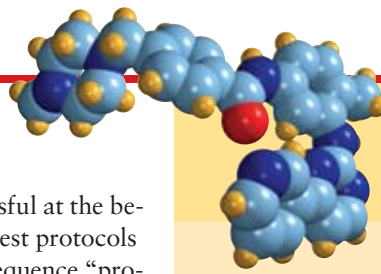
A recent study led by investigators at Johns Hopkins University illustrates both the power of large-scale genomics applied to the discovery of cancer genes and the tremendous undertaking a comprehensive cancer genome atlas will be. The group sequenced about 13,000 genes in tumor tissues taken from 11 colorectal cancer patients and 11 breast cancer patients and reported finding potentially significant mutations in nearly 200 different genes. Interestingly, only about a dozen genes had previously been linked to these two types of cancer, and most scientists had generally expected to find just a few more.

Among the major challenges encountered by researchers sequencing cancer cell genomes is the difficulty of distinguishing meaningless mutations in the tumor samples from those that are cancer-related. Somewhat surprisingly, early sequencing studies have also found very little overlap among the genetic mutations present in different types of cancer and even substantial variation in the pattern of genetic mutations among tumor samples from patients with the same type of cancer. Such findings underscore the idea that many different possible combinations of mutations can transform a normal cell into a cancer cell. Therefore, even among patients with cancers of the same body organ or tissue, the genetic profile of each individual's tumor can differ greatly.

To grasp the full scope of what TCGA hopes to achieve, one must consider the complexities identified in such early efforts and imagine extending the work to more than 100 types of cancer. It is enough to give even veterans of the Human Genome Project and seasoned cancer biologists pause. Yet TCGA participants and other scientific pioneers from around the world are forging ahead, because we are convinced that amid the intricacies of the cancer genome may lie the greatest promise for saving the lives of patients.

Although researchers will probably take many years to complete a comprehensive catalogue of all the genomic mutations that cause normal cells to become malignant, findings with the potential to revolutionize cancer treatment are likely to appear well before this compendium is finished, as the proofs of concept have shown. As each new type of cancer is studied and added to TCGA, investigators will gain another rich new set of genomic targets and profiles that can be used to develop more tailored therapies.

PHOTO RESEARCHERS, INC. (Boveri); DEPARTMENT OF CLINICAL CYTOGENETICS, ADDENBROOKE'S HOSPITAL Photo Researchers, Inc. (Philadelphia chromosome); STEVE GSCHMEISSNER Photo Researchers, Inc. (tumor cell); MARK J. WINTER Photo Researchers, Inc. (Gleevec molecule); CECIL H. FOX Photo Researchers, Inc. (Biospecimen Core Resource); AFFYMETRIX (Cancer Genome Characterization Centers); NATIONAL HUMAN GENOME RESEARCH INSTITUTE (Genome Sequencing Centers); © CDC/PHIL Corbis (Data Coordinating Center)



◀ Gleevec model

Compiling a Colossal Atlas

A PHASED-IN STRATEGY that proved successful at the beginning of the Human Genome Project was to test protocols and technology before scaling up to full DNA sequence “production.” Similarly, TCGA is beginning with a pilot project to develop and test the scientific framework needed to ultimately map all the genomic abnormalities involved in cancer.

In 2006 the National Cancer Institute and National Human Genome Research Institute selected the scientific teams and facilities that will participate in this pilot project, along with the cancer types they will begin examining. Over the next three years these two institutes will devote \$100 million to compiling an atlas of genomic changes in three tumor types: glioblastomas of the brain, lung cancer and ovarian cancer. These particular cancers were chosen for several reasons, including their value in gauging the feasibility of scaling up this project to a much larger number of cancer types. Indeed, only if this pilot phase achieves its goals will the NIH move forward with a full-fledged project to develop a complete cancer atlas.

The three malignancies that we selected for the pilot collectively account for more than 210,000 cancer cases in the U.S. every year and caused an estimated 191,000 deaths in this country in 2006 alone. Moreover, tumor specimen collections meeting the project’s strict scientific, technical and ethical requirements exist for these cancer types. Last September our institutes announced the selection of three biorepositories to provide such specimens, along with new tumor samples as needed, and normal tissue from the same patients for comparison. Those facilities will deliver materials to a central Biospecimen Core Resource, one of four major structural components in TCGA’s pilot project.

Cancer Genome Characterization Centers, Genome Sequencing Centers and a Data Coordinating Center constitute the project’s other three main elements [see illustration at right], and all these groups will collaborate and exchange data openly. Specifically, the seven Cancer Genome Characterization Centers will use a variety of technologies to examine the activity levels of genes within tumor samples and to uncover and catalogue so-called large-scale genomic changes that contribute to the development and progression of cancer. Such alterations include chromosome rearrangements, changes in gene copy numbers and epigenetic changes, which are chemical modifications of the DNA strand that can turn gene activity on or off without actually altering the DNA sequence.

Genes and other chromosomal areas of interest identified by the Cancer Genome Characterization Centers will become targets for sequencing by the three Genome Sequencing Centers. In addition, families of genes suspected to be important in cancer, such as those encoding enzymes involved in cell-cycle control known as tyrosine kinases and phosphatases, will be sequenced to identify genetic mutations or other small-scale changes in their DNA code. At present, we estimate that some 2,000 genes—in each of perhaps 1,500 tumor samples—will be sequenced during this pilot project. The exact numbers will, of course, depend on the samples obtained and what is discovered

1993

Preclinical testing starts on drug that would become Gleevec, the first therapy developed to target a known gene-based cause of a cancer.

1999

Gene-activity profiles are first shown to distinguish between cancer types and to predict chemotherapy response.



2001

Gleevec earns FDA approval.

2002

Wellcome Trust Sanger Institute tumor genome survey discovers a mutation in *B-RAF* gene common to 70 percent of melanomas.

2003

Human Genome Project is completed.



2005

The Cancer Genome Atlas (TCGA) pilot project is announced by the National Institutes of Health.

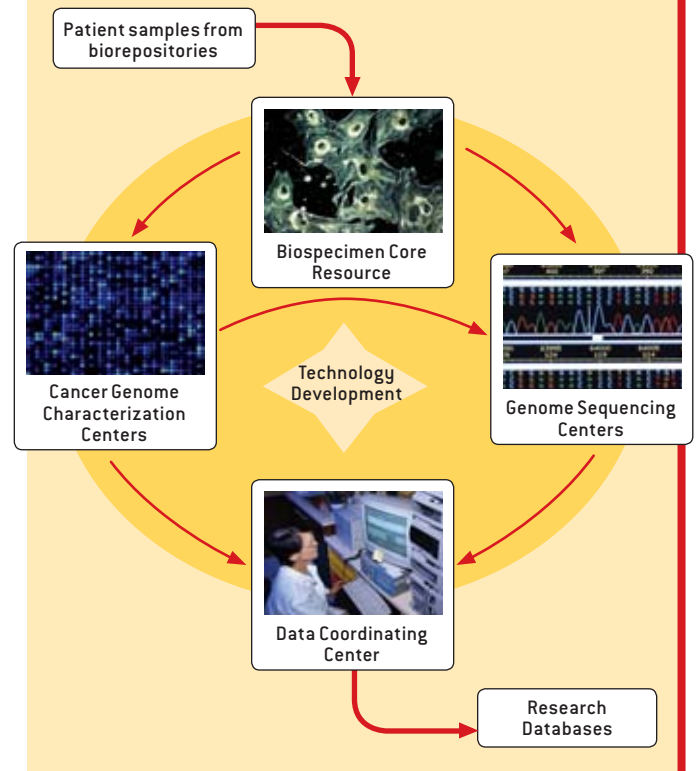
2006

TCGA names pilot project participants and three cancer types for sequencing and genetic analysis.

2007–2010

TCGA will collect and analyze tumor samples obtained from designated biorepository institutions treating patients with cancer. The project’s four primary components—a Biospecimen Core Resource, seven Cancer Genome Characterization Centers, three Genome Sequencing Centers and a Data Coordinating Center—will cooperate to test methods and technologies and to generate and manage data that will be made available to the wider research community.

How Will It Work?



about them by the Cancer Genome Characterization Centers.

Both the sequencing and genome characterization groups, many of which were participants in the Human Genome Project, can expect to encounter a far greater level of complexity than that in the DNA of normal cells. Once cells become cancerous, they are prone to an even greater rate of mutation as their self-control and repair mechanisms fail. The genomic makeup of individual cells can therefore vary dramatically within a single tumor, and the integrated teams will need to develop robust methods for efficiently distinguishing the “signal” of a potentially biologically significant mutation from

the “noise” of the high background rate of mutations seen in many tumors. Furthermore, tumors almost always harbor some nonmalignant cells, which can dilute the sample. If the tumor DNA to be sequenced is too heterogeneous, some important mutations may be missed.

Following the lead of the Human Genome Project and other recent medical genomics efforts, all these data will be made swiftly and freely available to the worldwide research community. To further enhance its usefulness to both basic and clinical researchers and, ultimately, health care professionals, TCGA will link its sequence data and genome analyses with

From Genome to Cancer—Why the Time Is Right

By Renato Dulbecco

When in 1986 I suggested a new project directed at identifying all human genes, one of my overriding goals was to find those genes involved in cancer development—a feat I hoped would lead to new tools for cancer research and, ultimately, to new therapies. That original human genome project has now been carried out and has demonstrated its usefulness for the discovery of genes involved in many diseases, including cancer. Moreover, the genome sequencing effort has been extended to other organisms—from bacteria to chimpanzees—and is showing the unity of life by revealing how many genes distant species share in common.

In the course of this work, new technologies have also provided a much more detailed understanding of the complicated processes by which genes give rise to a variety of functional molecules. An important outcome of this research is the realization that genes do not act alone but are participants in extensive networks of activity within cells. Any change in the functioning of one gene can therefore be accompanied by changes in the workings of multiple genes and proteins involved in the cells' self-maintenance.

The complexity of this system in normal cells is evident in what we already know about cancer—that it results from the stepwise loss of such cellular self-control, which becomes more and more complete as the disease progresses. That progression is caused only in part by physical alterations, or mutations, in specific genes; mostly it is the result of consequent changes in the activity of many other genes involved in cell regulation. Single genes may therefore be responsible in the initiation of cancer and so potential therapeutic targets. To reach the more advanced stages of these cancers (such as the acute phase

of myeloid leukemia or the metastatic phase of other cancers), however, the participation of many other genes is required. Most of them are still unknown.

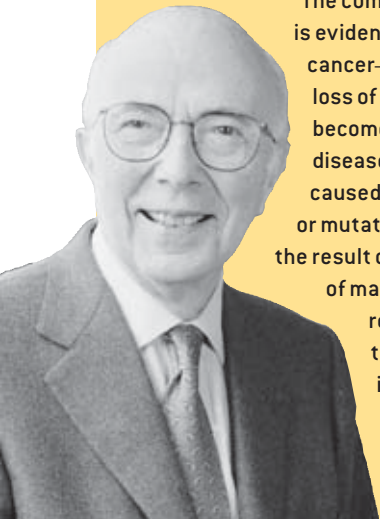
An exception is the recently observed phenomenon of oncogene addiction in certain tumor cells: despite the presence of numerous mutations to the cellular genome, turning off the activity of one so-called oncogene causes the cells to commit suicide via a mechanism known as apoptosis. But how generally this phenomenon occurs is also unknown. To approach these questions, it will be necessary to have a complete catalogue of the structural and functional alterations of genes and other cellular components that cause the loss of regulation in cancer cells. This process, in turn, will require a complete determination of their connections into networks by computational means—a task for the future.

On the way to this goal, however, many other unanswered questions can be explored by the research community. A possible role for stem cells in cancer, for example, is supported by similarities in the behavior of stem cells and cancer cells: both have an unlimited ability to divide; both are very sensitive to the cellular environment, or niche, in which they grow; and many of the genes known to be active in stem cells are also activated in cancer cells.

The advent of genomics has provided welcome insight into the mechanisms by which normal cells become cancerous, but our picture is still incomplete. The time has come to obtain a truly comprehensive catalogue of the genes involved in cancer, bringing to bear all the power of the new tools of genomics and molecular biology to the problem. The Cancer Genome Atlas project aims to do just that.

Renato Dulbecco is president emeritus of the Salk Institute for Biological Studies and co-recipient of the 1975 Nobel Prize in Physiology or Medicine for discoveries related to the interaction of tumor viruses and the genetic material of the cell.

Cancer results from the stepwise loss of cellular self-control.



information about observable characteristics of the original tumors and the clinical outcomes of the sample donors. Developing the bioinformatic tools to gather, integrate and analyze those massive amounts of data, while safeguarding the confidentiality of patient information, is therefore another hurdle that must be cleared to turn our vision into reality.

Uncharted Territory

THE ROAD AHEAD is fraught with scientific, technological and policy challenges—some of which are known and others as yet unknown. Among the uncertainties to be resolved: Will new sequencing technologies deliver on their early promise in time to make this effort economically feasible? How quickly can we improve and expand our toolbox for systematically detecting epigenetic changes and other large-scale genomic alterations involved in cancer, especially those associated with metastasis? How can we harness the power of computational biology to create data portals that prove useful to basic biologists, clinical researchers and, eventually, health care professionals on the front lines? How can we balance intellectual-property rights in a way that promotes both basic research and the development of therapies? When will Congress finally pass genetic nondiscrimination legislation so that knowledge gained through TCGA will have the maximum positive influence on Americans' health? The list goes on.

To avoid raising false expectations, we also must be clear about the questions this project will not attempt to answer. Although it will serve as a resource for a broad range of biological exploration, TCGA is only a foundation for the future of cancer research and certainly not the entire house. And we face the sobering issue of time—something that is in short supply for many cancer patients and their families. As we survey the considerable empty spaces that exist in our current map of genomic knowledge about cancer, the prospect of filling those gaps is both exhilarating and daunting. Scientists and the public need to know up front that this unprecedented foray into molecular cartography is going to take years of hard work and creative problem solving by thousands of researchers from many different disciplines.

Where all this work will lead can only be dimly glimpsed today. In this sense, our position is similar to that of the early 19th-century explorers Meriwether Lewis and William Clark. As they ventured up the Missouri River into the largely uncharted Northwest Territory in 1804, their orders from President Thomas Jefferson were to “take observations of latitude and longitude at all remarkable points. . . . Your observations are to be taken with great pains and accuracy; to be entered distinctly and intelligibly for others, as well as yourself.”

Although Lewis and Clark did not find the much-longed-for water route across the continent, their detailed maps proved valuable to their fledgling nation in myriad ways that Jefferson could never have imagined. For the sake of all those whose lives have and will be touched by cancer, we can only hope our 21st-century expedition into cancer biology exceeds even Renato Dulbecco's grandest dreams. SA

Targeting Gene Changes in Cancer

TCGA pilot project teams will examine the DNA of some 1,500 tumor samples from patients with cancers of the lung, ovaries or brain (glioblastoma), looking for genetic changes. Approximately 2,000 suspect genes in each sample will be sequenced to identify specific mutations. The list of target genes will be tailored to each cancer type and largely determined by what the Cancer Genome Characterization Centers find in the samples, although candidates will also be drawn from categories of genes already associated with cancer.



From left to right: Glioblastoma, lung cancer, ovarian cancer

GENE CATEGORIES	EXAMPLES
Genes identified by TCGA Cancer Genome Characterization Centers as having aberrant structure or activity in a significant number of tumor samples	In some brain tumor cell lines, a gene encoding the intracellular protein NF-KAPPA B is much more active than in normal brain tissue
Well-known oncogenes (genes whose overactivity or alteration is cancer-promoting)	<ul style="list-style-type: none"> • Growth factor receptor genes: <i>HER2</i> (breast and lung cancers), <i>EGFR</i> (lung and colon cancers) • Signaling protein genes: <i>BCR-ABL</i> (chronic myelogenous leukemia), <i>RAS</i> (many cancers), <i>B-RAF</i> (skin cancers) • Regulators of cell death: <i>BCL-3</i> (lymphoma)
Well-known tumor suppressors (genes that protect cells from malignant transformation, unless disabled by mutation)	<ul style="list-style-type: none"> • Controllers of cell division: <i>RB1</i> (retinoblastoma) • DNA repairers: <i>HNPCC</i> (colon cancer, endometrial cancer) • Promoters of programmed cell suicide: <i>P53</i> (lung, colon, breast and brain tumors)
Genes related to known oncogenes and tumor suppressor genes by similarity or pathway membership	The oncogenes <i>HER2</i> and <i>EGFR</i> are part of the epidermal growth factor receptor signaling pathway, which contains at least half a dozen other genes suspected of playing key roles in cancer development and progression

MORE TO EXPLORE

The New Era in Cancer Research. Harold Varmus in *Science*, Vol. 312, pages 1162–1165; May 26, 2006.

The Consensus Coding Sequences of Human Breast and Colorectal Cancers. Tobias Sjöblom et al. in *Science*, Vol. 314, pages 268–274; October 13, 2006. (Published online September 7, 2006.)

The Cancer Genome Atlas: <http://cancergenome.nih.gov>

“He often needed a portable, infallible, artificial memory.”

New systems may allow people to record everything they see and hear—and even things they cannot sense—and to store all these data in a personal digital archive

By Gordon Bell
and Jim Gemmell

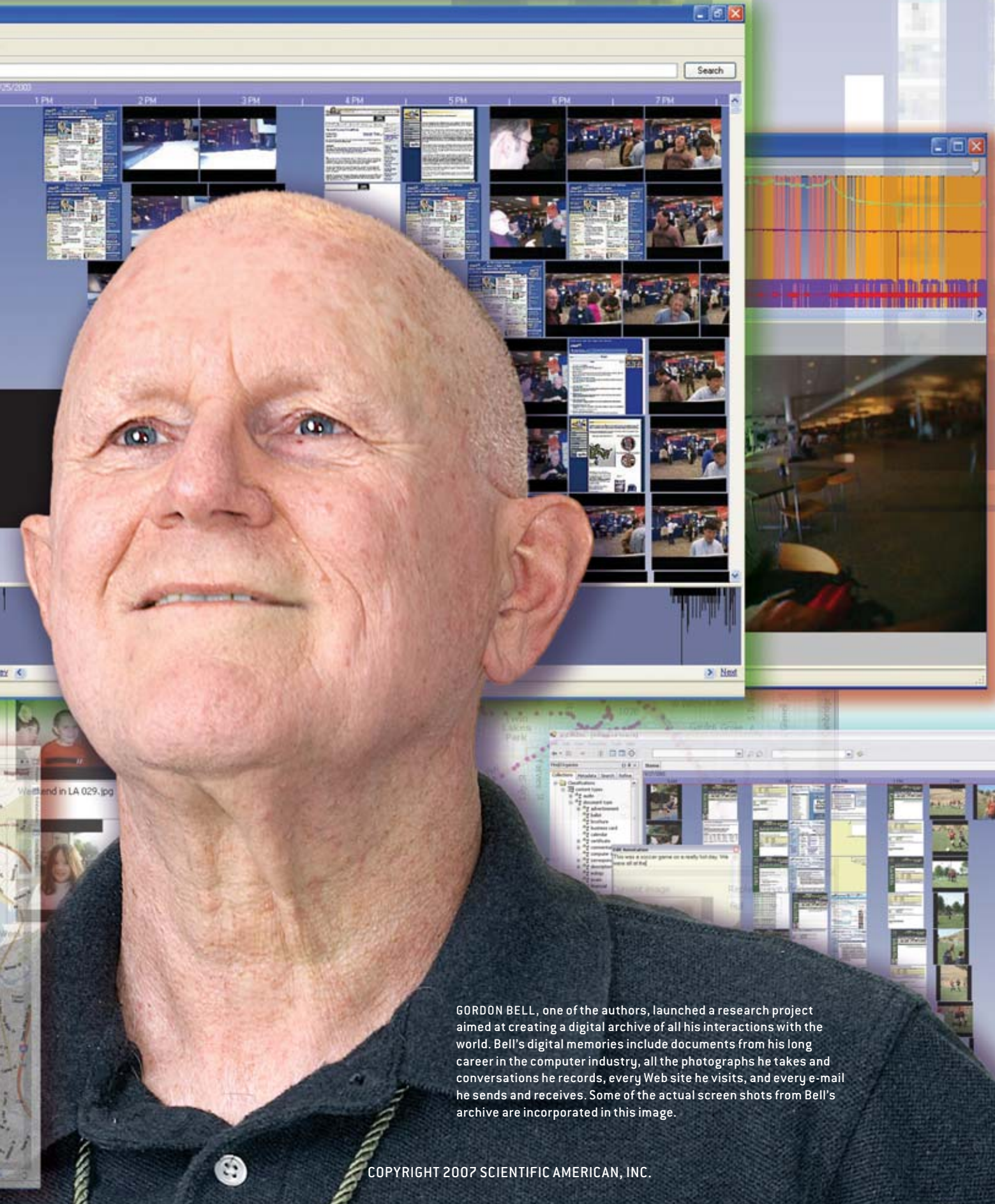
A Digital Life

Human memory can be maddeningly elusive. We stumble upon its limitations every day, when we forget a friend's telephone number, the name of a business contact or the title of a favorite book. People have developed a variety of strategies for combating forgetfulness—messages scribbled on Post-it notes, for example, or electronic address books carried in handheld devices—but important information continues to slip through the cracks. Recently, however, our team at Microsoft Research has begun a quest to digitally chronicle every aspect of a person's life, starting with one of our own lives (Bell's). For the past six years, we have attempted to record all of Bell's communications with other people and machines, as well as the images he sees, the sounds he hears and the Web sites he visits—storing everything in a personal digital archive that is both searchable and secure.

Digital memories can do more than simply assist the recollection of past events, conversations and projects. Portable sensors can take readings of things that are not even perceived by humans, such as oxygen levels in the blood or the amount of carbon dioxide in the air. Computers can then scan these data to identify patterns: for instance, they might determine which environmental conditions worsen a child's asthma. Sensors can also log the three billion or so heartbeats in a person's lifetime, along with other physiological indicators, and warn of a possible heart attack. This information would allow doctors to spot irregularities early, providing warnings before an illness becomes serious. Your physician would have access to a detailed, ongoing health record, and you would no longer have to rack your brain to answer questions such as “When did you first feel this way?”

Our research project, called MyLifeBits, has provided some of the tools needed to compile a lifelong digital archive. We have found that digital memories allow one to vividly relive an event with sounds and images, enhancing personal reflection in much the same way that the Internet has aided scientific investigations. Every word one has ever read, whether in an e-mail, an electronic document or on a Web site, can be found again with just a few keystrokes. Computers can analyze digital memories to help with

—Patrick O'Brian, *The Fortune of War*



GORDON BELL, one of the authors, launched a research project aimed at creating a digital archive of all his interactions with the world. Bell's digital memories include documents from his long career in the computer industry, all the photographs he takes and conversations he records, every Web site he visits, and every e-mail he sends and receives. Some of the actual screen shots from Bell's archive are incorporated in this image.

