The New Genetics

U.S. DEPARTMENT OF
HEALTH AND HUMAN SERVICES
National Institutes of Health
National Institute of General Medical Sciences

WHAT IS NIGMS? The National Institute of General Medical Sciences (NIGMS) supports basic research on genes, proteins and cells. It also funds studies on fundamental processes such as how cells communicate, how our bodies use energy and how we respond to medicines. The results of this research increase our understanding of life and lay the foundation for advances in the diagnosis, treatment and prevention of disease. The Institute's research training programs produce the next generation of scientists, and NIGMS has programs to increase the diversity of the biomedical and behavioral research workforce. NIGMS supported the research of most of the scientists mentioned in this booklet.

Produced by the Office of Communications and Public Liaison
National Institute of General Medical Sciences
National Institutes of Health
U.S. Department of Health and Human Services

The New Genetics

NIH Publication No.10-662
Revised April 2010
http://www.nigms.nih.gov

Contents

Foreword

Consider just three of Earth's inhabitants: a bright yellow daffodil that greets the spring, the single-celled creature called *Thermococcus* that lives in boiling hot springs, and you. Even a science-fiction writer inventing a story set on a distant planet could hardly imagine three more different forms of life. Yet you, *Thermococcus* and the daffodil are related! Indeed, all of the Earth's billions of living things are kin to each other.

And every living thing does one thing the same way: To make more of itself, it first copies its molecular instruction manual—its genes—and then passes this information on to its offspring. This cycle has been repeated for three and a half billion years.

But how did we and our very distant relatives come to look so different and develop so many different ways of getting along in the world? A century ago, researchers began to answer that question with the help of a science called genetics. Get a refresher course on the basics in **Chapter 1, "How Genes Work."**

It's likely that when you think of heredity you think first of DNA, but in the past few years, researchers have made surprising findings about

another molecular actor that plays a starring role. Check out the modern view of RNA in **Chapter 2, "RNA and DNA Revealed: New Roles, New Rules."**

When genetics first started, scientists didn't have the tools they have today. They could only look at one gene, or a few genes, at a time. Now, researchers can examine all of the genes in a living organism—its genome—at once. They are doing this for organisms on every branch of the tree of life and finding that the genomes of mice, frogs, fish and a slew of other creatures have many genes similar to our own.

So why doesn't your brother look like your dog or the fish in your aquarium? It's because of evolution. In **Chapter 3, "Life's Genetic Tree,"** find out how evolution works and how it relates to genetics and medical research.

Can DNA and RNA help doctors predict whether we'll get diseases like cancer, diabetes or asthma? What other mysteries are locked within the 6 feet of DNA inside nearly every cell in our bodies? **Chapter 4, "Genes Are Us,"** explains what researchers know, and what they are still learning, about the role of genes in health and disease.

Finally, in **Chapter 5, "21st-Century Genetics,"** see a preview of things to come. Learn how medicine and science are changing in big ways, and how these changes influence society.

From metabolism to medicines to agriculture, the science of genetics affects us every day. It is part of life … part of *your* life!

How Genes Work

People have known for many years that living things inherit traits from their parents. That common-sense observation led to agriculture, the purposeful breeding and cultivation of animals and plants for desirable characteristics. Firming up the details took quite some time, though. Researchers did not understand exactly how traits were passed to the next generation until the middle of the 20th century.

Now it is clear that **genes** are what carry our traits through generations and that genes are made of **deoxyribonucleic acid (DNA)**. But genes themselves don't do the actual work. Rather, they serve as instruction books for making functional molecules such as **ribonucleic acid (RNA)** and **proteins**, which perform the chemical reactions in our bodies.

Proteins do many other things, too. They provide the body's main building materials, forming the cell's architecture and structural components. But one thing proteins can't do is make copies of themselves. When a cell needs more proteins, it uses the manufacturing instructions coded in DNA.

The DNA code of a gene—the sequence of its individual DNA building blocks, labeled A (adenine), T (thymine), C (cytosine) and G (guanine) and collectively called **nucleotides**—spells out the exact order of a protein's building blocks, **amino acids**.

Occasionally, there is a kind of typographical error in a gene's DNA sequence. This mistake—which can be a change, gap or duplication—is called a **mutation**.

Genetics in the Garden

The monk Gregor Mendel first described how traits are inherited from one generation to the next.

In 1900, three European scientists independently discovered an obscure research paper that had been published nearly 35 years before. Written by Gregor Mendel, an Austrian monk who was also a scientist, the report described a series of breeding experiments performed with pea plants growing in his abbey garden.

Mendel had studied how pea plants inherited the two variant forms of easy-to-see traits. These included flower color (white or purple) and the texture of the peas (smooth or wrinkled). Mendel counted many generations of pea plant offspring and learned that these characteristics were passed on to the next generation in orderly, predictable ratios.

When he cross-bred purple-flowered pea plants with white-flowered ones, the next generation had only purple flowers. But directions for making white flowers were hidden somewhere in the peas of that generation, because when those purple-flowered

A mutation can cause a gene to encode a protein that works incorrectly or that doesn't work at all. Sometimes, the error means that no protein is made.

But not all DNA changes are harmful. Some mutations have no effect, and others produce new versions of proteins that may give a survival advantage to the organisms that have them. Over time, mutations supply the raw material from which new life forms evolve (see Chapter 3, "Life's Genetic Tree").

Beautiful DNA

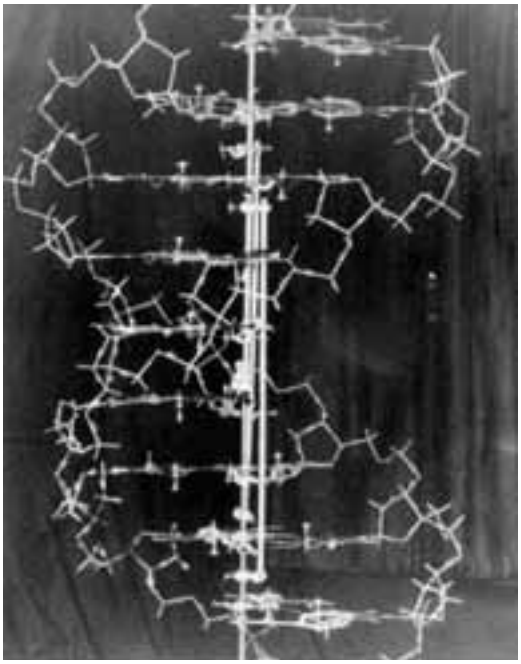
Up until the 1950s, scientists knew a good deal about heredity, but they didn't have a clue what DNA looked like. In order to learn more about DNA and its structure, some scientists experimented with using X rays as a form of molecular photography.

Rosalind Franklin, a physical chemist working with Maurice Wilkins at King's College in London, was among the first to use this method to analyze genetic material. Her experiments

plants were bred to each other, some of their offspring had white flowers. What's more, the second-generation plants displayed the colors in a predictable pattern. On average, 75 percent of the second-generation plants had purple flowers and 25 percent of the plants had white flowers. Those same ratios persisted, and were reproduced when the experiment was repeated many times over.

Trying to solve the mystery of the missing color blooms, Mendel imagined that the reproductive cells of his pea plants might contain discrete "factors," each of which specified a particular trait, such as white flowers. Mendel reasoned that the factors, whatever they were, must be physical material because they passed from parent to offspring in a mathematically orderly way. It wasn't until many years later, when the other scientists unearthed Mendel's report, that the factors were named genes.

Early geneticists quickly discovered that Mendel's mathematical rules of inheritance applied not just to peas, but also to all plants, animals and people. The discovery of a quantitative rule for inheritance was momentous. It revealed that a common, general principle governed the growth and development of all life on Earth.



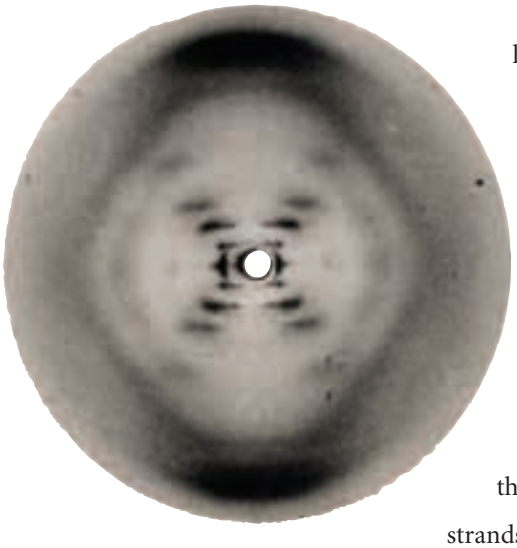
COLD SPRING HARBOR LABORATORY ARCHIVES

▲ In 1953, Watson and Crick created their historic model of the shape of DNA: the double helix.

produced what were referred to at the time as "the most beautiful X-ray photographs of any substance ever taken."

Other scientists, including zoologist James Watson and physicist Francis Crick, both working at Cambridge University in the United Kingdom, were trying to determine the shape of DNA too. Ultimately, this line of research revealed one of the most profound scientific discoveries of the 20th century: that DNA exists as a double helix.

The 1962 Nobel Prize in physiology or medicine was awarded to Watson, Crick and Wilkins for this work. Although Franklin did not earn a share of the prize due to her untimely death at age 38, she is widely recognized as having played a significant role in the discovery.



▲ Rosalind Franklin's original X-ray diffraction photo revealed the physical structure of DNA.

OREGON STATE UNIVERSITY LIBRARIES SPECIAL COLLECTIONS

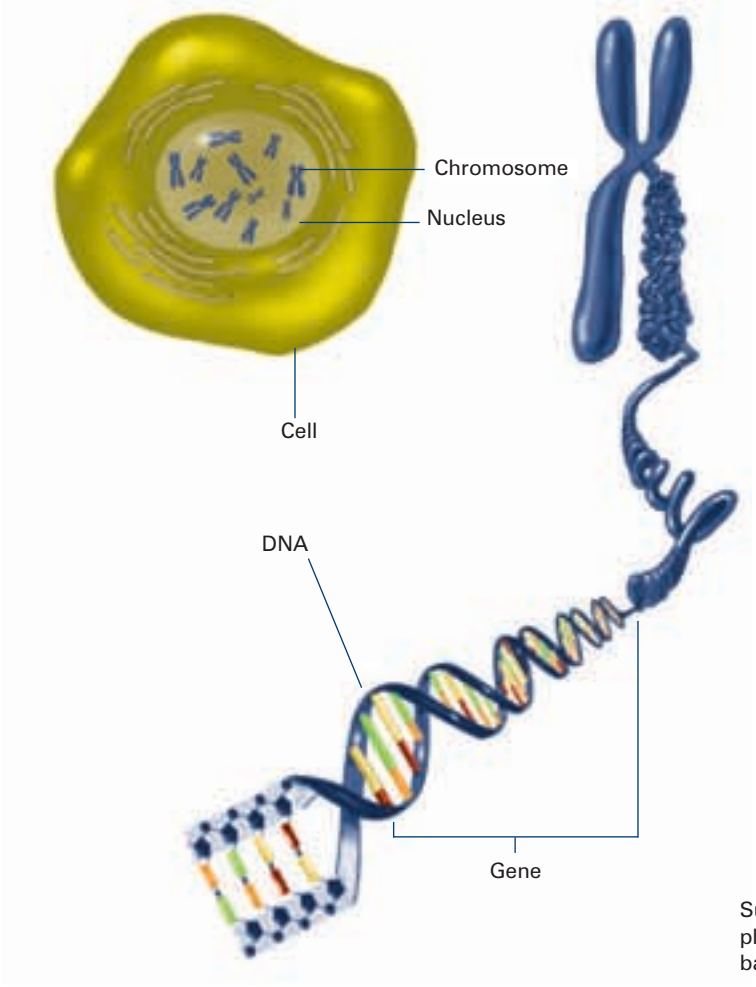
The spiral staircase-shaped double helix has attained global status as the symbol for DNA. But what is so beautiful about the discovery of the twisting ladder structure isn't just its good looks. Rather, the structure of DNA taught researchers a fundamental lesson about **genetics**. It taught them that the two connected strands—winding together like parallel handrails—were complementary to each other, and this unlocked the secret of how genetic information is stored, transferred and copied.

In genetics, complementary means that if you know the sequence of nucleotide building blocks on one strand, you know the sequence of nucleotide building blocks on the other strand: A always matches up with T and C always links to G (see drawing, page 7).

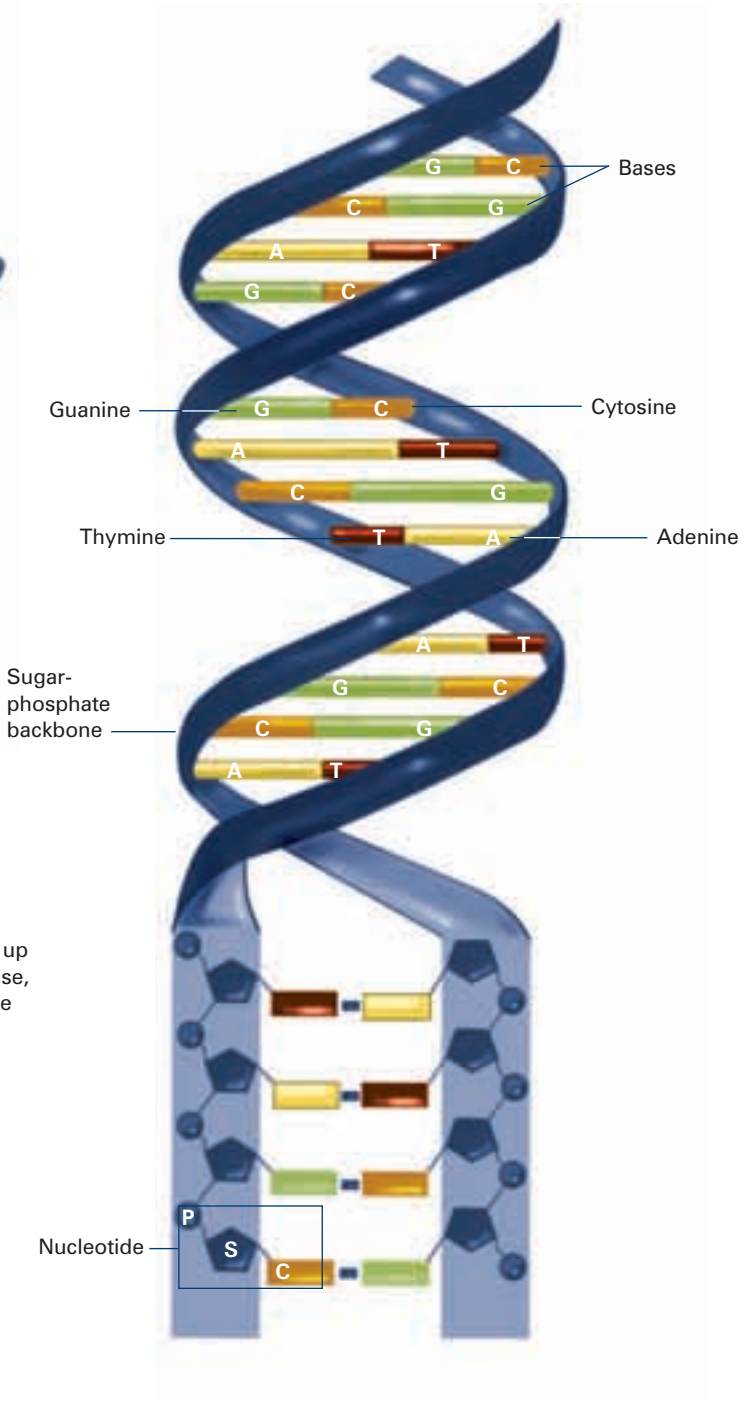
Long strings of nucleotides form genes, and groups of genes are packaged tightly into structures called **chromosomes**. Every cell in your body except for eggs, sperm and red blood cells contains a full set of chromosomes in its **nucleus**.

If the chromosomes in one of your cells were uncoiled and placed end to end, the DNA would be about 6 feet long. If all the DNA in your body were connected in this way, it would stretch approximately 67 billion miles! That's nearly 150,000 round trips to the Moon.

DNA Structure



The long, stringy DNA that makes up genes is spooled within chromosomes inside the nucleus of a cell. (Note that a gene would actually be a much longer stretch of DNA than what is shown here.)



DNA consists of two long, twisted chains made up of nucleotides. Each nucleotide contains one base, one phosphate molecule and the sugar molecule deoxyribose. The bases in DNA nucleotides are adenine, thymine, cytosine and guanine.

▲ Humans have 23 pairs of chromosomes. Male DNA (pictured here) contains an X and a Y chromosome, whereas female DNA contains two X chromosomes.

CYTOGENETICS LABORATORY, BRIGHAM AND WOMEN'S HOSPITAL

Copycat

It's astounding to think that your body consists of trillions of cells. But what's most amazing is that it all starts with one cell. How does this massive expansion take place?

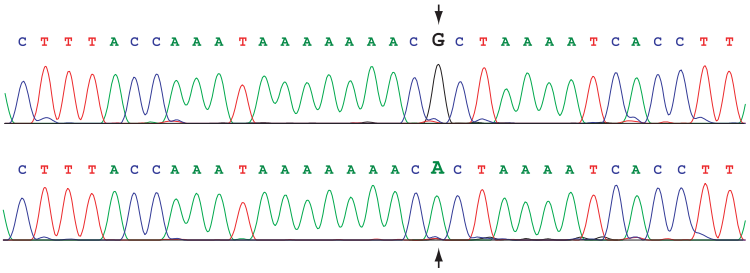
As an embryo progresses through development, its cells must reproduce. But before a cell divides into two new, nearly identical cells, it must copy its DNA so there will be a complete set of genes to pass on to each of the new cells.

To make a copy of itself, the twisted, compacted double helix of DNA has to unwind and separate its two strands. Each strand becomes a pattern, or template, for making a new strand, so the two new DNA molecules have one new strand and one old strand.

The copy is courtesy of a cellular protein machine called **DNA polymerase**, which reads the template DNA strand and stitches together the complementary new strand. The process, called **replication**, is astonishingly fast and accurate, although occasional mistakes, such as deletions or duplications, occur. Fortunately, a cellular spell-checker catches and corrects nearly all of these errors.

Mistakes that are not corrected can lead to diseases such as cancer and certain genetic disorders. Some of these include Fanconi anemia, early aging diseases and other conditions in which people are extremely sensitive to sunlight and some chemicals.

DNA copying is not the only time when DNA damage can happen. Prolonged, unprotected sun exposure can cause DNA changes that lead to skin cancer, and toxins in cigarette smoke can cause lung cancer.



▲ When DNA polymerase makes an error while copying a gene's DNA sequence, the mistake is called a mutation. In this example, the nucleotide G has been changed to an A.

It may seem ironic, then, that many drugs used to treat cancer work by attacking DNA. That's because these chemotherapy drugs disrupt the DNA copying process, which goes on much faster in rapidly dividing cancer cells than in other cells of the body. The trouble is that most of these drugs do affect normal cells that grow and divide frequently, such as cells of the immune system and hair cells.

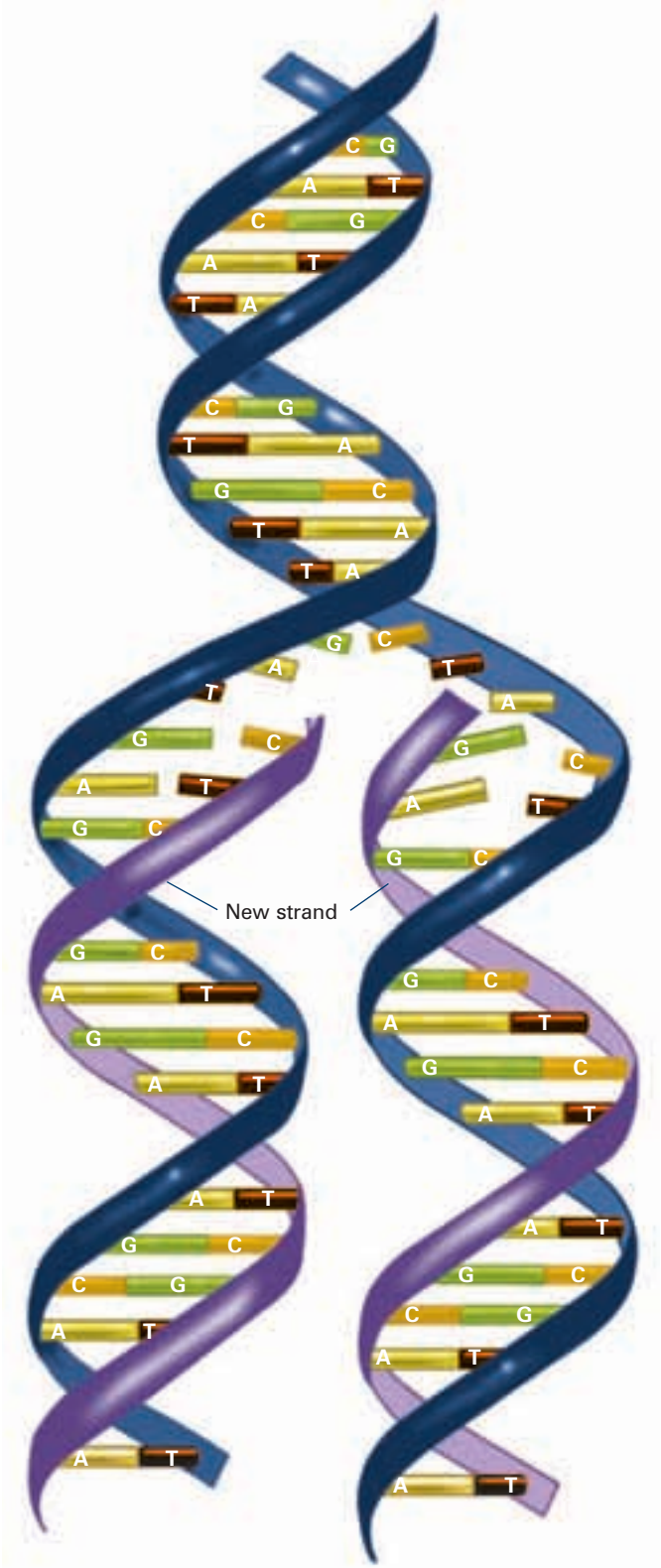
Understanding DNA replication better could be a key to limiting a drug's action to cancer cells only.

Let's Call It Even

After copying its DNA, a cell's next challenge is getting just the right amount of genetic material into each of its two offspring.

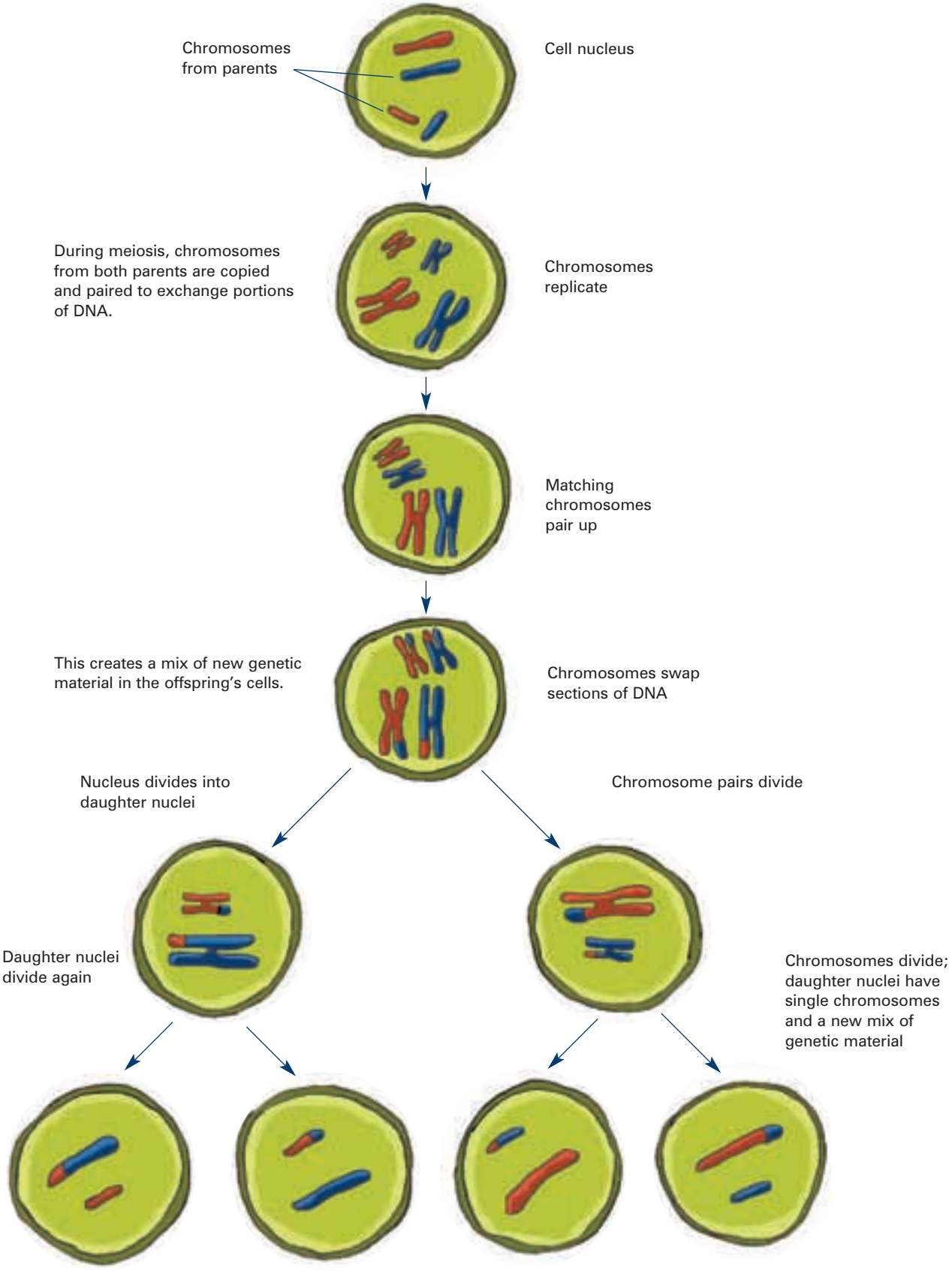
Most of your cells are called **diploid** ("di" means two, and "ploid" refers to sets of chromosomes) because they have two sets of chromosomes (23 pairs). Eggs and sperm are different; these are known as **haploid** cells. Each haploid cell has only one set of 23 chromosomes so that at fertilization the math will work out: A haploid egg cell will combine with a haploid sperm cell to form a diploid cell with the right number of chromosomes: 46.

Chromosomes are numbered 1 to 22, according to size, with 1 being the largest chromosome. The 23rd pair, known as the sex chromosomes, are called X and Y. In humans, abnormalities of chromosome number usually occur during **meiosis**, the time when a cell



▲ During DNA replication, each strand of the original molecule acts as a template for the synthesis of a new, complementary DNA strand.

Meiosis


Chromosomes from parents
Cell nucleus
During meiosis, chromosomes from both parents are copied and paired to exchange portions of DNA.
Chromosomes replicate
Matching chromosomes pair up
This creates a mix of new genetic material in the offspring's cells.
Chromosomes swap sections of DNA
Nucleus divides into daughter nuclei
Chromosome pairs divide
Daughter nuclei divide again
Chromosomes divide; daughter nuclei have single chromosomes and a new mix of genetic material

reduces its chromosomes from diploid to haploid in creating eggs or sperm.

What happens if an egg or a sperm cell gets the wrong number of chromosomes, and how often does this happen?

Molecular biologist Angelika Amon of the Massachusetts Institute of Technology in Cambridge says that mistakes in dividing DNA between daughter cells during meiosis are the leading cause of human birth defects and miscarriages. Current estimates are that 10 percent of all embryos have an incorrect chromosome number. Most of these don't go to full term and are miscarried.

In women, the likelihood that chromosomes won't be apportioned properly increases with age. One of every 18 babies born to women over 45 has three copies of chromosome 13, 18 or 21 instead of the normal two, and this improper balancing can cause trouble. For example, three copies of chromosome 21 lead to Down syndrome.

