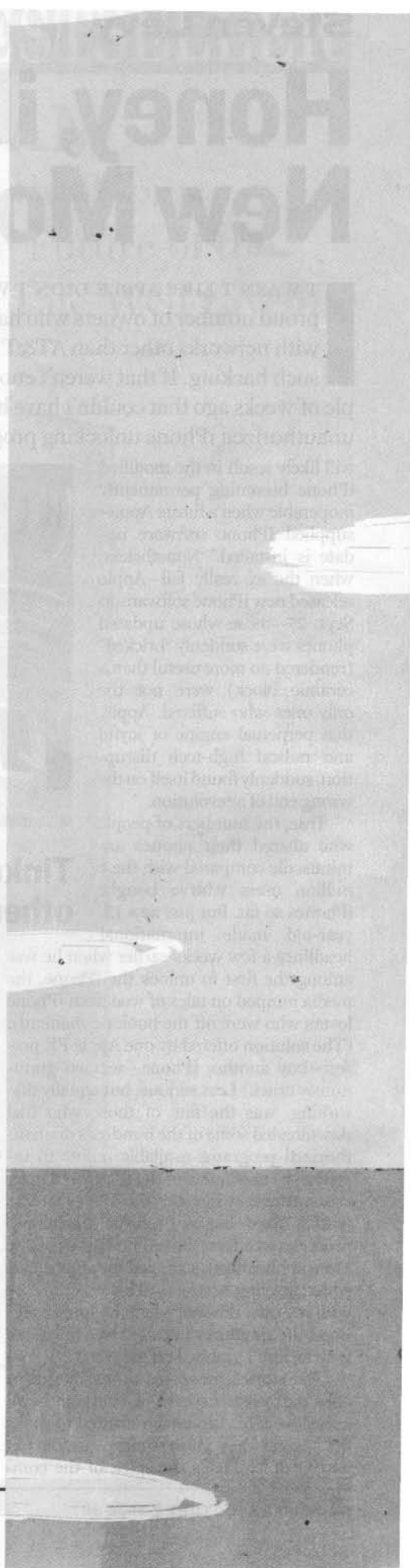


THE YEAR OF MIRACLES

■ BY LEE SILVER

THE YEAR 1905 WAS AN *ANNUS MIRABILIS*, OR MIRACLE YEAR—a rare historical moment in which key flashes of insight suddenly made the field of physics take off in new directions. That was the year Albert Einstein presented four papers that turned the conventional wisdom about how the universe works, from the infinitesimal realm of atoms to the vast reaches of the cosmos, upside down. During the next several decades, Einstein and a handful of other brilliant physicists went on to shape the 20th century and lay the foundation for all its technological accomplishments.

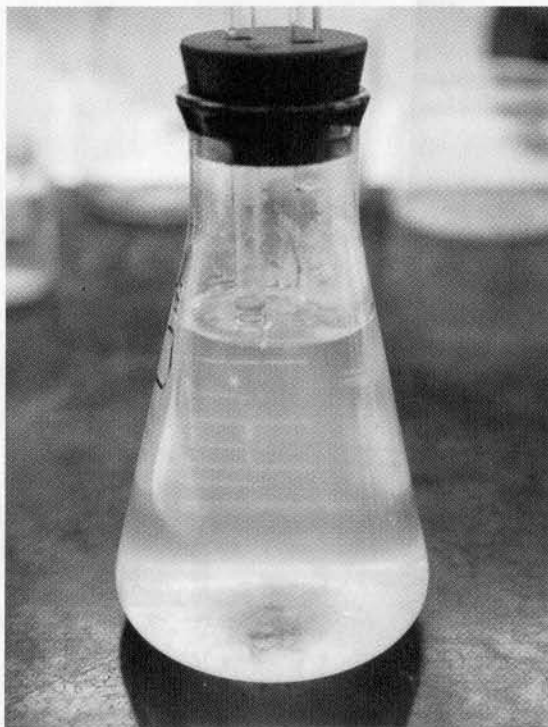
A century later, the year 2007 is shaping up to be another *annus mirabilis*. This time biology is the field in transition, and the ideas being shattered are old notions of genes and inheritance.