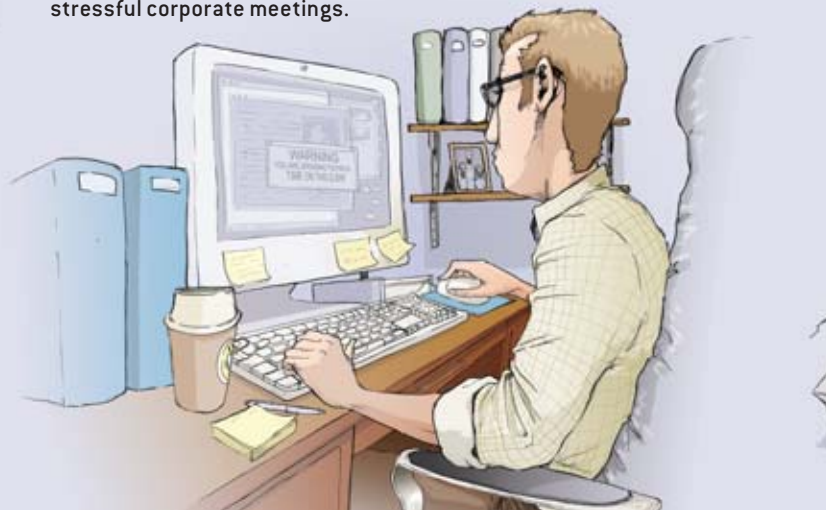
A DAY IN THE DIGITAL LIFE

Access to lifelong digital memories can improve job productivity, medical treatment, school performance and much else. A glimpse of the daily routines of the Digitals, a fictitious family living in the not so distant future, illustrates some of the potential benefits.



Anne Digital, a chemistry professor, needs to review an article written by a colleague but cannot recall the name of the author. Anne does, however, remember viewing the document during a phone call with a student last summer. She limits her document search to those viewed during phone calls with the student—which were automatically logged at the time—and immediately sees the one she needs.

Dave Digital, a stockbroker, archives all his documents, e-mails, phone calls and Web site visits over the course of the workday. As Dave composes an e-mail, his time management program warns that he is spending too many hours communicating with an unimportant customer. He forgoes polishing the message and sends it off as is, then goes to work on a higher-priority account. Dave also reviews recent records of his weight, heart rate and caloric input to determine the negative effects of two days of stressful corporate meetings.



Dave takes his seven-year-old daughter, Laura, to the doctor. Laura's health analysis program recommended a checkup because her weight gain over the past six months (as logged daily by her bathroom scale) is lower than expected.

The doctor says this symptom could be a side effect of Laura's asthma medication. Because her breathing has been excellent for several months, the doctor suggests taking her off the medication for a while.

time management, pointing out when you are not spending enough time on your highest priorities. Your locations can be logged at regular intervals, producing animated maps that trace your peregrinations. Perhaps most important, digital memories can enable all people to tell their life stories to their

descendants in a compelling, detailed fashion that until now has been reserved solely for the rich and famous.

A Web of Trails

THE VISION of machine-extended memory was first expounded at the end of World War II by Vannevar Bush, then

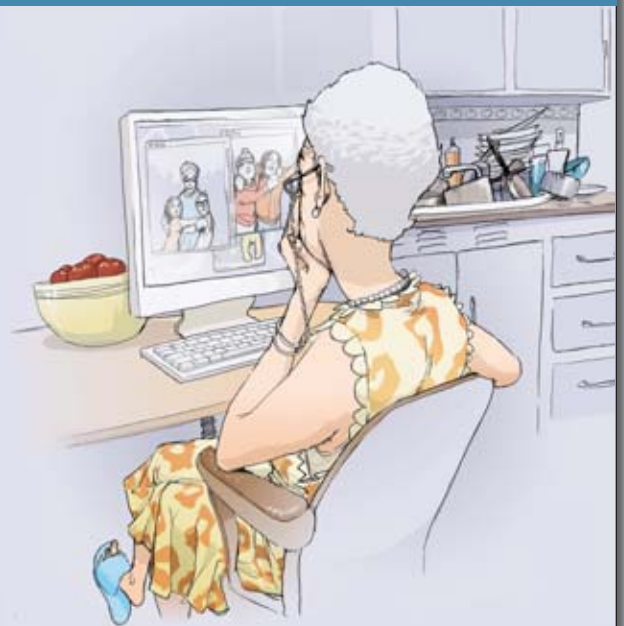
director of the U.S. government office that controlled wartime research. Bush proposed a device called the Memex (short for "memory extender")—a microfilm-based machine that would store all of an individual's books, records and communications. The Memex was to be built into a desk and equipped with a keyboard, a microphone and several display surfaces. The person behind the desk could use a camera to make microfilm copies of photographs and papers or create new documents by writing on a touch-sensitive screen. The Memex user could also mount a camera on his or her forehead to capture pictures while away from the desk. One of the most prescient of Bush's ideas was the suggestion that the Memex should be designed to imitate the associative thinking of the

Overview/Digital Memories

- Because human memory is fallible, researchers are striving to develop systems that can automatically record communications, documents, images and video, storing everything in a searchable archive.
- Ongoing advances in sensors and data storage promise to make digital recording easier. The bigger challenge is devising software that can organize the information.
- Digital memories may yield benefits in medical care, job productivity and other areas, but developers must ensure that the archives are secure.



At dinner Dave and Anne argue with their 14-year-old son, Steve. Anne is frustrated that Steve leaves his homework assignments to the last minute and wants him to start his next essay immediately. But Steve shows his parents the results of his education analysis program, which indicate that his grades are just as high when he does his homework late. Steve's digital memories also reveal that he is an auditory learner, benefiting more from group discussion than from reading.



Anne's mother, Jean, is a senior citizen living in an assisted-care home. The staff at the home have access to parts of Jean's digital memories; they will be automatically alerted if any heart or breathing irregularities occur or if her wearable monitors show that she is failing to take her usual walks. Anne has found that watching her mother's dishwasher log is insightful: when Jean feels run down, she often gets behind on the dishes. At bedtime Jean watches old photographs and videos from her digital archive, using an interactive display to take an extended walk down memory lane.



human mind, which he described in lively terms: “With one item in its grasp, it snaps instantly to the next that is suggested by the association of thoughts, in accordance with some intricate web of trails carried by the cells of the brain.”

Over the next half a century intrepid computer science pioneers, including Ted Nelson and Douglas Engelbart, developed some of Bush's ideas, and the inventors of the World Wide Web borrowed the concept of the “web of trails” to build their system of linking sites. But the Memex itself remained technologically out of reach. In recent years, however, rapid advances in storage, sensor and processor technologies have paved the way for new digital recording and retrieval systems that may ultimately go far beyond Bush's vision.

The growth of digital storage capacity has been staggering: today a \$600 hard drive can hold a terabyte (one trillion bytes) of data, which is enough to store everything you read (including e-mails, Web pages, papers and books), all the music you purchase, eight hours of speech and 10 pictures a day for the next 60 years [see table on page 64]. If current trends continue, within a decade you will be able to carry the same amount of information in your cell phone's flash memory, while connecting wirelessly to a \$100 four-terabyte drive on your PC. In 20 years \$600 will buy 250 terabytes of storage—enough to hold tens of thousands of hours of video and tens of millions of photographs. This capacity should be able to satisfy anyone's recording needs for more than 100 years.

At the same time, manufacturers are producing a new generation of inexpensive sensors that may soon become ubiquitous. Some of these devices can record a wealth of information about the user's health and physical movements. Others can gauge the temperature, humidity, air pressure and light level in the surrounding environment and even detect the presence of warm bodies nearby. Some monitors are meant to be worn, and others are designed to be placed in rooms or incorporated into appliances such as refrigerators. (A fridge sensor could keep track of your snacking habits by measuring the number of times the door is opened.) And microphones and cameras are now cheap enough to be installed virtually anywhere—particularly in cell phones, where camera inclusion is becoming the norm,



with voice recording soon to follow.

Finally, the dramatic increase in computing power over the past decade has led to the introduction of processors that can efficiently retrieve, analyze and visualize vast amounts of information. An ordinary notebook PC can run a database that is more powerful and almost 100 times as large as that of a major bank of the 1980s. An inexpensive cell phone can surf the Web, play videos and even understand some speech.

As the hardware for digital recording has improved, more and more people have started to create electronic chronicles of their lives. The advent of cheap, high-quality digital cameras (including those incorporated into cell phones) has triggered a boom in photography. Blogs that incorporate photographs are now becoming more common than personal Web sites. Young people in particular are embracing blogging and the use of mobile devices. The

DAILY JOURNEYS can be documented with the SenseCam (*hanging from Bell's neck in photograph above*), a camera that automatically takes pictures when its sensors indicate changes in light levels or the presence of people nearby. A Global Positioning System device continually tracks Bell's location, enabling the creation of visual diaries of his travels.

fact that this proliferation of digital chronicling is taking place with only very rudimentary tools demonstrates how deep the desire must be. And the interest will surely grow once the process of digital recording becomes easier and more comprehensive.

One Man's Memories

OUR OWN EXPERIENCE with digital memories began in 1998, when Bell decided to go paperless, doing away with a messy mountain of articles, books, cards, letters, memos, posters and photographs. To transfer this heap of memories to a digital record, Bell became obsessed with scanning all the documents and artifacts from his personal life and his long career in the computer business. (He even went so far as to scan the logos

on coffee mugs and T-shirts.) He also began digitizing home movies, videotaped lectures and voice recordings. Bell is now paperless, but the cost was high: it took a personal assistant working for several years to complete the task. (Archiving more recent items has not required such strenuous effort, because the great majority of documents, images and videos are now created in digital formats, so capture is automatic.)

After scanning all this information, however, Bell became frustrated with his inability to make real use of it with the software available to him at the time. This frustration led to the MyLifeBits project. When we started the project in 2001, the search tools that had been developed for desktop computers were cumbersome. We set out to create a database that would give us the ability not only to do full-text searches of our PCs (a capability that is now commonplace) but also to quickly retrieve digital memories using attributes called metadata: for example, the date, place and subject of a photograph or written or spoken comments that the database appends to the file. Metadata are frequent-

GORDON BELL and JIM GEMMELL have been working together on the MyLifeBits project at Microsoft Research since 2001. One of the pioneers of the computer industry, Bell oversaw the development of the famous VAX minicomputer for Digital Equipment Corporation in the 1970s. During the 1980s he became involved in public policy for computer science, and in 1995 he joined Microsoft as a principal researcher with its eSciences Research Group in San Francisco. Gemmell is a senior researcher in Microsoft's Next Media research group. His current focus is on personal lifetime storage, but his interests also include personal media management, telepresence and reliable multicast.

ly a crucial part of recall; a person seeking a specific e-mail, for instance, may remember that it was sent at a certain time of year. By linking these metadata, much of which are obtained automatically, to digital memories, the database allows users to efficiently comb through even the largest archives.

MyLifeBits has also provided Bell with a new suite of tools for capturing his interactions with other people and machines. The system records his telephone calls and the programs playing on radio and television. When he is working at his PC, MyLifeBits automatically stores a copy of every Web page he visits and a transcript of every instant message he sends or receives. It also records the files he opens, the songs he plays and the searches he performs. The system even monitors which windows are in the foreground of his screen at any time and how much mouse and keyboard activity is going on. When Bell is on the go, MyLifeBits continually uploads his location from a portable Global Positioning System device, wirelessly transmitting the information to his archive. This geographic tracking allows the software to automatically assign locations to Bell's photographs, based on the time each is taken.

To obtain a visual record of his day, Bell wears the SenseCam, a camera developed by Microsoft Research that automatically takes pictures when its sensors indicate that the user might want a photograph. For example, if the SenseCam's passive infrared sensor detects a warm body nearby, it photographs the person. If the light level changes significantly—a sign that the user has probably moved in or out of a room and entered a new setting—the camera takes another snapshot. A recent study led by researchers at Addenbrooke's Hospital in Cambridge, England, showed that a memory-impaired patient who reviewed SenseCam images every night was able to retain memories for more than two months. (In contrast, a nightly review of a written diary resulted in almost no improvement in memory retention.) Neuropsychologist Martin Conway of the University of Leeds in England speculates that the SenseCam could become

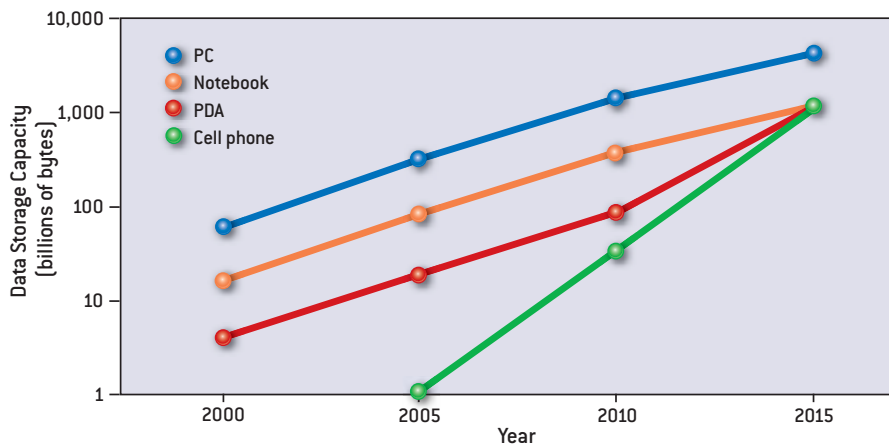
“the first truly powerful 21st-century memory stimulant.”

After six years, Bell has amassed a digital archive of more than 300,000 records, taking up about 150 gigabytes of memory. The information is stored on Bell's dual-disk notebook computer and his assistant's desktop PC, which are backed up locally and off-site. Video files grab the lion's share of the storage space—more than 60 gigabytes—whereas images take up 25 gigabytes and audio files (mostly music) occupy 18 gigabytes. The remainder is shared by 100,000 Web pages, 100,000 e-mails, 15,000 text files, 2,000 PowerPoint files, and so on. Bell has found the system particularly useful for contacting old acquaintances and finding other people with whom he needs to communicate. He has also employed MyLifeBits to retrieve Web sites for citations in his research papers, to provide doctors with records of a 25-year-old coronary bypass, and to obtain a photograph of a deceased friend for a newspaper obituary.

Some features of MyLifeBits, such as full-text search, have already been incorporated into commercial products. As a whole, though, the system requires more development to improve its ease of use and its management of the data. Better software for converting speech to text would greatly enhance the system by allowing users to search for words or phrases in phone conversations or other voice recordings. Similarly, automatic face recognition would solve the pesky problem of photograph labeling. And the retrieval of information could become easier if the system automatically identified the nature of each of the several hundred document types, perhaps by analyzing their form and content. But our research project has already dramatized the evolution of the PC from a word processor and number cruncher to a transaction processor that can log everything about the user's life in high-fidelity multimedia. Many experts have predicted the demise of the personal computer, but it is clear that the “P” in “PC” is not



FAMILY VACATIONS may become more memorable if they are digitally recorded and archived. Vacation snapshots are linked to GPS records (pink dots on map) in this diary of a trip to Los Angeles taken by Jim Gemmell, one of the authors, and his children.



ITEM	SIZE EACH (BYTES)	DAILY RATE	ANNUAL STORAGE (BILLIONS OF BYTES)	60-YEAR STORAGE (BILLIONS OF BYTES)
Books/reports	1,000,000	1	0.4	21.9
E-mails	5,000	100	0.2	11.0
Scanned pages	100,000	5	0.2	11.0
Web pages	50,000	100	1.8	109.5
Purchased songs (MP3)	4,000,000	1	1.5	87.6
Recorded speech	1,000/second	Eight hours	10.5	630.7
Photos (JPEG)	1,000,000	10	3.7	219.0
TOTAL			18.2	1,090.6

Numbers may not add up because of rounding.

STORAGE CAPACITY has grown tremendously in recent years and will continue to climb (graph). A typical desktop PC will hold more than a terabyte (a trillion bytes) of data by 2010; notebook computers, personal digital assistants and cell phones are expected to reach that milestone by 2015. A terabyte of storage is enough to hold all the books, e-mails, recorded conversations, songs and photographs that one is likely to accumulate over 60 years (table).

going away. If anything, PCs will become even more personal. What will change is the “C.” Our machines will evolve into computing ecosystems that encompass not just computers but storage services on the Internet, new access devices (such as cell phones and entertainment centers), and ubiquitous sensors. Most likely our LifeBits will eventually be housed in a home server connected to various Web services.

Realizing the Vision

TO ILLUSTRATE the potential impact of digital memories, we have imagined a day in the life of a fictitious family making full use of this technology in the not so distant future [see box on pages 60 and 61]. Various pieces of the family’s digital memories are stored in their personal devices—their cell phones, lap-

tops, home computers and so on—but all that information is also securely transmitted over the Internet to a host server run by a hypothetical company called LifeBits, Inc. This company manages the storage of the data, performs regular backups (so as to recover any inadvertently deleted material) and places copies of the archive in various locations to ensure that it is not destroyed in a natural or man-made disaster.

Because most of their information is available via secure Web access, the family members can retrieve it anywhere and at any time. Particularly sensitive information that might put someone in legal jeopardy can be kept in an offshore data storage account—a “Swiss data bank,” if you will—to place it beyond the reach of U.S. courts. The children in the family can encrypt their recordings,

but the LifeBits service will give the parents access to the data in case of an emergency. Likewise, some of the parents’ digital memories may be covered by employment contracts that stipulate that the data related to their jobs belong to their employers. When such employees leave their jobs, they may have to perform a “partial lobotomy” on their copies of the memories, expunging everything deemed to be company property.

Some of the scenarios we have described are not all that futuristic. Wearable sensor platforms that collect health data and monitor vital signs such as heart rate, breathing and the number of calories burned are already being commercialized by companies such as VivoMetrics in Ventura, Calif., and BodyMedia in Pittsburgh. In the meantime, Dust Networks in Hayward, Calif., has developed a wireless hub for relaying signals among a network of sensors. The Human Speechome Project, led by Deb Roy at the Massachusetts Institute of Technology Media Lab, is engaged in recording nearly every waking hour of the first three years of a child’s life—the child is Roy’s son, now a one-year-old—to study how people acquire language. And Kiyoharu Aizawa and his colleagues at the University of Tokyo are working on wearable video camera systems that would identify interesting moments to capture for posterity by monitoring the alpha waves in the user’s brain.

Microsoft Research is supporting 14 universities undertaking a variety of projects in the field of digital memories. One of them is MyHealthBits, led by Bambang Parmanto of the University of Pittsburgh; this effort is taking on the challenge of recording huge amounts of health data and managing the voluminous records that result. Recent studies at the University of Washington have shown the benefits of continuous health monitoring in diabetic patients and individuals with sleep disorders.

This early progress is encouraging, but the advent of the digital-memories era will not be trouble-free. Some countries and U.S. states currently impose restrictions on recording conversations or photographing people. Many individuals are

equally concerned about recording information that could be used against them in court. Digital memories, unlike those in our brains, would be fair game in a legal proceeding. Richard Nixon famously advised his aides to say “I can’t recall” when testifying before a grand jury, but tape recordings of his own conversations were his downfall. For those of us who view digital memories as an extension of our own minds, the use of such materials in court would feel like self-incrimination. New technologies, however, can help minimize the potential dangers. When recording others, for instance, it may be possible to obscure their images or speech to avoid illegal recording.

Guarding the privacy of digital memories will be critical. The prospect that identity thieves, gossipmongers or authoritarian states could gain access to such records is frightening. Most people, however, already have quite a lot of sensitive information on their PCs. Security is an important concern regardless of how far you go with the concept of digital memories (although storing a lifetime of personal data in a single archive does at least make the problem quantitatively worse, if not qualitatively). Furthermore, even if our computer systems can be made as secure as Fort Knox, users must be very careful when sharing their information; with a single errant keystroke, one’s medical records might inadvertently be distributed to the entire world. To prevent such mistakes, the user interfaces for digital memories must be better than the ones we have now, and we will need intelligent software to provide warnings when sharing data looks risky.

Another technical challenge will be ensuring that users are able to open their digital files decades after storing them. We have already run into cases where we could not access documents because their formats were obsolete. Digital archivists will have to constantly convert their files to the latest formats, and in some cases they may need to run emulators of older machines to retrieve the data. A small industry will probably emerge just to keep people from losing information because of format evolution.



WEARABLE MONITORS that constantly record and archive an individual’s vital signs can enable doctors to detect illnesses before they become serious. For example, the SenseWear armband, produced by a Pittsburgh company called BodyMedia, can calculate the number of calories burned by measuring the skin’s temperature, heat loss and impedance of electric current.

An even bigger challenge will be devising software that can enable computers to perform useful tasks by tapping into this gigantic store of collected knowledge. The ultimate goal is a machine that can act like a personal assistant, anticipating its user’s needs. At a minimum, computers must do a better job of organizing the information. Search strategies that work well for a few shelves of books may not work at all for a collection the size of the Library of Congress. Most of us do not want to be the librarians of our digital archives—we want the computer to be the librarian!

Consequently, our research group is very interested in applying artificial intelligence (AI) to digital memories. Although many experts are skeptical about AI efforts, we believe that such software may yield practical results if it can draw on the tremendous stores of data in personal archives. An AI system designed to work with a wealth of information is bound to perform better than one that has to make a recommendation based on very few data points. We have

begun work in this area, developing software that could organize files based on their content, but much remains to be done.

In a sense, the era of digital memories is inevitable. Even those who recoil at our vision will have vastly more storage on their computers in the coming years and will expect software to help them more and more in utilizing it. Although some may be frightened at the prospect of ubiquitous recording, for us the excitement far outweighs the fear. Digital memories will yield benefits in a wide spectrum of areas, providing treasure troves of information about how people think and feel. By constantly monitoring the health of their patients, future doctors may develop better treatments for heart disease, cancer and other illnesses. Scientists will be able to get a glimpse into the thought processes of their predecessors, and future historians will be able to examine the past in unprecedented detail. The opportunities are restricted only by our ability to imagine them. SA

MORE TO EXPLORE

As We May Think. Vannevar Bush in *Atlantic Monthly*, Vol. 176, No. 1, pages 101–108; July 1945.

A Personal Digital Store. G. Bell in *Communications of the ACM*, Vol. 44, No. 1, pages 86–91; January 2001.

Digital Memories in an Era of Ubiquitous Computing and Abundant Storage. Mary Czerwinski, Douglas W. Gage, Jim Gemmell, Catherine C. Marshall, Manuel A. Pérez-Quñones, Meredith M. Skeels and Tiziana Catarci in *Communications of the ACM*, Vol. 49, No. 1, pages 44–50; January 2006.

MyLifeBits: A Personal Database for Everything. Jim Gemmell, Gordon Bell and Roger Lueder in *Communications of the ACM*, Vol. 49, No. 1, pages 88–95; January 2006.

Information about MyLifeBits can be found at www.mylifebits.com

DOWN GO THE

Many dams are being torn down these days, allowing rivers and the ecosystems they support to rebound. But ecological risks abound as well. Can they be averted?

BY JANE C. MARKS



DAMS



KEITH PLUMMER Courtesy of NOAA Restoration Center

At the start of the 20th century, Fossil Creek was a spring-fed waterway sustaining an oasis in the middle of the Arizona desert. The wild river and lush riparian ecosystem attracted fish and a host of animals and plants that could not survive in other environments. The river and its surrounds also attracted prospectors and settlers to the Southwest. By 1916 engineers had dammed Fossil Creek, redirecting water through flumes that wound along steep hillsides to two hydroelectric plants. Those plants powered the mining operations that fueled Arizona's economic growth and helped support the rapid expansion of the city of Phoenix. By 2001, however, the Fossil Creek generating stations were providing less than 0.1 percent of the state's power supply.