To make her work easier, Amon—like many other basic scientists—studies yeast cells, which separate their chromosomes almost exactly the same way human cells do, except that yeast do it much faster. A yeast cell copies its DNA and produces daughter cells in about $1\frac{1}{2}$ hours, compared to a whole day for human cells.

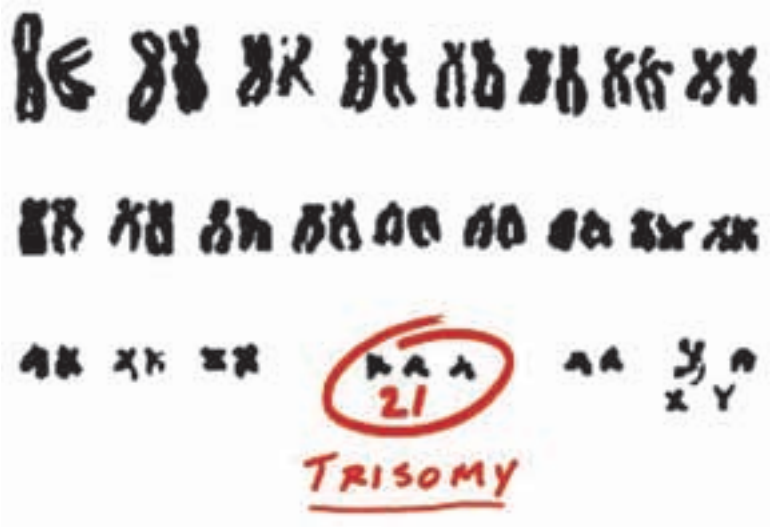
The yeast cells she uses are the same kind bakeries use to make bread and breweries use to make beer!

Amon has made major progress in understanding the details of meiosis. Her research shows how, in healthy cells, gluelike protein complexes called cohesins release pairs of chromosomes at exactly the right time. This allows the chromosomes to separate properly.

These findings have important implications for understanding and treating infertility, birth defects and cancer.

Getting the Message

So, we've described DNA—its basic properties and how our bodies make more of it. But how does DNA serve as the language of life? How do you get a protein from a gene?



▲ Trisomy, the hallmark of Down syndrome, results when a baby is born with three copies of chromosome 21 instead of the usual two.

There are two major steps in making a protein. The first is **transcription**, where the information coded in DNA is copied into RNA. The RNA nucleotides are complementary to those on the DNA: a C on the RNA strand matches a G on the DNA strand.

The only difference is that RNA pairs a nucleotide called uracil (U), instead of a T, with an A on the DNA.

A protein machine called **RNA polymerase** reads the DNA and makes the RNA copy. This copy is called messenger RNA, or mRNA, because it delivers the gene's message to the protein-producing machinery.

At this point you may be wondering why all of the cells in the human body aren't exactly alike, since they all contain the same DNA. What makes a liver cell different from a brain cell? How do the cells in the heart make the organ contract, but those in skin allow us to sweat?

Cells can look and act differently, and do entirely different jobs, because each cell "turns on," or expresses, only the genes appropriate for what it needs to do.

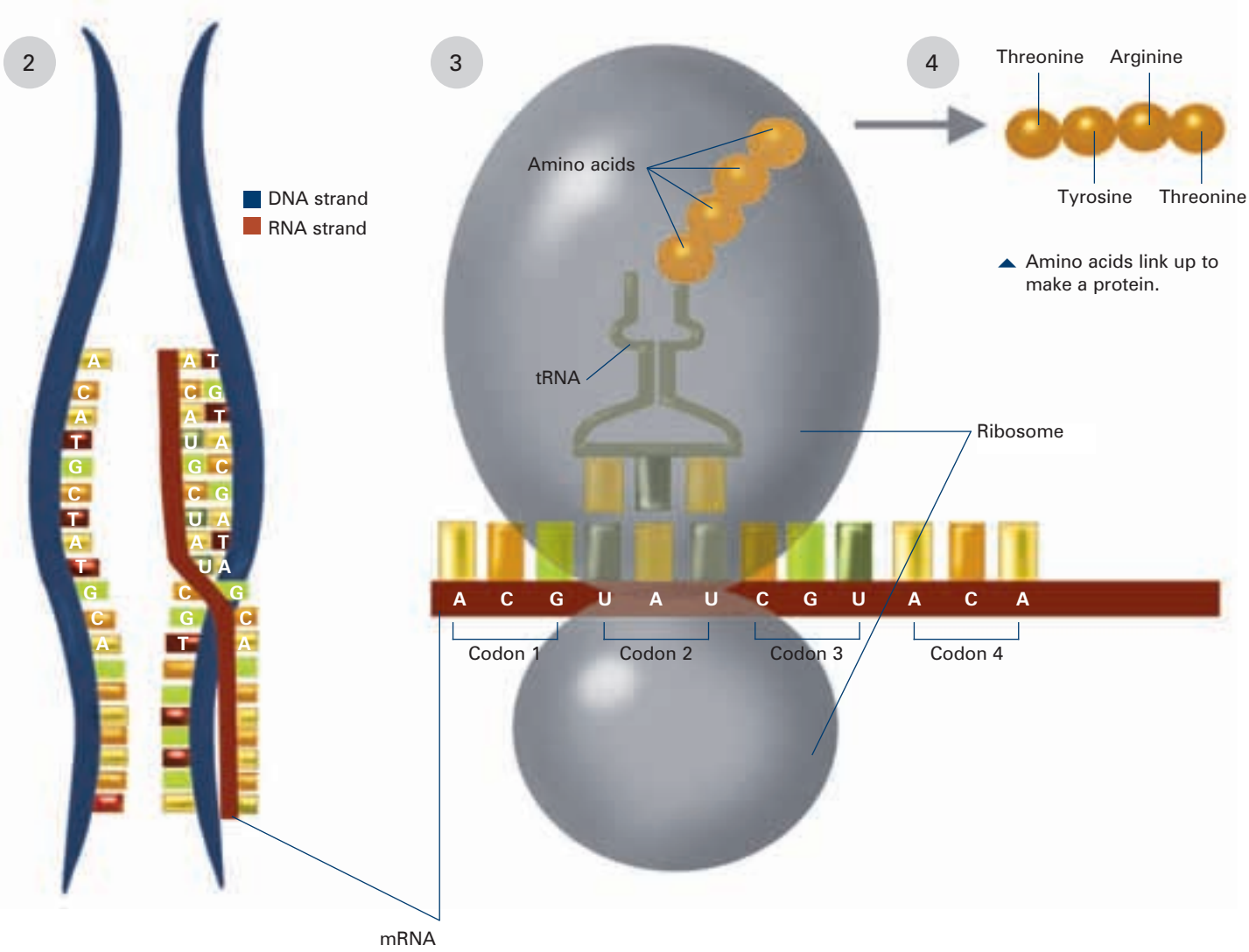
That's because RNA polymerase does not work alone, but rather functions with the aid of many helper proteins. While the core part of RNA polymerase is the same in all cells, the helpers vary in different cell types throughout the body.

You'd think that for a process so essential to life, researchers would know a lot about how transcription works. While it's true that the basics are clear—biologists have been studying gene transcribing by RNA polymerases since these proteins were first discovered in 1960—some of the details are actually still murky.

▲ RNA polymerase transcribes DNA to make messenger RNA (mRNA).

The biggest obstacle to learning more has been a lack of tools. Until fairly recently, researchers were unable to get a picture at the atomic level of the giant RNA polymerase protein assemblies inside cells to understand how the many pieces of this amazing, living machine do what they do, and do it so well.

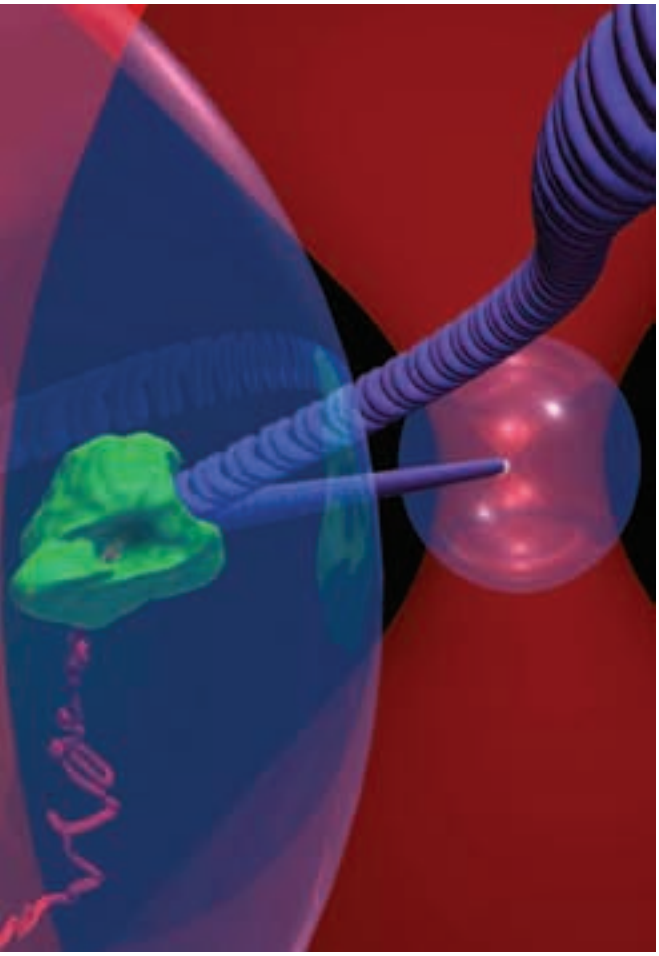
But our understanding is improving fast, thanks to spectacular technological advances. We have new X-ray pictures that are far more sophisticated than those that revealed the structure of DNA. Roger Kornberg of Stanford University in California used such methods to determine the structure of RNA polymerase. This work earned



▲ The mRNA sequence (dark red strand) is complementary to the DNA sequence (blue strand).

▲ On ribosomes, transfer RNA (tRNA) helps convert mRNA into protein.

▲ Amino acids link up to make a protein.



▲ RNA polymerase (green) and one end of a DNA strand (blue) are attached to clear beads pinned down in two optical traps. As RNA polymerase moves along the DNA, it creates an RNA copy of a gene, shown here as a pink strand.

STEVEN BLOCK

him the 2006 Nobel Prize in chemistry. In addition, very powerful microscopes and other tools that allow us to watch one molecule at a time provide a new look at RNA polymerase while it's at work reading DNA and producing RNA.

For example, Steven Block, also of Stanford, has used a physics technique called optical trapping to track RNA polymerase as it inches along DNA. Block and his team performed this work by designing a specialized microscope sensitive enough to watch the real-time motion of a single polymerase traveling down a gene on one chromosome.

The researchers discovered that molecules of RNA polymerase behave like battery-powered spiders as they crawl along the DNA ladder, adding nucleotides one at a time to the growing RNA strand. The **enzyme** works much like a motor, Block believes, powered by energy released during the chemical synthesis of RNA.

Nature's Cut-and-Paste Job

Several types of RNA play key roles in making a protein. The gene transcript (the mRNA) transfers information from DNA in the nucleus to the **ribosomes** that make protein. Ribosomal RNA forms about 60 percent of the ribosomes. Lastly, transfer RNA carries amino acids to the ribosomes. As you can see, all three types of cellular RNAs come together to produce new proteins.

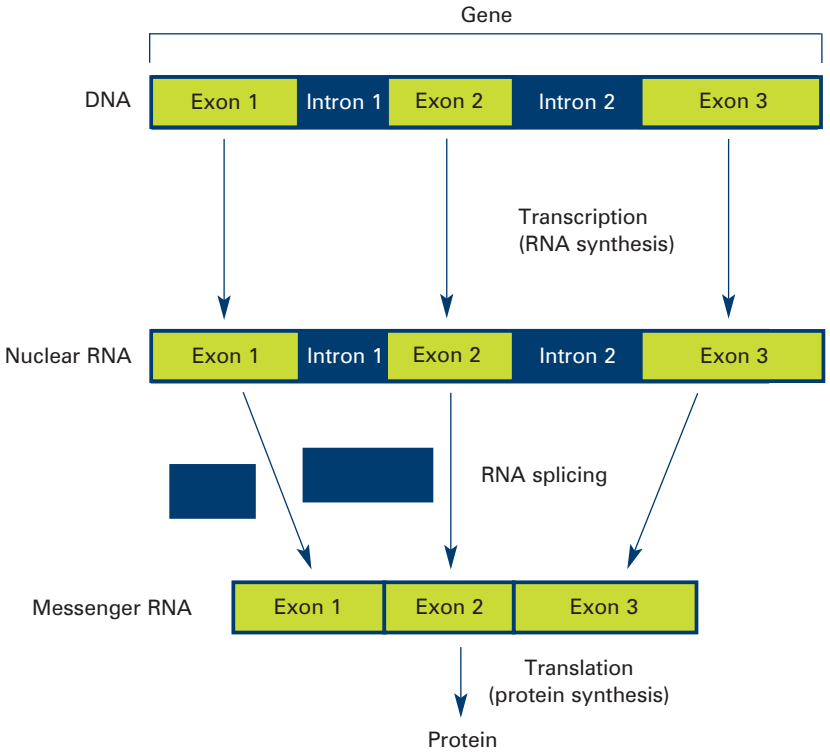
But the journey from gene to protein isn't quite as simple as we've just made it out to be. After transcription, several things need to happen to mRNA before a protein can be made. For example, the genetic material of humans and other **eukaryotes** (organisms that have a nucleus) includes a lot of DNA that doesn't encode proteins. Some of this DNA is stuck right in the middle of genes.

To distinguish the two types of DNA, scientists call the coding sequences of genes **exons** and the pieces in between **introns** (for intervening sequences).

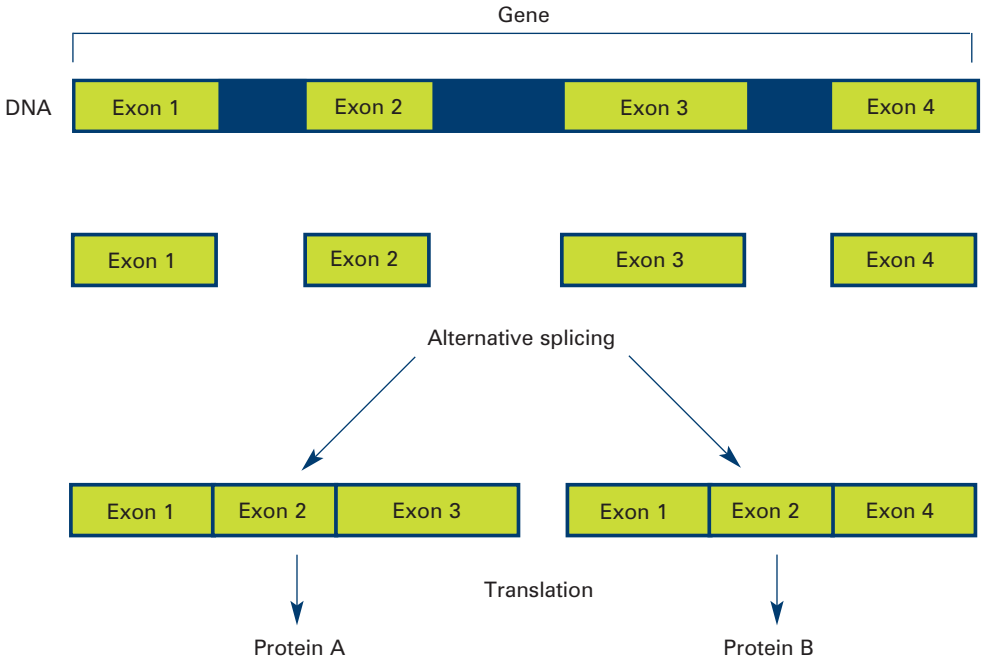
If RNA polymerase were to transcribe DNA from the start of an intron-containing gene to the end, the RNA would be complementary to the introns as well as the exons.

To get an mRNA molecule that yields a working protein, the cell needs to trim out the intron sections and then stitch only the exon pieces together (see drawing, page 15). This process is called **RNA splicing.**

RNA Splicing



Genes are often interrupted by stretches of DNA (introns, blue) that do not contain instructions for making a protein. The DNA segments that do contain protein-making instructions are known as exons (green).



Arranging exons in different patterns, called alternative splicing, enables cells to make different proteins from a single gene.

Splicing has to be extremely accurate. An error in the splicing process, even one that results in the deletion of just one nucleotide in an exon or the addition of just one nucleotide in an intron, will throw the whole sequence out of alignment. The result is usually an abnormal protein—or no protein at all. One form of Alzheimer's disease, for example, is caused by this kind of splicing error.

Molecular biologist Christine Guthrie of the University of California, San Francisco, wants to understand more fully the mechanism for removing intron RNA and find out how it stays so accurate.

She uses yeast cells for these experiments. Just like human DNA, yeast DNA has introns, but they are fewer and simpler in structure and are therefore easier to study. Guthrie can identify which genes are required for splicing by finding abnormal yeast cells that mangle splicing.

So why do introns exist, if they're just going to be chopped out? Without introns, cells wouldn't need to go through the splicing process and keep monitoring it to be sure it's working right.

As it turns out, splicing also makes it possible for cells to create more proteins.

Think about all the exons in a gene. If a cell stitches together exons 1, 2 and 4, leaving out exon 3, the mRNA will specify the production of a particular protein. But instead, if the cell stitches together exons 1, 2 and 3, this time leaving out exon 4, then the mRNA will be translated into a different protein (see drawing, page 15).

By cutting and pasting the exons in different patterns, which scientists call alternative splicing, a cell can create different proteins from a single gene. Alternative splicing is one of the reasons why human cells, which have about 20,000 genes, can make hundreds of thousands of different proteins.

All Together Now

Until recently, researchers looked at genes, and the proteins they encode, one at a time. Now, they can look at how large numbers of genes and proteins act, as well as how they interact. This gives them a much better picture of what goes on in a living organism.

Already, scientists can identify all of the genes that are transcribed in a cell—or in an organ, like the heart. And although researchers can't tell you, right now, what's going on in every cell of your body while you read a book or walk down the street, they can do this sort of "whole-body" scan for simpler, single-celled organisms like yeast.

Using a technique called genome-wide location analysis, Richard Young of the Massachusetts Institute of Technology unraveled a "regulatory code" of living yeast cells, which have more than 6,000 genes in their genome. Young's technique enabled him to determine the exact places where RNA polymerase's helper proteins sit on DNA and tell RNA polymerase to begin transcribing a gene.

Since he did the experiment with the yeast exposed to a variety of different conditions,

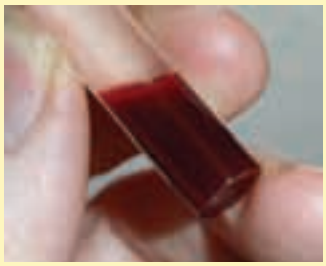
GENETICS AND YOU: *Nursery Genetics*

While most genetic research uses lab organisms, test tubes and petri dishes, the results have real consequences for people. Your first encounter with genetic analysis probably happened shortly after you were born, when a doctor or nurse took a drop of blood from the heel of your tiny foot.

Lab tests performed with that single drop of blood can diagnose certain rare genetic disorders as well as metabolic problems like phenylketonuria (PKU).

Screening newborns in this way began in the 1960s in Massachusetts with testing for PKU, a disease affecting 1 in 14,000 people. PKU is caused by an enzyme that doesn't work properly due to a genetic mutation. Those born with this disorder cannot metabolize the amino acid phenylalanine, which is present in many foods. Left untreated, PKU can lead to mental retardation and neurological damage, but a special diet can prevent these outcomes. Testing for this condition has made a huge difference in many lives.

Newborn screening is governed by individual states. This means that the state in which a baby is born determines the genetic conditions for which he or she will be screened. Currently, states test for between 28 and 54 conditions. All states test for PKU.



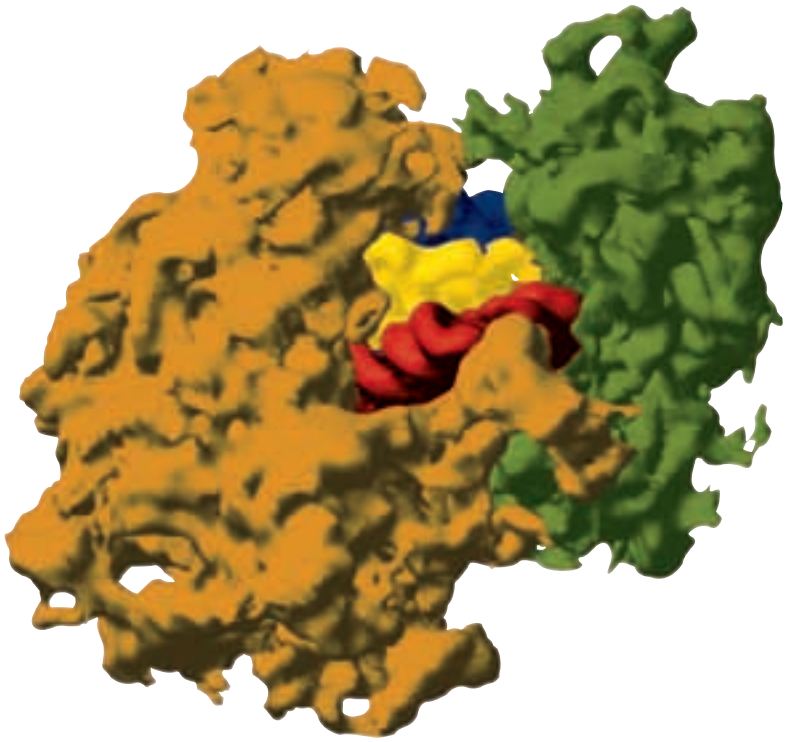
Although expanded screening for genetic diseases in newborns is advocated by some, others question the value of screening for conditions that are currently untreatable. Another issue is that some children with mild versions of certain genetic diseases may be treated needlessly.

In 2006, the Advisory Committee on Heritable Disorders in Newborns and Children, which assists the Secretary of the U.S. Department of Health and Human Services, recommended a standard, national set of newborn tests for 29 conditions, ranging from relatively common hearing problems to very rare metabolic diseases.

Young was able to figure out how transcription patterns differ when the yeast cell is under stress (say, in a dry environment) or thriving in a sugary-rich nutrient solution. Done one gene at a time, using methods considered state-of-the-art just a few years ago, this kind of analysis would have taken hundreds of years.

After demonstrating that his technique worked in yeast, Young then took his research a step forward. He used a variation of the yeast method to scan the entire human genome in small samples of cells taken from the pancreases and livers of people with type 2 diabetes. He used the results to identify genes that aren't transcribed correctly in people with the disease.

This information provides researchers with an important tool for understanding how diabetes and other diseases are influenced by defective genes. By building models to predict how genes respond in diverse situations, researchers may be able to learn how to stop or jump-start genes on demand, change the course of a disease or prevent it from ever happening.



▲ A ribosome consists of large and small protein subunits with transfer RNAs nestled in the middle.

RIBOSOME STRUCTURE COURTESY OF JAMIE CATE, MARAT YUSUPOV, GULNARA YUSUPOVA, THOMAS EARNEST AND HARRY NOLLER. GRAPHIC COURTESY OF ALBION BAUCOM, UNIVERSITY OF CALIFORNIA, SANTA CRUZ.

Found in Translation

After a gene has been read by RNA polymerase and the RNA is spliced, what happens next in the journey from gene to protein? The next step is reading the RNA information and fitting the building blocks of a protein together. This is called **translation**, and its principal actors are the ribosome and amino acids.

Ribosomes are among the biggest and most intricate structures in the cell. The ribosomes of bacteria contain not only huge amounts of RNA, but also more than 50 different proteins. Human ribosomes have even more RNA and between 70 and 80 different proteins!

Harry Noller of the University of California, Santa Cruz, has found that a ribosome performs several key jobs when it translates the genetic code of mRNA. As the messenger RNA threads through the ribosome protein machine, the

ribosome reads the mRNA sequence and helps recognize and recruit the correct amino acid-carrying transfer RNA to match the mRNA code. The ribosome also links each additional amino acid into a growing protein chain (see drawing, page 13).

For many years, researchers believed that even though RNAs formed a part of the ribosome, the protein portion of the ribosome did all of the work. Noller thought, instead, that maybe RNA, not proteins, performed the ribosome's job. His idea was not popular at first, because at that time it was thought that RNA could not perform such complex functions.

Some time later, however, the consensus changed. Sidney Altman of Yale University in New Haven, Connecticut, and Thomas Cech, who was then at the University of Colorado in Boulder, each discovered that RNA can perform work as complex as that done by protein enzymes. Their "RNA-as-an-enzyme" discovery turned the research world on its head and earned Cech and Altman the 1989 Nobel Prize in chemistry.

Noller and other researchers have continued the painstaking work of understanding ribosomes. In 1999, he showed how different parts of a bacterial ribosome interact with one another and how the ribosome interacts with molecules involved in protein synthesis. These studies provided near proof that the fundamental mechanism of translation is performed by RNA, not by the proteins of the ribosome.

▲ Some first-aid ointments contain the antibiotic neomycin, which treats infections by attacking ribosomes in bacteria.

RNA Surprises

But which ribosomal RNAs are doing the work? Most scientists assumed that RNA nucleotides buried deep within the ribosome complex—the ones that have the same sequence in every species from bacteria to people—were the important ones for piecing the growing protein together.

However, recent research by Rachel Green, who worked with Noller before moving to Johns Hopkins University in Baltimore, Maryland, showed that this is not the case. Green discovered that those RNA nucleotides are not needed for assembling a protein. Instead, she found, the nucleotides do something else entirely: They help the growing protein slip off the ribosome once it's finished.

Noller, Green and hundreds of other scientists work with the ribosomes of bacteria. Why should you care about how bacteria create proteins from their genes?

One reason is that this knowledge is important for learning how to disrupt the actions of disease-causing microorganisms. For example, antibiotics like erythromycin and neomycin work by attacking the ribosomes of bacteria, which are different enough from human ribosomes that our cells are not affected by these drugs.

As researchers gain new information about bacterial translation, the knowledge may lead to more antibiotics for people.

New antibiotics are urgently needed because many bacteria have developed resistance to the current arsenal. This resistance is sometimes the result of changes in the bacteria's ribosomal RNA. It can be difficult to find those small, but critical, changes that may lead to resistance, so it is important to find completely new ways to block bacterial translation.

Green is working on that problem too. Her strategy is to make random mutations to the genes in a bacterium that affect its ribosomes. But what if the mutation disables the ribosome so much that it can't make proteins? Then the bacterium won't grow, and Green wouldn't find it.

Using clever molecular tricks, Green figured out a way to rescue some of the bacteria with defective ribosomes so they could grow. While some of the rescued bacteria have changes in their ribosomal RNA that make them resistant to certain antibiotics (and thus would not make good antibiotic targets) other RNA changes that don't affect resistance may point to promising ideas for new antibiotics.

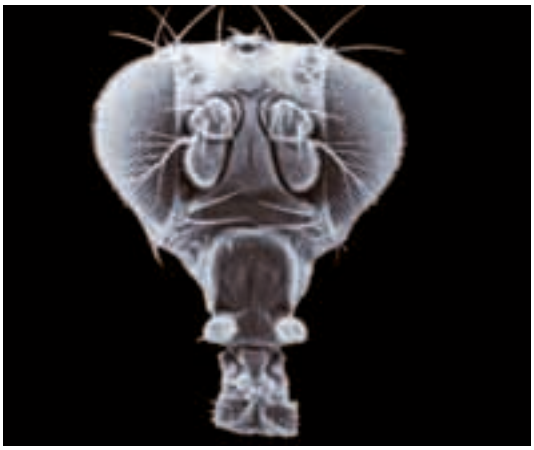
An Interesting Development

In the human body, one of the most important jobs for proteins is to control how embryos develop. Scientists discovered a hugely important set of proteins involved in development by studying mutations that cause bizarre malformations in fruit flies.

The most famous such abnormality is a fruit fly with a leg, rather than the usual antenna, growing out of its head (see page 21). According to Thomas C. Kaufman of Indiana University in Bloomington, the leg is perfectly normal—it's just growing in the wrong place.

In this type of mutation and many others, something goes wrong with the genetic program that directs some of the cells in an embryo to follow developmental pathways, which are a series of chemical reactions that occur in a specific order. In the antenna-into-leg problem, it is as if the cells growing from the fly's head, which normally would become an antenna, mistakenly believe that they are in the fly's thorax, and therefore ought to grow into a leg. And so they do.