Ever since 1900, when Gregor Mendel's work on peas and inheritance was rediscovered, scientists have regarded the "gene" as the fundamental unit of heredity (just as the atom was regarded as the bedrock of pre-Einsteinian physics). Crick and Watson's discovery of the DNA double helix as the carrier of hereditary information did little to disturb the status quo. In recent months, however, a perfect storm of new technology and research has blown apart 20th-century dogma. The notion of the Mendelian gene as a unit of heredity, scientists now realize, is a fiction.

What's taking its place? Many scientists now believe that heredity is the result of an incredibly complex interplay among the basic components of the genome, scattered among many different genes and even the vast stretches of "junk DNA" once thought to serve no purpose. Biology has been building up to this insight for years, but some big puzzle pieces have now fallen into place. Once scientists abandoned their preconceived notions of genes and looked instead at individual DNA "letters" in the genome—the four bases A, C, T and G—they immediately began to see cause-and-effect connections to myriad diseases and human traits.

The result of this seemingly modest conceptual breakthrough has been a torrent of new discoveries. In five months, from April through August, geneticists at the Harvard/MIT Broad Institute, founded by Eric Lander; at deCODE Genetics in Iceland, founded by Kari Stefansson, and several other institutions have published papers suggesting that the key to a deeper understanding of the human genome may finally be in hand. These scientists have identified specific alterations in the sequence of DNA that play causative roles in a broad range of common diseases, including type 1 and type 2 diabetes; schizophrenia; bipolar disorder; glaucoma; inflammatory bowel disease; rheumatoid arthritis; hypertension; restless legs syndrome; susceptibility to gallstone formation; lupus; multiple sclerosis; coronary heart disease; colorectal, prostate and breast cancer, and the pace at



GENETICS USED TO BE A COTTAGE INDUSTRY OF PROFESSORS EACH DIRECTING RESEARCH ON A GENE.

which HIV infection causes full-blown AIDS. Unlike so many previous "disease gene" discoveries, these findings are being replicated and validated. "The race to discover disease-linked genes reaches fever pitch," declared the leading British science journal, *Nature*. Its American counterparts at *Science* chimed in: "After years of chasing false leads, gene hunters feel that they have finally cornered their prey. They are experiencing a rush this spring as they find, time after time, that a new strategy is enabling them to identify genetic variations that likely lie behind common diseases." That the world's top two scientific journals still use the old language of "genes" to describe these discoveries shows how new the new thinking really is.

These findings are just a prelude to what's shaping up as a true conceptual and technological revolution. Just as physics

shocked the world in the 20th century, it is now clear that the life sciences will shake up the world in the 21st. In a handful of years, your doctor may be able to run a computer analysis of your personal genome to get a detailed profile of your health prospects. This goes well beyond merely making predictions. A new technology called RNA interference may also allow doctors to control how your DNA is "expressed," helping you circumvent potential health risks. Many common diseases that have preyed on humans for eons—devastating neurological conditions such as Alzheimer's, Parkinson's, cancer and heart disease—could be eradicated. If this sounds outrageously optimistic, so did the promise of eliminating smallpox and polio to previous generations.

Why is all this happening now? What has changed between this year and last? To answer these questions, we need to trace the story of how mainstream biomedical scientists tried to link the cause of diseases to single genes and, despite early success, hit a brick wall. Meanwhile, a handful of renegade scientists, pursuing their own pet projects, happened to develop exactly the intellectual tools needed to break through that wall. These biologists are now the leaders of the new revolution in biomedical science.

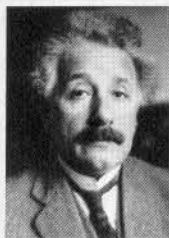
The seeds of our new understanding were first sown in the 1960s, when molecular biologists figured out how genetic information is organized, regulated and reproduced inside single-cell bacteria. In bacteria, a gene is a discrete segment of DNA that contains the "code" that tells the cell how to make a particular type of protein. Bacterial genes are arranged along a single DNA molecule, one after the other, with only tiny gaps in between. Since all organisms have DNA and work by essentially the same biochemistry, scientists assumed that a human genome would look

TWO REVOLUTIONS

Track major discoveries in science, and you'll see that although physics (red years) dominated the 20th century, biology and genetics (gray) rule the 21st.

■ physics ■ genetics

1866 Gregor Mendel, an Austrian monk, experiments with pea plants and publishes the theory that heredity tends to occur in discrete units (later to be attributed to genes).