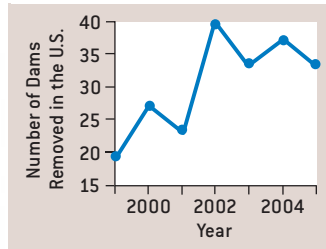
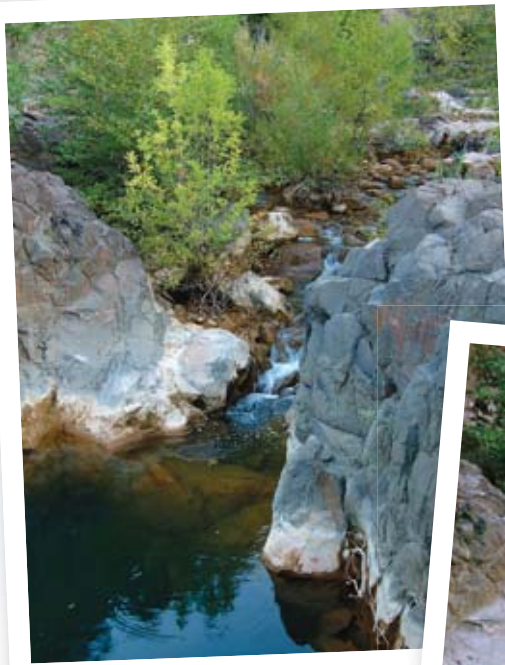
Nearly two years ago the plants were shut down, and an experiment began to unfold. In the summer of 2005 utility workers retired the dam and the flumes and in so doing restored most of the flow to the 22.5 kilometers of Fossil Creek riverbed that had not seen much water in nearly a century. Trickles became waterfalls, and stagnant shallows became deep turquoise pools. Scientists are now monitoring the ecosystem to see whether it can recover after being partially sere for so long, to see whether native fish and plants can again take hold. They are also on the lookout for unintended ecological consequences of the project.

Decommissioning dams (particularly small ones, as is the case in Fossil Creek) is becoming a regular occurrence as structures age, provide an inconsequential share of a region's power, become unsafe or too costly to repair, or as communities decide they want their rivers wild and full of fish again. But simply removing a dam does not automatically mean a long-altered ecosystem will flourish once more. As with all things natural, reality often proves far more complex and intricate than people anticipate. Those of us who have witnessed many of the unexpected consequences of dam removals are now using that knowledge to try to minimize negative results in the future.

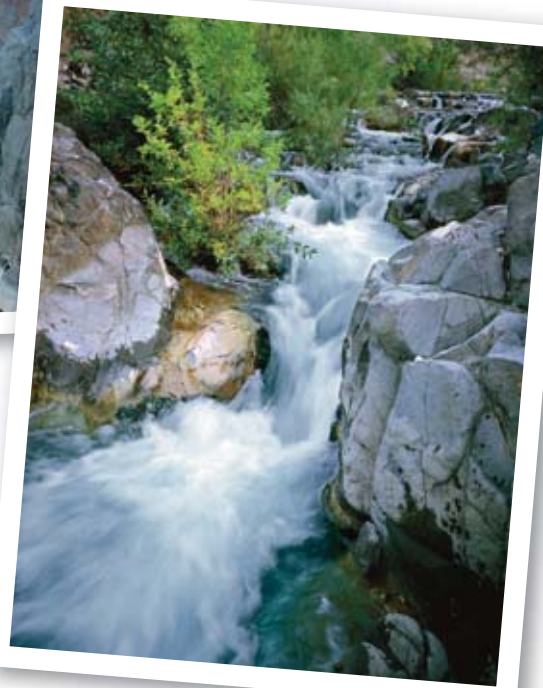
A Global Trend

TODAY ABOUT 800,000 dams operate worldwide, 45,000 of which are large—that is, greater than 15 meters tall. Most were built in the past century, primarily after World War II. Their benefits are clear. Hydroelectric power makes up 20 percent of the globe's electric supply, and the energy is largely clean and renewable, especially when contrasted with other sources. Dams control flooding, and their reservoirs provide a reliable supply of water for irrigation, drinking and recreation. Some serve to help navigation, by stabilizing flow.

SANDY RIVER DAM removal is part of the long-term restoration of Maine's Kennebec River. In 1999 the Edwards Dam on the Kennebec was taken down; soon after, many of the river's native fish returned and their populations grew dramatically. Unconstrained flow of the Sandy River, a tributary of the Kennebec, was restored last summer to ensure that no barriers prevent migratory fish from moving freely.



FOSSIL CREEK, which is fed by seven underground springs, went from merely a trickle (*above*) to a bubbling flow of 314 gallons a second after engineers redirected the water around an old hydropower dam (*right*). Scientists are now studying the creek to see how the food chain changes and to determine whether native species flourish. The dam was one of more than 30 removed in the U.S. in 2005 (*graph*).



Their costs are obvious as well. Dams displace people and as a result have become increasingly controversial in the developing world [see “The Himba and the Dam,” by Carol Ezzell; *SCIENTIFIC AMERICAN*, June 2001]. The structures ruin vistas, trap sediments (needed for deltas, riverbanks and beaches), stymie migratory fish and destroy ecosystems in and around waterways. Conservation-

ists have a long history of opposing dams: John Muir tried to block the dam in Yosemite’s Hetch Hetchy Valley; Edward Abbey’s novel *The Monkey Wrench Gang* targeted Arizona’s Glen Canyon Dam for guerrilla demolition. In recent years, as the downsides of dams have become more widely recognized, groups made up of several interested parties—utility officials, regulators,

policymakers, conservationists, native peoples, researchers and the public—have fought to decommission aging dams.

In the U.S., where hydropower dams must be relicensed every 30 to 50 years, the rate of dam removal has exceeded the rate of construction for the past decade or so. In the previous two years alone, about 80 dams have fallen, and researchers following the trend expect that dams will continue to come down, especially small ones. Although the U.S. is currently leading the effort, it is not alone. France has dismantled dams in the Loire Valley; Australia, Canada and Japan have also removed, or are planning to remove, dams.

Clear successes have driven much of this activity. In 1999 engineers took apart the Edwards Dam on Maine’s Kennebec River after a long battle waged by environmentalists culminated in the Federal Energy Regulatory Commission’s denial of a renewal permit. Within years, biologists observed with some surprise the return of scores of striped bass, alewives, American shad, Atlantic salmon, sturgeon, ospreys, kingfishers, cormorants and bald eagles. They also found that the water became well aerated and that populations of important food-chain insects such as mayflies, stoneflies and caddisflies grew.

In the Loire Valley, the story is similar. Salmon were abundant in the 19th century—about 100,000 would migrate each year—but by 1997, only 389 were counted making the trip. Despite the incorporation of fish ladders and elevators, the eight dams along the Loire and its major tributaries—as well as their turbines and pumps—had decimated the salmon population. Nongovernmental organizations, including the European Rivers Network, led a campaign to bring the salmon back. In response, the French government decommissioned four of the dams—two in 1998, one in 2003 and one in 2005. Within a few months of each dam removal, five species of fish, Atlantic salmon and shad

Overview/Restoring River Flows

- Some 800,000 dams exist around the world, but small ones—and even some large ones—are increasingly being removed so rivers and streams can recover.
- Ecologists are learning, however, that removing or lowering dams takes a great deal of careful planning and active intervention because sometimes the dams confer environmental benefits, such as holding back toxic sediments or blocking the progress of invasive species.
- Before decommissioning a dam on Fossil Creek in Arizona, managers poisoned exotic fish and airlifted native species to safety. Such strategies could prove key to the success of future dam removal projects.

among them, began to reestablish their historic migratory pathways.

In most places where dams have been eliminated, the stories of the Kennebec and the Loire have been repeated. Water clarity and oxygen levels increase as flow comes back, and aquatic insects thrive again. Warm stagnant water runs from behind the dam along with the fish, such as nonnative carp, that love it. As the water moves freely, its temperature falls and cold-loving fish species, such as trout, proliferate or return. The carp population, which tends to squeeze out others, dwindles, sometimes disappearing completely. People, in addition to flora and fauna, return to enjoy the rivers. Biologists have observed these benefits from Wisconsin—one of the U.S. leaders in small dam removal—to New South Wales in Australia. Even restoring some water to rivers without removing a dam has had positive effects [see “Experimental Flooding in Grand Canyon,” by Michael P. Collier, Robert H. Webb and Edmund D. Andrews; *SCIENTIFIC AMERICAN*, January 1997].

The Downsides

BIOLOGISTS have also recorded unexpected problems. The release of sediments trapped behind a dam’s walls can choke waterways, muddying the environment and wiping out insects and algae, which are important food for fish. This wave of turbidity can also eliminate habitat for sessile filter feeders, such as freshwater mussels. Sometimes the mud that had been held back by the structures is rife with contaminants. When engineers removed the Fort Edward Dam on the Hudson River in 1973, concentrations of PCBs rose in downstream fish and remained high for many years; even today the striped bass fishery remains closed because of high levels of PCBs.

Sediments that are not washed downstream can become problematic as well. As they dry out, they may provide fertile ground for potentially noxious exotic plants whose seeds they harbored. Eurasian reed canary grass—which homogenizes wetlands by outcompeting native plant species—grew explosively after Wisconsin’s Oak Street Dam fell, even

though restoration scientists had seeded the area with native prairie plant species.

In some cases, dams have blocked invasive species from moving upriver and into zones above the dam. The dam at Fossil Creek, for example, halted the advance of exotic fish such as bass and sunfish, creating a sanctuary above the structure for imperiled southwestern fish, including headwater chub and speckled dace. The reservoir also provided habitat for a locally threatened species, the lowland leopard frog.

And dam removal can pose dangers for people living nearby. In places where flood control is crucial, government organizations have had to devise safety

Often a big issue facing managers is what to do with ACCUMULATIONS OF DIRT AND DEBRIS.

strategies before dams could come down. In the case of the Loire basin, the government computerized data on weather patterns, rainfall and river levels so flood warnings could be released at least four hours before danger arrived. Engineers also redesigned riverbeds to be wider and deeper, so the waters of the Loire Valley could move more freely without overflowing the banks.

Delicate Decommissioning

THE FOSSIL CREEK restoration project offers a prime example of the kind of planning that could help minimize the damaging effects of dam removal. Researchers carefully planned to control possible disadvantages of the operation. Their principal concerns were what to do with the accumulated sediments, whether to manage the fishery as a native one (which would mean removing exotic species) and how to protect the

reservoir-resident frogs. Ultimately engineers decided to reroute water around the dam, keeping it as a barrier to exotics and permitting the frogs to survive in the backwater.

In addition, biologists decided to actively manage the native fish. They caught as many as they could from the creek itself and airlifted them to a holding tank. They then doused the creek with fish poison to kill exotic species and returned the natives to the water once the poison had dissipated. The U.S. Bureau of Reclamation built a fish barrier 12 kilometers below the existing dam to further impede exotics. Now managers are waiting to see how the Fossil Creek species

do. The dam’s fate will be decided in 2010: if the leopard frog becomes established downstream and exotic fish have not reinvaded the creek, the dam will come out. If not, it will be lowered but not eliminated.

Interestingly, restoring Fossil Creek involves the creation of many more dams—but these will be made of travertine, formed naturally as the calcium carbonate-rich water of the springs interacts with algae to form layers of limestone. These barriers create small, deep pools, the perfect habitat for a variety of fish and insects. They also trap leaf litter, a crucial food source for the river’s denizens—one that the presence of man-made dams often eliminates by trapping it permanently behind the barrier.

Wrangling Sediment

SEDIMENTS STUCK behind dams are proving crucial variables when dams are

THE AUTHOR

JANE C. MARKS owes her career as an ecologist to algae. After completing her undergraduate degree in English at the University of Michigan at Ann Arbor, Marks became fascinated with aquatic plants and earned an M.S. in biology from Bowling Green State University and then a doctorate from the University of California, Berkeley. In 1995 she began working for the U.S. Agency for International Development, advising the organization about conservation and resource management issues all over the world. In 1999 she joined the faculty at Northern Arizona University. Her work on Fossil Creek is being featured in a new documentary, *A River Reborn: The Restoration of Fossil Creek* (see www.riverreborn.org).

taken down. Often the biggest issue facing managers is how to contend with what can be a massive accumulation of dirt and debris. Because of the legacy of releasing PCBs downstream in the Hudson River, scientists now routinely test these materials for toxicity. If the sediments contain high levels of pollutants, the cost of removing them—especially from remote locations—has to be weighed against the ability of the waterway to wash them away. If the sediment load is very high and the river's flushing capacity low, engineers might opt to remove the dam in stages, allowing small amounts of sediment to be released at a time. Sometimes engineers build channels through reservoirs, planting vegetation to stabilize sediments or placing physical barriers such as rocks or temporary fencing to hold the dirt in place.

In Fossil Creek, where roughly 25,000 cubic yards of sediment are trapped behind the dam, geologists and others predicted that the river would naturally flush the sediments downstream within a decade, without any ad-

verse effects. So the project did not have to weigh the cost and negative environmental impacts of transporting heavy machinery into a wilderness area.

Sediments pose a much bigger problem in many other places, however. Six million cubic yards of dirt lie behind the Matilija Dam on the Matilija Creek in southern California. (So much sediment, in fact, that the dam no longer serves to store water for irrigation or drinking.) At the same time, the downstream beaches are starved of sediment: they badly need dirt and sand to stave off ongoing erosion from wind and rain.

Matilija Dam is scheduled to be decommissioned in 2009, and managers have devised an elaborate sediment plan. They intend to transport fine sediments from behind the dam through a slurry pipe to sites five to 11 kilometers downstream. From there, the river will do the work by redistributing these materials during flood events to form beaches and sandbars. The larger, or coarse-grained, sediments that have accrued upstream of the dam will be left in place, but en-

gineers will regrade the river channel there into a more naturally sinuous one, which will better protect against flooding by allowing sediments to settle and rebuild the banks.

Going Forward

AT FOSSIL CREEK and elsewhere, managers and scientists are using all available information about dam removal and restoration ecology, as well as what they know of the entire watershed, to make decisions. But many gaps in our knowledge about ecosystems remain, and those working on decommissioning dams recognize they are conducting long-term experiments that may have unanticipated results. Fossil Creek, for example, was the first such project in which exotic fish were removed. If successful, this strategy could become routine, especially in smaller streams where chemical treatment is feasible.

At Fossil Creek our research team will now document how the river recovers. Among many unanswered questions we hope to focus on in the next five to 10 years are: Will native fish prosper without intervention? Will exotic fish come back? One interesting but problematic twist in the Fossil Creek story is that the chemical used to eliminate the exotic fish does not harm exotic crayfish, which are notorious for wreaking havoc on the food chain. The exotic fish had consumed crayfish, thereby keeping the crustacean's population down. Perhaps we will have exchanged one adverse situation for



NATIVE FISH return to Fossil Creek in buckets (above), after having been airlifted out and placed in holding tanks. Biologists treated the river with fish poison to get rid of the exotic species before returning the natives. Nonnative species were also a problem after engineers removed the Oak Street Dam on the Baraboo River in Wisconsin. Eurasian reed canary grass dominated the riverbanks, even though managers had planted native species (right).



CARRI J. LEROY (left); FROM "TRADING OFF: THE ECOLOGICAL EFFECTS OF DAM REMOVAL," BY E. H. AND M. W. DOYLE, IN *FRONTIERS IN ECOLOGY AND THE ENVIRONMENT*, VOL. 1, PAGES 15-22, 2003; COURTESY OF CALVIN HUYCK ORR (right)

another. In addition, as Fossil Creek rebounds, so do the numbers of visitors to it. With more hiking trails in place along the river, managers now need to devise rules that can allow people access but also protect the fragile ecosystem.

To supplement the in situ experiments such as the one at Fossil Creek, researchers are using computer simulations and are conducting indoor studies. The National Center for Earth-surface Dynamics in Minnesota has created a model ecosystem of miniature streams, dams and reservoirs. Investigators there use time-lapse photography to determine how sediments move downstream as dams are removed in different ways and to different extents.

Many engineers who were once dedicated to building dams now find themselves instead working on decommissioning them. U.S. government agencies such as the Bureau of Reclamation and the Army Corps of Engineers, as well as their European counterparts, are studying not only how to remove dams but also how to provide the benefits of the structures without their injurious effects—for instance, how to extract water from rivers without building blockades. In response to a 2000 report by the World Commission on Dams, engineers are also trying to incorporate decommissioning into the original designs of future dams.

Societies will continue to balance the pros and cons of dams, weighing their utility and benefits against their destructive costs. And scientists must continue to learn about how best to remove dams so natural ecosystems and human communities both can thrive. In the next few years the decommissioning of several large dams will provide further important knowledge. In 2009 two dams will be removed from Washington State's Olympic National Park: the 210-foot-high Glines Canyon Dam and the 108-foot-high Elwha Dam. Scientists in both locations are now collecting baseline data about salmon and steelhead, as well as oxygen levels, insect populations and sediment loads. Japan's Arase Dam will come down in 2010 in response to a long campaign by citizen activists con-



RIVER MODEL at the National Center for Earth-surface Dynamics in Minnesota provides scientists with a way to study how sediments move. This research can help experts plan what to do with the dirt and other material that accumulates behind dams.

cerned about poor water quality and a decline in fisheries. Australia will transform the 19,500-acre Lake Mokoan into a wetlands again when its dam is removed, while France contemplates the fall of a fifth Loire Valley dam.

In most cases, controversy about decommissioning arises—and sometimes the debate is unexpected. In the Loire Valley, a father and son ended up on different sides of the divide. The father remembered the wild rivers and the salmon runs; the son had grown up swimming and boating in the reservoir. In the case of Fossil Creek, the local community wanted to preserve components of the generating station, the Childs-Irving facility. Built by one of the few female engineers of that era, Iva Tutt, and main-

tained by generations of engineers who lived at the site with their families, the plant was culturally significant, and, accordingly, its preservation became part of the restoration plan.

The same proved true of the Wellington Dam in New South Wales, Australia. In 2002 the State Water Corporation ensured that a one-meter-high footprint of the structure remained (minus one gap for flow) across Bushrangers Creek so the public could still appreciate the dam that was built in 1898. With compromises such as these, along with further ecological insights and more flexible engineering, it seems possible to think of the world's waterways as ultimately fulfilling their promise for all parties—from plants to people. SA

MORE TO EXPLORE

Undamming Rivers: A Review of the Ecological Impacts of Dam Removal. A. T. Bednarek in *Environmental Management*, Vol. 27, No. 6, pages 803–814; June 2001.

Dam Removal: Challenges and Opportunities for Ecological Research and River Restoration. D. D. Hart et al. in *BioScience*, Vol. 52, No. 8, pages 669–681; August 1, 2002.

Dam Removal Research: Status and Prospects. Heinz Center, 2003. Available at www.heinzcenter.org/NEW_WEB/PDF/Dam_Research_Full_Report.pdf

Trading Off: The Ecological Effects of Dam Removal. Emily H. Stanley and Martin W. Doyle in *Frontiers in Ecology and the Environment*, Vol. 1, No. 1, pages 15–22; 2003.

Interactions between Geomorphology and Ecosystem Processes in Travertine Streams: Implications for Decommissioning a Dam on Fossil Creek, Arizona. Jane C. Marks et al. in *Geomorphology*, Vol. 77, Nos. 3–4, pages 299–307; July 2006.

American Rivers: www.americanrivers.org

NOVEL CRYSTAL BALL: One day Y-shaped molecules called autoantibodies in a patient's blood may tell doctors whether a patient is "brewing" certain diseases and may even indicate roughly how soon the individual will begin to feel symptoms.

New PREDICTORS of Disease

Molecules called predictive autoantibodies appear in the blood years before people show symptoms of various disorders. Tests that detected these molecules could warn of the need to take preventive action

By Abner Louis Notkins

A middle-aged woman—call her Anne—was taken aback when one day her right hand refused to hold a pen. A few weeks later her right foot began to drag reluctantly behind her left. After her symptoms worsened over months, she consulted a neurologist. Anne, it turned out, was suffering from multiple sclerosis, a potentially disabling type of autoimmune disease. The immune system normally jumps into action in response to bacteria and viruses, deploying antibodies, other molecules and various white blood cells to recognize and destroy trespassers. But in autoimmune disorders, components of the body's immune system target one or more of the person's own tissues. In Anne's case, her defensive system had begun to turn against her nerves, eroding her ability to move.

Every story of autoimmune disease is sad—but collectively the impact of these illnesses is staggering. More than 40 autoimmune conditions have been identified, including such common examples as type 1 (insulin-dependent) diabetes, rheumatoid arthritis and celiac disease. Together they constitute the third leading cause of sickness and death after heart disease and cancer. And they afflict between 5 and 8 percent

of the U.S. population, racking up an annual medical bill in the tens of billions of dollars.

Recent findings offer a way to brighten this gloomy picture. In the past 10 years a growing number of studies have revealed that the body makes certain antibodies directed against itself—otherwise known as autoantibodies—years, and sometimes a decade, before autoimmunity causes clinical disease, damaging tissues so much that people begin showing symptoms. This profound insight is changing the way that doctors and researchers think about autoimmune conditions and how long they take to arise. It suggests that physicians might one day screen a healthy person's blood for certain autoantibodies and foretell whether a specific disease is likely to develop years down the line. Armed with such predictions, patients could start fighting the ailment with drugs or other available interventions, thereby preventing or delaying symptoms.

Those interventions may not be easy to find; most likely, preventive therapy would have to be tailored specifically for each condition. In certain disorders, such as myasthenia gravis, autoantibodies participate in the disease process, and so blocking the activity of the particular autoanti-

CARY WOLINSKY (photographs); JEN CHRISTIANSEN (photo/illustration)





COPYRIGHT 2007 SCIENTIFIC AMERICAN, INC.

bodies at fault could be therapeutic. Autoantibodies that presage certain other conditions, though, probably are more siren than fire, announcing brewing disease actually caused by other components of the immune system, such as cells known as T lymphocytes and macrophages. In those cases, preventive treatments would have to target the offending cells.

The revolution in predictive medicine and preventive care will take time and effort to effect. Many autoantibodies have been uncovered, but only a few large-scale trials have been conducted to evaluate how accurately they can predict disease. If inexpensive, quick tests for predictive autoantibodies can be developed, though, they could become as standard a part of routine checkups as cholesterol monitoring.

Early Insight from Diabetes

PEOPLE FAMILIAR with advances in genetics might wonder why researchers would want to develop tests for predictive autoantibodies when doctors might soon be able to scan a person's genes for those that put the individual at risk of various disorders. The answer is that most chronic diseases arise from a complex interplay between environmental influences and multiple genes (each of which makes but a small contribution to a disease). So detection of susceptibility genes would not necessarily reveal with any certainty whether or when an individual will come down with a particular autoimmune condition. In contrast, detection of specific autoantibodies would signal that a disease-causing process was already under way. Eventually, genetic screening for those with



Potential Uses

Predictive autoantibodies could:

PREDICT RISK

Project one's probability of coming down with a particular disease so the person can consider preventive therapy

ANTICIPATE TIMING

Reveal how soon a disease process is likely to cause symptoms

PROJECT COURSE OF DISEASE

Predict the severity and probable rate of progression of a disorder

SIMPLIFY HUMAN TRIALS

Make trials of preventives for rare disorders feasible by identifying subjects at high risk (thus avoiding the need for prohibitively large numbers of volunteers from the general population)

an inherited predisposition to a disease may help reveal those who need early autoantibody screening.

Studies of patients with type 1 diabetes provided the first clues that autoantibodies could be valuable for predicting later illness. In this condition, which typically arises in children or teenagers, the immune system ambushes the beta cells in the pancreas. These cells are the manufacturers of insulin, a hormone that enables cells to take up vital glucose from the blood for energy. When the body lacks insulin, cells starve and blood glucose levels soar, potentially leading to blindness, kidney failure, and a host of other complications.