Thinking about this odd situation taught scientists an important lesson—that the proteins made by some genes can act as switches. Switch genes are master controllers that provide each body part with a kind of identification card. If a protein that normally instructs cells to become an antenna is disrupted, cells can receive new instructions to become a leg instead.



▲ Normal fruit fly head.

FLYBASE; R. TURNER

▲ Fruit fly head showing the effects of the *Antennapedia* gene. This fly has legs where its antennae should be.

Scientists determined that several different genes, each with a common sequence, provide these anatomical identification card instructions. Kaufman isolated and described one of these genes, which became known as *Antennapedia*, a word that means "antenna feet."

Kaufman then began looking a lot more closely at the molecular structure of the *Antennapedia* gene. In the early 1980s, he and other researchers made a discovery that has been fundamental to understanding evolution as well as developmental biology.

The scientists found a short sequence of DNA, now called the **homeobox**, that is present not only in *Antennapedia* but in the several genes next to it and in genes in many other organisms. When geneticists find very similar DNA sequences in the genes of different organisms, it's a good clue that these genes do something so important and useful that evolution uses the same sequence over and over and permits very few changes in its structure as new species evolve.

Researchers quickly discovered nearly identical versions of homeobox DNA in almost every non-bacterial cell they examined—from yeast to plants, frogs, worms, beetles, chickens, mice and people.

Hundreds of homeobox-containing genes have been identified, and the proteins they make turn out to be involved in the early stages of development of many species. For example, researchers have found that abnormalities in the homeobox genes can lead to extra fingers or toes in humans.

The Tools of Genetics: Mighty Microarrays

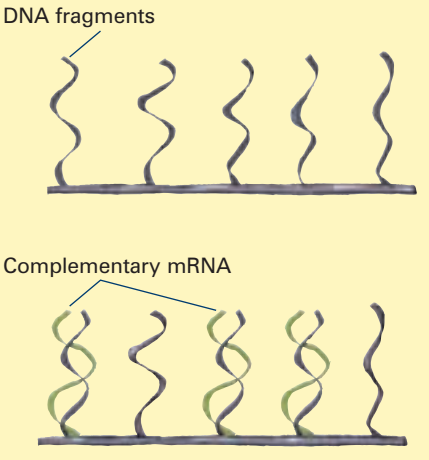
We now have the ability to attach a piece of every gene in a **genome** (all of an organism's genes) to a postage stamp-sized glass microscope slide. This ordered series of DNA spots is called a DNA **microarray**, a **gene chip** or a **DNA chip**.

Whichever name you prefer, the chip could also be called revolutionary. This technology has changed the way many geneticists do their work by making it possible to observe the activity of thousands of genes at once.

In recent years, microarrays have become standard equipment for modern biologists, but teachers and students are using them, too. The Genome Consortium for Active Teaching program (www.bio.davidson.edu/GCAT) provides resources and instructions for high school and college students to do gene-chip experiments in class.

Microarrays are used to get clues about which genes are expressed to control cell, tissue or organ function. By measuring the level of RNA production for every gene at the same time, researchers can learn the genetic programming that makes cell types different and diseased cells different from healthy ones.

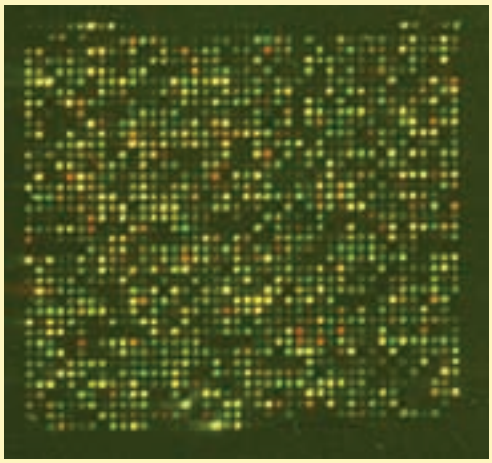
The chips consist of large numbers of DNA fragments distributed in rows in a very small space. The arrays are laid out by robots that can

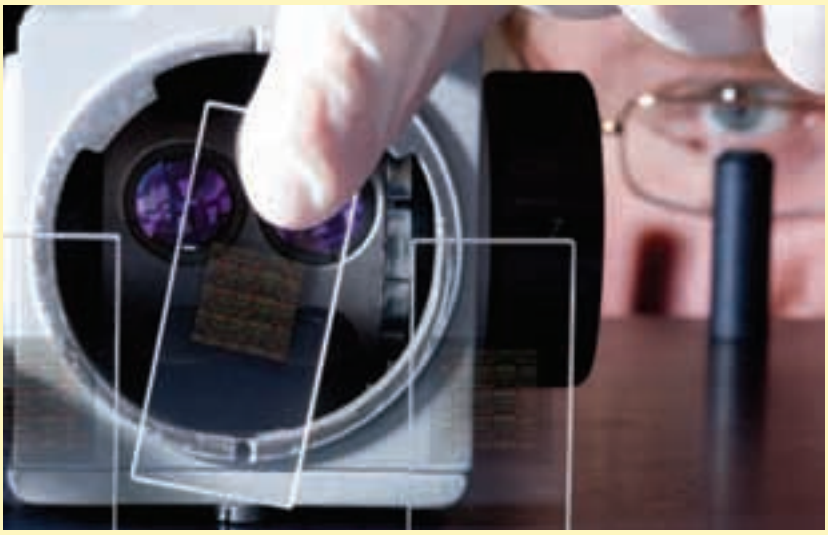


DNA fragments are attached to glass or plastic, then fluorescently tagged molecules are washed over the fragments.

Some molecules (green) bind to their complementary sequence. These molecules can be identified because they glow under fluorescent light.

▼ The resulting pattern of fluorescence indicates which genes are active.





position DNA fragments so precisely that more than 20,000 of them can fit on one microscope slide.

Scientists isolate mRNA from cells grown under two conditions and tag the two sources of RNA with different colors of fluorescent molecules. The two colors of RNA are then placed on the chip, where they attach to complementary DNA fragments anchored to the chip's surface.

Next, a scanner measures the amount of fluorescence at each spot on the chip, revealing how active each gene was (how much mRNA each gene produced). A computer analyzes the patterns of gene activity, providing a snapshot of a genome under two conditions (*e.g.*, healthy or diseased).

In December 2004, the U.S. Food and Drug Administration cleared the first gene chip for medical use. The Amplichip CYP450™, made by Roche Molecular Systems Inc. of Pleasanton, California, analyzes variations in two genes that play a major role in the body's processing of many widely prescribed drugs. This information can help doctors choose the proper dose of certain medicines for an individual patient.

Got It?

Why are some infections hard to treat with antibiotics? What are some things researchers might do to solve this public health problem?

How does DNA work as a form of information storage?

How can 20,000 human genes provide the instructions for making hundreds of thousands of different proteins?

What newborn tests does your area hospital routinely do?

RNA and DNA Revealed: New Roles, New Rules

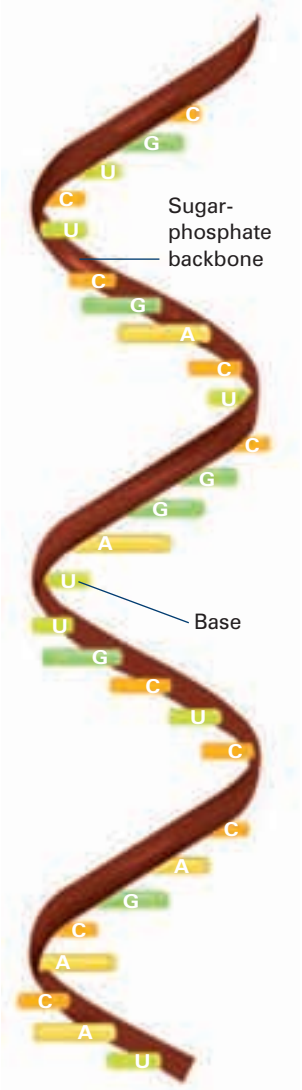
For many years, when scientists thought about heredity, DNA was the first thing to come to mind. It's true that DNA is the basic ingredient of our genes and, as such, it often steals the limelight from RNA, the other form of genetic material inside our cells.

But, while they are both types of genetic material, RNA and DNA are rather different.

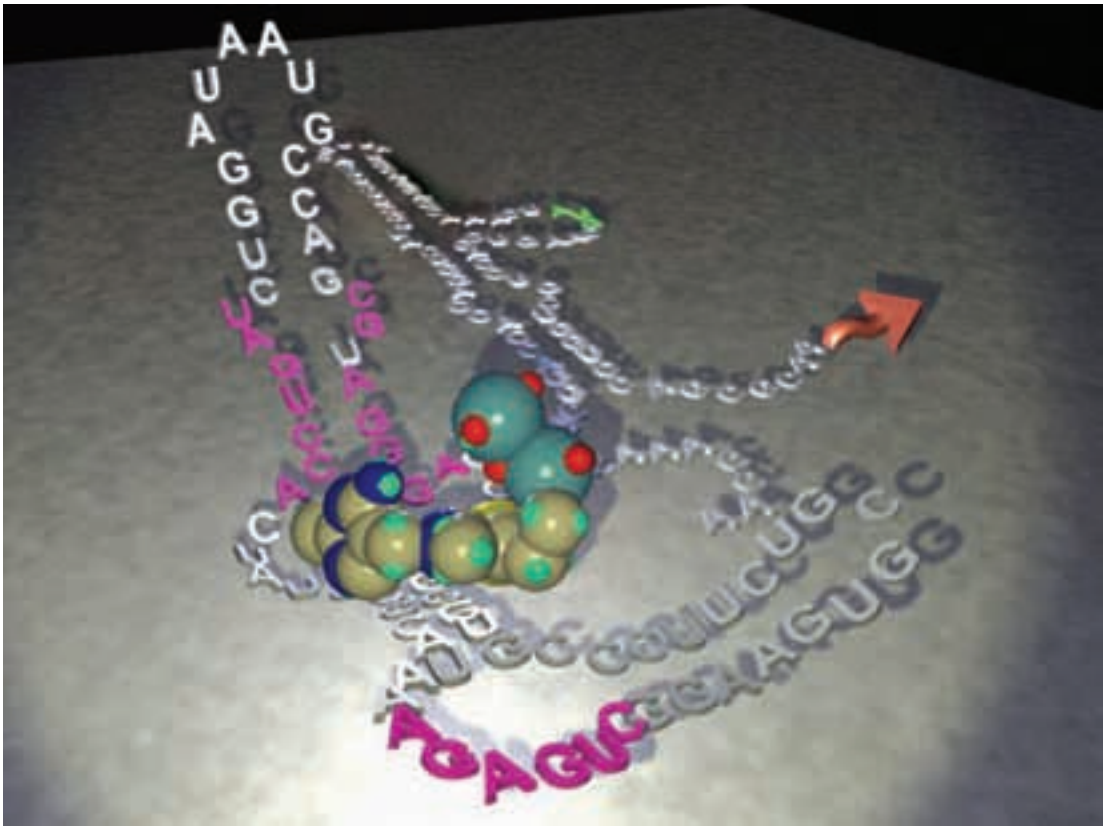
The chemical units of RNA are like those of DNA, except that RNA has the nucleotide uracil (U) instead of thymine (T). Unlike double-stranded DNA, RNA usually comes as only a single strand. And the nucleotides in RNA contain ribose sugar molecules in place of deoxyribose.

RNA is quite flexible—unlike DNA, which is a rigid, spiral-staircase molecule that is very stable. RNA can twist itself into a variety of complicated, three-dimensional shapes. RNA is also unstable in that cells constantly break it down and must continually make it fresh, while DNA is not broken down often. RNA's instability lets cells change their patterns of protein synthesis very quickly in response to what's going on around them.

Many textbooks still portray RNA as a passive molecule, simply a "middle step" in the cell's gene-reading activities. But that view is no longer accurate. Each year, researchers unlock new secrets about RNA. These discoveries reveal that it is truly a remarkable molecule and a multi-talented actor in heredity.



▲ Ribonucleic acid (RNA) has the bases adenine (A), cytosine (C), guanine (G) and uracil (U).



RONALD BREAKER

Riboswitches are RNA sequences that control gene activity. The riboswitch shown here bends into a special shape when it grips tightly onto a molecule called a metabolite (colored balls) that bacteria need to survive.

Today, many scientists believe that RNA evolved on the Earth long before DNA did. Researchers hypothesize—obviously, no one was around to write this down—that RNA was a major participant in the chemical reactions that ultimately spawned the first signs of life on the planet.

RNA World

At least two basic requirements exist for making a cell: the ability to hook molecules together and break them apart, and the ability to replicate, or copy itself, from existing information.

RNA probably helped to form the first cell. The first organic molecules, meaning molecules containing carbon, most likely arose out of random collisions of gases in the Earth's primitive atmosphere, energy from the Sun, and heat from naturally occurring radioactivity. Some scientists think that in this primitive world, RNA was a critical molecule because of its ability to lead a double life: to store information and to conduct chemical reactions. In other words, in this world, RNA served the functions of both DNA and proteins.

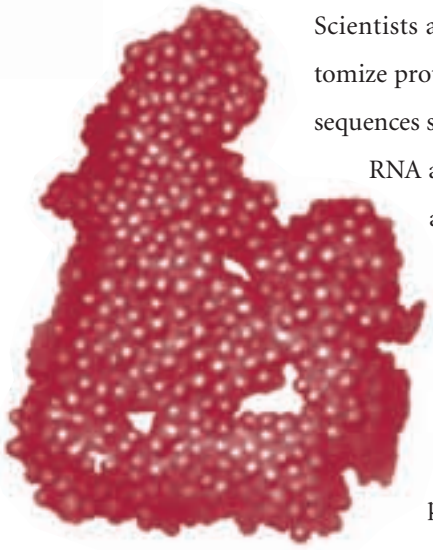
What does any of this have to do with human health? Plenty, it turns out.

Today's researchers are harnessing some of RNA's flexibility and power. For example, through a strategy he calls directed evolution, molecular engineer Ronald R. Breaker of Yale University is developing ways to create entirely new forms of RNA and DNA that both work as enzymes.

Breaker and others have also uncovered a hidden world of RNAs that play a major role in controlling gene activity, a job once thought to be performed exclusively by proteins. These RNAs, which the scientists named riboswitches, are found in a wide variety of bacteria and other organisms.

RNA comes in a variety of different shapes (above and right).

Double-stranded DNA (left) is a staircase-like molecule.



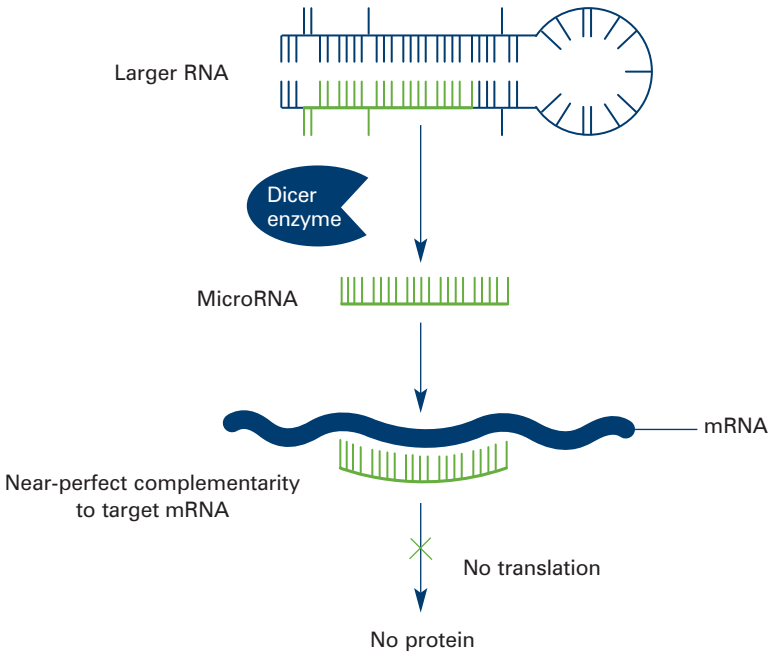
This discovery has led Breaker to speculate that new kinds of antibiotic medicines could be developed to target bacterial riboswitches.

Molecular Editor

Scientists are learning of another way to customize proteins: by RNA editing. Although DNA sequences spell out instructions for producing RNA and proteins, these instructions aren't always followed precisely. Editing a gene's mRNA, even by a single chemical letter, can radically change the resulting protein's function.

Nature likely evolved the RNA editing function as a way to get more proteins out of the same number of

Small But Powerful



The enzyme Dicer generates microRNAs by chopping larger RNA molecules into tiny Velcro®-like pieces. MicroRNAs stick to mRNA molecules and prevent the mRNAs from being made into proteins.

Recently, molecules called **microRNAs** have been found in organisms as diverse as plants, worms and people. The molecules are truly "micro," consisting of only a few dozen nucleotides, compared to typical human mRNAs that are a few thousand nucleotides long.

What's particularly interesting about microRNAs is that many of them arise from DNA that used to be considered merely filler material (see page 14).

How do these small but important RNA molecules do their work? They start out much bigger but get trimmed by cellular enzymes, including one aptly named Dicer. Like tiny pieces of

genes. For example, researchers have found that the mRNAs for certain proteins important for the proper functioning of the nervous system are particularly prone to editing. It may be that RNA editing gives certain brain cells the capacity to react quickly to a changing environment.

Which molecules serve as the editor and how does this happen? Brenda Bass of the University of Utah School of Medicine in Salt Lake City studies one particular class of editors called adenosine deaminases. These enzymes "retype" RNA letters at various places within an mRNA transcript.

They do their job by searching for characteristic RNA shapes. Telltale twists and bends in folded RNA molecules signal these enzymes to change the RNA sequence, which in turn changes the protein that gets made.

Bass' experiments show that RNA editing occurs in a variety of organisms, including people. Another interesting aspect of editing is that certain disease-causing microorganisms, such as some forms of parasites, use RNA editing to gain a survival edge when living in a human host. Understanding the details of this process is an important area of medical research.

AMY PASQUINELLI

Worms with a mutated form of the microRNA let-7 (right) have severe growth problems, rupturing as they develop.

Velcro®, microRNAs stick to certain mRNA molecules and stop them from passing on their protein-making instructions.

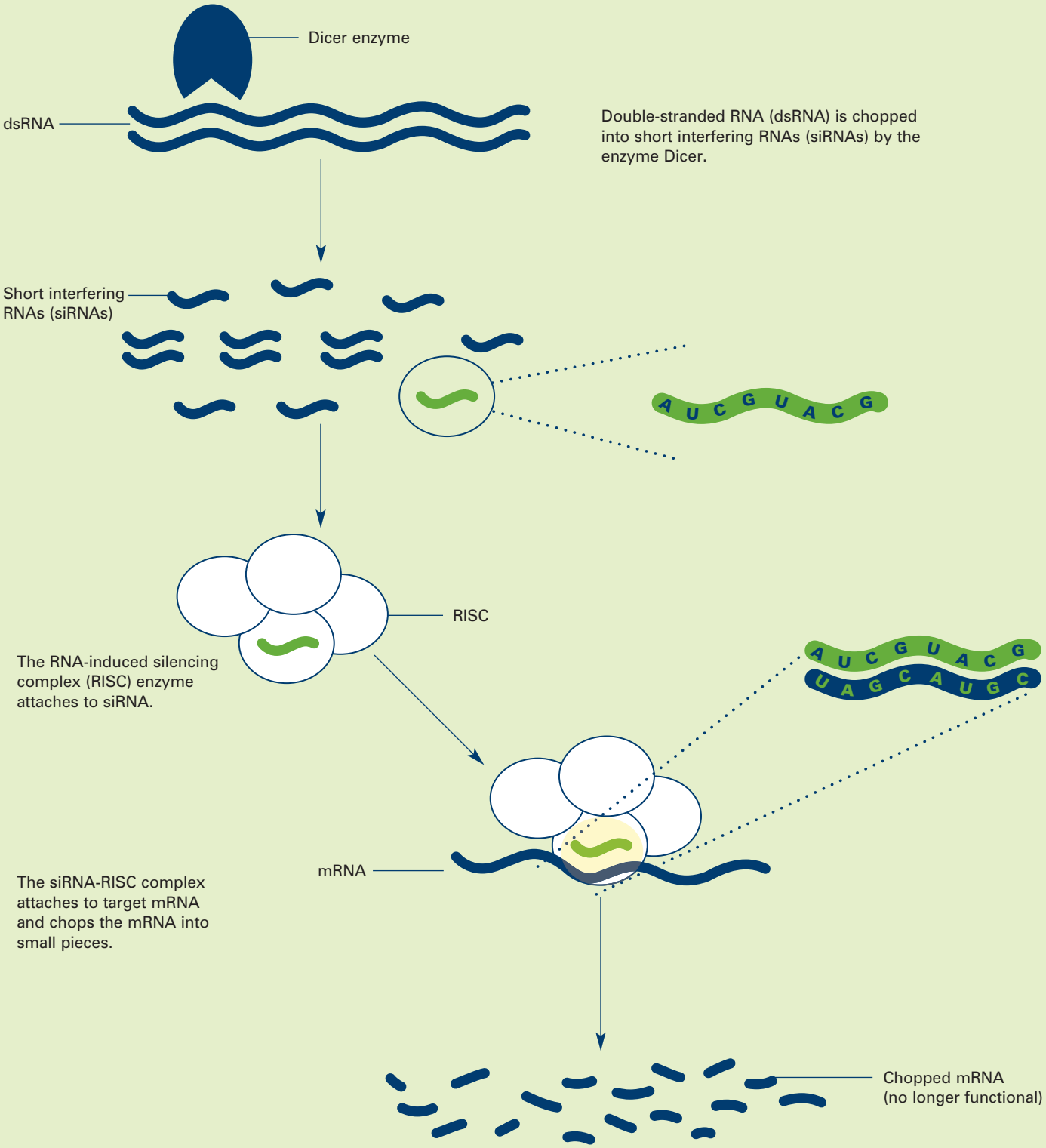
First discovered in a roundworm model system (see *Living Laboratories*, page 49), some microRNAs help determine the organism's body plan. In their absence, very bad things can happen. For example, worms engineered to lack a microRNA called let-7 develop so abnormally that they often rupture and practically break in half as the worm grows.

Perhaps it is not surprising that since microRNAs help specify the timing of an organism's developmental plan, the appearance of the microRNAs themselves is carefully timed inside a developing organism. Biologists, including Amy Pasquinelli of the University of California, San Diego, are currently figuring out how microRNAs are made and cut to size, as well as how they are produced at the proper time during development.

MicroRNA molecules also have been linked to cancer. For example, Gregory Hannon of the Cold Spring Harbor Laboratory on Long Island, New York, found that certain microRNAs are associated with the severity of the blood cancer B-cell lymphoma in mice.

Since the discovery of microRNAs in the first years of the 21st century, scientists have identified hundreds of them that likely exist as part of a large family with similar nucleotide sequences. New roles for these molecules are still being found.

RNA Interference (RNAi)



Healthy Interference

RNA controls genes in a way that was only discovered recently: a process called **RNA interference**, or **RNAi**. Although scientists identified RNAi less than 10 years ago, they now know that organisms have been using this trick for millions of years.

Researchers believe that RNAi arose as a way to reduce the production of a gene's encoded protein for purposes of fine-tuning growth or self-defense. When viruses infect cells, for example, they command their host to produce specialized RNAs that allow the virus to survive and make copies of itself. Researchers believe that RNAi eliminates unwanted viral RNA, and some speculate that it may even play a role in human immunity.

Oddly enough, scientists discovered RNAi from a failed experiment! Researchers investigating genes involved in plant growth noticed something strange: When they tried to turn petunia flowers purple by adding an extra "purple" gene, the flowers bloomed white instead.

This result fascinated researchers, who could not understand how adding genetic material could somehow get rid of an inherited trait. The mystery remained unsolved until, a few years later, two geneticists studying development saw a similar thing happening in lab animals.

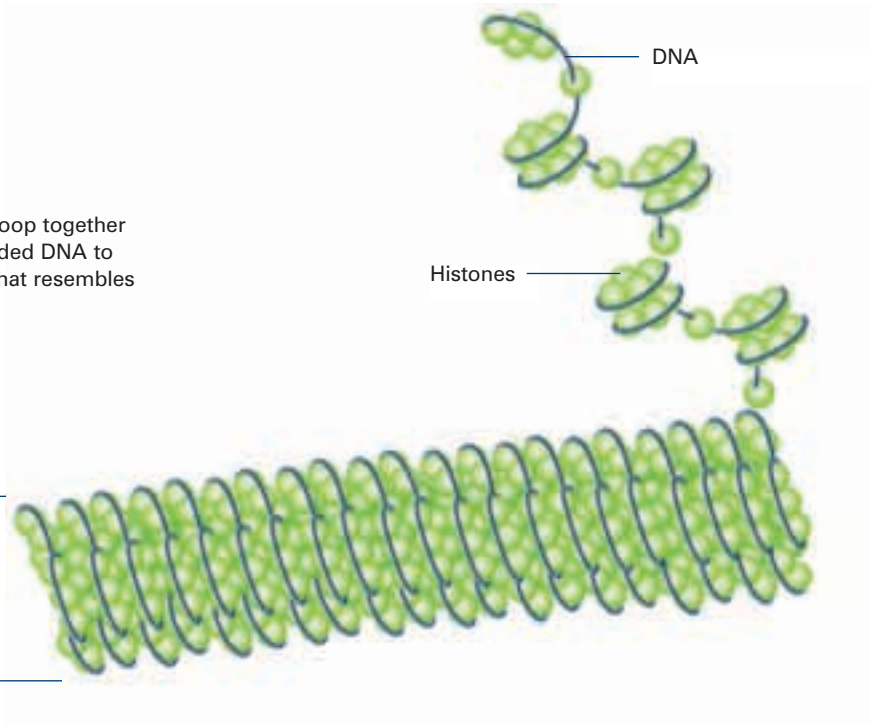
The researchers, Andrew Z. Fire, then of the Carnegie Institution of Washington in Baltimore and now at Stanford University, and Craig Mello of the University of Massachusetts Medical School in Worcester, were trying to block the expression of genes that affect cell growth and tissue formation in roundworms, using a molecular tool called antisense RNA.

To their surprise, Mello and Fire found that their antisense RNA tool wasn't doing much at all. Rather, they determined, a double-stranded contaminant produced during the synthesis of the single-stranded antisense RNA interfered with **gene expression**. Mello and Fire named the process RNAi, and in 2006 were awarded the Nobel Prize in physiology or medicine for their discovery.

Further experiments revealed that the double-stranded RNA gets chopped up inside the cell into much smaller pieces that stick to mRNA and block its action, much like the microRNA pieces of Velcro discussed above (see drawing, page 28).

Today, scientists are taking a cue from nature and using RNAi to explore biology. They have learned, for example, that the process is not limited to worms and plants, but operates in humans too.

Medical researchers are currently testing new types of RNAi-based drugs for treating conditions such as macular degeneration, the leading cause of blindness, and various infections, including those caused by HIV and the herpes virus.



Histone proteins loop together with double-stranded DNA to form a structure that resembles beads on a string.

Dynamic DNA

A good part of who we are is "written in our genes," inherited from Mom and Dad. Many traits, like red or brown hair, body shape and even some personality quirks, are passed on from parent to offspring.

But genes are not the whole story. Where we live, how much we exercise, what we eat: These and many other environmental factors can all affect how our genes get expressed.

You know that changes in DNA and RNA can produce changes in proteins. But additional control happens at the level of DNA, even though these changes do not alter DNA directly. Inherited factors that do not change the DNA sequence of nucleotides are called **epigenetic** changes, and they too help make each of us unique.

Epigenetic means, literally, "upon" or "over" genetics. It describes a type of chemical reaction that can alter the physical properties of DNA without changing its sequence. These changes make genes either more or less likely to be expressed (see drawing, page 31).

Currently, scientists are following an intriguing course of discovery to identify epigenetic factors that, along with diet and other environmental influences, affect who we are and what type of illnesses we might get.

Secret Code

DNA is spooled up compactly inside cells in an arrangement called **chromatin**. This packaging is critical for DNA to do its work. Chromatin consists of long strings of DNA spooled around a compact assembly of proteins called **histones**.