1905 An obscure physicist, Albert Einstein, presents four papers that redefine time, space, matter and energy.

1906 The term "genetics" is used for the first time.

1913 Niels Bohr uses quantum theory to explain why glowing hydrogen emits certain wavelengths of light.

1915 Einstein, now world-renowned, completes his theory of general relativity, which describes gravity as a property of four-dimensional space-time.

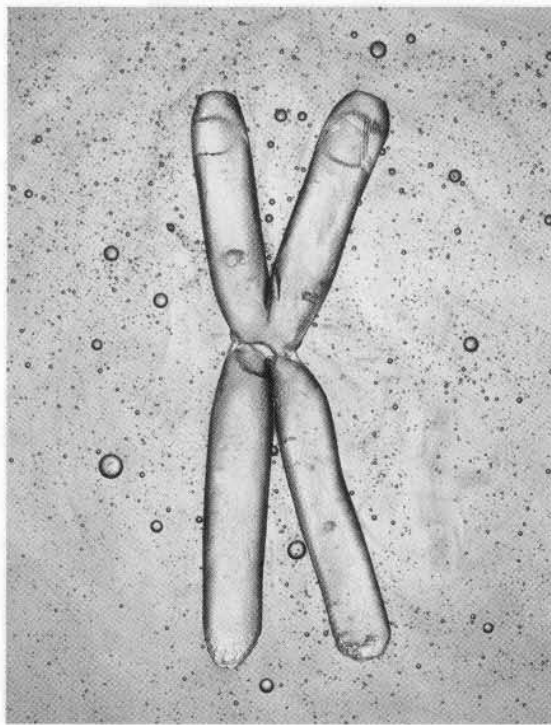


like a larger version of a bacterium's.

Clues that something was amiss came quickly with the development of DNA-sequencing methods in the 1970s. The first surprising result was that genes accounted for only 2 percent of the human genome—the rest of the DNA didn't seem to have any purpose at all. Biologists Phillip Sharp and Richard Roberts made things worse with a discovery that won them a Nobel Prize in 1993. If the gene were the basic unit of heredity, the DNA required to make any particular protein should be contained in its corresponding gene. But Sharp and Roberts found that DNA that codes for individual proteins is often split and scattered throughout the genome.

Scientists could ignore these signs largely because they seemed to be making progress. By combining new DNA-sequencing tools with studies of inherited diseases in large families, medical geneticists identified the genetic culprits responsible for cystic fibrosis, Huntington's disease, Duchenne muscular dystrophy and a host of other diseases. Each of these "all or none" diseases is caused by a mutation in a single protein-coding region of the DNA. Few diseases, unfortunately, work so neatly. In particular, the search for genetic bases of common diseases that affect large numbers of aging people came up empty.

DURING THIS LULL, A VISIONARY physician-scientist named Leroy Hood, now at the Institute for Systems Biology in Seattle, was growing impatient. Genetics, he recognized, was still a cottage industry of government-funded university professors, who each directed a small group of students and technicians to study an isolated



HEREDITY IS AN INCREDIBLY COMPLEX INTERPLAY AMONG THE MOST BASIC COMPONENTS OF THE GENOME.

gene. At the pace research was progressing, it would have required 100,000 worker-years of concerted effort to decipher just one complete human genome.

Hood thought it was absurd that genetic scientists spent nearly all their lab time performing tedious and repetitive mechanical and chemical procedures. At the same time, he grasped the far-reaching implications of a fundamental fact: while even the simplest organism is immensely complicated, the primary structures of its most complicated parts—DNA and proteins—are very simple. The alphabet of DNA contains only the four chemical letters (or bases) A, C, G and T, and proteins are made from just 21 amino acids. Hood saw that this simplicity would make it possible for robots and computers to read and write DNA and proteins more quickly, accurately and cheaply than human beings.

The rest of the biomedical community refused to believe that robots could analyze something as complex as a living system. And in any case, no practicing geneticist had the capacity to design such machines. Unable to obtain government grants, Hood secured private funding to bring together dozens of scientists, engineers and computer programmers (far larger and more diverse than any other genetics team). They proceeded to invent the first generation of molecular-biology machines. Two read and recorded information from DNA and proteins respectively (a process known as sequencing), and two others worked backward, converting digital electronic information into newly written sequences of DNA or protein.