Forty years ago type 1 diabetes was not yet recognized as an autoimmune disease, and no one knew what caused the beta cells to die. But in the 1970s Willy Gepts of Vrije University of Brussels in Belgium examined the pancreases of children who had died of the disease and found that the islets of Langerhans, where the beta cells reside, had been infiltrated by lymphocytes—a sign of probable autoimmune activity. Soon thereafter Franco Bottazzo of Middlesex Hospital Medical School in London

established that blood from patients with type 1 diabetes reacted to islets but that blood of nondiabetics did not, which suggested that autoantibodies targeted to the diabetics' own beta cells were circulating in the patients' blood. This finding set off a hunt for the autoantigens—the specific molecular targets of the autoantibodies—in the beta cells, because researchers hoped that discovery of the autoantigens would clarify how diabetes arises.

Intensive research over the past 20 years has uncovered three major pancreatic autoantigens produced in people with newly diagnosed type 1 diabetes: insulin itself, an enzyme called glutamic acid decarboxylase (GAD) and a protein known as islet antigen-2 (IA-2), which was discovered by my group at the National Institutes of Health and is a component of the tiny sacs that ferry insulin around in beta cells. Experts still do not know whether the autoantibodies that bind these proteins play a part in killing beta cells. But they do know, based on highly sensitive detection tests, that one or more are present at diagnosis in some 70 to 90 percent of patients with type 1 diabetes. Today research laboratories use these tests to diagnose type 1 diabetes and distinguish it from type 2 diabetes, which usually arises in overweight adults and does not stem from autoimmunity. (Surprisingly, such tests have uncovered autoantibodies in about 5 percent of patients otherwise diagnosed with type 2 diabetes, which suggests that those individuals have been misclassified or

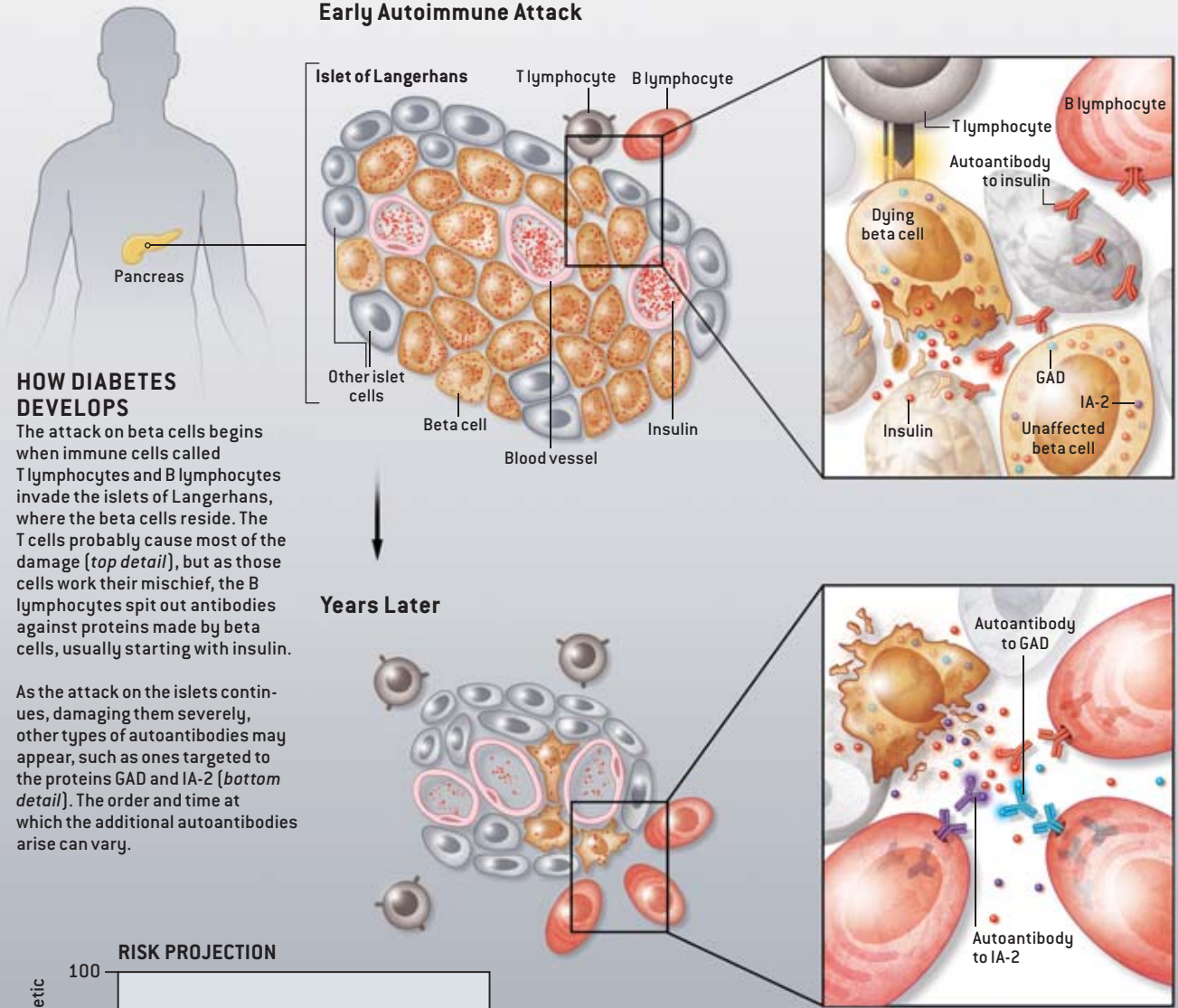
Overview/*Predictive Antibodies*

- In autoimmune diseases, such as type 1 diabetes, the immune system mistakenly manufactures antibodies that target the body's tissues.
- Certain of these "autoantibodies" appear many years before overt symptoms of disease, suggesting that screening for these molecules could be used to predict who is at risk of falling ill.
- Autoantibodies might also serve as guides to disease severity and progression and might even warn of risk for some nonimmune disorders.
- Screening for predictive autoantibodies could one day become routine, although a dearth of preventive treatments currently stands in the way.

LESSONS FROM DIABETES

The idea that autoantibodies (immune system molecules that mistakenly react to the body's own tissues) might serve as early warning signs of later disease came from research into how type 1 diabetes arises. The work revealed that this form of diabetes stems from an autoimmune attack on the beta cells of

the pancreas, which are the insulin producers (*diagrams*), and that autoantibodies targeted to substances made by beta cells appear years before symptoms do (*graph*). Symptoms arise when too few beta cells are left to meet the body's insulin requirements.



HOW DIABETES DEVELOPS

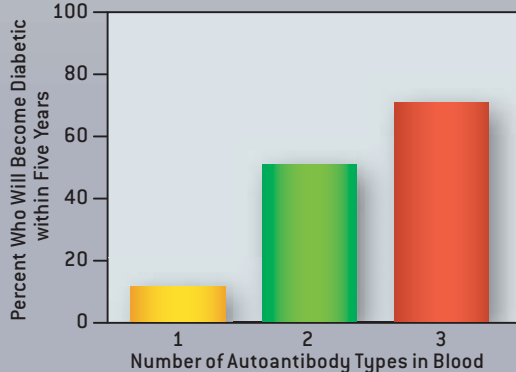
The attack on beta cells begins when immune cells called T lymphocytes and B lymphocytes invade the islets of Langerhans, where the beta cells reside. The T cells probably cause most of the damage (*top detail*), but as those cells work their mischief, the B lymphocytes spit out antibodies against proteins made by beta cells, usually starting with insulin.

As the attack on the islets continues, damaging them severely, other types of autoantibodies may appear, such as ones targeted to the proteins GAD and IA-2 (*bottom detail*). The order and time at which the additional autoantibodies arise can vary.

Early Autoimmune Attack

Years Later

RISK PROJECTION



AUTOANTIBODIES AND DIABETES RISK

Whether autoantibodies to insulin, GAD and IA-2 contribute to the beta cell killing is not known, but studies have shown that the molecules can signal greatly enhanced risk for diabetes. Risk increases with the number of diabetes-related autoantibody types in the blood.

TAMI TOLPA (Illustrations); JOHNNY JOHNSON, SOURCE: ABNER LOUIS NOTKINS (graph)

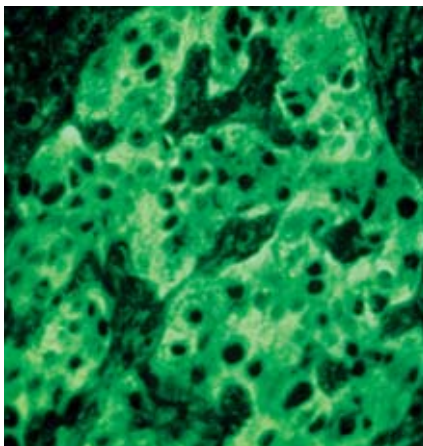
have a combination of type 1 and type 2 diabetes.)

Interest in the three autoantibodies escalated with the discovery that they appear long before the onset of diabetic symptoms. In studies conducted by various laboratories, investigators took blood samples from thousands of healthy schoolchildren and then monitored the youngsters' health for up to 10 years. When a child came down with type 1 diabetes, the researchers pulled the individual's blood sample out of storage to see whether it contained autoantibodies. The vast majority of children destined to become diabetic had one or more of the three signature diabetes-related autoantibodies in their blood as long as 10 years before any recognizable symptoms arose.

Before this work, some experts thought that type 1 diabetes developed suddenly, perhaps within a matter of weeks. The new data demonstrated, instead, that in most cases the immune system silently assails the pancreas for years until so many beta cells die that the organ can no longer make enough insulin for the body's needs. That is the point when the classic early symptoms of diabetes arise, such as excessive hunger, thirst and urination.

More important, these studies also raised the prospect that doctors might forecast whether a child is at risk for type 1 diabetes by testing blood for the presence of these autoantibodies. Clinical researchers found that an individual with one autoantibody has a 10 percent risk of showing symptoms within five years. With two autoantibodies, the chance of disease jumps to 50 percent; with three autoantibodies, the threat rockets to between 60 and 80 percent.

The ability to predict whether a person is likely to fall ill with type 1 diabetes has had major repercussions for medical researchers trying to better understand and prevent the disease. Before the discovery of predictive autoantibodies, for example, it was almost impossible to conduct clinical trials of new preventive therapies, because the disorder is relatively rare, affecting about one individual in 400. Such odds meant that more than 40,000 subjects would have to be entered into a trial in order to assess the effects of an intervention on the



AUTOANTIBODIES are revealed by green fluorescence in a micrograph of a human islet exposed to blood from a patient with type 1 diabetes. The surrounding area is dark because it lacks islets.

100 who would eventually be affected.

Now scientists can select for study those people whose blood shows two or more of the diabetes-related autoantibodies, because at least half the subjects, if untreated, will most likely come down with the disease within five years. Slashing the number of subjects who must be enrolled in a prevention trial has made such experiments feasible for the first time. In one investigation, doctors identified several thousand individuals at high risk of diabetes and tested whether injections of insulin could avert the disease. Sadly, this treatment proved unsuccessful; efforts to find useful interventions continue.

The discovery that autoantibodies frequently herald the onset of type 1 diabetes prompted scientists to examine whether the same might be true in other autoimmune diseases. One that has been the focus of especially intense research is rheumatoid arthritis, a debilitating condition that is highly prevalent, afflicting about 1 percent of the world's population. In those affected, the immune system attacks and destroys the lining of the joints, causing swelling, chronic pain and eventual loss of movement.

Predicting Other Diseases

IMMUNOLOGISTS HAVE RECENTLY unearthed an autoantibody that is present in 30 to 70 percent of patients diagnosed with rheumatoid arthritis. The antibody latches onto citrulline (a modified version of the amino acid arginine), which is present in certain proteins. Studies have now revealed that the autoantibody appears in the bloodstream before the first symptoms turn up, in some cases more than 10 years before. Further, the likelihood that the illness will develop is as much as 15 times greater in people carrying that antibody than in those who lack it.

The knowledge that the anticitrulline autoantibody might serve as a predictive marker is particularly exciting because, in contrast to the situation in type 1 diabetes, doctors already have medicines that might be delivered to prevent or slow the onset of arthritis. Rheumatologists know that quickly and aggressively treating newly diagnosed patients with certain drugs, such as ones that combat inflammation, can retard or sometimes stop the devastating loss of joint flexibility. It is not unreasonable to think therefore that earlier intervention might be even more protective. The hope now is that doctors will be able to screen the general population, or those with a family history of the condition, and then start treating those who make anticitrulline antibodies before autoimmunity irrevocably harms their tissues. First, however, further clinical trials must be carried out to confirm that these autoantibodies accurately predict the onset of joint symptoms. In addition, a reasonably priced test suitable for screening will have to be

ABNER LOUIS NOTKINS is chief of the Experimental Medicine Section of the National Institute of Dental and Craniofacial Research (NIDCR) at the National Institutes of Health (NIH). He received his B.A. from Yale University and his M.D. from New York University and trained in internal medicine at Johns Hopkins Hospital. Notkins has spent his entire professional career as a research physician at the NIH, where he was scientific director of the NIDCR before taking up his current post. He has done extensive research in virology, immunology and diabetes. In 1986 he received the Paul Ehrlich Prize. This is his third article for *Scientific American*.

introduced, along with protocols for deciding exactly who should be tested, when and how often.

For certain other autoimmune disorders, the detection of predictive autoantibodies could potentially enable people to shut down autoimmune activity by avoiding certain triggers in their environment. A case in point is celiac disease. In people with this condition, the gluten protein found in wheat, rye and barley incites the immune system to attack the lining of the small intestine, which then fails to absorb food properly; diarrhea, weight loss and malnourishment then ensue. Patients must eliminate gluten from their diet, bypassing most bread, pasta and cereal for the rest of their lives.

Investigations into the underpinnings of celiac disease have revealed that many patients make an autoantibody that reacts with tissue transglutaminase, an enzyme that modifies many newly made proteins. This autoantibody emerges up to seven years before symptoms do, suggesting that high-risk individuals might forestall the disease entirely by eliminating gluten from their diet. This idea has not yet been tested, however.

More Uses for Autoantibody Tests

IMMUNOLOGISTS ARE EXPLORING whether autoantibodies can serve as early warnings in other ways as well. For instance, some autoantibodies might help doctors to gauge the rate at which an already diagnosed autoimmune condition is likely to progress or how severe it will become.

Patients with multiple sclerosis often start off with relatively mild symptoms that then disappear for a while. Some people continue in remission for a long time or have manageable recurrences. But others grapple with more frequent or severe symptoms, and a few enjoy no remissions at all. Doctors struggle to discern which individuals with early symptoms will go on to

suffer from the harshest effects, so that they can counsel the patients accordingly. In 2003 a study of more than 100 individuals with newly identified multiple sclerosis revealed that those who made autoantibodies directed against two proteins that insulate nerve cells were almost four times more likely to suffer a relapse after the initial symptoms abated than were those without the autoantibodies. In addition, the antibody-positive patients relapsed more quickly than the others. These results suggest that testing for these autoantibodies could offer a quick way to predict whether, and how rapidly, multiple sclerosis will advance, although further study is needed before such testing can be put into practice and used to guide therapy.

In the past few years, researchers have made the intriguing finding that autoantibodies can also appear in people with certain disorders not typically thought of as autoimmune conditions, such as some cancers. These autoantibodies probably do not control tumor growth, but laboratories around the world are examining whether they can be useful for the early detection of cancer. In other conditions, such as atherosclerosis, investigators are looking into the possibility that autoantibodies might show which patients are more prone to a blockage in the arteries to the brain and therefore to stroke.

Scientific Challenges

SO FAR MUCH OF THE WORK I have discussed has been confined to a small number of academic laboratories and to a few of the major autoimmune diseases. Investigators and companies, however, are now beginning to recognize the potential value of these proteins for improving patient care. They are trying to extend the findings and unearth predictive autoantibodies linked to other autoimmune disorders.

This task is challenging, however, in part because research-

Disorders under Study

Investigators have found autoantibodies that might serve as predictors of risk or of progression for a number of autoimmune conditions beyond diabetes, including those listed below.

DISORDER	STATUS OF RESEARCH
Addison's disease (a disorder of the adrenal glands; results in low blood pressure, weakness and weight loss)	Autoantibodies to adrenal tissue and the enzyme 21-hydroxylase are highly predictive in children
Antiphospholipid syndrome (marked by recurrent clots in blood vessels and pregnancy loss)	Autoantibodies to various molecules appear to signal risk for complications of the disorder
Celiac disease (a digestive disorder triggered by gluten in foods)	Predictive autoantibodies have been identified that target an enzyme called tissue transglutaminase
Multiple sclerosis (neurological condition causing loss of movement)	Autoantibodies to proteins in the myelin sheath that insulates nerve cells appear to predict risk of relapse
Rheumatoid arthritis (chronic joint inflammation)	Autoantibodies to citrulline, a component of many modified proteins, have been found to appear as many as 10 years before symptoms occur
Systemic lupus erythematosus (can affect many organs, including the joints, kidneys and skin)	Several related autoantibodies have been found; one or more of these appear in up to 80 percent of patients before symptoms arise

ers will have to follow large populations for years to prove that particular autoantibodies can signal future disease. That is, many thousands of healthy people must be recruited to give blood samples and then tracked carefully for 10 years or more to see if they fall sick. Aside from posing logistical difficulties, these prospective studies can cost tens of millions of dollars.

An alternative to conducting prospective studies from scratch might be to tap into existing health databases and carry out retrospective studies. For example, blood samples and medical information have already been collected over many years from members of the U.S. military and from subjects in the Women’s Health Initiative, a vast, ongoing study of more than 100,000 women. Experts in autoimmunity could team up with investigators in these and other projects, identify individuals who have been diagnosed with an autoimmune disease and then examine their stored blood for the presence of predictive autoantibodies. This approach would be relatively inexpensive and could yield rapid results—and a few researchers have already embarked on such collaborations.

A second avenue of attack would involve identifying heretofore unrecognized autoantigens. One could search human genetic databases for the sequences that encode proteins and use this information to manufacture these proteins in the lab.

Scientists could pinpoint those that are autoantigens by mixing each of the manufactured proteins with blood from patients who have an autoimmune disease and allowing complexes of proteins and antibodies to form. Analyses of such complexes could identify both the autoantigens in the collection and the autoantibodies that recognize them. That done, the predictive value of the autoantibodies could be determined in a prospective or retrospective study.

This full-genome approach to isolating autoantigens is difficult. Nevertheless, a handful of research groups are now screening smaller batches of proteins in this way. In my laboratory, for example, we are hunting for new autoantigens involved in type 1 diabetes by manufacturing dozens of selected pancreatic proteins known to be involved in the secretion of insulin and testing whether autoantibodies in blood from diabetics bind these proteins.

Practical Challenges

MEDICINE AS WE KNOW IT is evolving from diagnosing and treating diseases after they develop to predicting and preventing them. Ten or 20 years from now autoantibody screening for at least some diseases will almost certainly become a familiar part of the standard medical examination.

CHECKUPS OF THE FUTURE

Someday physicals could routinely include screening for autoantibodies.

1 A blood sample would be given to a laboratory, which would extract the plasma (the acellular part).

2 The plasma would be washed over a chip containing an array of autoantigens—molecules known to be capable of eliciting autoimmune reactions—at known positions. Any autoantibodies in the blood would bind to the autoantigens, triggering signals indicating the identity and quantity of the bound autoantibodies.

3 This information would be translated into a prediction of the patient’s risk for becoming afflicted with specific conditions.

Disorder	Five-year risk
Type 1 diabetes	70%
Rheumatoid arthritis	No sign
Multiple sclerosis	No sign
Lupus	No sign
Thyroid disease	15%

(SCROLL FOR MORE)

In the future, patients visiting their doctors for a physical might have their blood tested for multiple predictive autoantibodies in a single test. In one plausible scenario, the doctor would send a blood sample to a lab for an autoantibody analysis, along with standard tests for cholesterol, blood glucose and other health indicators. There a machine would pass the blood over a tiny chip displaying an array of known autoantigens. Autoantibodies in the blood that bound to one or more of these antigens would trigger pulses of light that would be picked up by a detector. Within hours, the doctor would receive a readout translating this information into a health forecast. The presence of predictive antibodies would not mean that a patient will definitely get sick, but would give a percentage risk of diabetes and numerous other conditions developing over some number of months or years.

These tests might even be combined with other biological assays to give more accurate health predictions. In the case of type 1 diabetes, possession of certain forms of genes that regulate the immune system, called HLA genes, are also known to correlate with disease risk. A prognostic assay might combine tests for those HLA variants with tests for predictive autoantibodies.

The vision of prediction is an enticing one, but even after the challenges of identifying predictive autoantibodies have been overcome, other issues will need to be resolved in preparation for their use in the clinic. One critical question relates to cost. At present, lab screening for predictive autoantibodies is cumbersome and labor-intensive. Widespread population screening for multiple autoantibodies will become practical only when rapid, inexpensive automated methods for detecting them are designed. To date, only a few small biotechnology companies are trying to devise such methods.

Another issue to be decided is which people should be screened and how often. It is unreasonable to test children for diseases that occur only in adults, and the converse would also hold true. Similarly, the frequency of screening would have to depend in part on whether autoantibodies tend to arise many years or just a few months before the onset of clinical symptoms.

Problems of Prediction

Before autoantibodies are widely used to foretell a patient's risk of future disease, many tough ethical and practical issues must be considered.



Should doctors test for diseases that have no preventive treatment or cure?



What is the best way to make sure patients understand that a positive test does not mean disease will definitely develop but indicates a given probability of risk?



How can the risks of false positive or negative tests be minimized so that few patients are unnecessarily alarmed or mistakenly reassured?



Is the cost of routine screening justified by the number of patients who would be found to be at risk and able to benefit from early treatment?



For autoimmune diseases that run in families, should family members of afflicted individuals be tested, and will the worry over a result indicating high risk be easier to live with than the anxiety of not knowing?



Will a positive test lead to discrimination from employers, health insurers or society in general?

By far the most important factor controlling whether predictive autoantibody screening will become widespread is the availability of therapy. Some would argue that embarking on predictive testing for diseases makes little sense if patients can be offered no preventive or ameliorating treatment. A large and intensive research effort is under way to develop new therapies for autoimmune diseases, but because the conditions are so complicated and varied, progress may not come quickly.

Of course, the ability to forecast someone's life and death raises thorny ethical issues. Some people may choose not to know that they are likely to come down with a given disease, and doctors must be careful to respect that decision. Patients may also be concerned that insurers or employers could obtain medical information

and use it to discriminate against them, even while they are healthy. As is true of genetic testing, such issues call for in-depth discussion.

Forecasts of the future have always intrigued and frightened people. Handled properly, though, such knowledge could benefit the millions of patients and doctors destined to battle autoimmune diseases. By making early intervention possible, predictive autoantibodies have the potential to alleviate much misery and to help provide extra years of healthy life. SA

MORE TO EXPLORE

Immunologic and Genetic Factors in Type 1 Diabetes. Abner Louis Notkins in *Journal of Biological Chemistry*, Vol. 277, No. 46, pages 43545–43548; November 15, 2002.

Antimyelin Antibodies as a Predictor of Clinically Definite Multiple Sclerosis after a First Demyelinating Event. Thomas Berger et al. in *New England Journal of Medicine*, Vol. 349, No. 2, pages 139–145; July 10, 2003.

Development of Autoantibodies before the Clinical Onset of Systemic Lupus Erythematosus. Melissa R. Arbuckle et al. in *New England Journal of Medicine*, Vol. 349, No. 16, pages 1526–1533; October 16, 2003.