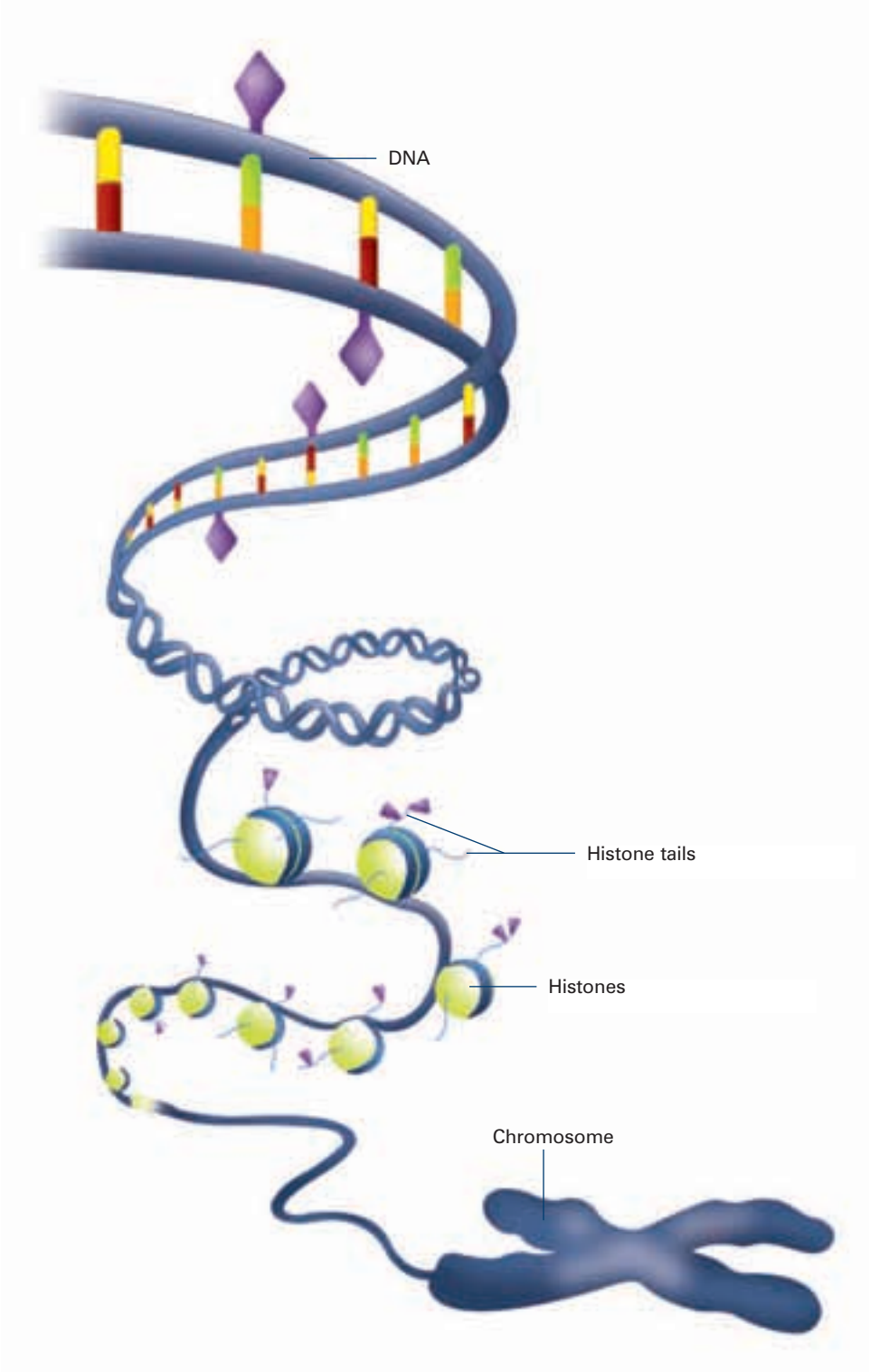
One of the key functions of chromatin is to control access to genes, since not all genes are turned on at the same time. Improper expression of growth-promoting genes, for example, can lead to cancer, birth defects or other health concerns.

Many years after the structure of DNA was determined, researchers used a powerful device known as an electron microscope to take pictures of chromatin fibers. Upon viewing chromatin up close, the researchers described it as "beads on a string," an image still used today. The beads were the histone balls, and the string was DNA wrapped around the histones and connecting one bead to the next.

Decades of study eventually revealed that histones have special chemical tags that act like switches to control access to the DNA. Flipping these switches, called epigenetic markings, unwinds the spooled DNA so the genes can be transcribed.

The observation that a cell's gene-reading machinery tracks epigenetic markings led C. David Allis, who was then at the University of Virginia Health Sciences Center in Charlottesville and now works at the Rockefeller University in New York City, to coin a new phrase, the "histone code." He and others believe that the histone code plays a major role in determining which proteins get made in a cell.

Flaws in the histone code have been associated with several types of cancer, and researchers are actively pursuing the development of medicines to correct such errors.



The "epigenetic code" controls gene activity with chemical tags that mark DNA (purple diamonds) and the "tails" of histone proteins (purple triangles). These markings help determine whether genes will be transcribed by RNA polymerase. Genes hidden from access to RNA polymerase are not expressed.

GENETICS AND YOU: *The Genetics of Anticipation*

Occasionally, unusual factors influence whether or not a child will be born with a genetic disease.

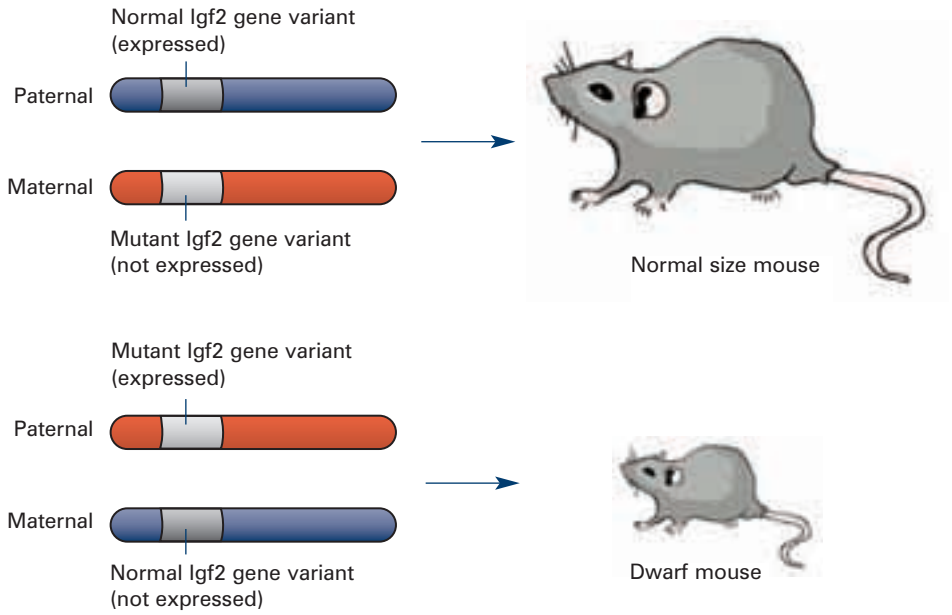
An example is the molecular error that causes Fragile X syndrome, a rare condition associated with mental retardation. The mutation leading to a fragile X chromosome is not a typical DNA typing mistake, in which nucleotides are switched around or dropped, or one of them is switched for another nucleotide. Instead, it is a kind of stutter by the DNA polymerase enzyme that copies DNA. This stutter creates a string of repeats of a DNA sequence that is composed of just three nucleotides, CGG.

Some people have only one repeat of the CGG nucleotide triplet. Thus, they have two copies of the repeat in a gene, and the extra sequence reads CGGCGG. Others have more than a thousand copies of the repeat. These people are the most severely affected.

The number of triplet repeats seems to increase as the chromosome is passed down through several generations. Thus, the grandsons of a man with a fragile X chromosome, who is not himself affected, have a 40 percent risk of retardation if they inherit the repeat-containing chromosome. The risk for great-grandsons is even higher: 50 percent.

Intrigued by the evidence that triplet repeats can cause genetic disease, scientists have searched for other examples of disorders associated with the DNA expansions. To date, more than a dozen such disorders have been found, and all of them affect the nervous system.

Analysis of the rare families in which such diseases are common has revealed that expansion of the triplet repeats is linked to something called genetic **anticipation**, when a disease's symptoms appear earlier and more severely in each successive generation.



Igf2 is an imprinted gene. A single copy of the abnormal, or mutant, form of the Igf2 gene (red) causes growth defects, but only if the abnormal gene variant is inherited from the father.

Battle of the Sexes

A process called **imprinting**, which occurs naturally in our cells, provides another example of how epigenetics affects gene activity.

With most genes, the two copies work exactly the same way. For some mammalian genes, however, only the mother's or the father's copy is switched on regardless of the child's gender. This is because the genes are chemically marked, or imprinted, during the process that generates eggs and sperm.

As a result, the embryo that emerges from the joining of egg and sperm can tell whether a gene copy came from Mom or Dad, so it knows which copy of the gene to shut off.

One example of an imprinted gene is insulin-like growth factor 2 (Igf2), a gene that helps a mammalian fetus grow. In this case, only the father's copy of Igf2 is expressed, and the mother's copy remains silent (is not expressed) throughout the life of the offspring.

Scientists have discovered that this selective silencing of Igf2 and many other imprinted genes occurs in all placental mammals (all except the platypus, echidna and marsupials) examined so far, but not in birds.

Why would nature tolerate a process that puts an organism at risk because only one of two copies of a gene is working? The likely reason, many researchers believe, is that mothers and fathers have competing interests, and the battlefield is DNA!

The scenario goes like this: It is in a father's interest for his embryos to get bigger faster, because that will improve his offspring's chances of survival after birth. The better an individual's

chance of surviving infancy, the better its chance of becoming an adult, mating and passing its genes on to the next generation.

Of course mothers want strong babies, but unlike fathers, mothers provide physical resources to embryos during pregnancy. Over her lifetime, a female is likely to be pregnant several times, so she needs to divide her resources among a number of embryos in different pregnancies.

Researchers have discovered over 200 imprinted genes in mammals since the first one was identified in 1991. We now know that imprinting controls some of the genes that have an important role in regulating embryonic and fetal growth and allocating maternal resources. Not surprisingly, mutations in these genes cause serious growth disorders.

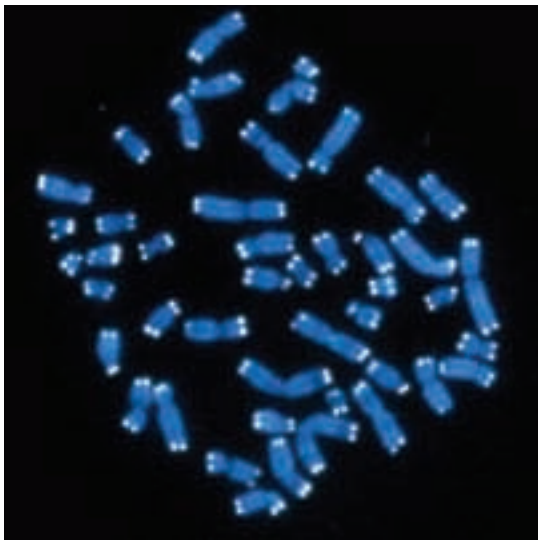
Marisa Bartolomei of the University of Pennsylvania School of Medicine in Philadelphia is trying to figure out how Igf2 and other genes become imprinted and stay silent throughout the life of an individual. She has already identified sequences within genes that are essential for imprinting. Bartolomei and other researchers have shown that these sequences, called insulators, serve as "landing sites" for a protein that keeps the imprinted gene from being transcribed.

Starting at the End

When we think of DNA, we think of genes. However, some DNA sequences are different: They don't encode RNAs or proteins. Introns, described in Chapter 1, are in this category.

Another example is **telomeres**—the ends of chromosomes. There are no genes in telomeres, but they serve an essential function. Like shoelaces without their tips, chromosomes without telomeres unravel and fray. And without telomeres, chromosomes stick to each other and cause cells to undergo harmful changes like dividing abnormally.

Researchers know a good deal about telomeres, dating back to experiments performed in the 1970s by Elizabeth Blackburn, a basic researcher who was curious about some of the fundamental events that take place within cells.



HESED PADILLA-NASH AND THOMAS RIED

Telomeres, repeated nucleotide sequences at the tips of chromosomes, appear white in this photo.

CAROL GREIDER

Molecular biologist Carol Greider discovered the enzyme telomerase. This license plate, which was on her car when she worked at Cold Spring Harbor Laboratory on Long Island, New York, advertises her research interest!

At the time, Blackburn, now at the University of California, San Francisco, was working with Joseph Gall at Yale University. For her experimental system, she chose a single-celled, pond-dwelling organism named *Tetrahymena*. These tiny, pear-shaped creatures are covered with hairlike cilia that they use to propel themselves through the water as they devour bacteria and fungi.

Tetrahymena was a good organism for Blackburn's experiments because it has a large number of chromosomes—which means it has a lot of telomeres!

Her research was also perfectly timed, because methods for sequencing DNA were just being developed. Blackburn found that *Tetrahymena's* telomeres had an unusual nucleotide sequence: TTGGGG, repeated about 50 times per telomere.

Since then, scientists have discovered that the telomeres of almost all organisms have repeated sequences of DNA with lots of Ts and Gs. In human and mouse telomeres, for example, the repeated sequence is TTAGGG.

The number of telomere repeats varies enormously, not just from organism to organism but in different cells of the same organism and even within a single cell over time. Blackburn reasoned that the repeat number might vary if cells had an enzyme that added copies of the repeated sequence to the telomeres of some but not all chromosomes.

With her then-graduate student Carol Greider, now at Johns Hopkins University, Blackburn hunted for the enzyme. The team found it and Greider named it telomerase. Blackburn, Greider and Jack Szostak of Harvard Medical School in Boston shared the 2009 Nobel Prize in physiology or medicine for their discoveries about telomeres and telomerase.

As it turns out, the telomerase enzyme consists of a protein and an RNA component, which the enzyme uses as a template for copying the repeated DNA sequence.

What is the natural function of telomerase? As cells divide again and again, their telomeres get shorter. Most normal cells stop dividing when their telomeres wear down to a certain point, and eventually the cells die. Telomerase can counteract the shortening. By adding DNA to telomeres, telomerase rebuilds the telomere and resets the cell's molecular clock.

The discovery of telomerase triggered new ideas and literally thousands of new studies. Many researchers thought that the enzyme might play important roles in cancer and aging. Researchers were hoping to find ways to turn telomerase *on* so that cells would continue to divide (to grow extra cells for burn patients, for example), or *off* so that cells would stop dividing (to stop cancer, for instance).

So far, they have been unsuccessful. Although it is clear that telomerase and cellular aging are related, researchers do not know whether telomerase plays a role in the normal cellular aging process or in diseases like cancer.

Recently, however, Blackburn and a team of other scientists discovered that chronic stress and the perception that life is stressful affect telomere length and telomerase activity in the cells of healthy women. Blackburn and her coworkers are currently conducting a long-term, follow-up study to confirm these intriguing results.

The Other Human Genome

Before you think everything's been said about DNA, there's one little thing we didn't mention: Some of the DNA in every cell is quite different from the DNA that we've been talking about up to this point. This special DNA isn't in chromosomes—it isn't even inside the cell's nucleus where all the chromosomes are!

So where is this special DNA? It's inside **mitochondria**, the organelles in our cells that produce the energy-rich molecule adenosine triphosphate, or ATP. Mendel knew nothing of mitochondria, since they weren't discovered until late in the 19th century. And it wasn't until the 1960s that researchers discovered the mitochondrial genome, which is circular like the genomes of bacteria.

In human cells, mitochondrial DNA makes up less than 1 percent of the total DNA in each of our cells. The mitochondrial genome is very small—containing only about three dozen genes. These encode a few of the proteins that are in the mitochondrion, plus a set of ribosomal RNAs used for synthesizing proteins for the organelle.

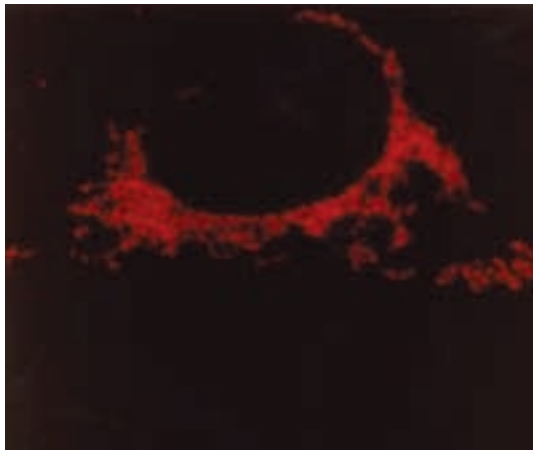
Mitochondria need many more proteins though, and most of these are encoded by genes in the nucleus. Thus, the energy-producing capabilities of human mitochondria—a vital part of any cell's everyday health—depend on coordinated teamwork among hundreds of genes in two cellular neighborhoods: the nucleus and the mitochondrion.

Mitochondrial DNA gets transcribed and the RNA is translated by enzymes that are very different from those that perform this job for genes in our chromosomes. Mitochondrial enzymes look and act much more like those from bacteria, which is not surprising because mitochondria are thought to have descended from free-living bacteria that were engulfed by another cell over a billion years ago.

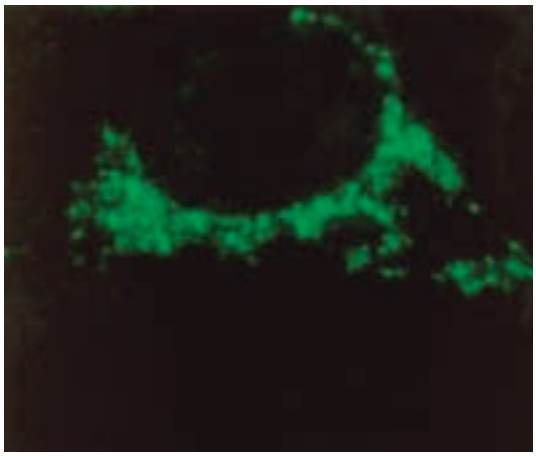
Scientists have linked mitochondrial DNA defects with a wide range of age-related diseases including neurodegenerative disorders, some forms of heart disease, diabetes and various cancers. It is still unclear, though, whether damaged mitochondria are a symptom or a cause of these health conditions.

Scientists have studied mitochondrial DNA for another reason: to understand the history of the human race. Unlike our chromosomal DNA, which we inherit from both parents, we get all of our mitochondrial DNA from our mothers.

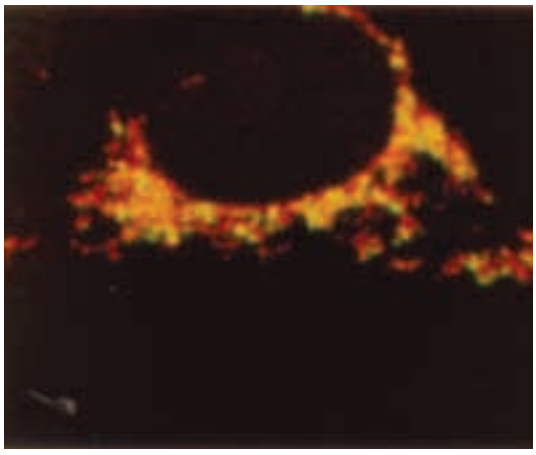
Thus, it is possible to deduce who our maternal ancestors were by tracking the inheritance of mutations in mitochondrial DNA. For reasons that are still not well understood, mutations accumulate in mitochondrial DNA more quickly than in chromosomal DNA. So, it's possible to trace your maternal ancestry way back beyond any relatives you may know by name—all the way back to "African Eve," the ancestor of us all!



Mitochondria (labeled with a red dye) are scattered throughout the cytoplasm of this human cancer cell.



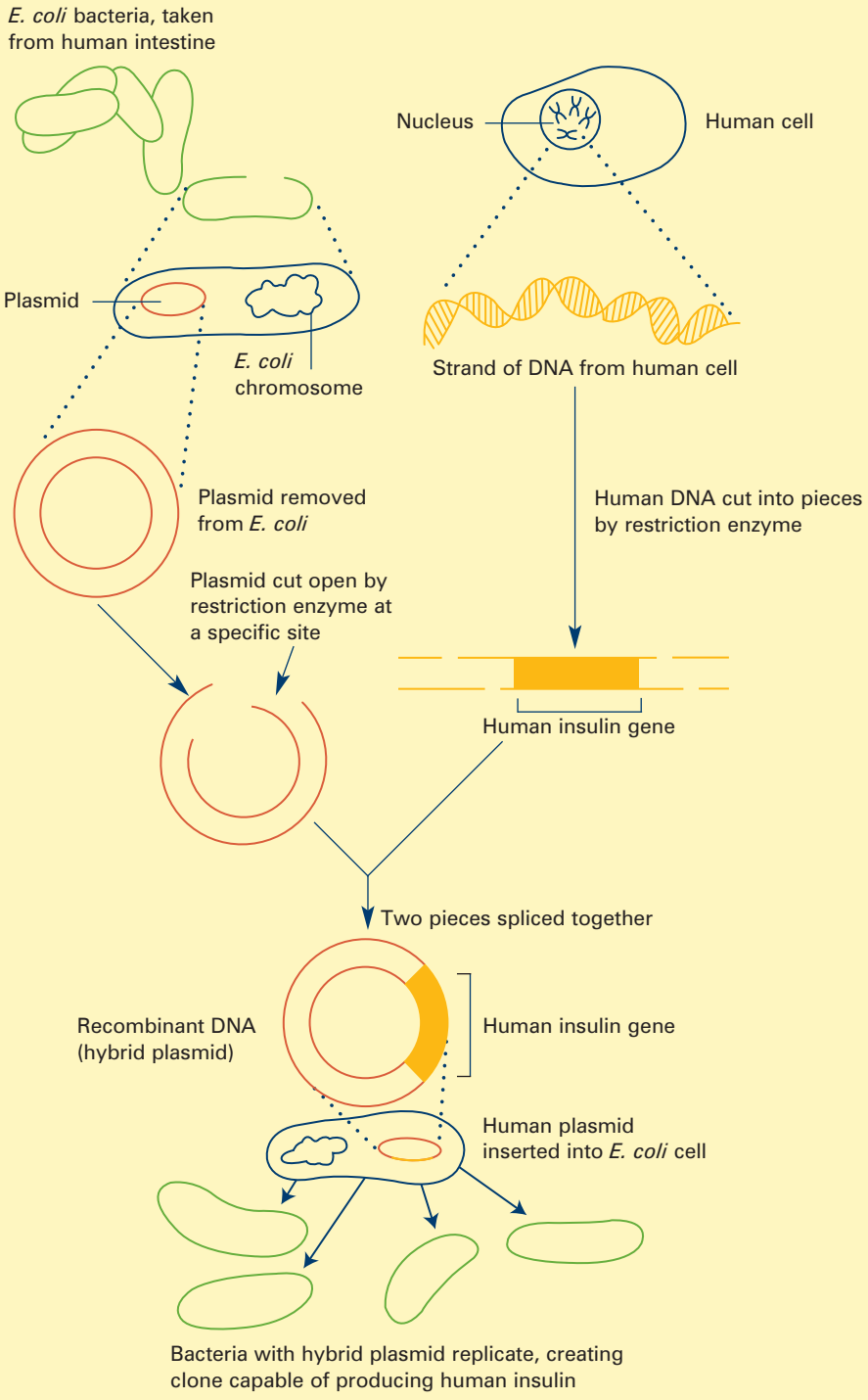
The cell has also been treated with a dye that colors the mitochondrial DNA green.



A computerized overlay of these two images of the same cell shows that mitochondria and its DNA appear together (yellow regions).

ALISON DAVIS

The Tools of Genetics: Recombinant DNA and Cloning



Recombinant DNA. To splice a human gene (in this case, the one for insulin) into a plasmid, scientists take the plasmid out of an *E. coli* bacterium, cut the plasmid with a restriction enzyme and splice in insulin-making human DNA. The resulting hybrid plasmid can be inserted into another *E. coli* bacterium, where it multiplies along with the bacterium. There, it can produce large quantities of insulin.

ROSLIN INSTITUTE, EDINBURGH

Scientists in Scotland were the first to clone an animal, this sheep named Dolly. She later gave birth to Bonnie, the lamb next to her.

In the early 1970s, scientists discovered that they could change an organism's genetic traits by putting genetic material from another organism into its cells. This discovery, which caused quite a stir, paved the way for many extraordinary accomplishments in medical research that have occurred over the past 35 years.

How do scientists move genes from one organism to another? The cutting and pasting gets done with chemical scissors: enzymes, to be specific. Take insulin, for example. Let's say a scientist wants to make large quantities of this protein to treat diabetes. She decides to transfer the human gene for insulin into a bacterium, *Escherichia coli*, or *E. coli*, which is commonly used for genetic research (see *Living Laboratories*, page 46). That's because *E. coli* reproduces really fast, so after one bacterium gets the human insulin gene, it doesn't take much time to grow millions of bacteria that contain the gene.

The first step is to cut the insulin gene out of a copied, or "cloned," version of the human DNA using a special bacterial enzyme from bacteria called a restriction endonuclease. (The normal role of these enzymes in bacteria is to chew up the DNA of viruses and other invaders.) Each restriction enzyme recognizes and cuts at a different nucleotide sequence, so it's possible to be very precise about DNA cutting by selecting one of several hundred of these enzymes that cuts at the desired sequence. Most restriction endonucleases make slightly staggered incisions, resulting in "sticky ends," out of which one strand protrudes.

The next step in this example is to splice, or paste, the human insulin gene into a circle of bacterial DNA called a plasmid. Attaching the cut ends together is done with a different enzyme (obtained from a virus), called DNA ligase. The sticky ends join back together kind of like jigsaw puzzle pieces. The result: a cut-and-pasted mixture of human and bacterial DNA.

The last step is putting the new, **recombinant DNA** back into *E. coli* and letting the bacteria reproduce in a petri dish. Now, the scientist has a great tool: a version of *E. coli* that produces lots of human insulin that can be used for treating people with diabetes.

So, what is cloning? Strictly speaking, it's making many copies. However, the term is more commonly used to refer to making many copies of a gene, as in the *E. coli* example above. Researchers can also **clone** entire organisms, like Dolly the sheep, which contained the identical genetic material of another sheep.

Got It?

Besides the sequence of nucleotides in genes, what are some other changes to DNA and RNA that can affect our health and who we are?