Hood completely transformed the biomedical enterprise. DNA-writing machines give genetic engineers an unlimited capacity to create novel genes that can be studied in test tubes or added to the genomes of living organisms. And protein-writing and -reading machines provided drug firms with the ability to create a new generation of protein-based drugs. The DNA-reading machines suddenly made it conceivable to crack the 3 billion-base sequence of an entire human genome. In 1990 the U.S. government embarked on

a 15-year, \$3 billion project to do just that.

Eight years later, however, the project—parceled out to many U.S. scientists—was still less than 10 percent complete. Now it was biotech entrepreneur Craig Venter who was frustrated. Convinced that government-funded workers were the problem rather than the solution, Venter enlisted private funding of \$200 million to build an enormous lab filled with hundreds of automated machines, working 24/7, overseen by a handful of technicians. Within three years, the first reading of a human genome was essentially complete.

Armed with data from the genome project, scientists figured they'd surely be able to crack the really hard diseases, like cancer and heart disease. But a funny thing happened when they began to look closely at this vast storehouse of genetic information. Geneticists Andrew Fire and

1927 Experimentalists prove that matter can act like a wave.

1929 Edwin Hubble discovers that distant galaxies are moving away faster than nearer ones, which suggests that the universe is expanding.

1942 The first nuclear reactor goes online, and the United States starts the Manhattan Project to build an atom bomb.



1953 James Watson and Francis Crick transform genes from abstract notions into the molecular realm of the DNA double helix.

1969 The first direct evidence is found that the "quarks" proposed in



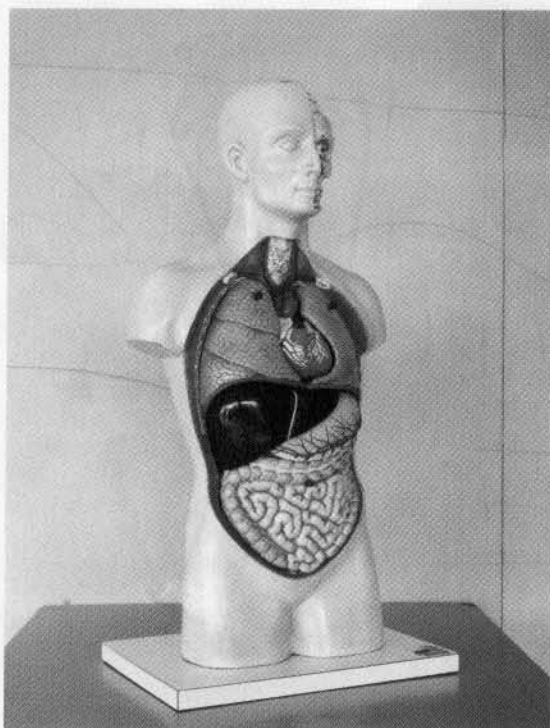
1964 actually exist inside protons and neutrons.

1980 Kary Mullis at Cetus Corp. invents a technique for making copies of DNA sequences,

Craig Mello galvanized the field by discovering a key mechanism that had been completely overlooked—the cellular process of RNA interference. (They shared a Nobel Prize in 2006 for the work.)

Finding evidence of extraterrestrial life couldn't have come as a bigger shock. Geneticists had taken for granted that the machinery of cells involved genes directing the production of proteins, and proteins doing the work of the cell. Here was a process that didn't involve proteins at all. Instead, tens of thousands of hitherto mysterious regions of the human genome—part of the so-called junk DNA—directed the production of specific molecules called microRNAs (consisting of bits of RNA, a well-known component of cells). These microRNAs then oversaw a whole new process, called RNA interference (RNAi), that served to modulate the expression of DNA.

The good news was that RNAi could open up a whole new approach to biomedical therapy (more on that later). But RNAi also made it clear that the fundamental unit of heredity and genetic function is not the gene but the position of each individual DNA letter.



SCIENTISTS HAVE FOUND VARIATIONS IN DNA LETTERS THAT MAY CAUSE A BROAD RANGE OF COMMON DISEASES.

TO MAKE IT ALL harder to fathom, each bit of DNA is susceptible to mutation and variation among individuals. Of the 3 billion DNA bases in the human genome, geneticists identified about one tenth of one percent (millions) that differ from one person to another. Variations in these particular letters—called “snips,” or SNPs, for single nucleotide polymorphisms—have replaced genes as the unit of heredity.

Many scientists responded to this devastating realization by going into a funk. “It will be difficult, if not impossible, to find the genes involved [in common diseases] or develop useful and reliable predictive tests for them,” Dr. Neil Holtzman, director of genetics and public policy at

Johns Hopkins University, said in 2001.