Autoantibodies as Diagnostic and Predictive Markers of Autoimmune Diseases. Edited by Abner Louis Notkins, Ake Lernmark and David Leslie in *Autoimmunity*, Vol. 37, No. 4, pages 251–368; June 2004.

Diagnostic and Predictive Value of Anti-Cyclic Citrullinated Protein Antibodies in Rheumatoid Arthritis: A Systematic Literature Review. J. Avouac, L. Gossec and M. Dougados in *Annals of Rheumatic Diseases*, Vol. 65, No. 7, pages 845–851; July 2006.

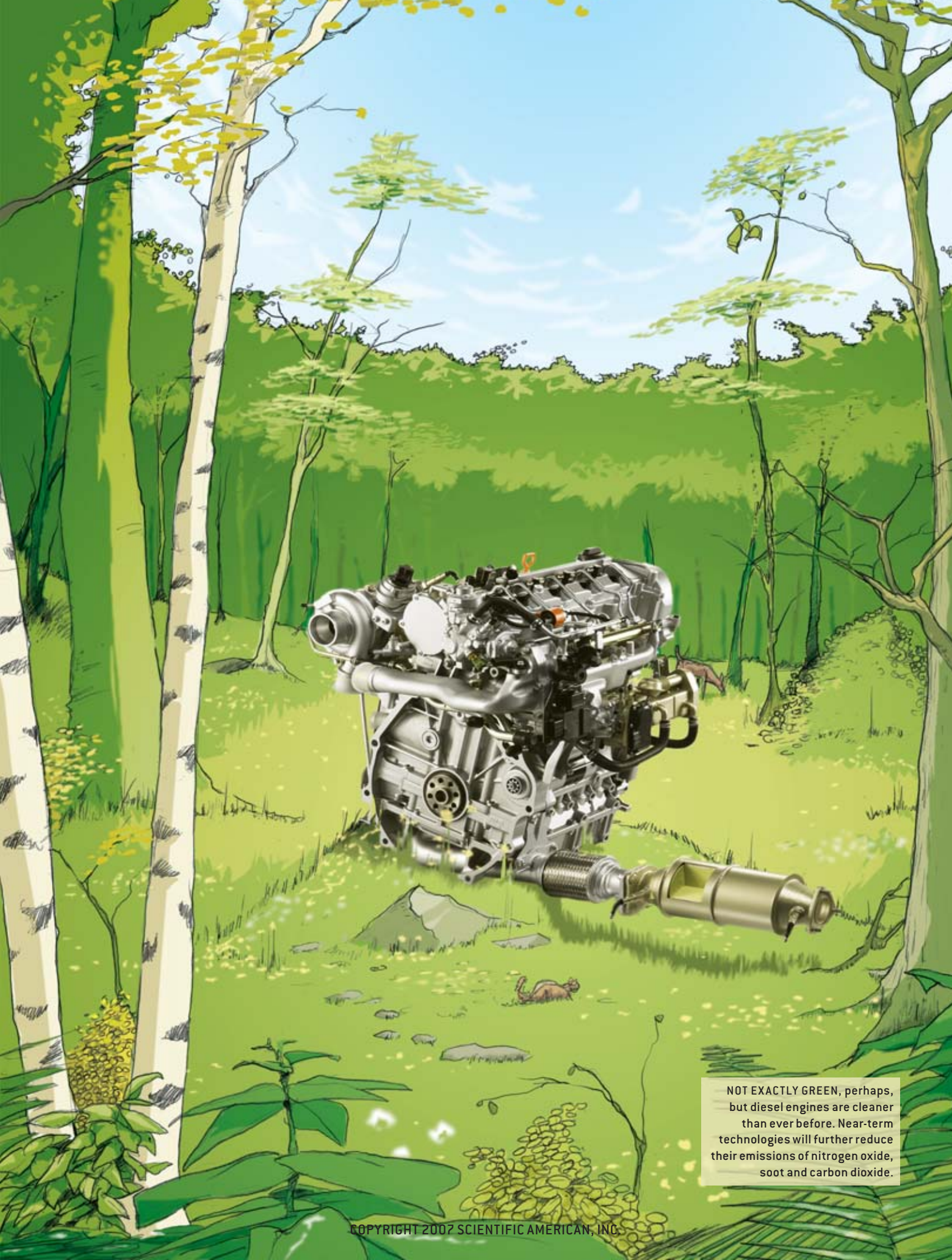
Diesels come Clean

Improved engines
and exhaust scrubbers,
combined with a new fuel, will
make energy-efficient diesels
nearly as green as hybrids

By Steven Ashley

Swinging his truck door open, the driver obligingly steps onto the cab seat and reaches for the roof. Extending himself upward, he slings a handkerchief over the exhaust stack of his late-model diesel rig. In mere moments, black fumes begrime a section of the white square with soot. “This good?” he asks, handing down the fluttering fabric. Nodding, I thank the man and retrieve the hankie. A short stroll away from his idling truck and its fellow 18-wheelers parked in this New Jersey Turnpike rest area sits their newborn brother, a Mercedes-Benz E320 Bluetec sedan. With a turn of the key, its diesel engine springs to life. Moments later, I kneel behind the car and cover its tailpipe with an unsoiled patch of cloth. It remains nearly spotless, even after a full minute.

As the so-called handkerchief test shows, the words “clean diesel” are no longer a contradiction in terms. Diesels have long been regarded as among the dirtiest of power plants, a reputation that lingers because so many decades-old examples of this durable technology still work the roads today. But the E320 is the vanguard of a new wave of diesel cars, SUVs and pickups that release far fewer air pollutants

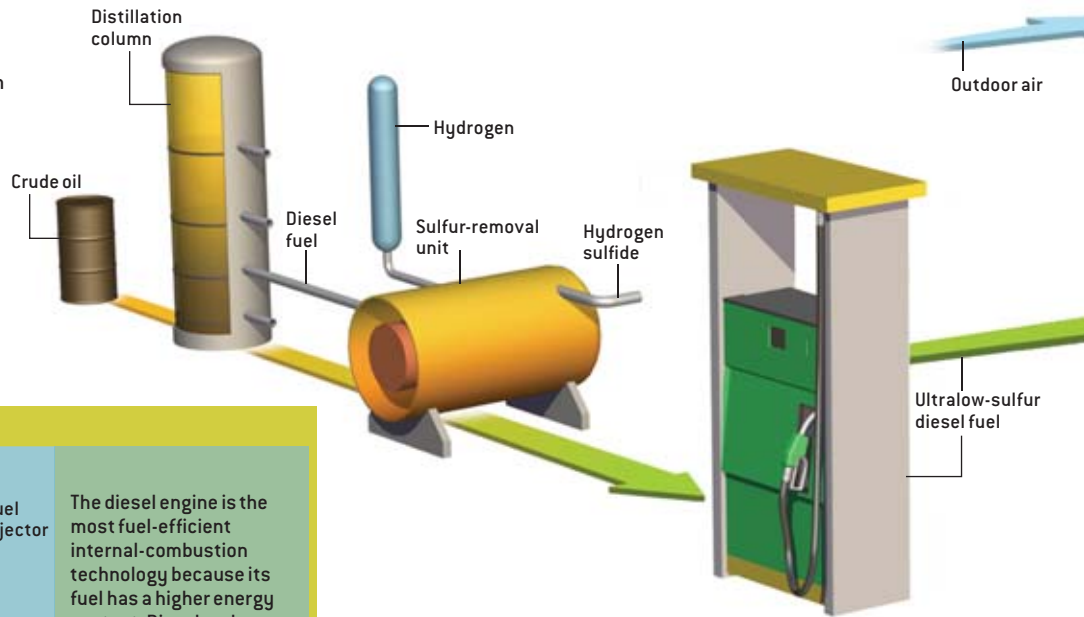


NOT EXACTLY GREEN, perhaps, but diesel engines are cleaner than ever before. Near-term technologies will further reduce their emissions of nitrogen oxide, soot and carbon dioxide.

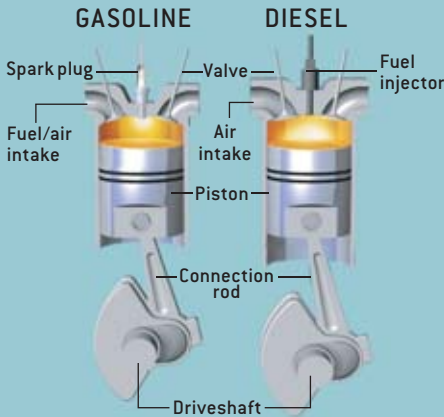
Elements of Clean Diesel Cars

ULTRALOW-SULFUR DIESEL FUEL

The fuel contains 97 percent less sulfur than previous diesel grades. Sulfur, a natural component of crude oil, would otherwise pollute the air and degrade the exhaust-treatment systems. Refiners add hydrogen to diesel distilled from crude, which bonds with the sulfur, allowing the resulting hydrogen sulfide to be separated out.



SPARK VS. COMPRESSION



The diesel engine is the most fuel-efficient internal-combustion technology because its fuel has a higher energy content. Diesels rely on compression rather than igniting fuel with spark plugs, as gasoline power plants do. The piston stroke in a diesel compresses the air so intensely that injected fuel explodes spontaneously.

without compromising the engine's traditionally excellent fuel economy. Powered by a 3.0-liter V-6 engine, the E320, for instance, gets 36 miles per gallon (combined) and can travel as much as 780 miles between fill-ups.

Overview/Greener Diesels

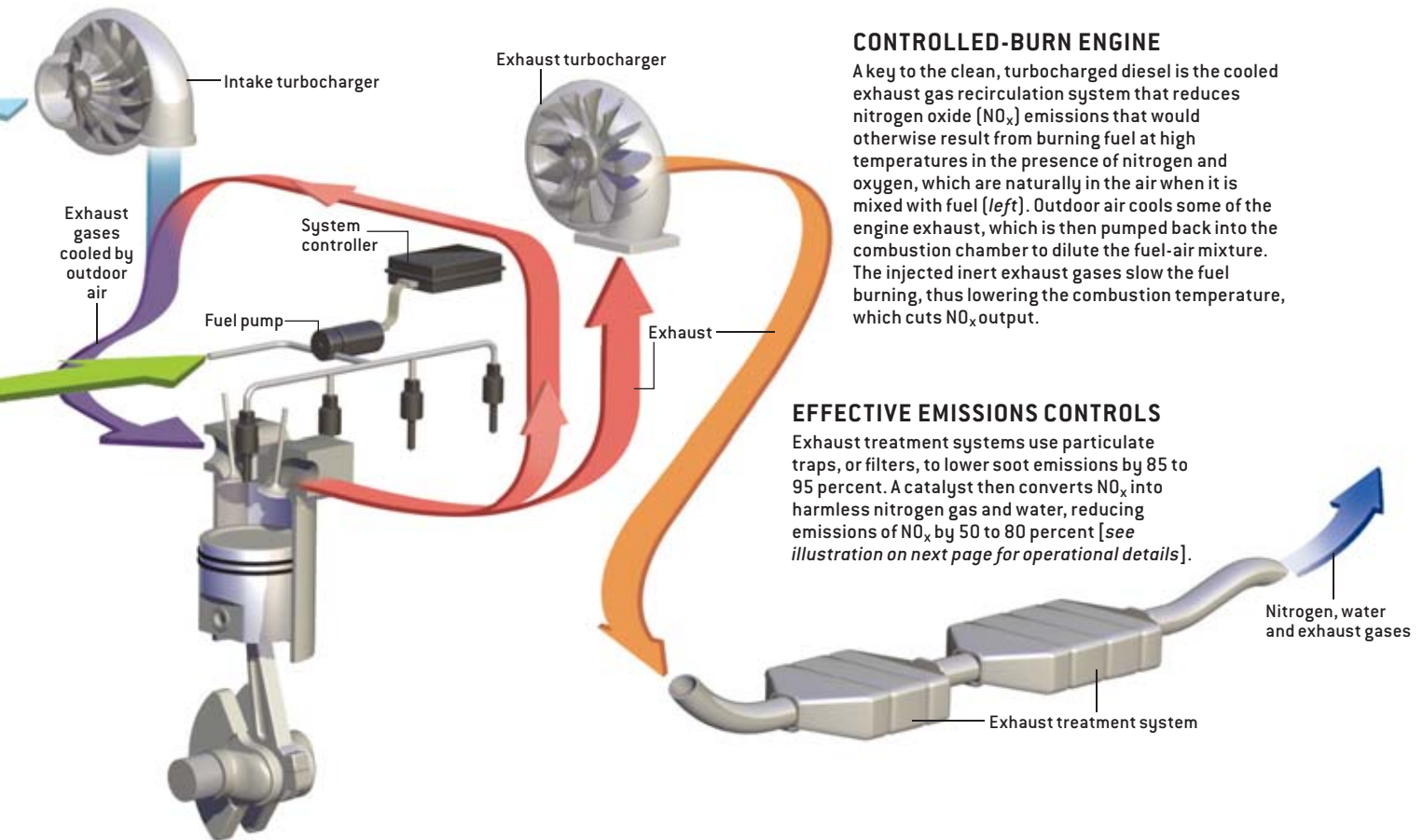
- Diesels have a bad reputation as dirty engines, but modern diesel automotive power plants put out significantly less pollution—principally carbon soot and nitrogen oxides—than their predecessors and get good gas mileage as well. Add cleaner fuel grades and new exhaust after-treatment systems, and diesel cars, sport utility vehicles and trucks can approach the environmental acceptability of hybrid vehicles.
- Ultralow-sulfur fuel, which recently became available in the U.S., is one key to the prospects of clean diesel technology. Sulfur otherwise would degrade the operation of exhaust-control systems over time.
- Tailpipe systems that scrub much of the soot and nitrogen oxide emissions from diesel exhaust are starting to appear in new models.

To bring about this transformation, modern diesels take advantage of advanced power plants that produce fewer pollutants, new exhaust systems that remove soot and convert tailpipe emissions into harmless gases, and cleaner low-sulfur fuels that just became available in North American markets this past fall. What is more, the improvements do not come with an exorbitant price tag. "The E320 Bluetec diesel costs only \$1,000 more than the gasoline version," says Thomas Weber, the board member in charge of research and development at DaimlerChrysler (manufacturer of the Mercedes).

Volkswagen, Audi, BMW, Honda, General Motors, Ford, PSA Peugeot Citroën and others are also planning to introduce cleaner diesels in the next few years. These vehicles will be almost as green as gasoline-electric hybrids but without much of the added cost and complexity of their extra drive systems. Automakers intend the new diesels, together with hybrids and advanced gasoline-powered vehicles, to bridge the gap between today's cars and the hydrogen fuel cell machines of the future [see "Hybrid Vehicles Gain Traction," by Joseph J. Romm and Andrew A. Frank; *SCIENTIFIC AMERICAN*, April 2006, and "On the Road to Fuel-Cell Cars," by Steven Ashley; *SCIENTIFIC AMERICAN*, March 2005].

The drive to more efficient, cleaner vehicles has grown in

GEORGE RETSECK



significance because cutting fossil-fuel consumption (and therefore carbon dioxide output) has become central to fending off global climate change as well as reducing our dependence on foreign oil. According to the Environmental Protection Agency (EPA), if diesels were to power one third of all light-duty vehicles (comprising cars, SUVs and small trucks) in the U.S., the shift would save about 1.4 million barrels of oil a day—equal to the daily shipments from Saudi Arabia (from which this country imports about 7 percent of its oil supply).

Compression Ignition

THE DIESEL ENGINE already has a long and storied past, having served as the backbone of modern heavy industry and transport since shortly after German engineer Rudolf Diesel invented it in the 1890s. Whereas the gasoline engine requires an electric spark to ignite a compressed mixture of fuel and air in a cylinder, a diesel relies on compression: the piston stroke squeezes the air in the cylinder more tightly, raising its temperature so that injected fuel explodes spontaneously. Today this so-called compression-ignition engine produces more energy per unit of fuel, and therefore 25 to 30 percent better mileage, than its gasoline-powered counterpart, explains Charles Freese, executive director of diesel engineering at GM

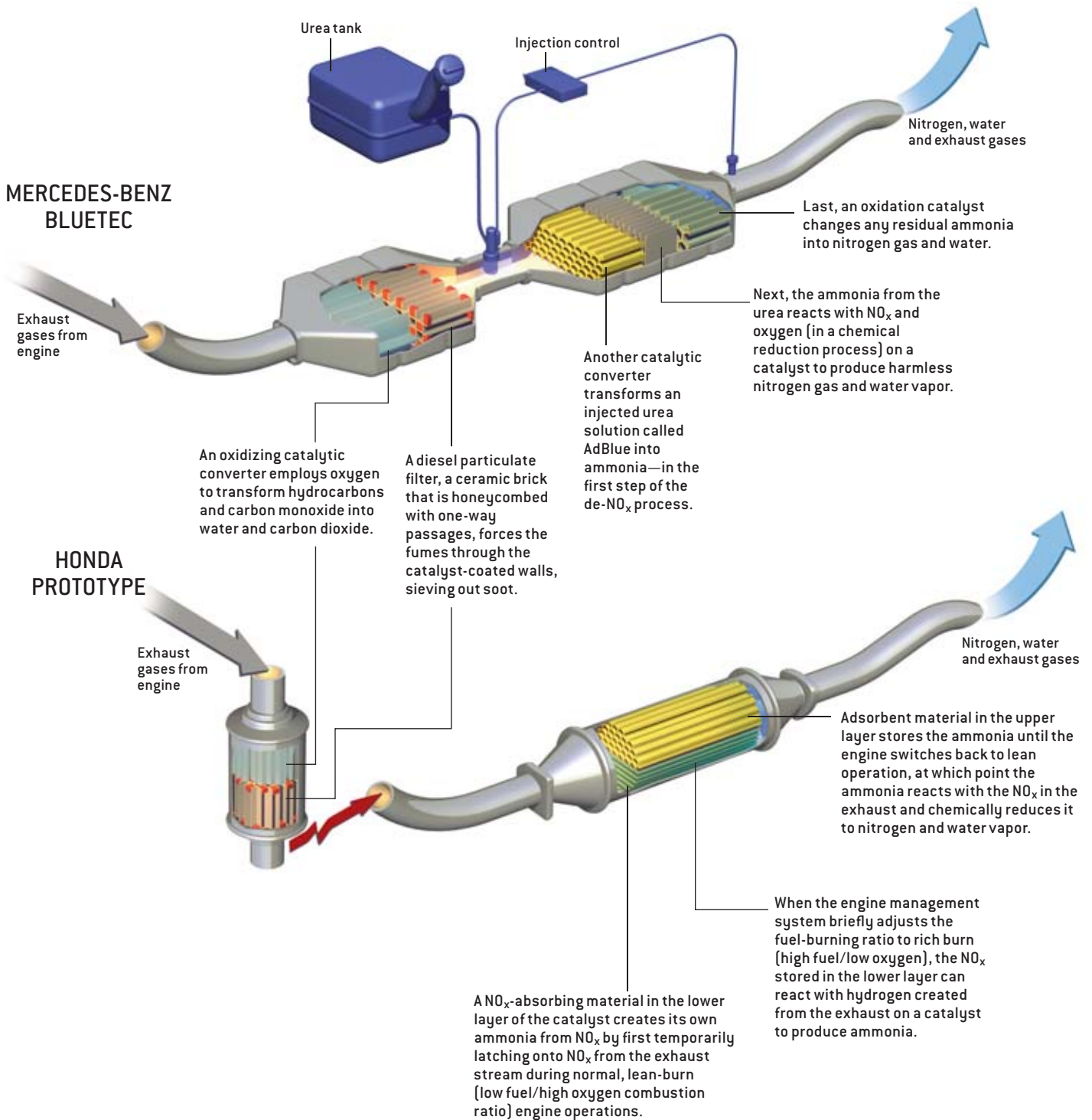
Powertrain. Compact diesel cars can get 40 to 50 miles to the gallon, for instance, whereas comparably sized gasoline-fueled vehicles rarely surpass 35 mpg.

The diesel provides “a lot of low-end torque [rotational force during starts], which, together with modern [power-boosting] turbochargers, gives diesels the good pickup and acceleration that makes them fun to drive,” Freese says. But the technology’s fuel characteristics and higher-temperature combustion conditions also yield more ultrafine carbon particulates (soot) and smog- and acid-rain-forming nitrogen oxide (NO_x) tailpipe pollutants than gasoline engines. Richard Kassel, director of the clean fuels and vehicles project at the Natural Resources Defense Council, estimates that diesel emissions from all sources—cars, trucks, locomotives and off-road equipment—annually cause as many as 25,000 premature deaths and 2.5 million lost workdays from cancer, asthma, bronchitis and other conditions in the U.S. Avoiding release of these aerial contaminants, he states, will save around \$150 billion a year in nationwide health costs by 2030.

Despite certain advantages, the diesel automobile has remained stuck in neutral in the U.S.—accounting for just 3 to 4 percent of sales, according to the J.D. Power market research firm. (Diesel’s worldwide share for cars and smaller

Exhaust Treatment Systems

Advanced diesel exhaust emissions-control systems from Mercedes-Benz and Honda can lower tailpipe output of hydrocarbons, carbon monoxide, nitrogen oxides (NO_x) and soot.



trucks is reportedly also relatively small: 13 percent are diesel-fueled, whereas 85 percent run on gasoline.) Diesels never sold well because of a poor reputation among U.S. buyers and unfavorable domestic environmental standards. Negative images conjured up by smoke-belching tractor-trailers and bad memories of American-made, diesel-powered lemons of the 1970s put most drivers off the engines. Meanwhile federal regulations placed harsher restrictions on the NO_x and soot emitted by diesels than on carbon monoxide and hydrocarbons, the principal air pollutants generated by gasoline engines. NO_x standards are particularly stringent in North

recirculation system in the early 1970s. This component dilutes the flame-supporting oxygen in the cylinder with injected (inert) exhaust, which slows combustion and therefore avoids NO_x-creating temperature spikes. Since its introduction, it has already cut raw NO_x production by about three quarters, according to industry claims. A relatively recent advance in this area is the cooled exhaust gas recirculation system, which lowers the temperature of exhaust fumes (by mixing them with outside air or cooling them in a heat exchanger) before they are pumped into the combustion cylinder, further reducing NO_x output.

THE WORDS “CLEAN DIESEL” ARE NO LONGER A CONTRADICTION IN TERMS.

America, where new rules require that diesels produce no more than one sixth the levels permitted in Europe.

The picture is quite different across the Atlantic, where somewhat lower-sulfur fuels have been available for several years because European refineries have incentives to produce more diesel than gasoline. In countries of that region, diesels make up more than 40 percent of the new cars purchased. European regulators, concerned more about fuel efficiency, have encouraged the use of diesels in recent decades with tax-based fuel-pricing strategies that make gasoline less attractive, and emissions standards that have deemphasized NO_x and soot output. For decades in Europe, these policies kept diesel cheaper than gasoline, although prices there have tended to level off in recent years—diesel currently costs about \$5 a gallon and gasoline is priced at around \$6, whereas in the U.S. they both go for \$2.25 to \$2.50. The tilt toward diesels in these countries has spurred auto engineers to develop smaller (2-liter) diesel engines that over the years have shed most of the technology's former undesirable characteristics, including loud engine noise and difficulty starting up in cold weather.

Engine and Fuel Upgrades

SO-CALLED CLEAN DIESELS get that way from engineers' efforts to improve the engines, fuels and exhaust systems. A series of engine innovations (most initially introduced in trucks) brought about the first category of improvements, Freese says. These upgrades include high-pressure fuel systems that reduce diesel injection time and boost cylinder compression, as well as optimized chamber and valve designs for good air-fuel mixing and hence uniform, stable and efficient combustion. Turbochargers—exhaust gas-driven compressors—boost power output by ramming air into the combustion chamber to raise the amount of oxygen available for burning. Meanwhile sophisticated control systems use sensors to monitor engine operations and adjust running parameters to fine-tune their workings.