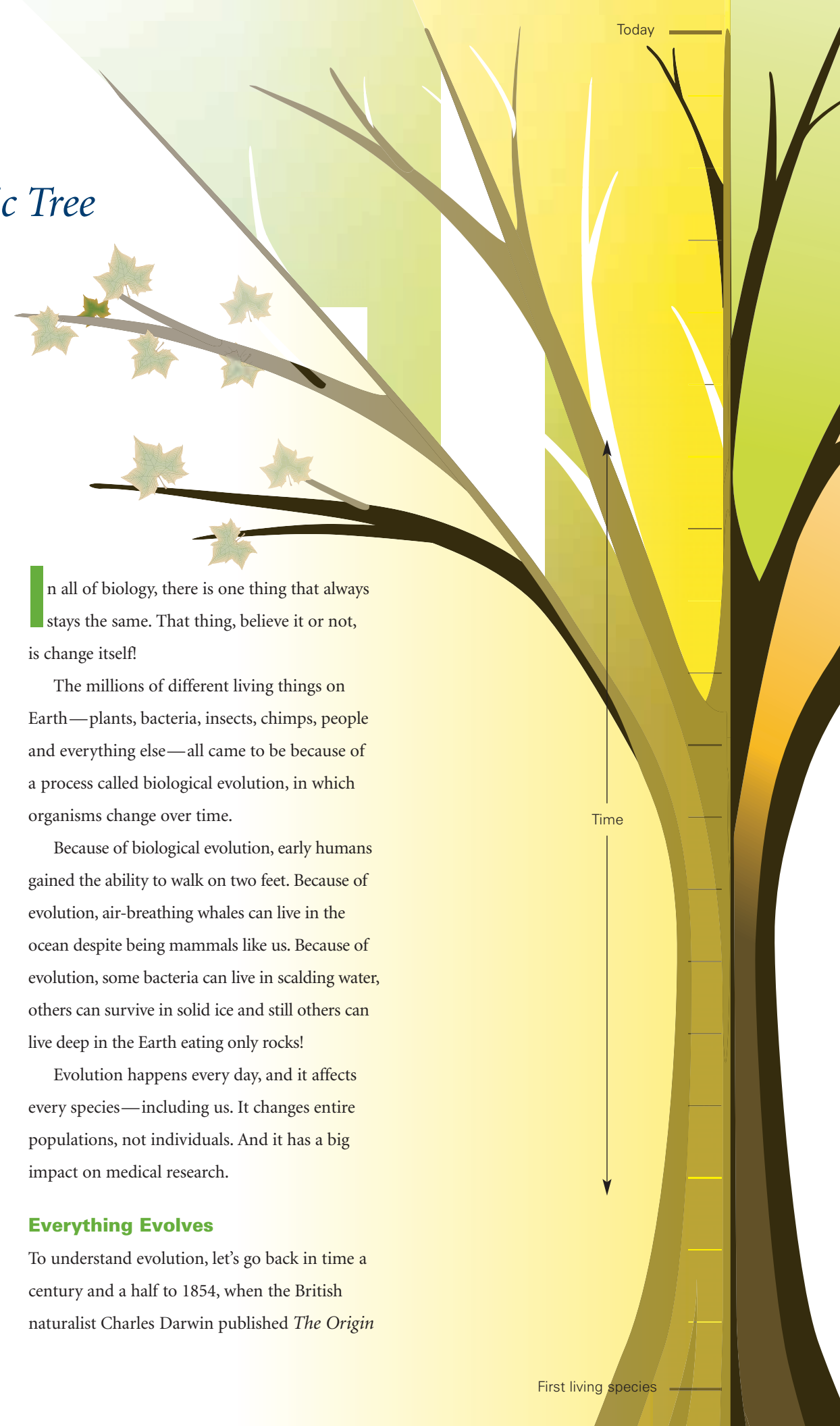
Can you imagine treat ments—other than vaccines and current medicines—crafted from genetic information and new molecular tools?

How is cloning a gene different from cloning an animal or a person? How do researchers use gene cloning to study health and disease?

Do you have any recurring illnesses in your extended family?

CHAPTER 3

Life's Genetic Tree



In all of biology, there is one thing that always stays the same. That thing, believe it or not, is change itself!

The millions of different living things on Earth—plants, bacteria, insects, chimps, people and everything else—all came to be because of a process called biological evolution, in which organisms change over time.

Because of biological evolution, early humans gained the ability to walk on two feet. Because of evolution, air-breathing whales can live in the ocean despite being mammals like us. Because of evolution, some bacteria can live in scalding water, others can survive in solid ice and still others can live deep in the Earth eating only rocks!

Evolution happens every day, and it affects every species—including us. It changes entire populations, not individuals. And it has a big impact on medical research.

Everything Evolves

To understand evolution, let's go back in time a century and a half to 1854, when the British naturalist Charles Darwin published *The Origin*


THE ORIGIN OF SPECIES

Charles Darwin described evolution in his classic text, *The Origin of Species*.

of Species, a book that proposed an explanation for how evolution works.

The main concept in evolution is that all living things share a common ancestor. The very earliest ancestor of all life forms on Earth lived about 4 billion years ago. From that early organism, millions of types of creatures—some living and some now extinct—have evolved.

Evolution requires diversity. You can tell that living things are diverse just by walking down the street and looking around you. Individual people are very different from one another. Chihuahuas are different from Great Danes, and Siamese cats are different from tabbies.

Evolution also depends on inheritance. Many of our unique characteristics are inherited—they are passed from parent to offspring. This is easy to see: Dalmatian puppies look like Dalmatians, not Chihuahuas. Petunias grow differently from pansies. Evolution works *only* on traits that are inherited.

Finally, as you probably already know, evolution favors the "fittest." Through a process called natural selection, only some offspring within a given generation will survive long enough to reproduce.

As an example, consider houseflies, each of which lays thousands of eggs every year. Why haven't they taken over the world? Because almost all of the baby houseflies die. The flies that survive are the ones that can find something to eat and drink … the ones that avoid being eaten, stepped on or swatted … and the ones that don't freeze, drown or land on a bug zapper.

The flies that survive all these ways to die have what it takes to outlive most of their brothers and sisters. These inherited traits give an organism a survival edge. Those who survive will mate with each other and will pass on to the next generation some of their DNA that encoded these advantageous traits.

Of course, not all aspects of survival are determined by genes. Whether a fly gets swatted

depends on genes that affect its reflexes—whether it's fast enough to avoid the swatter—but also on the environment. If there's no human around waving the swatter, the fly is quite likely to survive, regardless of its reflexes.

Evolution often takes a long time to make a difference. But it can also happen very quickly, especially in organisms with short lifespans. For example, as you read earlier, some bacteria have molecular features that let them survive in the presence of antibiotics. When you take an antibiotic medicine, antibiotic-resistant bacteria flourish while antibiotic-sensitive bacteria die.

Because antibiotic resistance is a growing public health threat, it's important to take the whole course of antibiotic medicine, not stop when you feel better. And you should take antibiotics only when they're needed, not for colds or other viral infections, which antibiotics can't treat.

Selective Study

Scientists doing medical research are very interested in genetic variants that have been selected by evolution. For example, researchers have discovered a rare genetic **variant** that protects people from getting AIDS. A genetic variant is a different version of a gene, one that has a slightly different sequence of nucleotides.

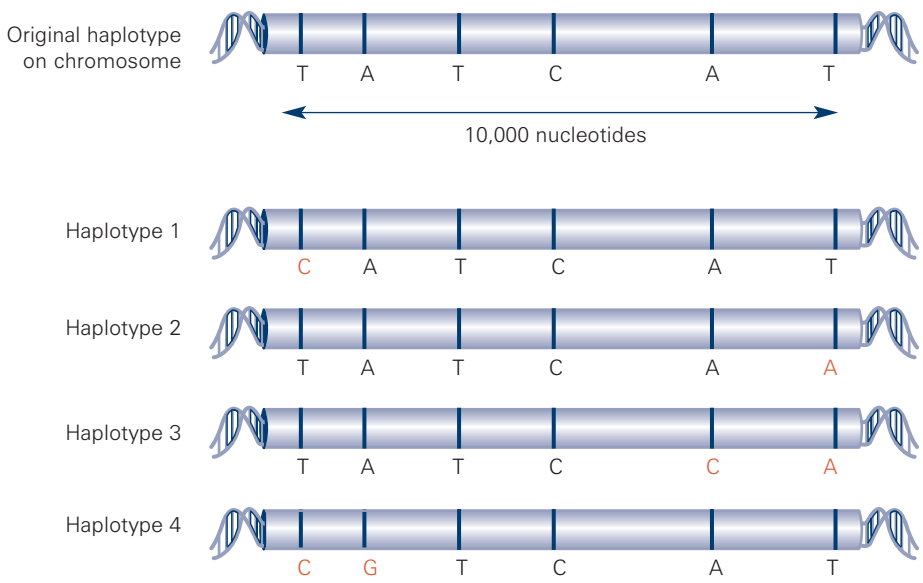
Scientists think that the rare variant of a gene called CCR5 originally may have been selected during evolution because it made people resistant to an organism unrelated to HIV.

Montgomery Slatkin of the University of California, Berkeley, has used mathematical modeling techniques to show that natural selection over time could explain the frequency of the CCR5 variant in human populations. The work indicates that the CCR5 gene variant's ability to protect against AIDS may contribute to keeping it in the human gene pool.

So, through evolution, living things change. Sometimes, that's good for us, as when humans understand HIV resistance in hopes of preventing AIDS. But sometimes the changes aren't so great—from a human perspective, anyway—as when bacteria become resistant to antibiotics.

Whether the consequences of evolutionary change are good or bad, understanding the

Different nucleotides (in this example, A or G) can appear in the DNA sequence of the same chromosome from two different individuals, creating a single-nucleotide polymorphism (SNP).



Haplotypes are combinations of gene variants, or SNPs, that are likely to be inherited together within the same chromosomal region. In this example, an original haplotype (top) evolved over time to create three newer haplotypes that each differ by a few nucleotides (red).

process can help us develop new strategies for fighting disease.

Clues from Variation

Scientists know quite a bit about how cells reshuffle genetic information to create each person's unique genome. But many details are missing about how this genetic variation contributes to disease, making for a very active area of research.

What scientists do know is that most of the human genome is the same in all of us. A little bit of genetic variation—differences that account for much less than 1 percent of our DNA—gives each of us a unique personality, appearance and health profile.

The parts of the human genome where the DNA sequences of many individuals vary by a single nucleotide are known as single-nucleotide polymorphisms (abbreviated SNPs and pronounced "snips").

For example, let's say that a certain nucleotide in one of your genes is A. In your uncle, however, the nucleotide in the same place on the same gene might be G. You and your uncle have slightly different versions of that gene. Scientists call the different gene versions alleles.

If two genes sit right next to each other on a chromosome, the SNPs in those genes tend to be inherited together. This set of neighboring SNPs is called a **haplotype** (see drawing above).

Most chromosome regions have only a few, common haplotypes among all humans. As it turns out, these few haplotypes—in different combinations in each person—appear to account for most of the variation from person to person in a population.

Scientists can use haplotype information to compare the genes of people affected by a disease with those of unaffected people. For example, this approach revealed a genetic variation that substantially increases the risk of age-related macular degeneration, the leading cause of severe vision loss in the elderly. Scientists discovered that a single SNP—one nucleotide in the 3 billion-nucleotide human genome—makes some people more likely to get this eye disease. The discovery paves the way for better diagnostic tests and treatments.

What about other diseases? In 2007, an international scientific team completed a catalog of common human haplotypes. Since then, researchers have been using the catalog to identify genes associated with susceptibility to many common diseases, including asthma, diabetes, cancer and heart disease.

But not all SNPs are in genes. Scientists studying genetic variation have also found SNPs in DNA that doesn't encode proteins. Nonetheless, some of these SNPs appear to affect gene activity.

Some researchers suspect that the "cryptic" (hidden) variation associated with SNPs in non-coding DNA plays an important role in determining the physical characteristics and behaviors of an organism.

Loren Rieseberg of Indiana University in Bloomington is one scientist who would love to take the mystery out of cryptic variation. He wants to know how this non-coding genetic variation can help organisms adapt to new

environments. He's also curious about whether it can create problems for some individuals.

You might be surprised to learn that Rieseberg's principal research subject is the sunflower. Although many plants produce only one generation a year, plants like sunflowers can be very useful tools for researchers asking fundamental questions about genetics. Because their genetic material is more malleable than that of many animals, plants are excellent models for studying how evolution works.

Wild sunflowers appealed to Rieseberg because there are several species that live in different habitats. Two ancient species of wild sunflowers grow in moderate climates and are broadly distributed throughout the central and western United States.

Three recently evolved sunflower species live in more specialized environments: One of the new species grows on sand dunes, another grows

in dry desert soil and the third species grows in a salt marsh.

To see how quickly new plant species could evolve, Rieseberg forced the two ancient sunflowers to interbreed with each other, something plants but not other organisms can do. Among the hybrid progeny were sunflowers that were just like the three recently evolved species! What that means is that Rieseberg had stimulated evolution in his lab, similar to what actually happened in nature some 60,000 to 200,000 years ago, when the newer species first arose.

That Rieseberg could do this is pretty amazing, but the really interesting part is how it happened. Scientists generally assume that, for a new species with very different characteristics to evolve, a lot of new mutations have to occur.

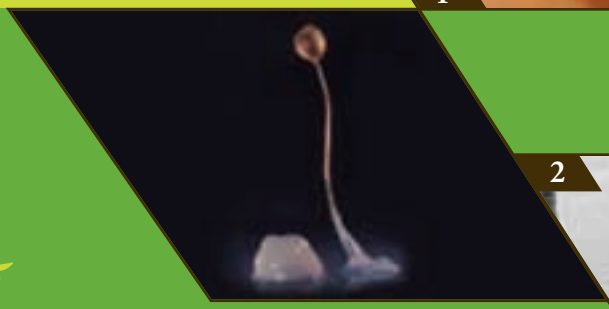
But when Rieseberg looked at the genomes of his hybrid sunflowers, he was surprised to find that they were just cut-and-pasted versions of the ancient sunflower species' genomes: large chunks had been moved rather than many new SNPs created.

Rieseberg reasons that plants stash away unused genetic material, giving them a ready supply of ingredients they can use to adapt quickly to a new environment. It may be that human genomes can recycle unused genetic material to confront new challenges, as well.

ALISON DAVIS

Plants like these sunflowers make great models for studying how evolution works.

1



2

REX L. CHISHOLM

Living Laboratories

Like most people, you probably think of fruit flies as kitchen nuisances. But did you know that scientists use these organisms for medical research?

Fruit flies and other model organisms—as different as mice, plants and zebrafish—permit scientists to investigate questions that would not be possible to study in any other way. These living systems are, relatively speaking, simple, inexpensive and easy to work with.

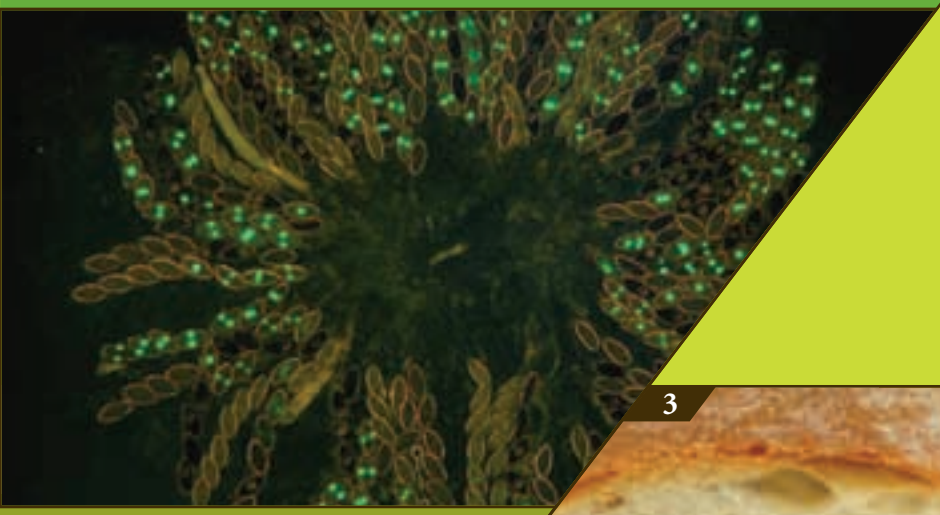
Model organisms are indispensable to science because creatures that appear very different from us and from each other actually have a lot in common when it comes to body chemistry. Even organisms that don't have a body—mold and yeast, for example—can give scientists clues to the workings of the tissues and organs of people.

This is because all living things process the nutrients they consume into the same chemicals, more or less. The genes for the enzymes involved in metabolism are similar in all organisms.

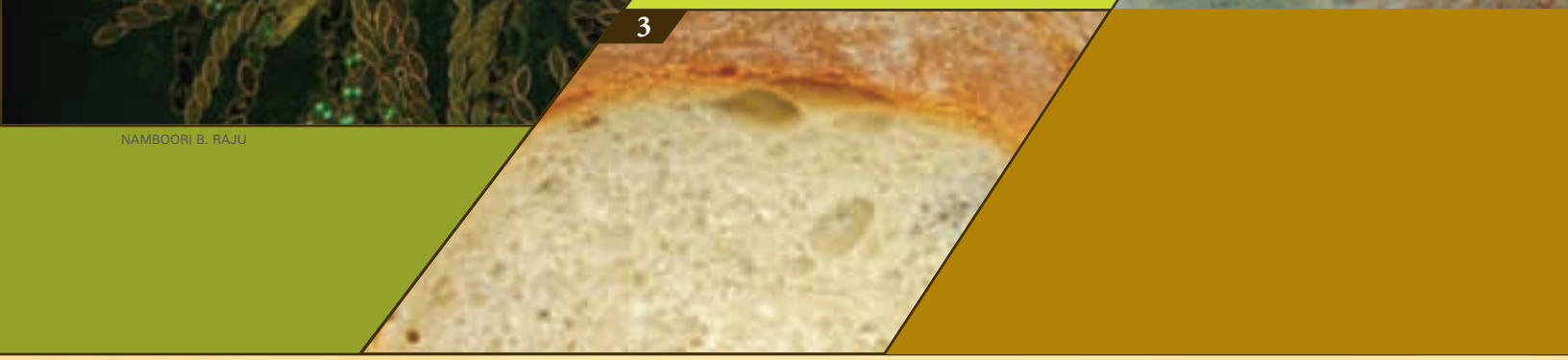
Below is a sampling of the wide variety of living laboratories that scientists are using to advance human health.

1 *Escherichia coli:* Bacterium

"Once we understand the biology of Escherichia coli, *we will understand the biology of an elephant."* So said Jacques Monod, a French scientist who won the 1965 Nobel Prize in physiology or medicine for his work on gene regulation. Monod was an early proponent of the value of experimenting with simple organisms like bacteria. Are all bacteria bad? If all you've ever heard about *E. coli* is its notorious link to tainted hamburger meat, you may not realize that non-disease-causing strains of the bacterium live in the intestinal tracts of humans and other animals, helping them in a variety of ways. For one thing, these bacteria are a main source of vitamin K and B-complex vitamins. They also aid digestion and protect against infection by harmful bacteria.



NAMBOORI B. RAJU



Scientists all over the world have banded together to sequence different versions of the *E. coli* genome. Among other things, these studies will help distinguish the genetic differences between bacteria in a healthy human gut and those that cause food poisoning.

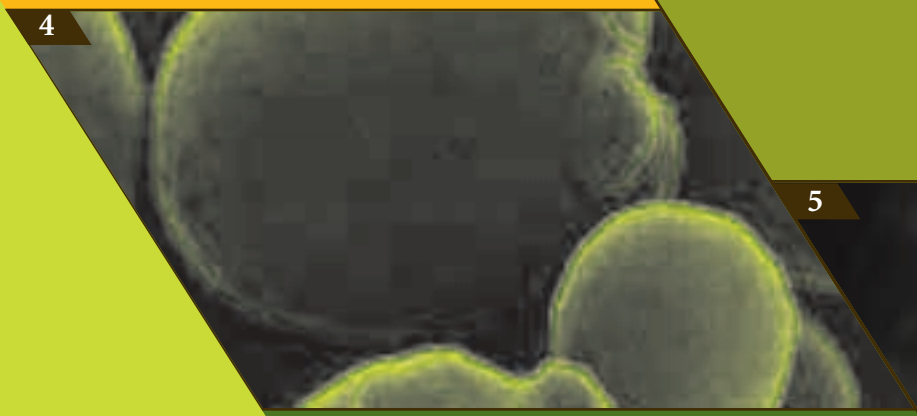
2 *Dictyostelium discoideum:* Amoeba

This microscopic amoeba—100,000 of them form a mound as big as a grain of sand—is an important tool for health studies. Scientists have determined that *Dictyostelium discoideum (Dicty)* has somewhere between 8,000 and 10,000 genes, many of which are close copies of those in people and animals but are missing in another single-celled organism, yeast. *Dicty* was first discovered in the 1930s in a North Carolina forest and has since been found in many similar habitats around the world.

Dicty normally grows as separate, independent cells. However, when food is limited, neighboring cells pile on top of each other to create a large, multicelled structure containing up to 100,000 cells. This blob ambles along like a slug, leaving a trail of slime behind. After migrating to a more suitable environment, the blob firms up into a towerlike structure that disperses spores, each capable of generating a new amoeba. Because of its unusual properties and ability to live alone or in a group, *Dicty* intrigues researchers who study cell division, movement and various aspects of organ and tissue development.

3 *Neurospora crassa:* Bread Mold

Chances are you don't think of a moldy bread crust as a potential science experiment, but thousands of researchers around the world do! *Neurospora crassa (Neurospora)*, which is a species of mold that thrives on bread, is a widely used model organism for genetic research.


4

ALAN WOELZ


5

GARY DITTA

Biologists like to use *Neurospora* because it is simple to grow and has features that make it very suitable for answering questions about how species arise and adapt, as well as how cells and tissues change their shape in different environments. Since *Neurospora* produces spores on a 24-hour cycle, the organism is also useful for studying the biological clocks that govern sleep, wakefulness and other rhythms of life.

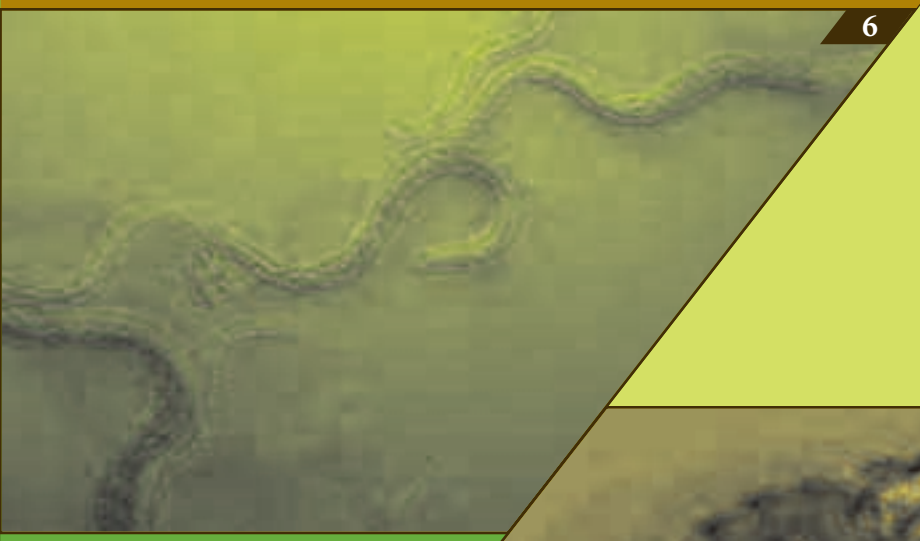
4 *Saccharomyces cerevisiae:* Yeast

There are hundreds of different kinds of yeast, but *Saccharomyces cerevisiae*, the one scientists study most often, is an important part of human life outside the lab, too. It is the yeast that bakers use to make bread and brewers use for beer.

Like *Neurospora*, yeast is actually a fungus—not a plant, not an animal, but related to both. It is also a eukaryote (as is *Neurospora*)—a "higher" organism with an organized, protective nucleus that holds its chromosomes. Researchers like yeast because it grows fast, is cheap to feed and safe to handle, and its genes are easy to work with. We know a lot about mammalian genes because scientists can easily insert them into yeast and then study how they work and what happens when they don't work.

5 *Arabidopsis thaliana:* Mustard Plant

Researchers who study plant growth often use *Arabidopsis thaliana (Arabidopsis*), a small, flowering plant related to cabbage and mustard. This organism is appealing to biologists because *Arabidopsis* has almost all of the same genes as other flowering plants and has relatively little DNA that does not encode proteins, simplifying the study of its genes. Like people and yeast, plants are also eukaryotes. *Arabidopsis* grows quickly, going from seed to mature plant in only 6 weeks—another plus for researchers who study how genes affect biology.



What do you have in common with a mustard plant? Plant cells, and parts of plant cells, communicate with each other in much the same way that human cells do. For that reason, plants are good models for genetic diseases that affect cell communication.

6 *Caenorhabditis elegans:* Roundworm

Caenorhabditis elegans (C. elegans) is a creature that is a lot smaller than its name! Several of these harmless roundworms would fit on the head of a pin, although their usual habitat is dirt. In the lab, they live in petri dishes and eat bacteria. *C. elegans* contains just 959 cells, almost a third of them forming its nervous system. Researchers know the fate of every one of these cells!

This worm is particularly prized by biologists because it is transparent, so what goes on in its tiny body is in plain view under a microscope. But for such a small, simple animal, *C. elegans* has a lot of genes—more than 19,000 (humans have about 20,000). Decoding the *C. elegans* genome was a huge milestone for biology, since it was the first animal genome to be sequenced completely. Scientists quickly learned that a vast number of genes in *C. elegans* are very similar to genes in other organisms, including people.

7 *Drosophila melanogaster:* Fruit Fly

The fruit fly species most commonly used for research is named *Drosophila melanogaster (Drosophila)*. A geneticist's fruit fly is pretty much the same as the ones that fly around your overripe bananas. In the lab, though, some of the flies are exposed to damaging chemicals or radiation, which changes the sequence of their DNA. Researchers allow the flies to mate, then search among the offspring for flies with abnormalities. The mutant flies are then mated to produce more offspring with the abnormality, enabling researchers to close in on the defective genes involved.

9

MONTE WESTERFIELD

8

Fruit flies have been a favorite experimental organism among geneticists since early in the 20th century. Hundreds of them can live in a pint-sized milk bottle or even a small vial, and they reproduce so quickly that keeping track of a particular gene as it passes through a couple of *Drosophila* generations takes only about a month. It's also relatively easy to create flies with mutations in many genes, enabling scientists to study how the genes work together.

8 *Danio rerio:* Zebrafish

Zebrafish were originally found in slow streams, rice paddies and the Ganges River in East India and Burma. They can also be found in most pet stores and are a home aquarium favorite.

Although the fish have been used by some geneticists for research since the early 1970s, in recent years they have become an especially popular model organism.

Many researchers are drawn to zebrafish because their eggs and embryos are transparent, making it possible to watch development unfold. In a span of 2 to 4 days, zebrafish cells split and form different parts of the baby fish's body: eyes, heart, liver, stomach and so on. Sometimes, researchers will move a cell to another spot to see if it will still go on to form the same part of the body or if it will do something different. This research has taught scientists about a range of health-related matters in people, including birth defects and the proper development of blood, the heart and the inner ear.

9 *Mus musculus:* Mouse

The branches of life's genetic tree that led eventually to mice and to human beings split off from each other 75 million years ago, back in the dinosaur age. But we are both mammals, and we share 85 percent of our genes. Because some mouse diseases are very similar—sometimes

identical—to human diseases, mice are exceptionally valuable for research.

Since the late 1980s, researchers have been able to engineer mice with missing genes. Scientists make these "knockout" mice to learn what goes wrong when a particular gene is removed. This gives them valuable clues about the gene's normal function. Identifying these genes in humans has helped define the molecular basis for many illnesses.

10 *Rattus norvegicus:* Rat

The Norway rat, or lab rat, was the first animal domesticated for use in scientific research. Currently, they make up about one-fourth of all research animals in the United States. Lab rats have been used for many decades for testing drugs, and much of what we know about cancer-causing molecules was learned in basic research with rats.

Although rats are mammals just like mice, they differ in important ways. Rats are much bigger than mice, making it easier for scientists to do experiments that involve the brain. For example, rats have taught scientists a lot about substance abuse and addiction, learning, memory and certain neurological diseases. Rats are also much better models than mice for studying asthma and lung injury. And since, in people, the disease arthritis is more common in women, studying rats makes more sense because female rats appear to be more susceptible to arthritis than male rats. The opposite is true with mice.

This *Living Laboratories* section is available as a poster. To order a free copy, visit http://publications.nigms.nih.gov/order.

The Genome Zoo

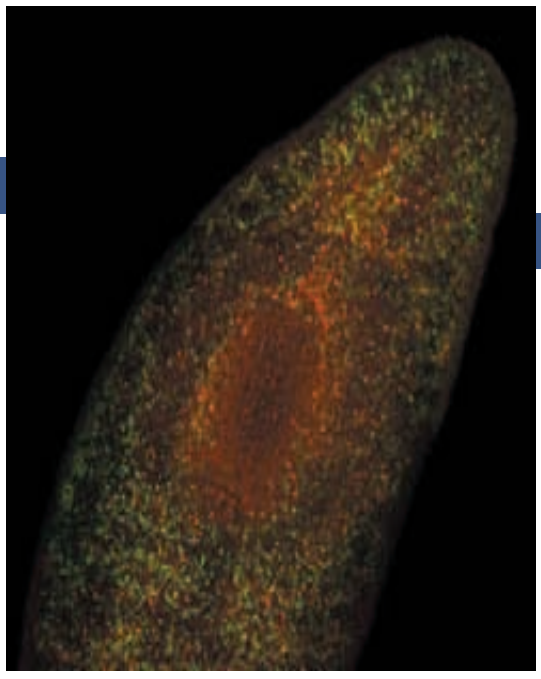
Scientists often use an image of a tree to depict how all organisms, living and extinct, are related to a common ancestor. In this "tree of life," each branch represents a species, and the forks between branches show when the species represented by those branches became different from one another. For example, researchers estimate that the common ancestor of humans and chimpanzees lived about 6 million years ago.