Fortunately, another visionary scientist, Kari Stefansson of Iceland, was already blazing a trail out of this thicket. If the genome was far more complex than scientists had thought, they would need to test for many more variables, and to do that they would need more test subjects. To find the cause of diseases would now require the participation of very large groups of *genetically related* people.

Like Hood and Venter, Stefansson was originally motivated by frustration with the pace of research. In the United States, where most of the disease-gene-discovery projects were being conducted, most people cannot trace their ancestors back more than a few generations, and the largest families consist of a few hundred living subjects at most. Subject panels of this size failed to provide sufficient data to identify the genetic bases

for complicated and variable common diseases. Stefansson decided to solve this problem by taking aim at the largest well-documented extended family that he knew—his own.

Nearly all the 300,000 citizens of Iceland can trace their ancestors back, through detailed, public genealogical records, to the Vikings who settled this desolate European island more than 1,000 years ago. Stefansson gave up his faculty position at Harvard Medical School to return to Iceland, where he founded the company deCODE Genetics in 1996. He persuaded the Icelandic government to provide deCODE with exclusive access to the health records of its citizens in return for bringing investment capital and high-tech jobs to the capital, Reykjavik. So far, more than 100,000 Icelandic volunteers have donated their DNA to deCODE.

Stefansson's project was roundly criticized by international bioethicists and other geneticists for violating the privacy of Icelanders (even though 90 percent of the population approved). Nevertheless, he persevered, placing “the genealogy of the entire nation on a computer database,” together with the health and DNA records of still-living individuals. The power of large numbers was soon apparent. In a study of obesity, he directed his software to look for SNPs associated with subsets of the population who were either extremely overweight or very thin. Within just a few hours, it began finding evidence that variations among particular DNA letters indeed played a causative role, confirming SNPs as the new unit of inheritance.

As of September, deCODE has made progress in identifying SNPs that may play a role in 28 common diseases, including glaucoma, schizophrenia, diabetes, heart disease, prostate cancer, hypertension and stroke. In some cases, such as glaucoma and prostate cancer, deCODE's findings

Join in a forum on the biocentury online at extra.Newsweek.com

for which he wins a Nobel Prize in 1993.
1984–86 Physicists realize that string theory could be developed to describe all known particles and their interactions.
1989 The National



Center for Human Genome Research is created to oversee the \$3 billion U.S. effort to map and sequence the human genome.
1995 Craig Venter at Celera completes the first sequencing of the full genome of a living organism other than a virus, for the bacterium *Hemophilus influenzae*.

1997 Researchers at Scotland's Roslin Institute report the first successful cloning of a sheep, by transferring a cell nucleus from an adult ewe into an embryonic sheep cell. The result is Dolly.
2000 A rough draft of



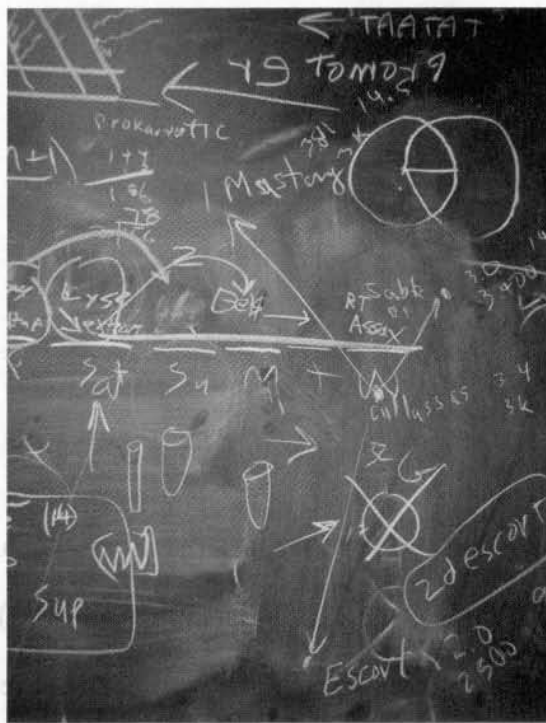
could lead to diagnostic tests for identifying people at risk of developing the disease. In other instances, such as schizophrenia, links to particular proteins have led to insight about the cause of the disease, which could lead to therapies.