Of particular importance to diesel emissions control was the development of an engine component called the exhaust gas

These improved technologies cut emissions production, but they were not enough by themselves to clean up diesel's act. Another key was the availability of an ultralow-sulfur diesel fuel that made their operation possible. Sulfur, which is naturally present in crude oil, not only corrodes engine components but “poisons,” or renders ineffective, catalytic converters because it bonds preferentially to catalytic compounds and clogs particulate traps, causing exhaust systems to back up. To extract sulfur, refineries process diesel with hydrogen, which bonds preferentially with the sulfur. The new federally mandated ultralow-sulfur fuel contains 97 percent less sulfur than prior grades; levels have dropped from 500 parts per million to 15, which is equivalent to just an ounce of sulfur in an entire tanker truck of diesel.

Cleansing the Tailpipe

THE LAST PIECES of the clean diesel puzzle are the superior exhaust-scrubbing techniques that are just now being introduced in cars. When exhaust emerges from a modern diesel engine, it typically enters the first stage of the emissions control system, the oxidizing catalytic converter, which strips it of a significant portion of its carbon monoxide and unburned hydrocarbon constituents by chemically binding some of these compounds to oxygen in the exhaust flow.

The gases then often run into a diesel particulate filter, or particle trap, which strains out their carbon soot. A trap usually contains a brick of heat-resistant ceramic (such as silicon carbide or cordierite) honeycombed with passages that are alternately blocked at opposite ends. Any exhaust that flows into a passage is forced to pass through substrate walls that are coated with a catalyst, which sieves out the ultrafine carbon particles. Such particle filters can reduce soot emissions by as much as 98 percent.

When clogged, some of these filters operate “something like a self-cleaning oven,” says GM's Freese. If sensors indicate a slowdown in exhaust flow, the engine controls raise the amount of fuel injected into the cylinder for a short time, which increas-



CLEANER DIESEL CAR, the 2007 Mercedes-Benz E320 sedan.

es hydrocarbon output. These chemicals get caught in the oxidation catalytic converter and then ignite, which elevates the exhaust temperature to around 650 degrees Celsius, enough to burn off the carbon and regenerate the filter's capabilities.

NO_x Removal

IN THE NEW CLEAN DIESEL vehicles, the exhaust next goes into a de-NO_x after-treatment system, which has been a main focus of recent engineering efforts in the diesel industry. Some of these components, called lean-NO_x traps or lean-NO_x catalysts, accomplish their functions during lean-engine operations. "Lean" refers to the concentration of fuel in the combustion mixture in the engine cylinder. An engine runs lean when fuel is injected sparingly, which means there is relatively more oxygen in the mix. "Rich" operation occurs when more fuel is burned. To maintain good fuel efficiency, engineers design diesels to run lean, but the consequent increased availability of oxygen promotes the formation of NO_x during fuel combustion.

Several basic approaches to de-NO_x systems exist, says Ben Knight, vice president of R&D for American Honda. One is known as continuous hydrocarbon selective catalytic reduction. This method involves running the diesel engine rich to place hydrocarbons in the exhaust stream, where they work with a catalyst to chemically reduce NO_x—that is, to ready the NO_x for conversion into nitrogen gas by adding an electron in a partial reaction. The method has cut NO_x output by as much as 40 percent during certain European test driving cycles but also adds a 5 percent fuel penalty, which is not considered cost-effective. In addition, it yields lower conversion efficiency at low temperatures.

Another de-NO_x technique is NO_x storage and conversion, which is currently used in vehicles with gasoline direct-injection engines. This so-called discontinuous technology stores NO_x on a barium-containing catalyst under normal lean-burn conditions; it then releases and chemically reduces the stored NO_x during the interval when the cylinder mixture is temporarily enriched with fuel, which leads to a dearth of oxygen in the exhaust. Applying this approach to diesel power plants requires major alterations in their design as well as in how they

are run. Like the hydrocarbon selective catalytic reduction technique, it exhibits reduced NO_x conversion efficiencies at low temperatures.

A third, more common method is urea-injection selective catalytic reduction. It employs a reducing agent, usually an ammonia-containing fluid called urea, to transform NO_x into nitrogen on a catalytic substrate. (Chemists formulate the usually strong-smelling additive to have a relatively benign aroma.) The principal drawbacks to this technique are that it uses a urea-injection system, needs a new commercial distribution infrastructure for the liquid and requires maintenance to keep the onboard reservoir filled.

The next generation of Bluetec technology (which will be used by DaimlerChrysler, Volkswagen and others) will rely on injecting a urea solution, which makers have dubbed AdBlue [see box on page 84]. Weber says that DaimlerChrysler engineers worked with researchers at Bosch to develop the system, which can cut NO_x emissions by as much as 80 percent. And when fitted with this next-generation exhaust after-treatment technology, the next version of the E320 has every prospect of satisfying the strict federal Tier 2, Bin 5 levels (which in part mandate 0.07 gram of NO_x and 0.01 gram of particulate per mile). These standards constitute more than a 10-fold reduction for soot and a twofold cut for NO_x over the previous regulations. Standards almost identical to Bin 5 levels are due to come into effect this year in California (whose huge auto market tends to set de facto national standards), whereas Bin 5 itself will be applied countrywide in 2009. Both are expected to continue to be enforced for the foreseeable future. The new E320 sedan, he notes, uses an average of 0.1 liter of the urea reducing agent per 100 kilometers traveled. The urea tank is to be sized large enough that it can span the standard oil change interval.

Solid-State Catalyst

HONDA RECENTLY announced an elegant new approach to de-NO_x technology, a lean-NO_x catalyst that demonstrates NO_x conversion rates of 90 percent yet requires no added reducing agents [see box on page 84]. This innovative technique converts some NO_x into ammonia and then recombines it with the remaining NO_x to make nitrogen gas. The prototype process, which was developed by a team led by chief engineer Hiroshi Ohno, is expected to first appear in a diesel-powered 2009 Honda Accord, according to news reports.

The Honda system centers around a compact dual-layer catalyst, Knight says. The top layer contains zeolites, microporous materials that act as high-surface-area, solid-state acid substrates that facilitate chemical reactions. The bottom layer incorporates two other common catalysts, cerium oxide and platinum.

"During normal lean engine operations," Knight explains, "the top layer simultaneously absorbs incoming NO_x and converts part of it into nitrogen. In the brief periods when the diesel is set to run rich, the bottom catalytic layer generates ammonia from the exhaust gases. But rather than passing this

ammonia directly back into the flow, it is stored in the upper layer's zeolite substrate until the engine flips back over to lean-burn, at which point the ammonia reduces the NO_x into nitrogen gas." The lower layer forms ammonia in two ways: it chemically combines ambient NO_x and hydrogen to yield the compound ammonia, and it also performs what chemists call a water-gas shift reaction with carbon monoxide and water vapor to form ammonia. Because the NO_x storage burden is lessened (compared with other methods), he says, there is less

produce gasoline rather than diesel, improved diesel fuel could require substantial capital investment in equipment.

In the meantime, despite the recent media buzz, industry observers expect that sulfur-free biodiesel (which can be derived domestically from crops such as soybeans) will remain a niche product. Says NRDC's Kassel: "biodiesel supplies will grow in the coming decades but will be a tiny piece of the entire diesel pie." He also cites "the need for national technical specifications for biodiesel grades" to ensure even performance

NEW DIESELS WILL HELP BRIDGE THE GAP TO FUTURE FUEL-CELL VEHICLES.

NO_x to trap during lean operation, which keeps the system compact and cuts down on rich operation time, thereby lowering the fuel penalty.

Knight emphasizes that the lean-NO_x catalyst's advanced control system is important because it determines how much ammonia is made. The controls realize this goal by monitoring and modifying the engine operation to set the durations of its lean and rich periods. "The controls are adaptive and optimize the system even if the catalysts degrade over time," Knight says—a key to enabling future Honda diesels to achieve EPA Tier 2, Bin 5 output over the required 125,000 miles of driving.

A Diesel Future?

EXACTLY WHERE this new chapter in the diesel story will lead remains uncertain. Although companies including DaimlerChrysler, Honda and VW are bullish on the diesel's prospects in the North American market, others are less sanguine about its future there, particularly in smaller vehicles. Both GM and Ford, for example, believe that light-duty diesel engines are better suited, at least initially, to pickup trucks and larger SUVs, where their extra torque output and high fuel efficiency can clearly shine. Toyota is even gloomier regarding domestic diesels. Says company spokesman John Hanson: "Toyota is not looking to promote diesels soon in North America. Although we continue to develop advanced diesels for the European and Asian markets, we don't think that the U.S. market is receptive to the technology and won't be for some time." Toyota thinks of the latest generation as "cleaner diesels," Hanson adds, but "none of the new models will rank high with regard to the EPA and California criteria for environmental friendliness... Even the cleanest among them will only barely qualify for use in California."

Questions about fuels also exist. Although ultralow-sulfur fuel is here, it typically has a lower and a wider range of cetane values than that available in Europe (cetane is the diesel equivalent of gasoline's octane). Better, more consistent fuel would help promote clean diesel technology. But because North American refineries are designed to

when engines are fueled from different sources. Elsewhere, European diesel manufacturers and some energy companies say they are interested in recent efforts to produce sulfur-free diesel fuel from suitable coal and natural gas grades using variants of Fischer-Tropsch chemistry, a fuel-synthesis method first developed during World War II.

One issue that is clear is that most of today's diesel engines will stay on the road for another two and a half decades or more. That reality means that the gross benefits of clean diesel technology are unlikely to start to show up in the atmosphere for at least a decade or so. Nevertheless, it should have a positive impact before too long. The EPA estimates that by 2030, when the entire diesel-powered vehicle fleet (on- and off-highway equipment) in the U.S. will finally turn over as a result of retirement, NO_x emissions will be reduced by four million tons annually and cancer-causing particulates by 250,000 tons per year. The benefits in the fight against climate change should also be significant, although perhaps not as apparent.

"The recent availability of low-sulfur fuel is a sea change in the car world," Kassel says. "It resembles the period in the early 1970s when the government removed lead from gasoline." This action not only halted lead emissions but also permitted engineers to use the catalytic converter to deal with the other exhaust pollutants. "It took more than 30 years to go from leaded gas and no effective emissions controls to today's extremely clean and increasingly fuel-efficient cars," he states. "By next year, in less than a decade, the car world will have undergone a similar change with diesels." SA

Steven Ashley is a staff editor and writer.

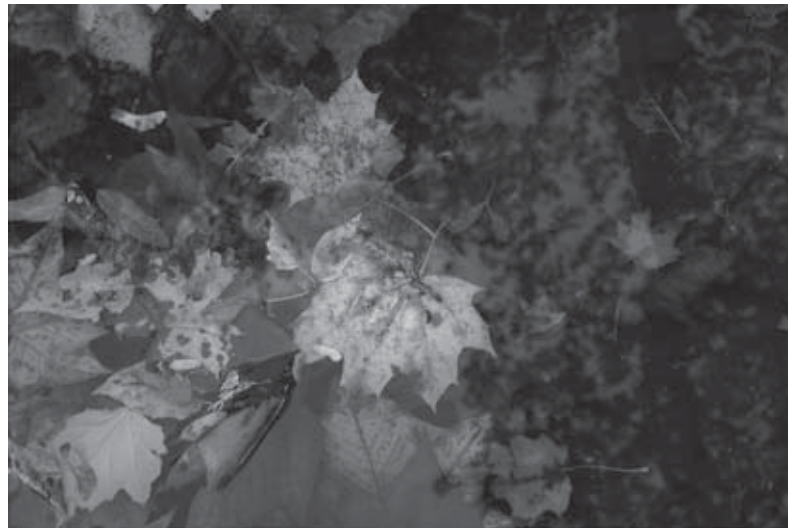
MORE TO EXPLORE

Diesel Technology Forum discussion of clean diesels:
www.dieselforum.org/meet-clean-diesel/what-is-clean-diesel/
U.S. Environmental Protection Agency's National Clean Diesel Campaign:
www.epa.gov/cleandiesel/
Mercedes-Benz Bluetec technology:
www4.mercedes-benz.com/specials/scr/en/index_nocom_en.htm
Honda clean diesel technology:
<http://world.honda.com/news/2006/c060925DieselEngine/>

Illusory Color & the Brain

Novel illusions suggest that the brain does not separate perception of color from perception of form and depth

By John S. Werner, Baingio Pinna and Lothar Spillmann



AUTUMN LEAVES and reflections in a fountain highlight the way color contributes to perception. Much of the depth and detail disappears in a black-and-white version of the scene.

A world without color appears to be missing crucial elements. And indeed it is. Colors not only enable us to see the world more precisely, they also create emergent qualities that would not exist without them. The color photograph on the opposite page, for example, reveals autumnal leaves in the placid water of a fountain, along with the reflections of trees and of a dark-blue afternoon sky behind them. In a black-and-white picture of the same scene, the leaves are less distinct, the dark-blue sky is absent, the reflections of the light are weak, the water itself is hardly visible, and the difference in apparent depth among the sky, trees and floating leaves is all but gone.

Yet this role for color, and even the true nature of color, is not well recognized. Many people believe that color is a defining and essential property of objects, one depending entirely on the specific wavelengths of light reflected from them. But this belief is mistaken. Color is a sensation created in the brain. If the colors we perceived depended only on the wavelength of reflected light, an object's color would appear to change dramatically with variations in illumination throughout the day and in shadows. Instead patterns of activity in the brain render an object's color relatively stable despite changes in its environment.

Most researchers who study vision agree that color helps us discriminate objects when differences in brightness are insufficient for this task. Some go so far as to say that color is a luxury and not really needed: after all, totally color-blind people and many species of animals seem to do well without the degree of color perception that most humans have. The pathway in the brain that serves navigation and movement, for example, is essentially color-blind. People who become color-blind after a stroke appear to have normal visual perception otherwise. Such observations have been taken as support for the insular nature of color processing, suggesting it has no role in processing depth and form—in short, that color is only about hue, saturation and brightness.

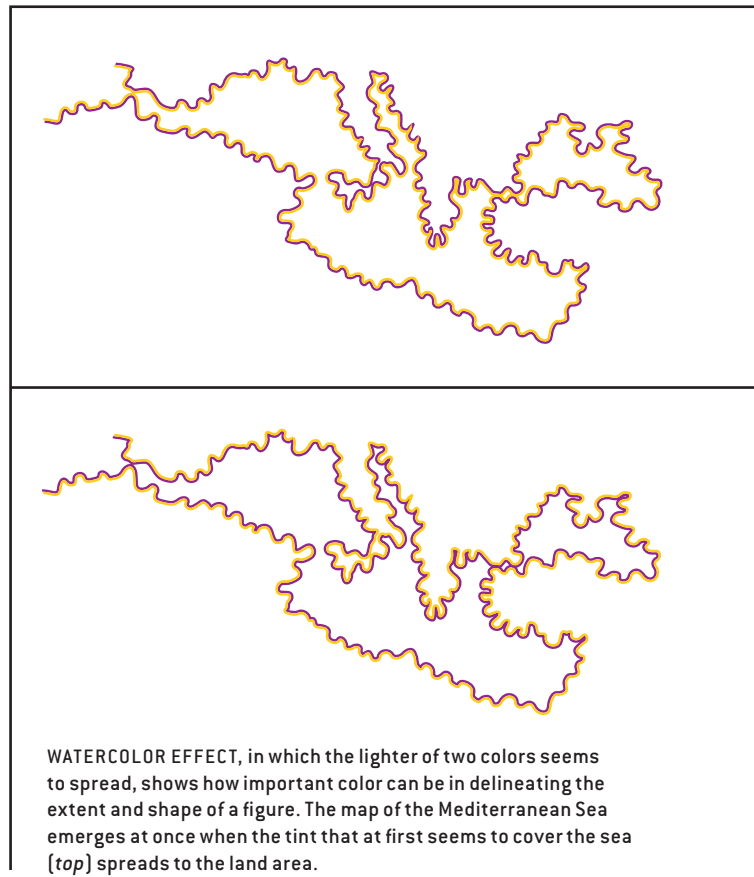
But the study of illusory colors—colors that the brain is tricked into seeing—demonstrates that color processing in the brain occurs hand in hand with processing of other properties, such as shape and boundary. In our decade-long attempt to discern how color influences perceptions of other properties in objects, we have considered a number of novel illusions, many created by us. They have helped us understand how the neural processing of color results in emergent properties of shape and boundary. Before we begin our discussion of these

illusions, however, we need to recall how the human visual system processes color.

Pathways to Illusions

VISUAL PERCEPTION BEGINS with the absorption of light—or, more precisely, the absorption of discrete packets of energy called photons—by the cones and rods located in the retina [*see box on next page*]. The cones are used for day vision; rods are responsible for night vision. A cone photoreceptor responds according to the number of photons it captures, and its response is transmitted to two different types of neurons, termed on and off bipolar cells. These neurons in turn provide input to on and off ganglion cells that sit side by side in the retina.

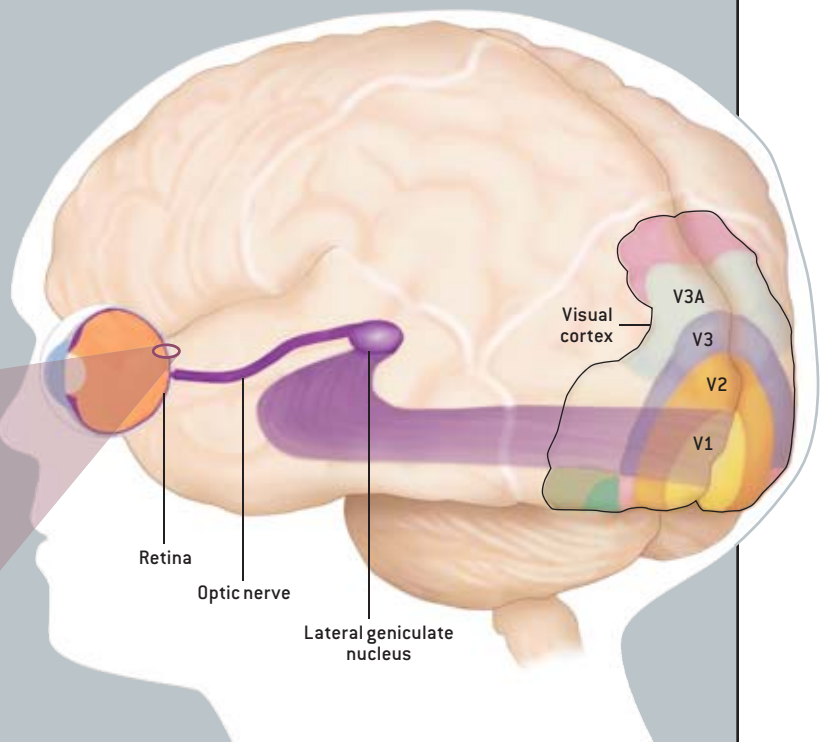
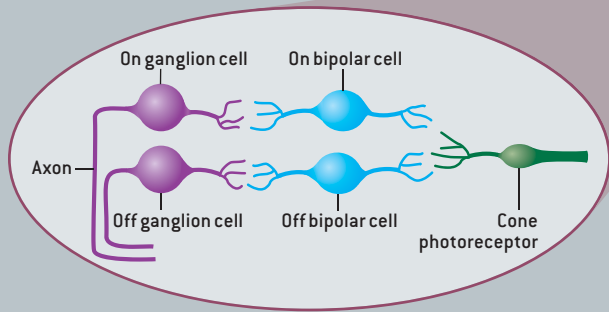
The ganglion cells have what is called a center-surround receptive field. The receptive field of any vision-related neuron is the area of space in the physical world that influences the activity of that neuron. A neuron with a center-surround receptive



WATERCOLOR EFFECT, in which the lighter of two colors seems to spread, shows how important color can be in delineating the extent and shape of a figure. The map of the Mediterranean Sea emerges at once when the tint that at first seems to cover the sea (top) spreads to the land area.

Seeing Color

Perception of color begins with the absorption of light by the cone cells in the retina (*detail below*). A cone photoreceptor responds in only one way, but its activity is conveyed by two different types of neurons, called on and off bipolar cells. The bipolar cells in turn provide input to on and off ganglion cells. The ganglion cell axons relay their signals to the brain, first to the lateral geniculate nucleus and from there to the visual cortex.



field responds differently depending on the relative amount of light in the center of the field and the region around the center.

An on ganglion cell fires maximally (at a high rate) when the center is lighter than the surround, firing minimally when the receptive field is uniformly illuminated. Off cells behave in the opposite way: they fire maximally when the center is darker than the surround and minimally when the center and surround are uniform. This antagonism between center and surround means that ganglion cells respond to contrast and in this way sharpen the brain's response to edges and borders.

Most of the ganglion cell axons (fibers) relay their signals to the brain, specifically to the lateral geniculate nucleus of the thalamus (near the center of the brain) and from there to the visual cortex (at the back of the brain). Different populations of ganglion cells are sensitive to somewhat different features of stimuli, such as motion and form, and their fibers conduct signals at different velocities. Color signals, for example, are carried by the slower fibers.

About 40 percent or more of the human brain is thought to be involved in vision. In the areas stimulated early in visual processing (parts of the visual cortex called V1, V2 and V3), neurons are organized into maps that provide a point-to-point representation of the visual field. From there, visual signals disperse to more than 30 different areas, interconnected by more than 300 circuits. Each of the areas has specialized functions, such as processing color, motion, depth and form, although no area mediates one perceptual quality exclusively. Somehow all this information is combined, in the end, into a unitary perception of an object having a particular shape and color. Neuroscientists do not yet understand the details of how this comes about.

Interestingly, bilateral damage to certain visual areas leads to deficits in the perception of form as well as color, which offers another piece of evidence that color is not disembodied from the other properties of an object. The intermingling of

color signals in the brain with signals carrying information about the form of objects can result in perceptions not expected from an analysis of the wavelengths of light reflected from those objects—as our illusions make startlingly clear.

The Watercolor Effect

ONE OF OUR EARLY EXPERIMENTS with illusory color illustrates how important color can be in delineating the extent and shape of a figure. Under certain conditions, color changes in response to the surrounding color; it can become more different (called contrast) or more similar (called assimilation). The spreading of similar color has been described only over rather narrow areas, in agreement with the finding that most connections among visual neurons in the brain are relatively short range. We were therefore surprised to find that when an uncolored area is enclosed by two differently colored boundary contours—with the inner contour lighter than the outer contour—tint emanates from the inner contour, spreading across the entire area, even over rather long distances [*see illustration on preceding page*].

Because the color resembles a faint veil such as that seen in watercolor painting, we call this illusion the watercolor effect.

Overview/Color Vision

- Vision researchers have long held that color processing in the brain is separate from the processing of other features, such as depth and form.
- The study of illusory colors, however, demonstrates that the perception of color generates emergent properties of form and depth.
- In particular, the authors have adapted a figure called the Ehrenstein illusion to reveal how color, shape and form are linked in the brain's perception of the visual world.

fect. We found that the spreading requires the two contours to be contiguous so that the darker color can act as a barrier, confining the spreading of the lighter color to the inside while preventing it from spreading to the outside. The figure defined by the illusory watercolor appears dense and slightly elevated. When the colors of the double contour are reversed, the same region appears a cold white and slightly recessed.