While it is obvious just by looking that people have a lot in common with our closest living relatives, chimpanzees, what about more distant species? If you look at an evolutionary tree, you'll see that humans are related to mice, worms and even bacteria. The ancestral species that gave rise to both humans and bacteria was alive a lot longer ago than the ancestor of humans and chimpanzees, yet we still share hundreds of genes with bacteria.

Scientists use the term **comparative genomics** to describe what they're doing when they compare the genomes of different species to see how similar (or how different!) the species' DNA sequences are. Sequences that the species have in common are the molecular footprints of an ancestor of those species.

Why are "old" DNA sequences still in our genomes? It turns out that nature is quite economical, so DNA sequences that are responsible for something as complicated and important as controlling gene activity may stay intact for millions of years.

Comparative genomic studies also have medical implications. What would you do if you wanted to develop new methods of preventing, diagnosing or treating a human disease that animals don't get?



PHILLIP NEWMARK

Starting All Over Again

Stem cells—what embryos are made up of just days after an egg is fertilized by a sperm—have the amazing ability to develop into any kind of cell in the body, from skin to heart, muscle and nerve.

Intrigued by the potential of these masterful cells, researchers want to know what gives stem cells their ability to change into a specific cell type upon the body's request, but stay in the "I can do anything" state until asked.

Some researchers are trying to figure out how stem cells work by using a unique model system: tiny, freshwater worms called planarians. These worms are like stem cells in the sense that they can regenerate. You can cut a planarian into hundreds of pieces, and each piece will grow into a complete worm.

Planarians' resemblance to stem cells isn't just coincidental. Scientists have discovered

If people have a gene that influences their risk for a disease, and mice have the gene too, you could study some aspect of the disease in mice, even though they don't ever have the symptoms of the disease. You could even study the disease in yeast, if it has the gene, as well.

Genes Meet Environment

If toxins from the environment get into our bodies, they don't always make us sick. That's because liver enzymes come to our rescue to make the chemicals less harmful. The genes that encode those enzymes are under constant evolutionary pressure to adapt quickly to new toxins.

For example, certain liver enzymes called cytochrome P450 proteins metabolize, or break down, hormones that our bodies make as well as many of the foreign substances that we encounter. These include harmful molecules like cancer-causing agents as well as beneficial ones, like medicines. In fact, just two genes within the cytochrome P450 family, abbreviated 3A4 and 3A5, encode proteins that process more than half of all of the medicines that are sold today.

Since the chemicals to which people are exposed vary so widely, a scientist might predict that there would be different variants of cytochrome P450 genes in different human populations. Using comparative genomics, researchers such as Anna Di Rienzo of the University of Chicago have shown that this is indeed the case. Di Rienzo has found many sequence differences within these genes in people living throughout the world.

It turns out that one variant of the gene that encodes the cytochrome P450 3A5 protein makes this enzyme very efficient at breaking down cortisol, a hormone that raises salt levels in the kidneys and helps the body retain water. Di Rienzo compared the DNA sequences of the 3A5 gene in DNA samples taken from more than 1,000 people

---

that planarians can perform the amazing act of regeneration due to the presence of, yes, specialized stem cells in their bodies.

Developmental biologist Alejandro Sánchez Alvarado of the University of Utah School of Medicine in Salt Lake City used the gene-silencing technique RNAi (see page 28) to identify planarian genes essential for regeneration. He and his team hope to figure out how these genes allow the specialized stem cells to travel to a wounded site and "turn into" any of the 30 or so cell types needed to recreate a mature worm.

Although humans are only distantly related to planarians, we have many of the same genes, so these findings could reveal strategies for regenerating injured body parts in people, too.

Scientists have also learned how to genetically reprogram human skin cells (and other easily obtained cells) to mimic the stem cells of embryos. In theory, these so-called induced pluripotent stem cells could generate any type of cell and be used to treat diseases. But to realize this potential, we need a much better understanding of the properties of these cells and how to efficiently produce cells that are safe for therapeutic uses.

▲ Scientists have discovered that some African populations near the equator have a high frequency of a genetic variant that helps the body conserve water.

representing over 50 populations worldwide. She was amazed to find a striking link between the existence of the gene variant and the geographic locale of the people who have it.

Di Rienzo discovered that African populations living very close to the equator were more likely than other populations to have the salt-saving version of the 3A5 gene. She suggests that this is because this gene variant provides a health advantage for people living in a very hot climate, since retaining salt helps ward off dehydration caused by intense heat.

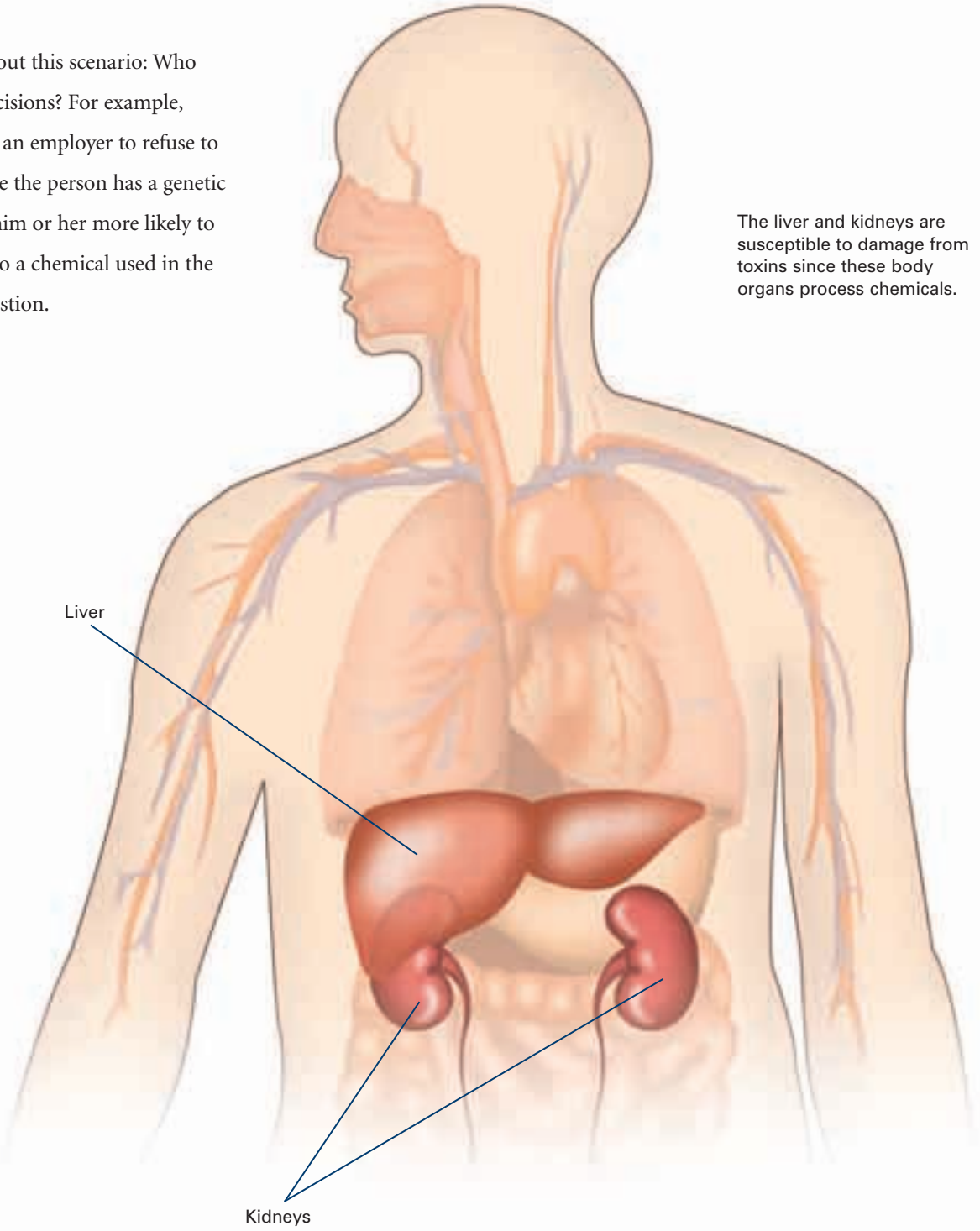
However, there seems to be a cost associated with that benefit—the 3A5 gene variant raises the risk for some types of high blood pressure. That means that in environments in which retaining salt is not beneficial, evolution selects against this gene variant.

Another scientist who studies interactions between genes and the environment is Serrine Lau of the University of Arizona in Tucson. She studies a class of harmful molecules called polyphenols, present in cigarette smoke and car exhaust, that cause kidney cancer in rats, and perhaps, in people.

Lau discovered that rats and humans who are more sensitive to some of the breakdown products of polyphenols have an unusual DNA sequence—a genetic signature—that increases their risk of developing cancer. She suspects that the gene that is affected encodes a tumor suppressor: a protein that prevents cancer from developing. In people and rats with the genetic signature, she reasons, the tumor suppressor doesn't work right, so tumors grow.

Taking this logic one step further, it may be that certain people's genetic make-up makes them unusually susceptible to DNA damage caused by exposure to carcinogens. If doctors could identify those at risk, Lau says, such people could be forewarned to avoid contact with specific chemicals to protect their health.

However, think about this scenario: Who should make those decisions? For example, would it be ethical for an employer to refuse to hire somebody because the person has a genetic signature that makes him or her more likely to get cancer if exposed to a chemical used in the workplace? Tough question.



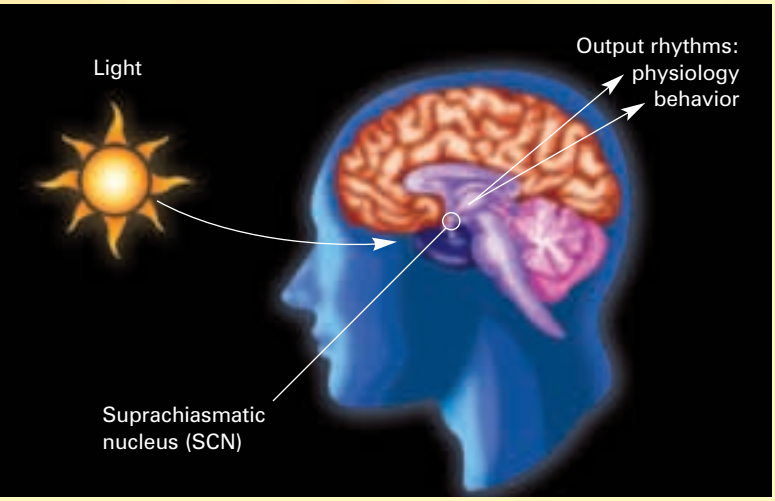
The liver and kidneys are susceptible to damage from toxins since these body organs process chemicals.

GENETICS AND YOU: *You've Got Rhythm!*

What do waking, sleeping, eating, reproducing and birds flying south for the winter have in common? These are all examples of nature's amazing sense of rhythm. All living things are equipped with molecular timepieces that set the pulse of life.

If you've ever crossed the country or an ocean by plane, you know about the importance of these clocks. You probably experienced that traveler's misery called jet lag, where the body is forced to adapt quickly to a new time zone.

But did you know that certain forms of insomnia and manic-depressive illness are associated with biological clocks not working properly? And biological rhythms may be the reason why some medicines and surgical treatments appear to work best at certain times of day.



The human body keeps time with a master clock called the suprachiasmatic nucleus or SCN. Situated inside the brain, it's a tiny sliver of tissue about the size of a grain of rice, located behind the eyes. It sits quite close to the optic nerve, which controls vision, and this means that the SCN "clock" can keep track of day and night. Given enough time, your SCN can reset itself after you fly in an airplane from one time zone to another.

The SCN helps control sleep by coordinating the actions of billions of miniature "clocks" throughout the body. These aren't actually clocks, but rather are ensembles of genes inside clusters of cells that switch on and off in a regular, 24-hour cycle—our physiological day.

Scientists call this 24-hour oscillation a **circadian** rhythm. ("Circadian" comes from the Latin words meaning "approximately a day.") Researchers have discovered that all living things—plants, animals and bacteria—have circadian rhythms. Many researchers working with insect and other model systems have identified genes that are critical for keeping biological time.

Understanding circadian rhythms will help scientists better understand sleep disorders. If we have the opportunity, most of us sleep 7 or 8 hours at night, and if we don't get enough rest we may have a hard time getting things done the next day. Some people,

however, routinely get by with only 3 to 4 hours of sleep. Researchers have noted that this trait seems to run in families, suggesting a genetic link.

As it turns out, fruit flies need even more sleep than people. Neuroscientist Chiara Cirelli of the University of Wisconsin-Madison did a genetic search for fruit fly mutants that don't sleep much. She discovered that flies with a variant of a gene called *shaker* sleep only 3 to 4 hours per night.

Although the *shaker* flies don't appear sleep-deprived, Cirelli found that they have a different problem: They don't live as long as flies without the mutation. She is now studying this new connection between sleep and lifespan.

Her work may also pave the way for improved sleep aids and effective remedies for jet lag.

Animals Helping People

Using technology that grew out of the Human Genome Project, scientists have read the sequences of the genomes of hundreds of organisms: dogs, mice, rats, chickens, honeybees, fruit flies, sea urchins, pufferfish, sea squirts, roundworms and many bacteria and fungi. Next in line are dozens of additional species, including a marmoset, a sea skate, an alpaca, an anteater and many reptiles.

What effect will all this gene sequence information have on medical research? We've already talked about the fact that people share many of their genes with other species. This means that when scientists read the sequence of another species' genome, they're likely to discover that the organism has many of the genes that, in humans, cause disease or raise disease risk when mutated.

Take fruit flies as one example. According to biologist Ethan Bier of the University of California, San Diego, 30 percent of the currently identified human disease genes most likely have functional counterparts in none other than *Drosophila melanogaster*, a fruit fly species widely used in genetic research (see *Living Laboratories*, page 49).

Currently, Bier and other scientists are using experimental flies to investigate a wide range of genes involved in conditions such as blindness, deafness, mental retardation, heart disease and the way in which bacterial toxins cause illness.

By reading the DNA sequences of many other species, researchers hope to find model systems that are even better than fruit flies for studying some aspects of human disease.

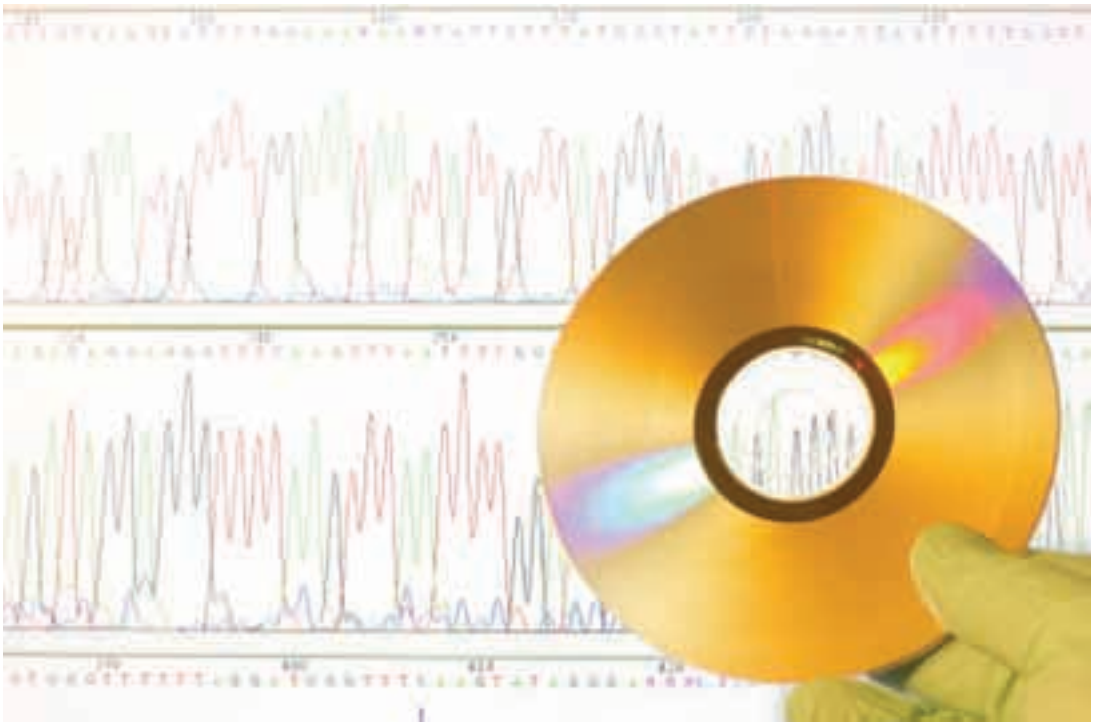
Sometimes, the genes that we don't have in common with other species are as important as the genes we share. For example, consider the fact that humans and chimpanzees have remarkably different abilities and physical features. But the chimpanzee genome is 99 percent identical to our own.

And did you know that chimpanzees don't get malaria or AIDS?

So a tiny portion of our genome determines whether we look and behave like a person or a chimp, and whether we are susceptible to malaria or AIDS.

My Collaborator Is a Computer

We've made the case that comparing genomes can offer fresh insight on the basic genetic ingredients for health and the causes of disease. But what does a scientist actually do when he or she compares gene sequences? Does this mean staring at thousands of pages of genetic letters, looking for those that are the same or different?



Computers are an essential tool for scientists who store and analyze huge amounts of genomic data. Read more about computers and biology at http://publications.nigms.nih.gov/computinglife.

Yes and no. Comparative genomics does involve looking for similarities and differences, but it isn't something that scientists do by hand. Certainly not for thousands of genes at a time.

Rather, the gigantic task of comparing the nucleotides that make up the genomes of two or more species is the perfect job for a computer, a natural multitasker. If you consider that the human genome contains 3 billion nucleotides, you can easily see why this is work well suited to a machine (with a human operator, of course).

Researchers called computational biologists help analyze genomic data. These scientists develop software programs that enable computers to perform genome comparisons. Among other things, the programs can figure out where in the DNA sequences a gene starts and stops: its "boundaries."

Other researchers who work in the field of **bioinformatics** mine genomic information hidden in the masses of data. They are looking for scientific treasure in the form of new biological knowledge. These experiments can zero in on previously hidden patterns and reveal links between different fields of research.

Bioinformaticists and computational biologists are in high demand because they play a very important role in 21st-century medical science. These scientists must be fluent in both computer science and biology.

The Tools of Genetics: Unlimited DNA

You might be amazed to learn that a microbe that lives in a boiling hot spring in Yellowstone National Park is the essential ingredient for one of the most important biological research tools ever invented.

Thermus aquaticus is a bacterium that makes a heat-resistant enzyme, which is why it can thrive in hot springs. The enzyme, Taq polymerase, is essential to a laboratory technique called the polymerase chain reaction, or PCR. And PCR is essential to lots of things that life scientists do—and to many other fields, too. PCR's inventor, Kary Mullis, won the 1993 Nobel Prize in chemistry.

PCR is a quick, easy method for generating unlimited copies of tiny amounts of DNA. Words

▲ A microbe that lives in hot springs, like this one in Yellowstone National Park, is home to the enzyme that makes the polymerase chain reaction, or PCR, possible.

PCR machine.

APPLIED BIOSYSTEMS

like "revolutionary" and "breakthrough" are not an exaggeration of its impact.

PCR is at the heart of modern DNA **sequencing** methods. It is essential for pinpointing mutations in genes, so it is the basis for much of the research discussed in this booklet. PCR has done for genetic material what the invention of the printing press did for written material. It makes copying easy, inexpensive and widely available.

PCR underlies many diagnostic techniques, like testing individuals for genes that cause breast cancer. It can also help diagnose diseases other than cancer, such as infections by HIV and hepatitis C.

PCR is a key element of "genetic fingerprinting," which has helped free prisoners who relied on it to prove that they were innocent of the crimes that got them locked up. Conversely, it has provided scientific evidence that helped convict criminals.

PCR has even revolutionized archaeology by helping to analyze badly damaged ancient DNA—sometimes thousands of years old—which can reveal new information about past people and cultures.

Scientists predict that future uses of PCR technology will enhance medical treatment, enabling better diagnosis and more accurate subtyping of disease.

Got It?

Discuss reasons why research studies with identical twins can provide valuable informa tion about health and disease.

Humans and mice share over 80 percent of the same genetic material: for chimps and humans, it s more than 99 percent. Why are people and animals so different, if their genes are so similar?

You are a scientist and you want to learn more about how humans age. Is there a way you can address your research question without spending many decades studying people?

Can you think of an experi ment using fruit flies that could help researchers better understand jet lag?

CHAPTER 4

Genes Are Us

For science, the sequencing of the human genome was a groundbreaking achievement, one that made a lot of news. But what does it actually mean? Will any of this information make a difference in your life?

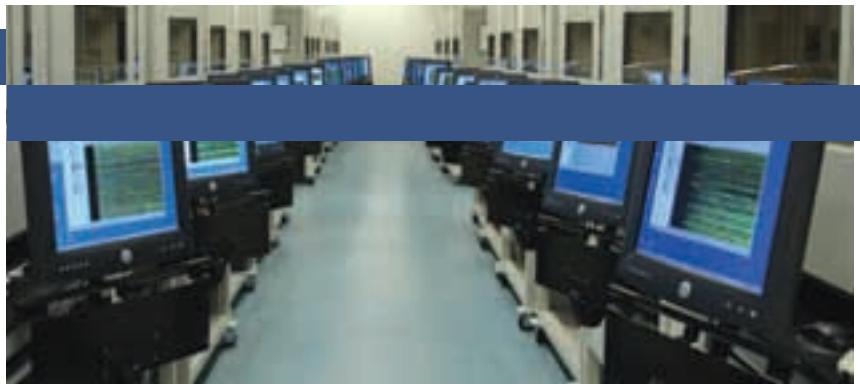
A genome is all of the genetic material that an individual (or a species) has. The human genome differs from the gorilla genome, which differs from the rice genome, and so on. And while every person has a "human genome," it is not exactly the same in all people. Sequence variations *within* your genes makes your DNA different from that of your mother, your cousin or a complete stranger.

Think of the human genome as a long story that contains roughly 20,000 words (the genes). With few exceptions, each person has the same number of words, but certain words have slightly different spellings. In some cases, the spelling changes create words with new meanings—genes that code for different proteins. Other spelling changes appear to have no effect whatsoever, at least not ones that today's scientists know how to measure.

Researchers are beginning to use knowledge learned from genome sequencing research to figure out how being healthy and being sick are different at the level of molecules. And doctors are starting to use genetic information to make treatment choices.

For example, a diagnostic test can search for differences in the level of expression of a particular gene in breast cancer cells and predict whether a person will respond to a drug called Herceptin®.

The cancerous cells of some people who have breast cancer make an abundance of "HER2" proteins that are targeted by Herceptin. For those people, Herceptin is a miracle drug because it



L. BARRY HETHERINGTON

Many DNA sequencing centers joined efforts to form the Human Genome Project, completed in 2003. Now the centers, like this one at the Broad Institute of MIT and Harvard University in Cambridge, Massachusetts, are working to better understand the human genome and to sequence the genomes of other organisms.

Reading the Book of Human Genes

In April 2003, researchers across the world celebrated a milestone and an anniversary. Almost 50 years to the day after James Watson, Francis Crick and Maurice Wilkins unveiled their Nobel Prize-winning description of the DNA double helix, scientists completed the sequencing of the human genome, a momentous achievement in biology.

The day was long in coming. In the 1980s, geneticists realized that they had both the need and the ability to learn the complete layout of the human genome. They wanted to map the location of every gene within chromosomes and decipher the complete, letter-by-letter sequence of the genome's 3 billion nucleotides.

reduces the risk that their breast cancer will come back, and it also decreases their odds of dying from the disease.

For cancer patients whose tumor genes do not express HER2, Herceptin won't do a thing, though, so it shouldn't be prescribed. Research is proceeding quickly to develop other genetic tests that may help diagnose and treat a wide range of health problems beyond cancer.

With that information in hand, scientists reasoned, it would eventually be possible to learn exactly what job each gene performs as well as how genes contribute to human health and disease.

Soon, thousands of scientists in labs all over the world got into the act. Critical to their success were new tools and technologies that made the work go faster and helped the researchers manage and analyze the flood of data.

Although the Human Genome Project is done, related genome sequencing efforts have continued. One involves sequencing the genomes of many other species (see page 58).

Another is roughly sequencing the genomes of 2,000 people to produce a detailed haplotype map showing both common and rare patterns of genetic variation. Researchers can link these variations to disease risk and health-related traits, such as individual reactions to medicines and environmental chemicals.

Individualized Prescriptions

One way variations in our genes make a difference in our health is by affecting how our bodies react to medicines. The unsettling truth is that medicines work as expected in fewer than half of the people who take them.

While environmental and lifestyle factors can explain some of this, a good part of the individual variability in response to medicines can be attributed to variants in the genes that make cytochrome P450 proteins (see page 53). These proteins process many of the drugs we take.

Did you know that medicines work like they're supposed to in fewer than half of the people who take them? Genetic differences among people are one reason.

Because each person's set of genes is a little different, the proteins that the genes encode are also slightly different. These changes can affect how the cytochrome P450 proteins (and many other types of proteins) work on drugs.

Doctors first realized this in the 1950s, when some patients had bad—sometimes fatal—reactions to an anesthetic medicine used in surgery. Experiments revealed that those who reacted poorly had a genetic variation in the enzyme that breaks down and disposes of the anesthetic after it's been in the body for a while.

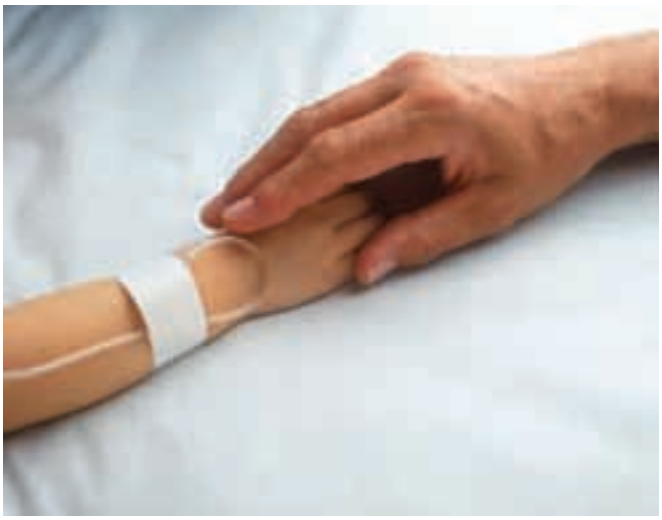
People whose genes encode the variant enzyme had no trouble at all until they needed surgery that required general anesthesia. In the operating room, a normal human genetic variation suddenly led to a medical crisis!

Fortunately, this type of serious reaction to an anesthetic is very rare. But many reactions to medicines aren't so unusual. Researchers know that genetic variations can cause some common medicines to have dangerous side effects. For example, some people who take the colon cancer drug Camptosar® (also known as irinotecan) can develop diarrhea and a life-threatening infection if they have a variant form of the gene for the protein that metabolizes Camptosar.

Genetic variations can also cause drugs to have little effect at all. For example, in some people, pain medicines containing codeine, like Tylenol® with Codeine Elixir, offer no relief because their bodies break it down in an unusual way.

The use of genetic information to predict how people will respond to medicines is called **pharmacogenetics**. The ultimate goal of this field of study is to customize treatments based on an individual's genes.

With this kind of approach, every patient won't be treated the same, because doctors will have the molecular tools to know ahead of time which drug, and how much of it, to prescribe—or whether to prescribe it at all.



Pharmacogenetic researchers have discovered that a gene test can predict which children with acute lymphoblastic leukemia will be cured by chemotherapy.

The Healing Power of DNA

Pharmacogenetics is advancing quickly since scientists have a lot of new information from the Human Genome Project and new computer tools that help them analyze the information. One disease for which progress has been rapid is cancer.