Buoyed by Stefansson's success, other geneticists were eager to perform large-scale family studies, yet few had similar access to ancient genealogical records. But serendipity would deliver an epiphany: it's possible to study the entire human population as a single extended family, provided scientists collect enormous amounts of data.

Eric Lander, an MIT professor and the intellectual leader of the U.S. government effort to sequence the first human genome, realized scaling up would require a new approach. In 2004, Lander persuaded MIT and Harvard to combine their enormous resources toward the creation of the Broad Institute. Backed by \$200 million from billionaire philanthropists Eli and Edythe Broad, the institute is driving the development of ever more advanced genetic technologies. One technology, based on computer-chip fabrication, can identify DNA base letters present at 500,000 SNPs in the genomes of 40,000 or more people.

Think of this as a spreadsheet with 500,000 columns (each representing a specific SNP) and 40,000 rows (one for each person). To hunt for a genetic basis for, say, bipolar disease, the computer searches rows of people who have the disorder, checking column by column for an unusually high frequency of particular letters in comparison with people without the disease. As it turns out, a collaboration of American and German researchers has done this work—and found that variations of DNA letters in 20 different positions are influential in bipolar disease.

Incredibly, most disease-causing variants are the most common ones present in



JUST AS PHYSICS SHOCKED THE 20TH CENTURY, THE LIFE SCIENCES WILL SHAKE UP THE WORLD IN THE 21ST.

the human population: the strongest-acting one, for instance, exists in 80 percent of people without bipolar disease and 85 percent of people with the disease. The implication is that these variants are beneficial in some way, and cause problems only when their number exceeds a threshold.

To make sense of this complexity, scientists would like ultimately to build a vast international database that contains the complete sequence of DNA bases in the genomes of hundreds of millions of people. Ideally, such a database would be available for analysis by all biomedical researchers and would provide the foundation for understanding the genetic components of all human traits. That sounds like a lot of data—think of a spreadsheet with 3 billion columns and 100 million rows—but computing power is getting cheaper by the year. Within a decade, the cost of obtain-

ing a sequence of all 3 billion DNA letters in an individual's genome will drop from \$2 million now to \$1,000. It will be a routine part of a person's health record, enabling physicians to prescribe genome-specific preventions and treatments.

The discovery of RNAi, meanwhile, suggests a completely new personalized form of disease therapy. Whereas drugs act on proteins, RNAi therapy would act on the expression of DNA itself, potentially preventing or reversing diseases such as Alzheimer's, Parkinson's, Huntington's, bipolar disorder, schizophrenia and others. Old-school pharmaceutical firms have taken notice. The largest ones are betting heavily on the gene-targeted RNAi therapeutic approach to fill product pipelines, as the discovery of traditional chemical drugs becomes more elusive. Novartis and Roche have both signed nonexclusive licensing deals with the biotech firm Alnylam (founded by Phillip Sharp) for new therapeutic techniques that are valued at up to \$700 million and \$1 billion respectively; Merck paid \$1.1 billion to buy another biotech company outright, solely to obtain its contested portfolio of RNAi intellectual property, and the London-based drug firm AstraZeneca has a \$405 million licensing deal with Alnylam's competitor Silence Therapeutics.

The explosion of genetic discoveries shows no sign of letting up any time soon. New diseases are being added to the list every month, and biologists are rapidly parlaying gene- and SNP-disease links into a deeper understanding of how proteins and other molecules can misbehave to cause different medical problems in different people. And other scientists are working to advance the biology revolution (accompanying interviews). As a result of their efforts, many children born this year could very well be alive and healthy at the dawn of the next century, when they may look back in awe at the *annus mirabilis* of biomedical genetics in 2007.

SILVER is a professor of molecular biology at Princeton. He is the author of "Challenging Nature." He has no financial ties to any biotech or drug firm.

the human genome is completed and published by the Human Genome Project and Celera. Originally planned to take 15 years, it was accelerated by rapid technological advances.

2004 Biologist Eric Lander persuades MIT and Harvard to combine their resources to form the Broad Institute, which studies the genetic basis for a wide range of physiological, infectious and psychiatric diseases.

2007 Geneticists publish a spate of papers that identify specific, inherited alterations in DNA that play causative roles in both type 1 and type 2 diabetes; schizophrenia;



bipolar disorder; lupus; multiple sclerosis; coronary heart disease; glaucoma; inflammatory bowel disease; hypertension; colorectal, prostate and breast cancers, and other common ailments.