The watercolor effect defines what becomes figure and what becomes ground even more powerfully than the properties discovered by the Gestalt psychologists at the turn of the 20th century, such as proximity, smooth continuation, closure, symmetry, and so on. The side of the double contour that has the lighter color fills in with watercolor and is perceived as figure, whereas the side that has the darker color is perceived as ground. This asymmetry thus helps to counteract ambiguity. The phenomenon is reminiscent of the notion of Edgar Rubin, one of the pioneers of figure-ground research, that the border belongs to the figure, not the ground.

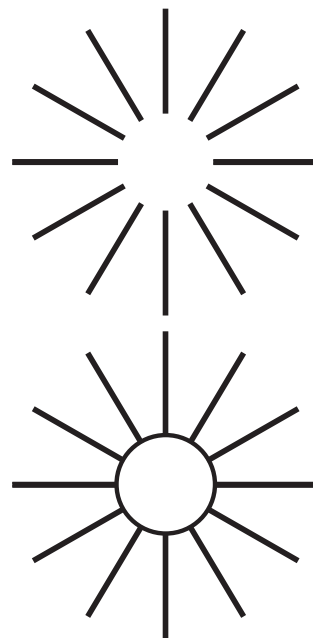
A possible neural explanation for the watercolor illusion is that the combination of a lighter contour flanked by a darker contour (on an even lighter background) stimulates neurons that respond only to a contour that is lighter on the inside than the outside or to a contour that is darker on the inside than the outside, but not to both. Border ownership most likely is encoded at early stages of processing in the visual cortex, such as in brain areas V1 and V2. In experiments with monkeys, neurophysiologists have found that approximately half the neurons in the visual cortex respond to the direction of contrast (whether it gets lighter or darker) and therefore could delineate the border. These same neurons have a role in depth perception that might contribute to figure-ground segregation.

Our investigations showed that wiggly lines produce stronger watercolor spreading than straight ones do, probably because the undulating borders engage more neurons responsive to orientation. The color signaled by these uneven edges must be propagated across regions of cortex that serve large areas of the visual field, continuing the spread of color until border-sensitive cells on the other side of the enclosed area provide a barrier to the flow. Color and form are thus bound together inextricably in the brain and perception at this level of cortical analysis.

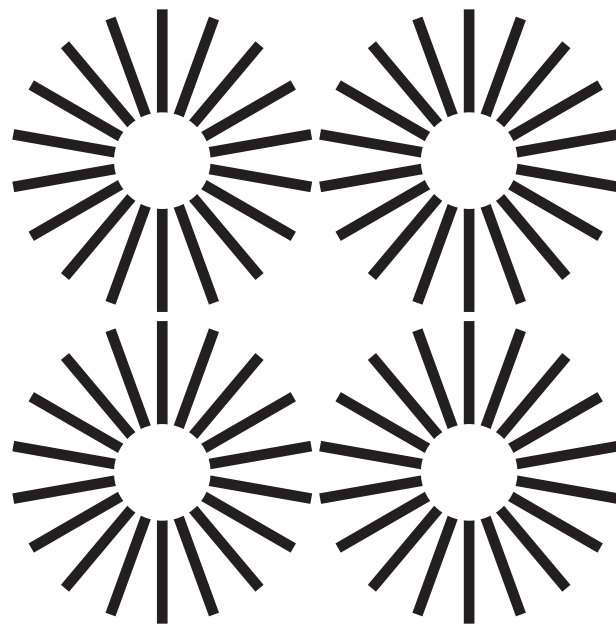
Radial Lines

THE RADIAL LINE ILLUSION offers further evidence of the role color plays in distinguishing figure from ground. In 1941 German psychologist Walter Ehrenstein demonstrated that a bright circular patch conspicuously fills the central gap between a series of radial lines. The patch and the circular border delineating it have no correlate in the physical stimulus; they are illusory. The bright illusory surface seems to lie slightly in front of the radial lines [see *top illustration on this page*].

The length, width, number and contrast of the radial lines determine the strength of this phenomenon. The spatial configuration of the lines necessary for the illusion to take effect implies the existence of neurons that respond to the termina-



EHRENSTEIN FIGURE, developed by German psychologist Walter Ehrenstein in 1941, provides the foundation for the illusions that follow. Adding a circle (*bottom*) destroys the illusion of a bright central disk.



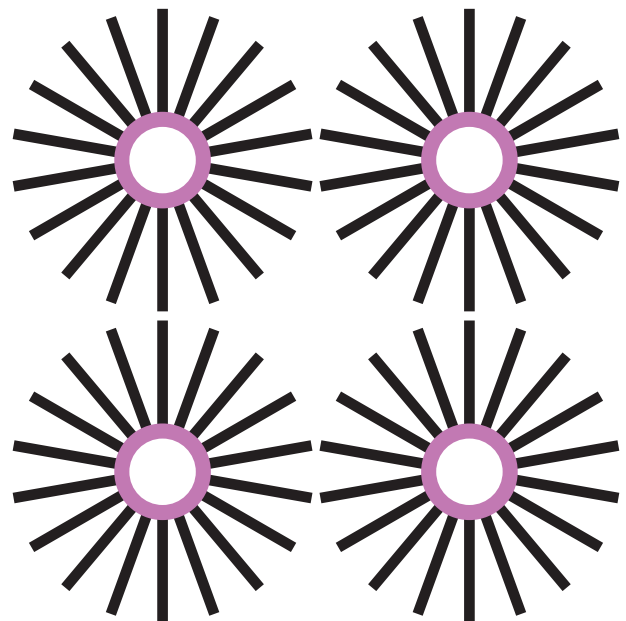
1 Bright circular patches fill the central gap of an Ehrenstein figure modified to enhance that illusion.

tion of a line. Such cells, called end-stopped neurons, have been identified in the visual cortex, and they may explain this effect. These local signals combine and become inputs to another (second-order) neuron, which fills in the central area with enhanced brightness.

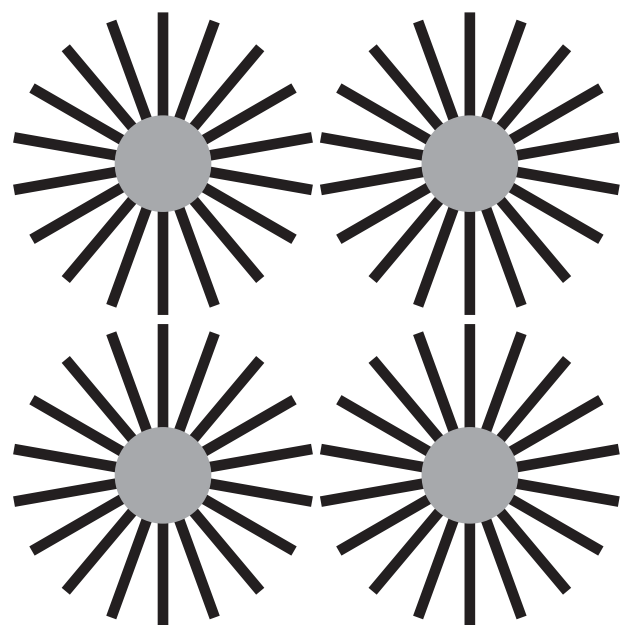
In our studies of the Ehrenstein illusion, we evaluated variations in the number, length and width of the radial lines, and the examples we present in this article use the most striking combination that we found [see numbered illustrations]. We show four copies of each pattern, arranged as a quartet, to enhance the effects. Once we determined the characteristics for the radial lines that produced the brightest central circle **1**, we experimented with variations in the chromatic properties of the central gap. First we added a black annulus, or ring, to the Ehrenstein figure, and the brightness of the central gap disappeared entirely—the illusion was destroyed, as Ehrenstein had already noticed. We suspect that this effect arises because the ring silences the cells that signal line terminations.

If the annulus is colored, however, other cells may be excited by this change. When we added a colored annulus, the white disk not only appeared much brighter (self-luminous) than it did in the Ehrenstein figure, it also had a dense appearance, as if a white paste had been applied to the surface of the paper **2**. This phenomenon surprised us; self-luminosity and surface qualities do not ordinarily appear together and have even been considered opposing, or mutually exclusive, modes of appearance. We call this phenomenon anomalous brightness induction. As with the watercolor effect, cells in early cortical areas are candidates for causing this illusion.

Next we inserted a gray disk into the central gap of an Ehrenstein figure **3**. Another phenomenon, called scintillating luster, arose, in which illusory brightness gives way to the perception of a glossy shimmer that occurs with each movement of the pattern or of the eye. The scintillation, or flashing, may come about by a competition between the on and off systems: line-induced brightness (illusory increment) competes with the dark gray of the disk (physical decrement). When we replaced the central white disks within the colored ring with black disks and used a black surround **4**, the disks looked even darker than the physically identical surrounding area. Instead of appearing self-luminous, as white disks do, black-



2 Anomalous brightness induction: the addition of colored rings makes the illusory patches appear even whiter.

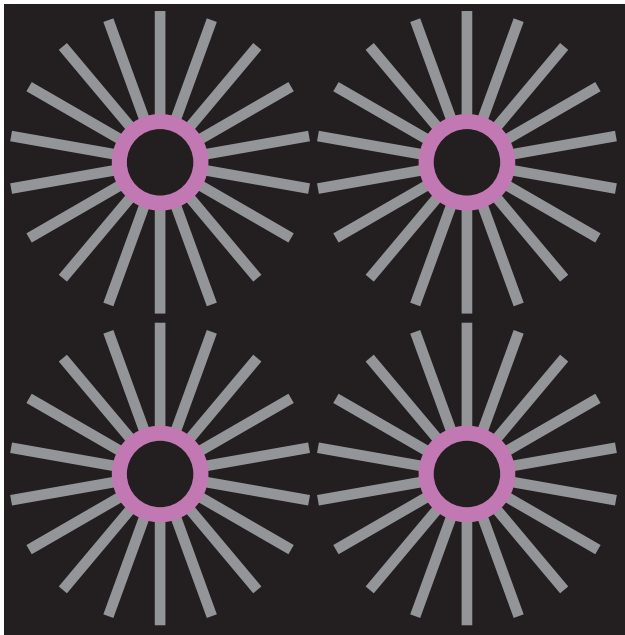


3 Scintillating luster: gray disks cause shimmering circular patches to fill the central gap.

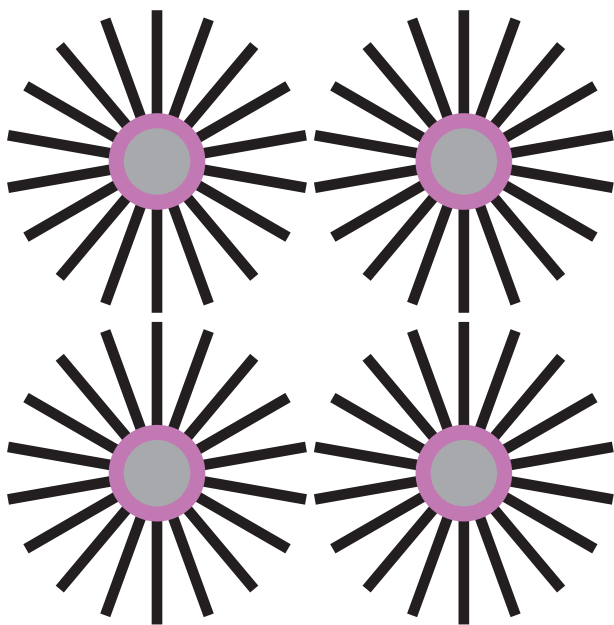
THE AUTHORS

JOHN S. WERNER, BAINGIO PINNA and LOTHAR SPILLMANN have worked on the illusions presented in this article over the past decade. Werner received a Ph.D. in psychology from Brown University and conducted research at the Institute for Perception-TNO in the Netherlands. He is a professor at the University of California, Davis. Pinna, a professor at the University of Sassari in Italy, received his undergraduate and graduate education at the University of Padua. Spillmann, who is head of the Visual Psychophysics Laboratory at Freiburg University in Germany, spent two years at the Massachusetts Institute of Technology and five years at the Retina Foundation and Massachusetts Eye and Ear Infirmary. Both Pinna and Spillmann have visual illusions on display at the Exploratorium in San Francisco.

JOHN S. WERNER, BAINGIO PINNA AND LOTHAR SPILLMANN



4 Anomalous darkness induction: black disks within a colored ring appear much blacker than the physically identical surrounding black.



5 Flashing anomalous color contrast: gray disks ringed in purple appear as greenish-yellow flashing lights when the pattern or the eye moves back and forth.

ness seems to generate a void, or a black hole, that absorbs all the light.

When the central disk within the chromatic ring was gray instead of white or black, the disk appeared to become tinted with the complementary color of the annulus—for example, greenish-yellow when the surrounding ring was purple (5). Furthermore, the disk appeared to flash with each eye movement, or when the pattern moved back and forth, and to move in relation to its surround. Flashing anomalous color contrast depends on radial lines and a chromatic annulus the way the other effects do, but it also has unique qualities that do not appear to be a simple combination of the other known effects. In this illusion, the induced color appears both self-luminous and scintillating. Strikingly, it appears to float above the rest of the image. The surface color and the self-luminous color do not mix; instead one belongs to the disk on the page, and the other emerges from a combination of the other characteristics of the stimuli.

In flashing anomalous color contrast, the radial lines may activate local end-stopped neurons, as has been proposed for the filling in of gaps by illusory contours, but activity by those cells does not account completely for the combined flashing and complementary color. It is not clear whether the radial lines have a direct effect on color contrast or whether the vividness of the color is derived indirectly from the luster and scintillation caused by the combination of radial lines and the gray center.

Current understanding of the brain cannot explain all the things going on in this illusion. The complexity of the illusion suggests that it is unlikely to result from a single unitary process but may represent an attempt by the brain to reconcile competing signals from multiple specialized pathways. Scientists clearly have much more to learn about how the brain perceives the physical world. Fortunately, ongoing work on illusory colors will continue to offer a tantalizing portal into the complexities of the human visual system. SA

MORE TO EXPLORE

- Sensory Experience, Adaptation and Perception.** Edited by Lothar Spillmann and Bill R. Wooten. Lawrence Erlbaum Associates, 1984.
- Visual Perception: The Neurophysiological Foundations.** Edited by Lothar Spillmann and John S. Werner. Academic Press, 1989.
- Neon Color Spreading: A Review.** P. Bressan, E. Mingolla, L. Spillmann and T. Watanabe in *Perception*, Vol. 26, No. 11, pages 1353–1366; 1997.
- The Watercolor Effect: A New Principle of Grouping and Figure-Ground Organization.** B. Pinna, J. S. Werner and L. Spillmann in *Vision Research*, Vol. 43, No. 1, pages 43–52; January 2003.
- The Visual Neurosciences.** Edited by L. M. Chalupa and J. S. Werner. MIT Press, 2004.
- Figure and Ground in the Visual Cortex: V2 Combines Stereoscopic Cues with Gestalt Rules.** F. T. Qiu and R. von der Heydt in *Neuron*, Vol. 47, No. 1, pages 155–166; July 7, 2005.
- The Watercolor Illusion and Neon Color Spreading: A Unified Analysis of New Cases and Neural Mechanisms.** B. Pinna and S. Grossberg in *Journal of the Optical Society of America*, Vol. 22, No. 10, pages 2207–2221; 2005.

WORKINGKNOWLEDGE

LEFT VENTRICULAR ASSIST DEVICES

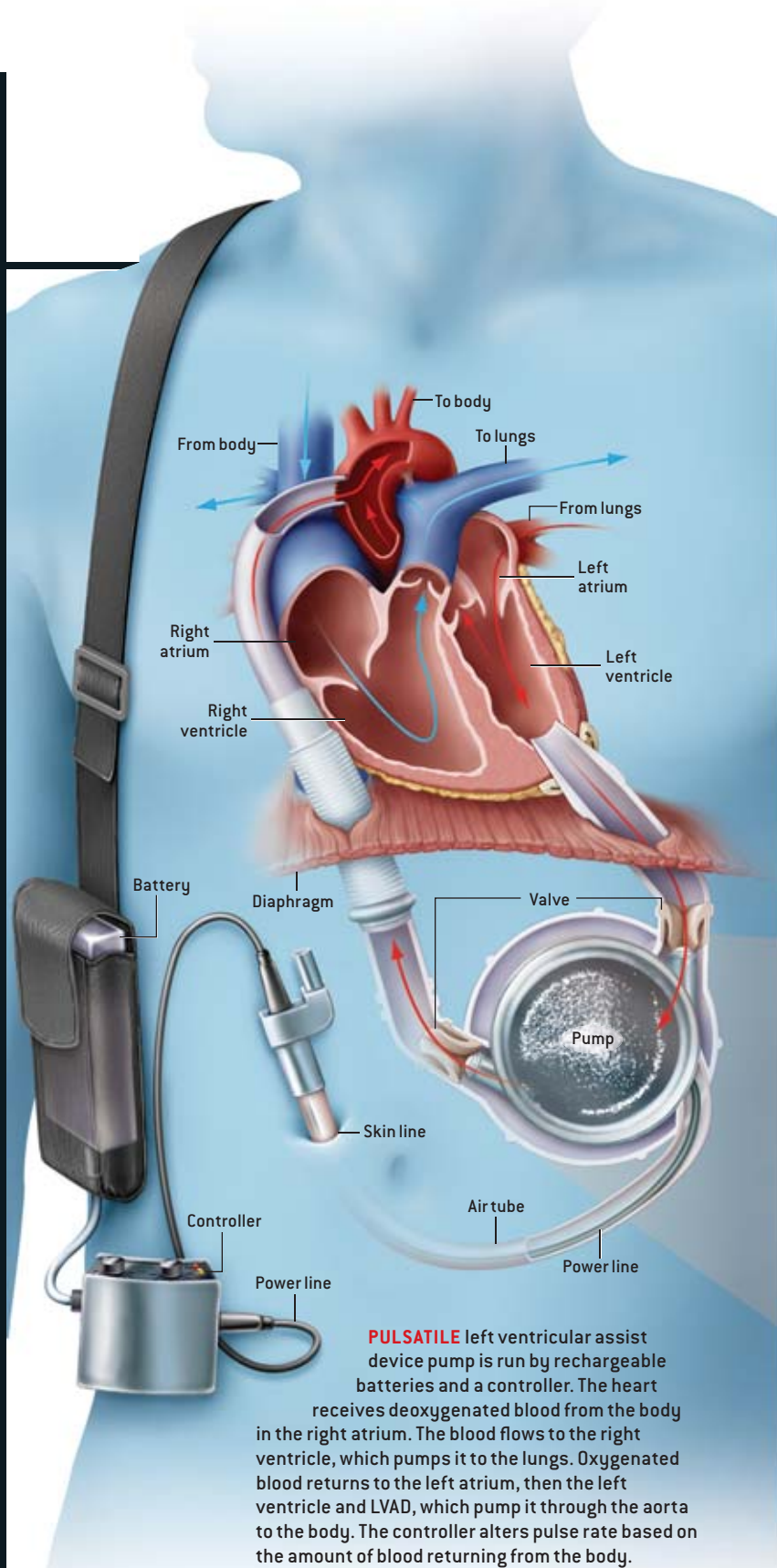
Restoring Flow

Although artificial hearts are stymied by complications, left ventricular assist devices (LVADs) are extending lives. Doctors began implanting them two decades ago to keep heart failure patients alive while they waited for weeks or months for an available transplant organ. Today improved designs are being installed as final fixes. Indeed, the distinction between an LVAD used as a bridge to transplant and as a permanent aid “is disappearing,” says Kiyotaka Fukamachi, head of the Cleveland Clinic’s Cardiovascular Dynamics Laboratory. “Some patients who received an LVAD as a bridge have been living with it for two or three years.”

A healthy left ventricle pumps freshly oxygenated blood through the aorta to the body. LVADs help the ventricle or take over its operations if the chamber is weak or has stopped functioning. First-generation designs, which still prevail, are pulsatile: an implanted pump pushes blood in pulses like a natural heart. Second-generation LVADs are smaller, relying on a rotor that continuously streams blood. Engineers are evaluating experimental, third-generation devices that use magnetically levitated rotors, reducing moving parts.

Yet “no one approach is necessarily better than the others,” Fukamachi says. “The choice depends on a patient’s circumstances.” The pulsatile machines, including Thoratec Corporation’s HeartMate I and World Heart Corporation’s Novacor, may still provide the best option if a patient needs a full takeover. Continuous-flow models such as MicroMed Cardiovascular’s DeBakey can be smaller and simpler because they do not require valves or a vent tube. Levitated machines may show less wear over time. (In the U.S., HeartMate I is approved for bridge and permanent therapy; Novacor is approved for bridge. Other models are in trials.)

Complications are involved, of course. A wire must protrude from the body to a controller and batteries, leading to infection in up to 15 percent of patients. Blood clots can form inside pumps, so patients must live on anticoagulants, which increase the chance for problematic bleeding. Device failure occurs, too. But doctors are likely to implant more LVADs because heart donors remain scarce. Only 2,100 transplants are performed in the U.S. every year, whereas 3,500 to 4,000 people are perennially on the waiting list. —Mark Fischetti



PULSATILE left ventricular assist device pump is run by rechargeable batteries and a controller. The heart receives deoxygenated blood from the body in the right atrium. The blood flows to the right ventricle, which pumps it to the lungs. Oxygenated blood returns to the left atrium, then the left ventricle and LVAD, which pump it through the aorta to the body. The controller alters pulse rate based on the amount of blood returning from the body.

DID YOU KNOW...

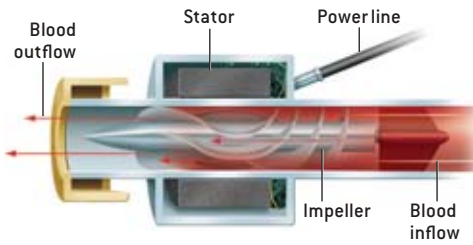
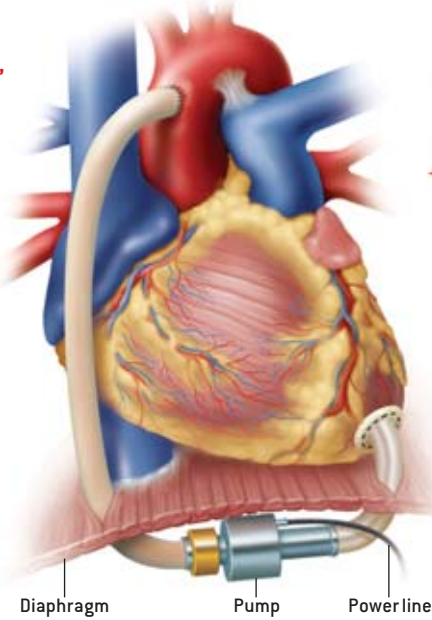
▣ **RIGHT VAD, TOO:** Patients who receive left ventricular assist devices often also have weak right ventricles; as many as 40 percent eventually suffer right ventricle failure. Engineers are testing right ventricular assists that could be implanted with a left assist. Kiyotaka Fukumachi's lab at the Cleveland Clinic is testing one model in animals, as are researchers elsewhere. Few designers are trying to combine both units into one, however, because that would, in effect, constitute yet another attempt at a total artificial heart.