Consider the fact that cancer is often treated with a chemotherapy "cocktail," a combination of several different medicines. Each of the drugs in the mixture interacts with different proteins that control how well that particular drug works and how quickly it is metabolized in the body. What's more, each drug may have its own set of unpleasant—even potentially life-threatening—side effects.

For these reasons, individually targeted, gene-based prescriptions for chemotherapy may offer a real benefit to people with cancer.

Currently, chemotherapy cures about 80 percent of the children who have been diagnosed with acute lymphoblastic leukemia, the most common childhood cancer. The remaining 20 percent are at risk of the cancer coming back.

Mary Relling, a research clinical pharmacist at St. Jude Children's Research Hospital in Memphis, Tennessee, discovered that variations in two genes can predict which patients with acute lymphoblastic leukemia are likely to be cured by chemotherapy. Her research team also identified more than 100 genes expressed only in cancer cells that can be used to predict resistance to chemotherapy drugs.

By taking patient and cancer cell genetic profiles into account, Relling says, researchers can develop more effective treatments for the disease.

Genetic variation produces different individual responses to the blood-thinning drug Coumadin®. A genetic test could lead to more accurate doses.

Other pharmacogenetic scientists are studying the effects of gene variants on patients' responses to drugs used to treat AIDS, allergies, infections, asthma, heart disorders and many other conditions.

For example, researchers recently identified two different genetic variants that play a central role in determining the body's response to Coumadin® (also known as warfarin), a widely prescribed medicine given to people who are at risk for blood clots or heart attacks. Although 2 million Americans take this blood-thinning drug every day, it is very difficult to administer, since its effects vary widely in different people who are taking the same dose. Giving the right dose is essential, because too much Coumadin can cause excessive bleeding, while too little can allow blood clots to form.

Allan Rettie, a medicinal chemist at the University of Washington in Seattle, discovered that genetic variation among people influences the activity of a protein in the blood that is Coumadin's molecular target. He and other scientists are now trying to translate these findings into a genetic test that could help doctors predict what dose of Coumadin is appropriate based on each patient's DNA profile.

Genes Can Do That?

ZACHARY HUANG, HTTP://CYBERBEE.MSU.EDU

Honeybees are social animals and they work together to keep their hive healthy. The forager bee (on the left) is about a month old and hunts for food. The 14-day-old undertaker bee (on the right) removes dead bees from the hive.

Did you know that, in addition to traits you can see like hair color and physique, genes also contribute to how we behave? It may come as a surprise that many researchers are answering basic questions about the genetics of behavior by studying insects.

For example, Gene Robinson, an entomologist at the University of Illinois at Urbana-Champaign, works with honeybees. Robinson says that if you look at honeybees in their natural hive environment, you'll quickly see that they are very outgoing. In fact, according to Robinson, honeybees can't survive without the social structure of their community within the hive.

This characteristic makes them a perfect species in which to study the genetics of behavior.

What's particularly interesting about bees is that rather than being stuck in a particular job, they change jobs according to the hive's needs. Robinson has identified certain genes whose activity changes during a job shift, suggesting that the insects' environment helps to shape their gene expression.

Researchers who are beginning to understand these connections are working in a brand-new field of investigation named by Robinson himself: sociogenomics.

What does all of this mean for humans, you wonder? It underscores the fact that, far from being set in stone, our genomes are influenced by both heredity and environment, fine-tuned and sculpted by our social life and the things we do every day.

Knowing about diseases that run in your family can help you guard against illness in the future.

Cause and Effect

What more do we need to know about how genes shape who we are and what we become?

"A lot," says Harvard's Richard Lewontin, who warned against oversimplifying the role of genes in health in his 2001 book, *The Triple Helix*. Lewontin's main point is that context plays an enormous role in determining how organisms grow and develop, and what diseases they get. A unique combination of genetic and environmental factors, which interact in a way that is very hard to predict, determines what each person is like.

Very few, if any, scientists would argue with this. Whether a gene is expressed, and even whether the mRNA transcript gets translated into a protein, depends on the environment. Few diseases—most of which are very rare—are caused completely by a mutated gene.

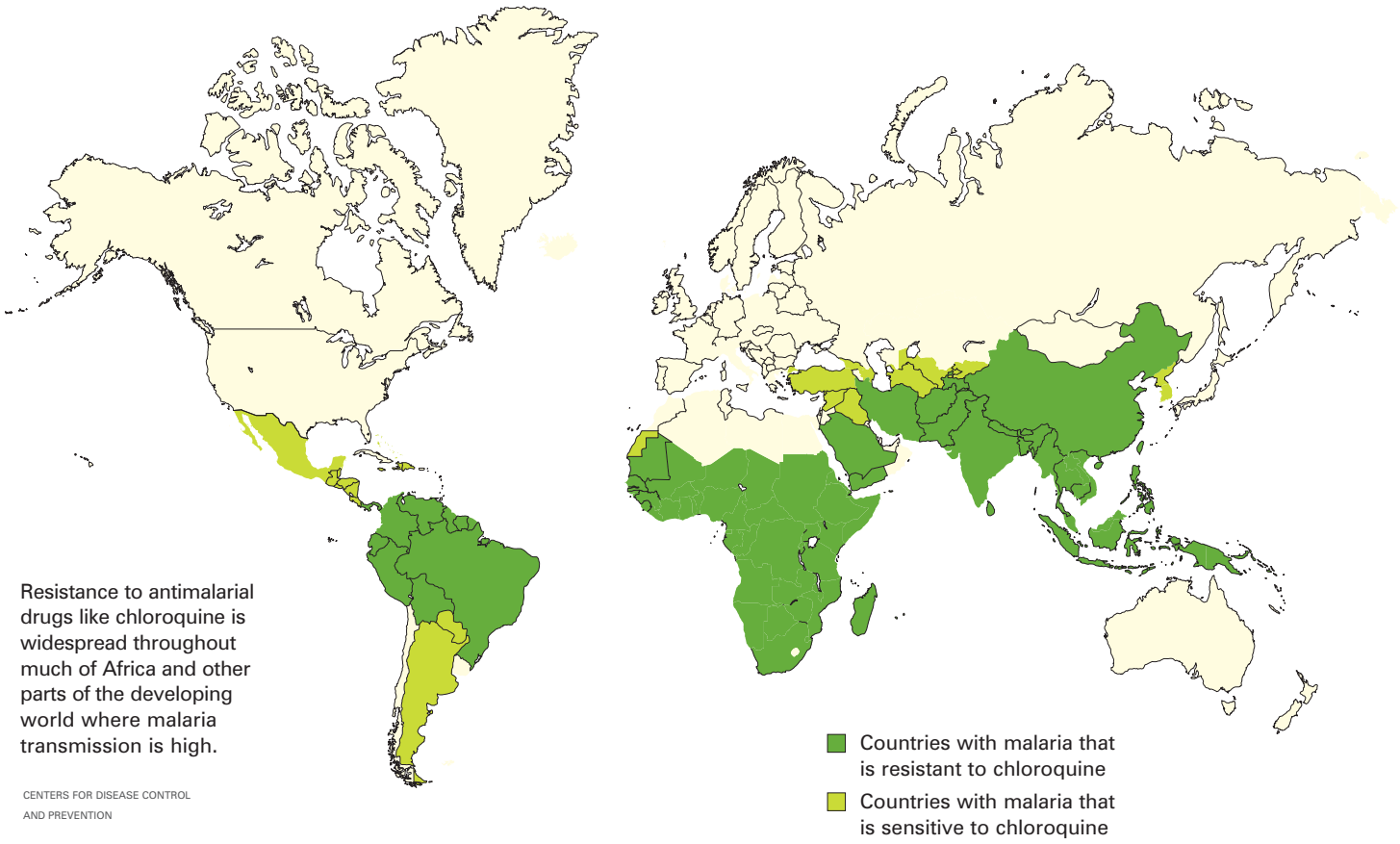
In most cases, getting or avoiding a disease depends not just on genes but on things within your control, such as diet, exercise and whether or not you smoke.

It will be many years before scientists clearly understand the detailed meaning of our DNA language and how it interacts with the environment in which we live. Still, it's a great idea to find out as much as you can about your family's health history. Did any of your relatives have diabetes? Do people in your family tree have cancer or heart disease?

Keep in mind that diseases such as these are relatively common, so it's pretty likely that at least one relative will have one of them. But if heart disease, diabetes or particular types of cancer "run in your family," especially if a lot of your relatives get the condition when they are fairly young, you may want to talk with your doctor about your own risk for developing the disease.

In 2005, the U.S. Surgeon General developed a Web-based tool for organizing family health information. Called "My Family Health Portrait" (see http://www.hhs.gov/familyhistory), this tool arranges information into a printout that you can carry to the doctor's office. The information can help you and your doctor determine your risks for various conditions.

If you do discover that you are at higher-than-usual risk for a disease like breast cancer or heart disease, you may be able to prevent the disease, or delay its onset, by altering your diet, exercising more or making other lifestyle changes. You may also be able to take advantage of screening tests like mammograms (breast X rays that detect signs of cancer) colonoscopies (imaging tests for colon cancer) or blood sugar tests for diabetes. Screening tests can catch diseases early, when treatment is most successful.



Resistance to antimalarial drugs like chloroquine is widespread throughout much of Africa and other parts of the developing world where malaria transmission is high.

CENTERS FOR DISEASE CONTROL AND PREVENTION

Us *vs.* Them

Many scientists focus on human genes, most of which have counterparts in the genomes of model organisms. However, in the case of infections caused by microorganisms, understanding how the genomes of bacteria, viruses and parasites differ from ours is a very important area of health research.

Most of the medicines we take to treat infections by bacteria and viruses have come from scientists' search for molecular weak points in these tiny organisms. As mentioned in Chapter 1, for example, some antibiotics kill bacteria by disarming their protein-making ribosomes.

So why don't they kill human cells, too? The answer is that human and bacterial ribosomes are different. Genome sequencing is a powerful tool for identifying differences that might be promising targets for new drugs.

Comparing genetic sequences in organisms that are resistant and non-resistant to drugs can reveal new approaches to fighting resistance. Drug resistance is a worldwide problem for a number of diseases, including malaria.

Although researchers have developed several different types of medicines to treat this disease—caused by parasites carried by mosquitoes, not by a bacterium or a virus—malaria is rampant, especially in the developing world.

GENETICS AND YOU: *Eat Less, Live Longer?*

Would you consume an extremely low-calorie diet if it meant you would live longer?

The kind of diet we're talking about isn't just cutting back here and there. It involves severely reducing calorie intake to about 60 percent of what we normally eat, enough to make most people ravenously hungry.

A 19th-century French doctor, Maurice Gueniot, thought the tradeoff would be worth it. Throughout his adult life, he ate very little. He died at the ripe old age of 102!

Later, in the 1930s, researchers followed up on this observation by showing that rats on a diet containing 20 percent indigestible fiber—calories that can't be used—lived much longer than their normally fed peers.

Intrigued by the health connection, scientists are continuing to investigate potential links between diet and aging, and genetic studies are starting to turn up some clues.

For example, geneticist David Sinclair of Harvard Medical School has found that proteins known as sirtuins may be able to stall aging. As yeast cells age, they accumulate extra DNA, which eventually kills them. Sinclair discovered that sirtuins become more active in yeast cells that are on a low-nutrient "diet." He reasons that by restricting the formation of extra DNA, sirtuins keep the yeast young.

Not so fast, say other scientists like geneticist Stanley Fields of the University of Washington. His experiments have turned up other, unrelated genes linked to lifespan in yeast. He argues that while calorie restriction is the only intervention that has been shown to extend lifespan in a wide range of organisms, including mammals, the accumulation of extra DNA does not always appear to play a role in this process.

What's the final answer, you ask? It's probably a bit of both.

Molecules like sirtuins, which are involved in cellular metabolism, may protect cells against the harmful effects of stress, extending lifespan. Other molecules that affect different aspects of cell health may be just as important.

Lifespan in complex, multicellular organisms like people is affected by many different factors, most of which we know very little about. For sure, understanding more about these mystery molecules could have a considerable benefit—perhaps providing you a chance to add years to your life without starving!

CDC/ JAMES GATHANY

Mosquitoes spread malaria by picking up parasites from blood and spreading them to the next person they bite. Resistance spreads this way, too.

This is partly because not all people have access to treatment, or to simple preventive measures like bed nets, which protect sleeping people from mosquito bites. But another problem is the malaria parasite itself, which has rapidly evolved ways to avoid the effects of antimalarial drugs.

Scientists are trying to counter this process by studying microbial genetic information. In the case of malaria, geneticists like Dyann Wirth of the Harvard School of Public Health compare the genomes of drug-resistant parasites and those that can still be killed by antimalarial medicines.

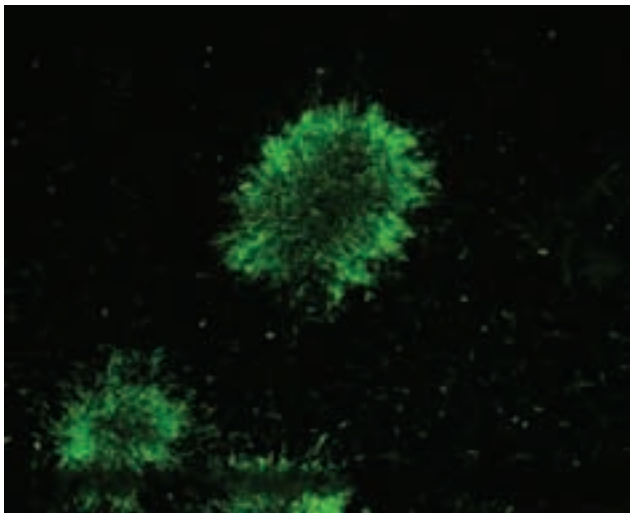
Wirth's research suggests that it should be possible to develop a simple, inexpensive genetic test that could be given to people with malaria, anywhere in the world. This test would identify drugs that are likely to be most effective and help decrease the rate at which parasites become resistant to the antimalarial medicines we already have.

Gang Warfare

Did you know that scientists are using genetics to break up gangs … of microbes, that is? These gangs, known as **biofilms**, are layers of slime that develop naturally when bacteria congregate on surfaces like stone, metal and wood. Or on your teeth: yuck!

Biofilms grow in all sorts of conditions. For example, one biofilm known as "desert varnish" thrives on rocks, canyon walls or, sometimes, entire mountain ranges, leaving a reddish or other-colored stain. It is thought that petroglyphs left on boulders and cave walls by early desert dwellers were often formed by scraping through the coating of desert varnish formations with a hard object.

Sometimes, biofilms perform helpful functions. One of the best examples of the use of biofilms to solve an important problem is in the cleaning of wastewater.



P. SINGH AND E. PETER GREENBERG

Biofilms, like the one shown in this fluorescent microscopic photo, are bacterial communities.

DENISE APPLEWHITE

Bonnie Bassler (right) uses glow-in-the dark bacteria to study the genetics of biofilms.

But biofilms can be quite harmful, contributing to a wide range of serious health problems including cholera, tuberculosis, cystic fibrosis and food poisoning. They also underlie many conditions that are not life-threatening but are nonetheless troublesome, like tooth decay and ear infections.

Bacteria form biofilms as a survival measure. By living in big groups rather than in isolation, the organisms are able to share nutrients and conserve energy. How do they do it?

A biofilm is not just a loose clump of cells—it's a highly sophisticated structure. As in any community, the individuals in biofilms communicate with each other.

Beyond that, many aspects of biofilms are poorly understood. Bacterial geneticist Bonnie Bassler of Princeton University in New Jersey is working to understand biofilms better, with the goal of being able to use this knowledge to break up bacterial "gang meetings."

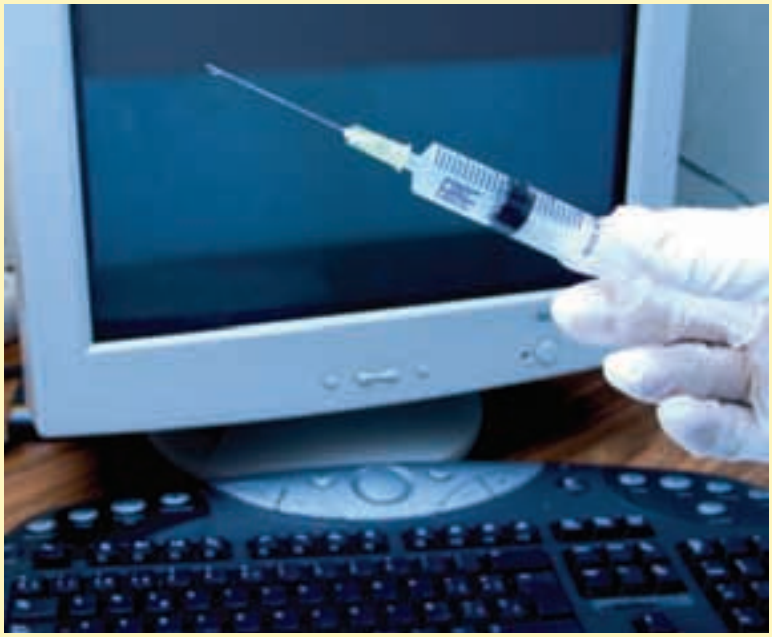
Bassler's research subjects have a definite visual appeal. They glow in the dark, but only when they are part of a group. The bioluminescence, as the glow is called, arises from chemical reactions taking place within the biofilm. It provides a way for the bacteria to talk to each other, estimate the population size of their community and distinguish themselves from other types of microorganisms.

Through her studies, Bassler has identified a set of molecules that biofilm-forming microorganisms use to pass messages to each other. By devising genetically based methods to cut off the chatter, Bassler reasons, she may be able to cause bacterial communities to fall apart. This approach would provide a whole new way to treat health problems linked to harmful biofilms.

The Tools of Genetics: Mathematics and Medicine

What if public health officials had a script for what to do in the face of an infectious disease outbreak that had never been seen before? One thing that would help them prepare for this sort of scenario is the ability to know, ahead of time, how an epidemic develops and spreads.

Toward this goal, some scientists are using mathematical tools to create simulations, or models, of infectious disease outbreaks. They can then use the models to test the effects of various intervention strategies. Part of the work involves plugging in genetic information about how infectious organisms evolve over time and how fast they change as they interact with human populations.



Computer simulations are helping scientists understand how infectious diseases spread.

Since 2005, the Models of Infectious Disease Agent Study (MIDAS), a team of biologists, computer scientists, statisticians, mathematicians, social scientists and others, has been modeling a flu pandemic—a huge, global epidemic.

Initially, the models focused on avian influenza, a type of disease occurring naturally among wild birds. At the time, health experts worldwide worried that the virus' genetic material could mutate, making it much easier for the so-called "bird flu" to pass between humans.

To simulate the potential disease spread, the scientists wrote computer programs that incorporated information about the bird flu virus and actual communities. Including details about people—not just their ages and genders, but also where they live, work or go to school—let the researchers create a synthetic population that could mirror how a real one might get sick and spread disease.

The scientists ran the programs on large computers to see how the flu could spread with and without different interventions. The results indicated that to successfully contain an epidemic, health officials would need to find the first flu cases fast and implement a combination of public health measures very quickly.

This early work helped MIDAS scientists develop similar models of H1N1 or "swine flu," the first actual pandemic flu strain since 1968. Starting in April 2009, they gathered incoming public health data to simulate the potential spread of this global flu, identify the groups most likely to get sick and evaluate the usefulness of different public health measures, such as vaccination and quarantine. Their models suggested that vaccinating schoolchildren early in an outbreak could reduce overall disease spread and that people at risk of serious complications should be given antiviral medications to take at the first signs of illness.

During both the bird and swine flu modeling efforts, the MIDAS scientists worked closely with public health officials to address specific questions. The answers informed U.S. pandemic flu preparedness planning.

Influenza, however, is not the only infectious disease making people sick. MIDAS scientists are also modeling other major health threats, including cholera, dengue fever, malaria, tuberculosis and methicillin-resistant *Staphylococcus aureus* (MRSA).

Got It?

Discuss how mathematics can help scientists ask questions about human health.

Would you contribute a sample of your DNA for genetic research on common diseases like heart disease, depression or cancer—even if you didn't have any of these health problems? Why or why not?

Drugs work like they're supposed to in only half the people who take them, so scientists are trying to make "personalized medicines" that work very well in an individual because they match his or her genetic make-up. Are there economic, social or other issues that the development of such medicines might raise?

21st-Century Genetics

Medicine has evolved tremendously since the earliest human civilizations, when the diagnosis and treatment of disease were far from scientific. Medieval medicine, for example, relied heavily on supernatural beliefs. Limited scientific knowledge led to seemingly bizarre practices like opening the vein of a sick person and draining off quarts of precious blood!

Later, in the Renaissance period of the 15th and 16th centuries, scholars centered on anatomy. One of them, the Italian artist-inventor Leonardo da Vinci, created beautiful and accurate

RARE BOOK AND SPECIAL COLLECTIONS DIVISION, LIBRARY OF CONGRESS

By the end of the 16th century, anatomy was a common focus for scientific scholars.

illustrations of the human body. His work and that of other scientists of his day focused on the practice of dissection, providing never-before-seen details of the body's architecture of limbs, joints, muscles, nerves and vessels.

Modern medicine got its real start during the 19th century, after the microscope was invented. Medical school subjects like physiology, pathology and microbiology were born. During this time, scientists discovered that bacteria—not evil spirits or other imaginary entities—caused human diseases like cholera, anthrax and tuberculosis.

The birth of modern genetics, which occurred in the 20th century, accelerated the study of all these areas of science. Now, at the start of the 21st century, opportunities have never been greater for turning scientific knowledge into better health for all.

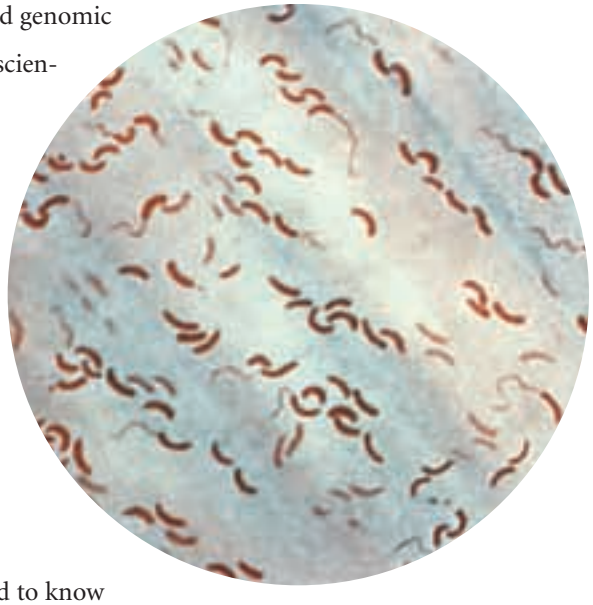
We often take for granted the amazing complexity of the human body. Without even thinking, we sweat to maintain body temperature, get hungry when we need energy and feel tired when we need to sleep.

These seemingly simple actions require a sophisticated coordination of many different organs and the millions of molecules that work together inside them. Thousands of networks of interacting genes underlie these actions in our bodies. But these systems are proving to have far more fluctuation than scientists originally suspected.

19th-century scientists discovered that bacteria can cause disease. *Bacillus anthracis* (left) causes anthrax and *Vibrio cholerae* (below) causes cholera.

PAUL KEIM (ANTHRAX), CDC/WILLIAM A. CLARK (CHOLERA)

One of today's challenges is to map the actions and interactions of all these molecules, a focus of the new field called **systems biology**. Genetic and genomic research is helping scientists tackle many questions in this area. By building models of cells, tissues and organs in action, scientists hope to learn how these complex, dynamic systems work.



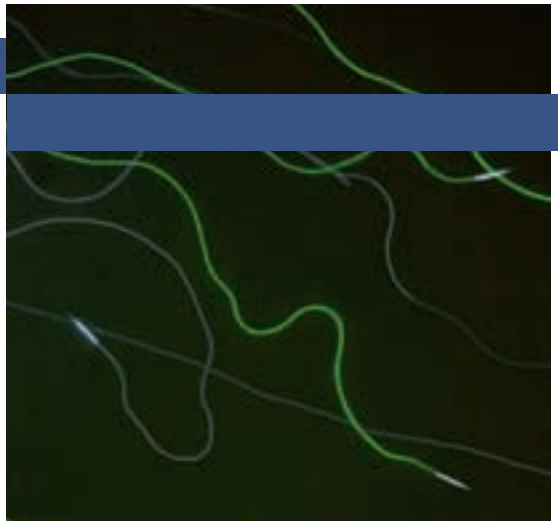
Researchers need to know these basics in order to understand how the systems fail, when disease strikes. An essential tool in this research is the computer.

No Lab? No Problem!

Those who work at the intersection of computer science and biology often combine and analyze data from many different sources, looking for informative patterns.

Andrey Rzhetsky of the University of Chicago is one of these people. Through an approach known as knowledge engineering, Rzhetsky and his team write computer programs that scan the contents of thousands of published scientific papers. The "knowledge mining" tool they use, called GeneWays, focuses mainly on research literature about changes in genes and proteins.

The program first scans scientific papers using pre-set search terms, much like a Google™ search of the Web. Next, it evaluates the search results and makes sure they don't overlap. For example, if a molecule has 16 different names in different papers, the program simplifies it to just one.



CATHERINE FERNANDEZ AND JERRY COYNE

▲ Fruit fly sperm cells glow bright green when they express the gene for green fluorescent protein.

Green Fluorescent Protein

Here's an interesting news flash: "*Glow-in-the-dark jellyfish revolutionizes genetic research!*"

Although it may sound bizarre, the claim is true. A jellyfish protein is essential to modern cell biology experiments that track the movements, quantities and interactions of the millions of proteins inside cells.

Called green fluorescent protein, or GFP, this natural protein is found in specific parts of the jellyfish. Those parts glow because the protein absorbs energy from light in the environment and then produces a different color of light.

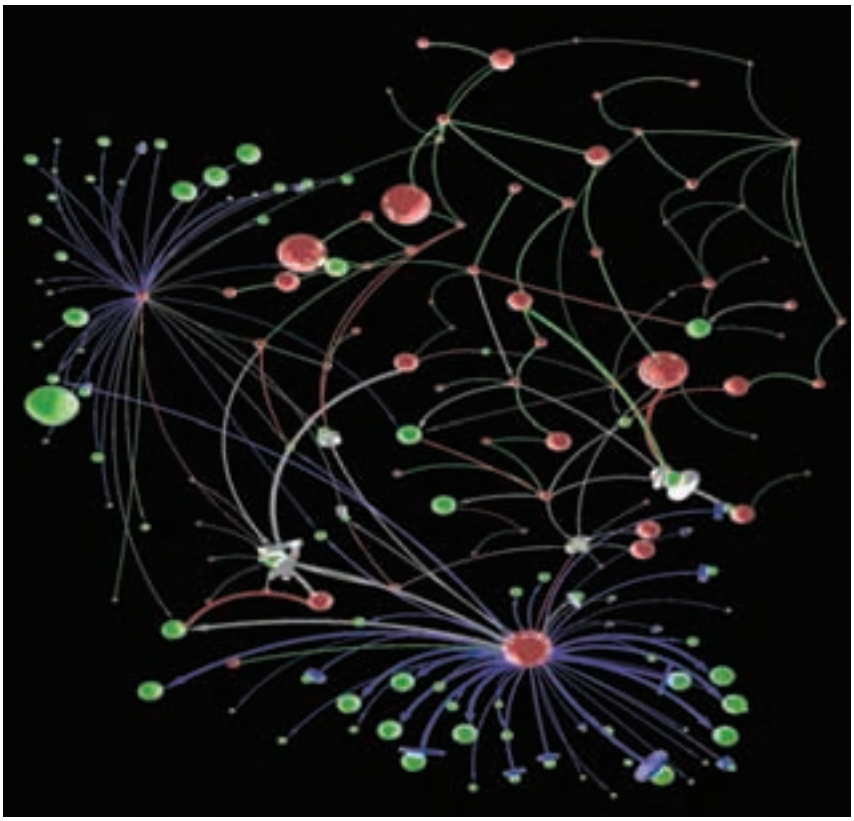
Scientists don't really know how and why jellyfish use their glow. They do know that jellyfish don't flash at each other in the dark, nor do they glow continuously. And the glow is rarely seen in undisturbed animals.