▣ **LEFT OUT:** Heart failure patients may not qualify for an LVAD if they are very thin or very short, because the implanted pump can be four to six inches in diameter and the connection tubes may be

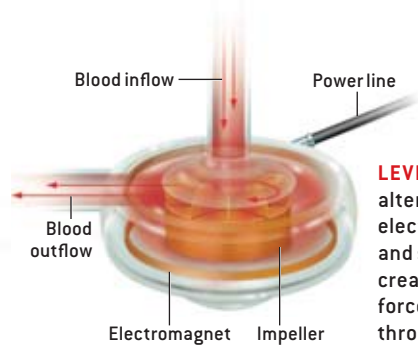
five to eight inches long. [Size is one reason engineers are pursuing rotor-based devices, which can be made smaller.] Patients who also have kidney, liver or lung failure—a common occurrence for many—are often excluded as well, because trials have shown that compromised organs typically cannot stand the stress of the LVAD surgery and recovery.

▣ **NO MORE AIR:** Engineers used external pneumatic power supplies for the earliest LVAD models because they were reliable and allowed smaller implanted pumps. But the approach has fallen out of favor in part because a large air-supply tube must pass through the abdomen's skin, leading to infection in 30 to 50 percent of recipients.

CONTINUOUS-FLOW LVAD, now in trials, has a rotor or levitated pump that circulates a continuous stream of blood through the heart's ventricles and the body's blood vessels; the person has a greatly reduced pulse.

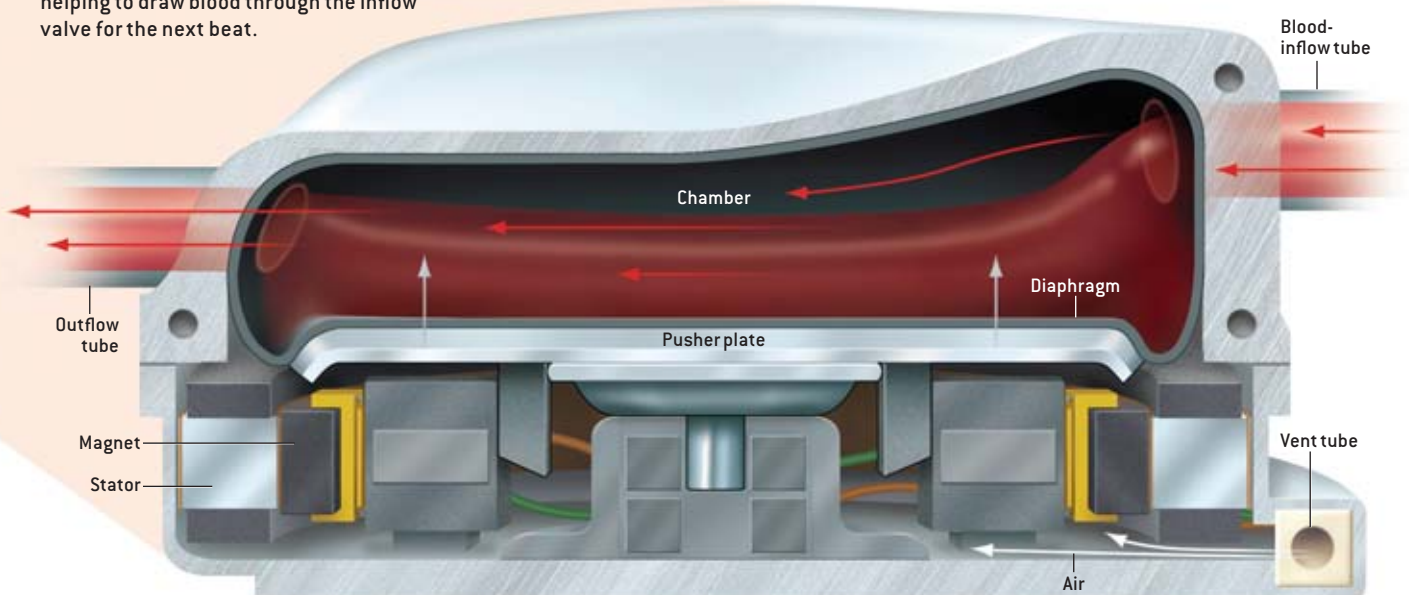


ROTOR has a stationary stator that creates a magnetic field, which pulls on magnets within the impeller's blades, spinning them.



LEVITATED PUMP is an alternative design; electromagnets suspend and spin an impeller, creating centrifugal force that pushes blood through vessels.

PUMP stator turns magnets so they revolve once per natural heartbeat, driving a pusher plate upward against blood entering the chamber from the left ventricle. Ambient air from the vent tube fills the growing void below the plate. When the chamber pressure exceeds that of the left ventricle, the inflow valve closes. Pressure continues to build until it is beyond that in the aorta; the outflow valve then opens, and blood is ejected. The air is exhausted, and the plate drops, helping to draw blood through the inflow valve for the next beat.



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

A New Journey into Hofstadter's Mind

THE ETERNAL GOLDEN BRAID EMERGES AS A STRANGE LOOP BY GEORGE JOHNSON

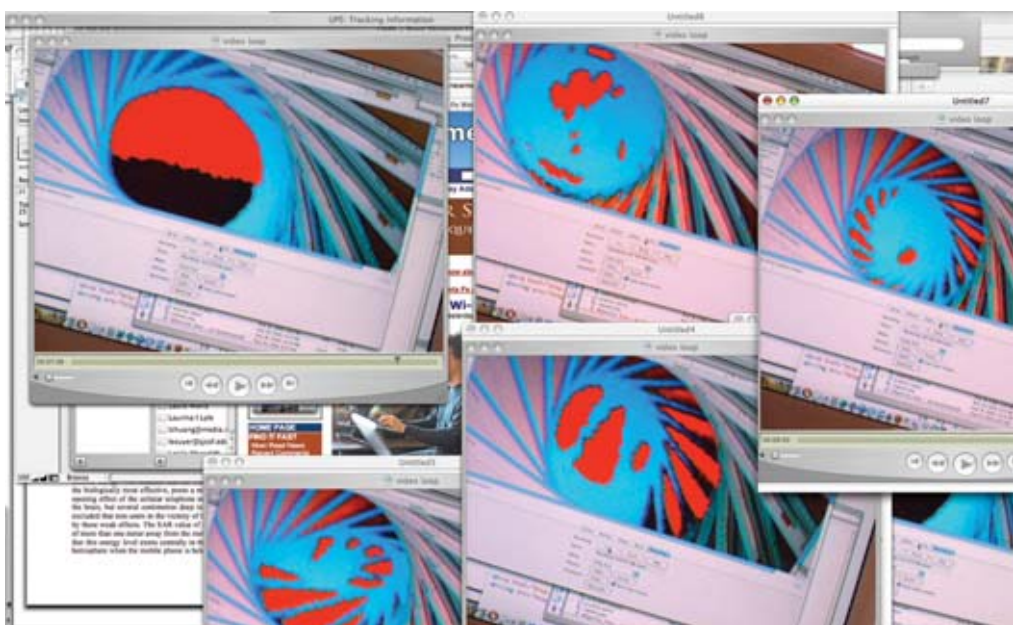
I AM A STRANGE LOOP

by Douglas R. Hofstadter
Basic Books, 2007 (\$26.95)

To get into a properly loopy mind-set for Douglas R. Hofstadter's new book on consciousness, I plugged a Webcam into my desktop computer and pointed it at the screen. In the first instant, an image of the screen appeared on the screen and then the screen inside the screen. Cycling round and round, the video signal rapidly gave rise to a long corridor leading toward a patch of shimmering blue, beckoning like the light at the end of death's tunnel.

Giving the camera a twist, I watched as the regress of rectangles took on a spiraling shape spinning fibonaccily deeper into nowhere. Somewhere along the way a spot of red—a glint of sunlight, I later realized—became caught in the swirl, which slowly congealed into a planet of red continents and blue seas. Zooming in closer, I explored a surface that was erupting with yellow, orange and green volcanoes. Like Homer Simpson putting a fork inside the microwave, I feared for a moment that I had ruptured the very fabric of space and time.

In *I Am a Strange Loop*, Hofstadter, a cognitive and computer scientist at Indiana University, describes a more elaborate experiment with video feedback that he did many years ago at Stanford University. By that time he had become obsessed with the paradoxical nature of Gödel's theorem, with its formulas that speak of themselves. Over the years this



VIDEO FEEDBACK experiment by the reviewer opens a window to the mind's self-referential power.

and other loopiness—Escher's drawings of hands drawing hands, Bach's invioluted fugues—were added to the stew, along with the conviction that all of this had something to do with consciousness. What finally emerged, in 1979, was *Gödel, Escher, Bach: An Eternal Golden Braid*, one of the most captivating books I have ever read.

I still remember standing in the aisle of a bookstore in Washington, D.C., where I had just finished graduate school, devouring the pages. *GEB*, as the author calls it, is not so much a "read" as an experience, a total immersion into Hofstadter's mind. It is a great place to be, and for those without time for the scenic route, *I Am a Strange Loop* pulls out the

big themes and develops them into a more focused picture of consciousness.

Think of your eyes as that video camera, but with a significant upgrade: a mechanism, the brain, that not only registers images but abstracts them, arranging and constantly rearranging the data into mental structures—symbols, Hofstadter calls them—that stand as proxies for the exterior world. Along with your models of things and places are symbols for each of your friends, family members and colleagues, some so rich that the people almost live in your head.

Among this library of simulations there is naturally one of yourself, and that is where the strangeness begins. "You make decisions, take actions,

REVIEWS

Go ahead...
choose your
Belize adventure.



Belize... a great place to observe wildlife. Explore thousands of miles of rivers, marshes and lagoons. Visit Maya villages and ancient temple sites. Seek out the elusive Jaguar, if you can. Snorkel on the longest barrier reef in the Western Hemisphere. All in a relaxing, peaceful country where the people are as warm and friendly as the climate. Experience the diversity of Belize, your English-speaking neighbor on the Caribbean coast of Central America, only 2 hours from the U.S.

Your
Caribbean
Gateway
to Central
America.



Call 1-800-624-0686 or visit our
website: www.travelbelize.org



affect the world, receive feedback from the world, incorporate it into yourself, then the updated ‘you’ makes more decisions, and so forth, round and round,” Hofstadter writes. What blossoms from the Gödelian vortex—this symbol system with the power to represent itself—is the “anatomically invisible, terribly murky thing called I.” A self, or, to use the name he favors, a soul.

It need know nothing of neurons. Sealed off from the biological substrate, the actors in the internal drama are not things like “serotonin” or “synapse” or even “cerebrum,” “hippocampus” or “cerebellum” but abstractions with names like “love,” “jealousy,” “hope” and “regret.”

And that is what leads to the grand illusion. “In the soft, ethereal, neurology-free world of *these* players,” the author writes, “the typical human brain perceives its very own ‘I’ as a pusher and a mover, never entertaining for a moment the idea that its star player might merely be a useful shorthand standing for a myriad infinitesimal entities and the invisible chemical transactions taking place among them.”

Thinking of souls this way makes me feel better about the fly I just swatted. Whatever repertoire of symbols it may have possessed was surely too constricted for Gödelian self-representation to arise. The same would probably go for an amphibian or a fish or, for that matter, a human ovum that had just been fertilized by a sperm. But somewhere along the line—maybe with parakeets or cats—it becomes harder to deny the glint of some kind of “who” inside.

Souls, as Hofstadter puts it, come in “different sizes.” In a whimsical moment, he even suggests that soulness might be measured—in units called “hunekers,” after an American music critic, James Huneker, who once wrote of a certain

Chopin étude that “small-souled men” should not attempt it. The scale might start with a mosquito, with a tiny fraction of a huneker, ascending to 100 for an average human and upward to maybe 200 for Mahatma Gandhi.

Hofstadter’s fans may find some of this familiar, but *I Am a Strange Loop* is much more than the condensed version of *Gödel, Escher, Bach*. In the 28 years since that book appeared, Hofstadter has lived with these ideas, working out their implications. From being a semivegetarian (fish and chicken were okay), he became, just recently, a strict one. Most significantly, in this time he learned what happens when another soul becomes entwined with your own. Her name was Carol, and as they became absorbed one into the other, consciousness spilling beyond its containers, he sometimes thought of them as “one individual with two bodies,” sharing “exactly the same dreads and dreams and hopes and fears.”

Then, when she was not yet 43, Carol died without warning from a brain tumor. Even though I knew that was coming, it did not diminish the blow. It is heart-wrenching to read how the author has tried to come to grips with her death, agonizing over how much “Carolness” and even “Carol-consciousness”—how much of her “interiority”—still lives in his brain and in those of the others who knew her.

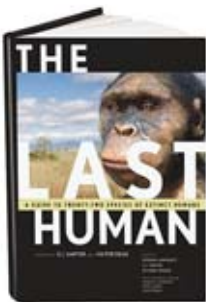
Consciousness is a pattern. The substrate is not supposed to matter. And yet it does. I finished the book with a sense of the desperation that must come from losing, in body if not in spirit, one half of a 400-huneker soul. SA

George Johnson's books include In the Palaces of Memory: How We Build the Worlds Inside Our Heads. A short movie of his video feedback experiment has been posted at www.talaya.net

THE EDITORS RECOMMEND

THE LAST HUMAN: A GUIDE TO TWENTY-TWO SPECIES OF EXTINCT HUMAN ANCESTORS

created by G. J. Sawyer and Viktor Deak.
Text by Esteban Sarmiento, G. J. Sawyer
and Richard Milner. Yale University Press,
2007 (\$39.95)



This unusual book draws on three-dimensional re-creations to bring to life 22 of our long-vanished ancestors. The startlingly lifelike re-constructions are the work of artist Viktor Deak and anthropologist Gary Sawyer of the

American Museum of Natural History in New York City, in collaboration with experts from around the world. Imaginative descriptions of the scenes by Esteban Sarmiento, also at the museum, convey both scientific information and the sense that these were once thinking, feeling creatures.

Several pages of information on teeth, diet, gait, age, range, tools, and so on follow the more dramatic introductions to each species. These give the book scientific heft, as does the excellent introductory essay by Ian Tattersall, a curator at the museum. "It is only today," he writes, "that we find a single kind of hominid in the world. The most important message encoded in this fact is that there is something very special and unusual about *Homo sapiens* that makes it a uniquely dangerous competitor."

It is, he continues, almost certainly our ability for symbolic reasoning, "together with our evident intolerance for competition (and we are still in the process of cleaning out our closest relatives, the great apes, from the rain forests of Africa and island Asia), that has led to our present (and entirely atypical) eminence as the only hominid on earth."

Families Have Saved Up To 50% On Heating Costs

And never have to buy fuel — wood, oil, gas, kerosene — ever again!



Lifetime Warranty

Hydro-Sil is a duct-free/furnace-free high performance room-by-room heating system that can save you hundreds of dollars in home heating costs by replacing old and inefficient heating. It can replace or supplement your electric heat, gas or oil furnace and woodstoves.

Hydro-Sil represents economy in heating: inside the heater case is a sealed copper tube filled with a harmless silicone fluid. It's permanent. You'll never run out. "Hydro-Sil Energy Star" thermostat controls a variable watt hydro element that is only being supplied a proportional amount of power on an as-needed basis. The silicone liquid is quickly heated, and with its heat retention qualities, continues to heat after the Hydro element shuts off. Hydro-Sil's "Energy Star" proportional control technology greatly increases energy savings and comfort.

Your Benefits with Hydro-Sil:

- Slash heating cost with Energy Star tech.
- Lifetime warranty. No service contracts
- Safe, complete peace of mind
- Clean, no fumes, environmentally safe
- U.L. listed
- Preassembled — ready to use
- No furnaces, ducts, or chimneys
- Portable (110V) or permanent (220V)
- Whole house heating or single room



Proportional "Energy Star" thermostat

220 VOLT PERMANENT	Approx. Area to Heat	Discount Price	Quantity
8' 2000 watts	250-300 s.f.	\$299	
6' 1500 watts	180-250 s.f.	\$279	
5' 1250 watts	130-180 s.f.	\$249	
4' 1000 watts	100-130 s.f.	\$229	
3' 750 watts	75-100 s.f.	\$199	
2' 500 watts	50-75 s.f.	\$169	
Thermostats		Call for options & exact heater needed	

110 VOLT PORTABLES (Thermostat Included)	Discount Price	Quantity
5' Hydro-Max 750-1500 watts	\$219	
4' Convector — Dual watt	\$179	
3' 750 watts — Silicone	\$179	
\$18.00 shipping per heater	\$	
Total Amount	\$	

Order today or contact us for more info.

Check ■ MasterCard ■ Visa

1-800-627-9276

www.hydrosil.com

Hydro-Sil, P.O. Box, 662, Fort Mill, SC 29715

Name _____
Address _____
City _____ St _____ Zip _____
Phone _____
MasterCard or Visa Account Information:
Acct # _____
Expiration Date _____



To renew your subscription to Scientific American:

www.SciAm.com/renew/

To give a gift subscription:

www.SciAm.com/gift/

To subscribe to Scientific American Mind:

www.SciAmMind.com

To subscribe to Scientific American Digital:

www.SciAmDigital.com

SCIENTIFIC AMERICAN

For any additional information please e-mail customer services at: sacust@sci.am

The results are in. Membership in Science Connection is strongly correlated with your "love life happiness quotient". Why not join and add to our success statistics?

Science Connection
www.sciconnect.com



Tough to Swallow

LOOK OUT STOMACH, HERE IT COMES BY STEVE MIRSKY

“**Sword swallowing** is said to be dangerous.” Thus begins a short report in the *British Medical Journal* of November 5, 2005. One danger, according to the report’s author, radiologist Brian Witcombe, was a death sentence. That outcome occurred during the Inquisition, when sword swallowing was associated with mysticism and mysticism was associated with execution. Imagine shoving a long metal blade down your throat being the highlight of your day.

Although one might assume that sword swallowing carries serious medical risks other than capital punishment, “few deaths related to sword swallowing have been described,” according to Witcombe. “A Canadian sword swallower did die, but that was after swallowing an umbrella.” As is well known, it is bad luck to open an umbrella inside.

The 2005 article notes that anecdotal accounts of injuries had circulated throughout the sword-swallowing community but that a detailed analysis of common risks associated with the practice remained to be delineated by the medical community. That gap in the scientific literature has now been filled. Witcombe and co-author Dan Meyer, the executive director of Sword Swallowers’ Association International (SSAI), performed a survey among sword swallowers, published in the December 23, 2006, issue of *BMJ*. The article includes an eye-and-epiglottis-catching photograph of “one of the authors” swallowing seven swords simultaneously (which prevents him from saying that three times fast) and then helpfully identifies that author

as being Meyer rather than Witcombe.

Anyway, Witcombe and Meyer contacted 110 SSAI members and got usable responses from 46 individuals who had together swallowed more than 2,000 swords in the previous three months. The authors “excluded cases in which injury was related to swallowing items other than swords, such as glass, neon tubes, spear guns or jack hammers.”

One common complaint was sore throats, known in the community as, of



course, sword throats. Less frequent but more serious problems included perforations of the pharynx or esophagus, which would be the very definition of an occupational hazard. One individual, perhaps trying to work up to the full-size sword, wound up having “a bread knife removed transabdominally.” Which easily beats having a bread knife inserted transabdominally.

“Some experienced artistes add embellishments that increase danger,” the authors write of the sword-swallowing elite. “One sometimes performs on a unicycle.” Which could eventually lead to the best flat tire story ever. Another swal-

lower works underwater, in what may be a futile attempt to whet the blade.

Actually, a little lubrication is key to the entire épée ingestion enterprise. “The clean sword is usually lubricated at least with saliva,” the authors say. “One performer used butter,” clearly unaware of the health dangers of saturated fat. “One had to retire because of a dry mouth caused by medication.” Look for litigation if the medication did not include “may end career as sword swallower” in the list of possible side effects.

Although the medical literature has, until now, featured only scattered case reports of sword swallowing–related injury, sword swallowers in fact played an important role in medical history. “For example,” Witcombe wrote in the 2005 *BMJ* paper, “a sword swallower helped Dr. Kussmaul of Freiberg to develop a rigid endoscope in 1868 using a straight tube, mirrors and a gasoline lamp.” So thanks to sword swallowers, doctors learned how to shove things down a patient’s throat other than the bill.

Speaking of which, the 2006 article points out that although no SSAI members have died for their art, they might have succumbed when opening their medical bills, which in one case reached \$70,000. Well, *you* try getting insurance after putting down “sword swallower” for occupation on your application.

Deep down, however, we all relate to the sword swallower. Because everyone has to swallow unpleasant stuff every day. But I’m happy to be doing it as a writer. Because the pen is mightier than the sword. And much smaller. ■

Do the **virtual particles** in quantum mechanics really exist?

—J. FLEMING, MADISON, WIS.

Gordon Kane, director of the Michigan Center for Theoretical Physics at the University of Michigan at Ann Arbor, answers:

Virtual particles are indeed real—they have observable effects that physicists have devised ways of measuring. Quantum theory predicts that every particle spends some time as a combination of other particles in all possible ways. Quantum mechanics allows, and indeed requires, temporary violations of conservation of energy. So one particle can become a pair of heavier “virtual” particles, which quickly rejoin into the original particle, almost as if they had never existed.

While the virtual particles are briefly part of our world, they can interact with other, less exotic particles, and researchers can study these interactions to test predictions about virtual particles.

In a hydrogen atom, photons (elementary particles of electromagnetic radiation) bind together a negatively charged electron and a positively charged proton. Every photon will spend some time as a virtual electron plus its antiparticle, the virtual positron, as described above. The hydrogen atom can be in more than one state; in one of those states, the atom interacts a little differently with the virtual electron and positron than when it is in another, so quantum theory predicts the two states’ properties (which would otherwise be identical) to diverge slightly as a result of those interactions. That divergence was measured in 1947 by Willis Lamb, who later received a Nobel Prize in Physics for this work.

Another phenomenon involves elementary particles known as quarks—specifically the “top” quark, the heaviest of the six varieties. In the early 1990s the European laboratory CERN produced millions of particles called Z bosons and measured their mass very accurately. The measured value deviated a little from the mass apparently predicted by the Standard Model of particle physics, but the difference could be explained by the time the Z spent as a virtual top quark if such a quark had a certain mass. The mass of the top quark,

directly measured a few years later at Fermi National Accelerator Laboratory in Batavia, Ill., agreed with that obtained from the CERN analysis, providing another dramatic confirmation of our understanding of virtual particles.


When you **lose weight**, where does it go?

Lora A. Sporny, adjunct associate professor of nutrition education at Columbia University, explains:

When you lose weight, the fat that disappears has been broken down into usable fuel for bodily activities.

Fats exist in chemical form as triglycerides—roughly E-shaped macromolecules with a glycerol molecule linked to three fatty acid chains. When trimming calories or increasing exercise, hormone-sensitive lipase, an enzyme within fat cells, responds to hormonal messages and disassembles triglycerides into their component parts, which then slip into the bloodstream. The liver preferentially absorbs the glycerol and some of the fatty acids; muscle takes in the remainder.

Inside the muscle and liver cells, the triglyceride pieces are further taken apart, eventually yielding large quantities of a compound called acetyl-CoA. Within mitochondria—the powerhouses of the cells—the acetyl-CoA combines with the compound oxaloacetate to form citric acid. This synthesis kicks off the citric acid cycle, or Krebs cycle, a set of chemical reactions that creates usable energy from fat, protein and carbohydrates.

These mitochondrial activities produce numerous products and by-products: carbon dioxide, which the lungs discharge during exhalation; water, which is expelled in urine or perspiration; heat, which helps to maintain a comfortable body temperature; and the energy-carrying molecule adenosine triphosphate (ATP). The ATP powers cellular activities—moving muscles, maintaining the heart’s 100,000-plus daily beats, digesting food and processing nutrients into bodily tissues. 

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