Taken out of the jellyfish, GFP has played a major role in advancing the study of genes and the proteins they encode. The story of how GFP

Finally, after applying specific rules, sort of like "biological grammar," the computer program identifies associations, which are possible links between molecules. The information then goes to a database that Rzhetsky and other scientists use to build large networks of molecular interactions.

Rzhetsky and his team used GeneWays to identify risk genes for Alzheimer's disease, a complex condition thought to be caused by many factors. In analyzing the data, Rzhetsky found important "nodes," molecules that play key roles in the disease gene network that GeneWays modeled.

These predicted molecular interactions were later confirmed by other researchers working in a lab, underscoring the value of computer modeling as a way to learn more about the molecular basis of disease.



ANDREY RZHETSKY AND KEVIN P. WHITE

▲ Andrey Rzhetsky uses the computer program GeneWays to locate important "hubs" of activity (large spheres) within massive gene networks. This particular network represents embryonic developmental pathways in a fruit fly.

became a research tool began in 1992, when Martin Chalfie of Columbia University showed that the gene that makes GFP produced a fluorescent protein when it was removed from the jellyfish genome and transferred to the cells of other organisms (see page 38). Chalfie, a developmental biologist, first put the gene into bacteria and roundworms, creating glowing versions of these animals.

Since then, researchers have transferred the GFP gene into many other organisms, including fruit flies, mice and rabbits—and even human cells growing in a lab dish. Recently, scientists used the GFP gene to create green-glowing zebrafish. Although the fish were created for the purpose of scientific research, they've also become an "exotic" species for home aquariums.

Thanks to GFP and related technologies, scientists can now view living cells and their constantly moving contents. GFP is also used in diagnostic tests for drugs, foods, herbicides and hazardous chemicals.

Chalfie and two other scientists received the 2008 Nobel Prize in chemistry for the discovery and development of GFP.

MARTIN CHALFIE

▲ Scientists engineered this experimental worm to express green fluorescent protein in two of its nerve cells (bright green spots).

Hard Questions

While the task of sorting through large volumes of genomic data remains a central challenge in modern biology and medicine, one of the knottiest dilemmas to emerge from this research is a social and ethical one. That is, how should people make use of information about their own genes?

Because genetic information is both powerful and incredibly personal, there are deep societal concerns regarding its use. These concerns include the potential for discrimination on the basis of a person's risk of disease or susceptibility to toxicity from an environmental chemical.

Some laws are already in place to protect individuals from the misuse of their genetic information. When you visit a new doctor, nurse practitioner, or dentist, you'll be asked to read and sign a form that outlines your medical privacy rights under the Health Insurance Portability and Accountability Act, or HIPAA. This law protects your genetic and other personal health information from being used or shared without your knowledge.

Another law, the Genetic Information Nondiscrimination Act, or GINA, prohibits discrimination in health coverage and employment based on genetic information.

It's important to realize that, in most cases, genetic information cannot offer definitive proof that a disease will occur. But if you have a very strong family history of breast cancer, for example, there may be a faulty gene in your family that increases your risk of getting the disease.

Doctors can now test for two known gene variants associated with inherited forms of breast cancer, BRCA1 and BRCA2. If you carry either of these gene variants, your lifetime risk of getting breast cancer is significantly higher than it would be for someone without either variant. But some people who have BRCA gene variants never get breast cancer.

Only about 5 percent of all breast cancer can be traced to a known, inherited gene variant. Since so many breast cancers are not linked to BRCA1 or BRCA2, genetic testing for these variants is irrelevant for the vast majority of people who do not have a family history of breast cancer.

But let's say you do have a relative who tested positive for BRCA1 or 2. Should you get tested, too?

A difficult question, for sure, but consider this: Knowing about this risk ahead of time might save your life. For example, you might want to begin getting mammograms or other screening tests at an early age. If cancer is found

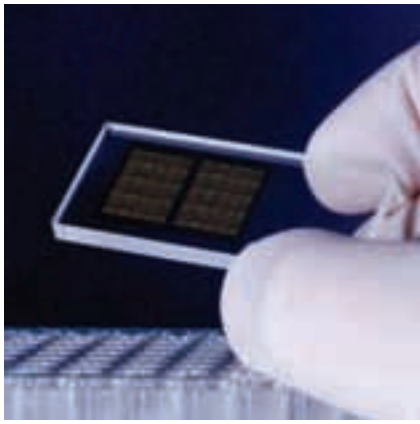
very early, it is usually more treatable, and the odds for a cure are much higher.

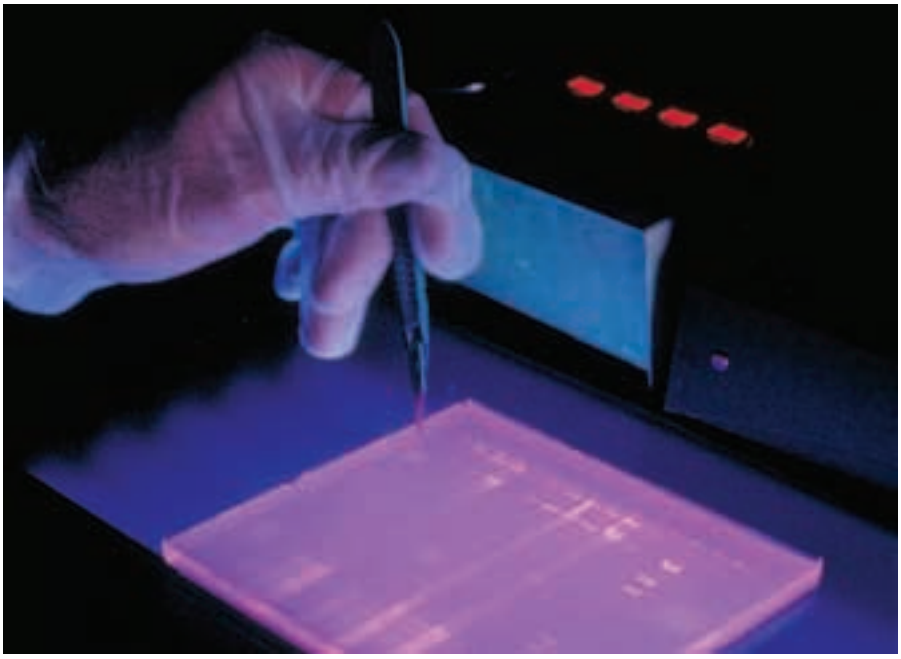
Currently, diagnostic laboratories across the United States offer genetic tests for almost 2,000 disorders. Some of these tests detect problems with entire chromosomes, not just individual genes. Perhaps the most well-known example of a chromosome problem is Down syndrome, in which cells have an extra copy of chromosome 21 (see page 11).

Most genetic diseases aren't caused by a chromosome abnormality, or even by one gene variant. Cystic fibrosis, for example, is due to a faulty gene, but more than 30 different variants of this gene can cause the disease, and those are just the ones researchers know about!

How can there be 30 different variants of one gene? Remember that a gene is a long DNA sequence, consisting of hundreds of nucleotides. A change in one of those nucleotides produces one variant, a change in another produces another variant, and so on.

Because there are so many possibilities, it's hard to tell whether a person has a variant form of the cystic fibrosis gene. So the standard genetic screening test for this disease scans for all of the more than 30 variants known to cause cystic fibrosis.





▲ Scientists are developing genetic tests that will help doctors diagnose and treat diseases.

Doctors usually order a genetic test only if a person has a strong family history of a disease. But even so, deciding to have such a test is not a simple choice. Think about what you would do with the information.

One thing you might consider is whether you *could* do something with what you learn from a genetic test.

You've already read about what you could do if you discovered that you were at high risk for developing breast cancer. But what about a condition that shows up in middle-aged or older people—or one for which there is currently no cure?

As a teen or young adult, would you want to know that you'd get a serious, perhaps incurable, disease later in life?

Patients and doctors face these tough issues every day. Even years from now, when researchers know more about the molecular roots of disease, genetic tests will rarely provide easy answers. In most cases, they won't even provide "yes" or "no" answers.

Rather, much like a cholesterol test, they will predict whether a person's risk of getting a disease is relatively high, low or somewhere in between. This is because many factors besides genes, including lifestyle choices such as diet and exercise, also play a role in determining your health.

Good Advice

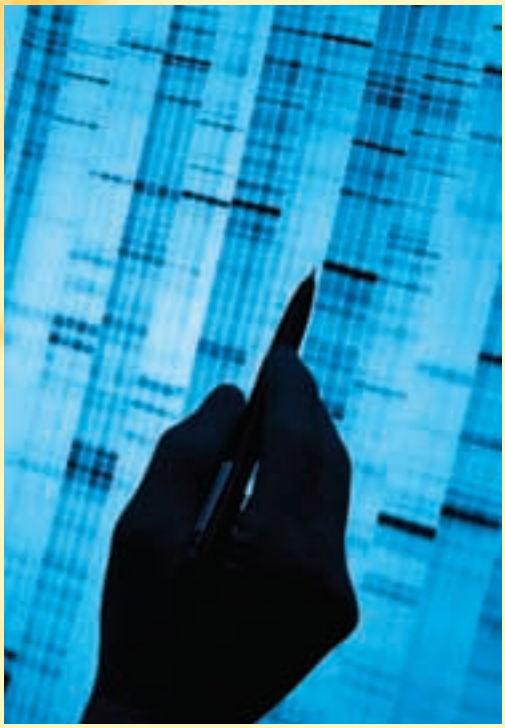
Since the story of genes and health is so complicated and is likely to stay that way for a while, it is very important to consider genetic information in context. Health care professionals known as genetic counselors can be a big help to people who are thinking about getting a genetic test.

As a profession, genetic counseling has been around since the mid-1900s. However, only a few specialty clinics offered counseling at that time. Now, genetic counseling is much more widely available.

GENETICS AND YOU: *Crime-Fighting DNA*

Like your thumbprint, your genes are unique, unless you have an identical twin. As such, DNA "fingerprinting" has become a powerful crime-fighting tool. DNA forensics is a fast-growing specialty that has applications beyond putting criminals behind bars.

In addition to identifying suspects who leave traces at the scene of a crime (for example, strands of hair, drops of blood or skin cells), DNA forensic technology can identify victims in a natural disaster, such as the December 2004 tsunami that ravaged Indonesia and other Asian countries. DNA fingerprinting can also match a transplant patient to an organ donor or establish paternity and other family relationships.



Genetic fingerprinting is not limited to people. It can find small but potentially deadly traces of disease-causing bacteria in food or water, determine whether an expensive horse was sired by a Kentucky Derby winner or figure out whether a puppy's parents were first cousins.

DNA fingerprinting techniques work by looking for differences among gene sequences that are known to vary between people (or between individuals from any species). Scientists read the sequence in a dozen or so places to create a molecular profile. The chances of a molecular fingerprint being the same in two people or two organisms are vanishingly small.

Today's genetic counselors have gone through a rigorous training process in which they earn a master's degree and learn genetics, medicine, laboratory procedures, counseling, social work and ethics. Genetic counselors do their work in many different settings, including hospitals, private clinics, government agencies and university laboratories.

An interesting aspect of the job is that genetic counselors address the needs of entire families, rather than just individual patients. To evaluate genetic risk and its potential consequences, these professionals gather a family medical history covering generations.

Genetics, Business, and the Law

Can a scientist claim rights to a gene that he discovered in worms and that has a nearly identical counterpart in humans?

Is a person who gave a blood or tissue sample entitled to profits from a company that develops a drug based on genetic information in her sample, or to a lifetime supply of the drug?

Can a blood or tissue sample that was donated for one purpose be used for an entirely different study several years later, without asking the donor if that's OK?

These and other issues are hotly debated in ethics and legal circles. Many of the most

Field Study

The word most often used to refer to applications of genetic research, especially those leading to products for human use, is **biotechnology**. It involves techniques that use living organisms—or substances derived from those organisms—for various practical purposes, such as making a biological product.

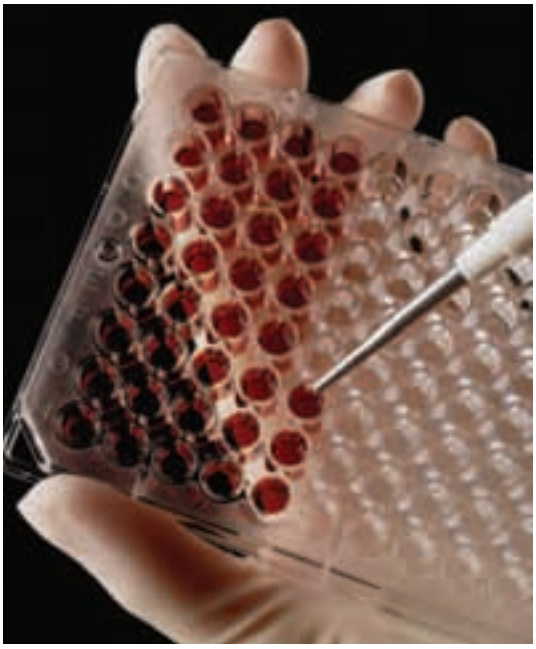
One major application of biotech - nology is in agriculture. Actually, this is hardly new: Humanity has engaged in agricultural biotechnology for 10,000 years or more. Many traditional farming practices, from plant breeding to animal husbandry, are really forms of biotechnology.

But in today's agricultural industry, biotechnology generally means the use of molecular biology, recombinant DNA technology, cloning and other recent scientific approaches to produce plants and animals with new traits.

This usually involves transferring genetic material from one kind of organism into another. Using the same techniques that were developed for putting genes into animals for research purposes, scientists can create crop plants with desirable traits, such as improved flavor or better resistance to insect pests. Transferring specific genes is faster and more efficient than traditional breeding approaches.

The United States is home to far more genetically modified crops than anywhere else in the world. In 2009, 85 percent of the country's corn, 88 percent of its cotton and 91 percent of its soybeans were cultivated from seeds genetically modified to resist plant pests and certain herbicides used to control weeds.

Many believe that agricultural biotechnology is an important driver for improving world health. They say that genetic modifications may be the only hope for pest-ravaged crops, such as bananas, that are essential to the economies of poor countries. The creation of edible plants that contain medicine, serve as a form of vaccination



controversial topics have to do with the idea of patenting life forms.

Traditionally, when an inventor comes up with a new idea and wants to sell it—whether it's a radio-controlled toy boat or a customized laboratory chemical—he or she submits an application to the U.S. Patent and Trademark Office.

By issuing patents, the Federal Government gives an inventor ownership of his or her creation. Patents give inventors time to optimize their products and control how their inventions are used, allowing them to make money from their creativity.

or deliver extra nutrients—such as the recently developed rice that makes vitamin A—could also contribute in major ways to global health.

But opposition from farmers and consumers within and outside the United States has clouded agricultural biotechnology's future. Some object to the development of plants that are naturally resistant to herbicides, partly out of concern that the trait might jump to weeds, making them impossible to destroy.

Environmental advocacy groups worry that genetically modified plants may impact the future biodiversity of our planet by harming beneficial insects and possibly other organisms. However, the U.S. Environmental Protection Agency has stated that there is no evidence to date that indicates that biotech crops have any adverse effects on non-targeted wildlife, plants or beneficial insects.

Of course, careful field tests of newly created, genetically modified plants and animals are essential to be sure that they cause no harm to other organisms or to the environment.

▲ Biotechnology helps agricultural scientists create crops with desired traits. The majority of cotton and soybeans in the United States are grown with genetically modified seeds that resist viruses and other plant pests.

However, nobody invented a gene, a naturally occurring chemical or a protein, so why should a person or a company be able to own it and control its destiny in the marketplace?

Patent laws in the United States and Europe prohibit anyone from patenting a gene as it exists in the human body. But patents have been issued for specific medical uses of genetic information.

Patents can be great for business, and they can help make the results of research widely available through commercial ventures, but they also have the potential to slow research because patent-holders control how information related to the patent is used. For example, researchers who wish to use patented genetic information may need to acquire a license first. This can be time-consuming and expensive.

Concerned about possible negative effects of patenting genes, the U.S. National Institutes of Health has worked with the U.S. Patent and Trademark Office to establish guidelines for what kind of genetic information can be patented. Since this area of medical research is an ever-moving target, government scientists, policymakers and the courts continue to clarify patent and licensing issues in the hope of keeping data that is valuable for research in the public domain.

Careers in Genetics

Opportunities to be part of genetic and genomic research have never been greater or more exciting. In addition to studying human genes, scientists are gathering information about the genes of many other living things, from microbes that cause disease to model organisms like mice and *Drosophila*, livestock and crop plants.

Although computers do some of the work, this avalanche of information has to be analyzed by thousands and thousands of human brains. In addition to identifying genes, scientists must figure out what the genes do and—even more complicated—how they do it.

We need laboratory scientists, doctors to do clinical research and treat patients, genetic counselors to help people understand the information in their genes, and lawyers and ethical specialists who can address legal and policy concerns about the use of genetic information.

In especially high demand are people with expertise in mathematics, engineering, computer science and physics. The field of bioinformatics, which develops hardware and software to store and analyze the huge amounts of data being

generated by life scientists, is especially short of qualified workers. As a result, bioinformatics scientists are in high demand.

Many careers in genetics and genomics require advanced degrees such as a Ph.D. or M.D. But people with master's or bachelor's degrees are also needed to fill thousands of rewarding jobs as genetic counselors, research assistants and lab technicians.

For more career information, see http://www.ornl.gov/sci/techresources/Human_genome/education/careers.shtml or http://science.education.nih.gov/LifeWorks.

The Tools of Genetics: Informatics and Databases

For most of its history, biology managed to amass its data mostly with the help of plain old arithmetic. Gregor Mendel did genetic analysis by simply counting the different kinds of offspring produced by his peas. By contrast, today's genetic research creates too much data for one person, or even a scientific team, to understand. New technologies are needed to manage this huge amount of data.

Consider this: Gene-sequencing machines can read hundreds of thousands of nucleotides a day. Gene chips are even faster. The information in GenBank®, a widely used database of all known DNA sequences, now doubles in just 3 years. A single laboratory doing cutting-edge genetic research can generate hundreds of gigabytes of data a day, every day. For comparison, 100 gigabytes could hold an entire floor of journals in an academic library.

How can anyone make sense of all this information? The only way is to enlist the aid of computers and software that can store the data and make it possible for researchers to organize, search and analyze it. In fact, many of today's challenges in biology, from gene analysis to drug discovery, are really challenges in information technology. This is not surprising when you remember that DNA is itself a form of information storage.

Where are genetic and genomic data stored? One of the first biological databases was created to store the huge volume of data from experiments with the fruit fly *Drosophila melanogaster*.

Called FlyBase, it has grown into a huge, comprehensive, international electronic repository for information on *Drosophila* genetics and molecular biology, run by scientists for scientists. The information spans a century's worth of published scientific literature on *Drosophila melanogaster* and its relatives, including their complete genome sequences.

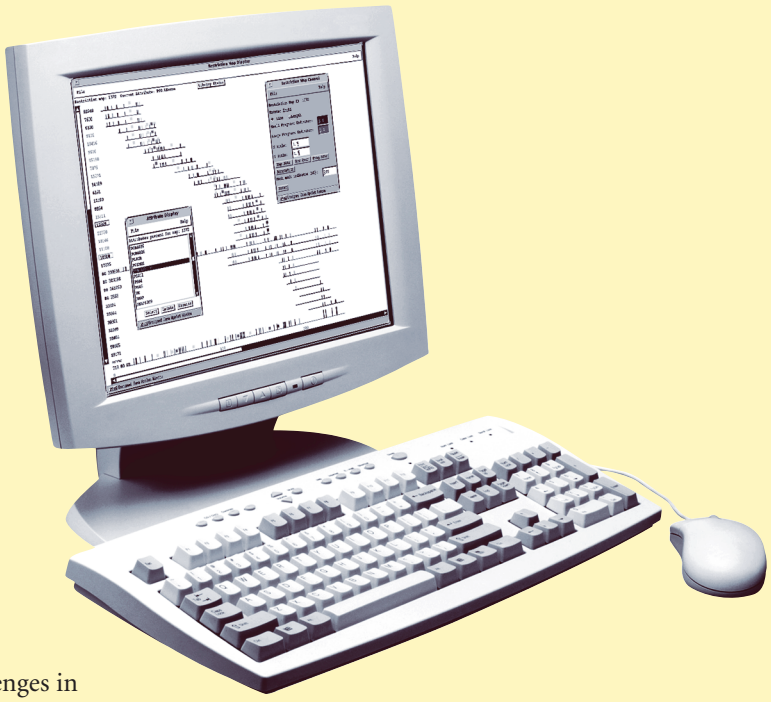
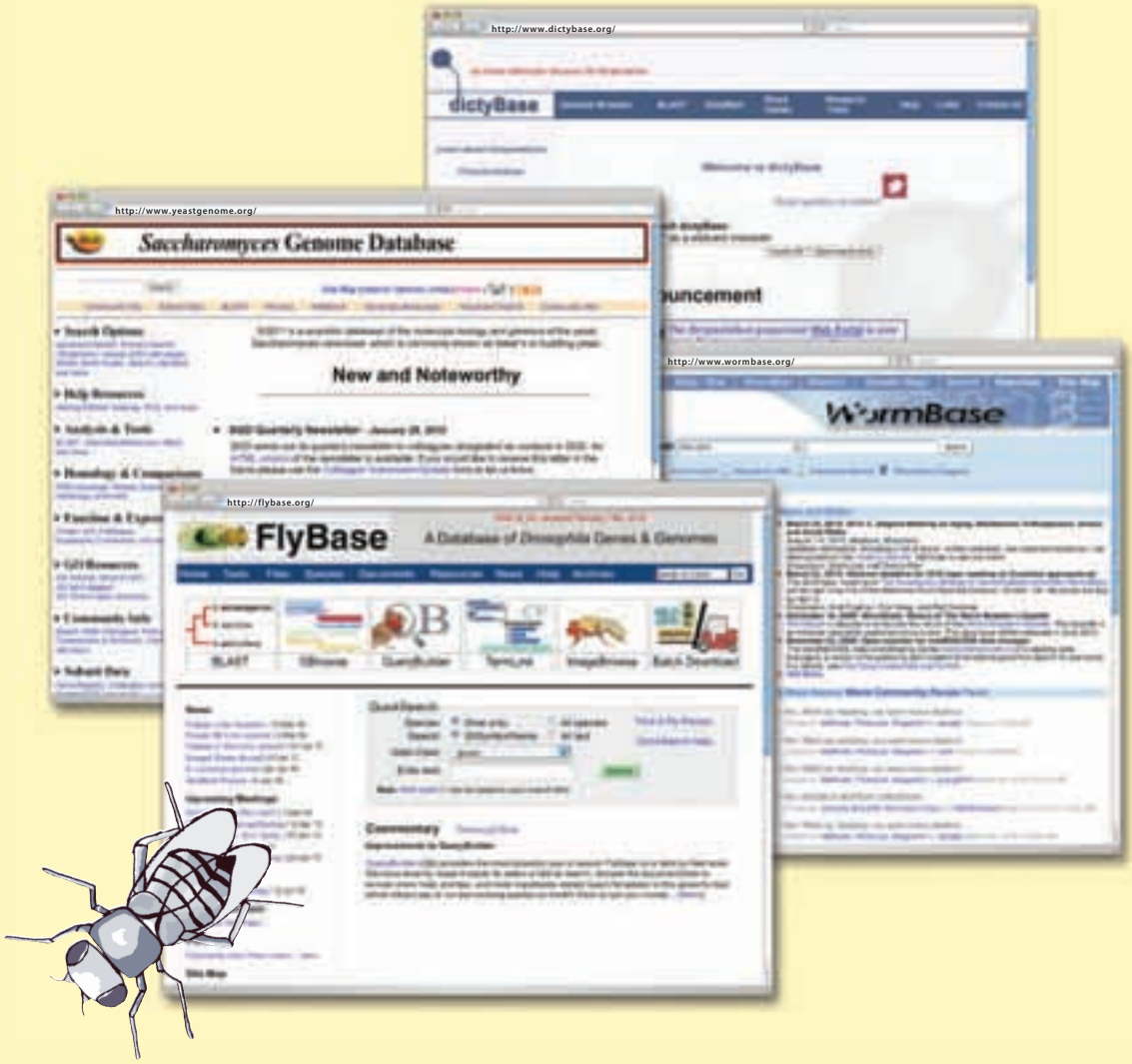


IMAGE ON COMPUTER SCREEN COURTESY OF TOM SLEZAK, LAWRENCE LIVERMORE NATIONAL LABORATORY



Databases like FlyBase are also useful to scientists working with other organisms, like mice or humans. A researcher who discovers a new mammalian gene may consult FlyBase to see if fruit flies have a similar gene and if the database contains hints about what the gene does. Since the functions of many genes are retained during evolution, knowing what a gene does in one organism often provides valuable clues about what it does in another organism, even if the two species are only distantly related.

Several other communities of researchers have created their own databases, including those dedicated to the investigation of the roundworm *Caenorhabditis elegans* (WormBase), the soil-dwelling amoeba *Dictyostelium discoideum* (DictyBase) and the strain of yeast used for many laboratory studies (*Saccharomyces* Genome Database).

A key goal is to make sure that all of these databases can "talk" to each other. That way, similar discoveries in different organisms—the important, common threads of all biology—can be identified quickly and analyzed further.

For this database communication to work, researchers in different fields must use the same terms to describe biological processes. The development and use of such a universal "ontology"—a common language—is helping scientists analyze the complex network of biology that underlies our health.

Got It?

Do you think modern research tools derived from genomics and bioinformatics will change the practice of medicine? How?

If a genetic test revealed that you had a 1 in 100 chance of develop ing a disease like type 2 diabetes, which can be prevented with lifestyle changes like eating a healthier diet and exercising more, would you change your behavior? What if the risk were 1 in 10?

How is genetic engineering similar to traditional farming? How is it different?

A biotechnology company uses genetic information from a patient volunteer and develops an effec tive, profitable medicine. Should the patient know that he or she was part of this process? Why or why not? What if the research did not lead to any medical advance?

Glossary

Amino acid | A building block of proteins. There are 20 amino acids, each of which is coded for by three adjacent nucleotides in a DNA sequence.

Anticipation | The disease process in which symptoms show up earlier and are increasingly severe in each generation.

Biofilm | A slime layer that develops naturally when bacteria congregate on surfaces.

Bioinformatics | The field of biology specializing in developing hardware and software to store and analyze the huge amounts of data being generated by life scientists.

Biotechnology | The industrial use of living organisms or biological methods derived through basic research; examples range from genetic engineering to making cheese or bread.

Chromatin | The organization and dense packaging of DNA in the nucleus of cells.

Chromosome | A cellular structure containing genes. Chromosomes are composed of DNA and proteins. Humans have 23 pairs of chromosomes in each body cell, one of each pair from the mother and the other from the father.

Circadian | Pertaining to a period of about 24 hours; applied especially to rhythmic biological repetition like the sleep-wake cycle.

Clone | In genetics, the process of making many copies of a gene or a whole organism. The term also refers to the isolation and manipulation of a gene.

Comparative Genomics | The study of human genetics by comparisons with the genetics of other organisms.

Diploid | Having two copies of each chromosome.

DNA | Abbreviation for deoxyribonucleic acid, the molecule that contains the genetic code for all life forms except for a few viruses. It consists of two long, twisted chains made up of nucleotides. Each nucleotide contains one base, one phosphate molecule and the sugar molecule deoxyribose. The bases in DNA nucleotides are adenine, thymine, guanine and cytosine.

DNA chip | See microarray.

DNA polymerase | An enzyme that copies DNA.

Enzyme | A substance (often a protein) that speeds up, or catalyzes, a chemical reaction without being permanently altered or consumed.

Epigenetics | The study of heritable changes in gene function that occur without a change in the DNA sequence.

Eukaryote | An organism whose cells have a membrane-bound nucleus.

Exon | A DNA sequence in a gene that codes for a gene product.

Gene | A segment of a DNA molecule that contains information for making a protein or, sometimes, an RNA molecule.

Gene chip | See microarray.

Gene expression | The process by which genes are first converted to messenger RNA and then to proteins.

Genetics | The scientific study of genes and heredity—of how particular qualities or traits are transmitted from parents to offspring.

Genome | All of an organism's genetic material.

Genomics | A "scaled-up" version of genetic research in which scientists can look at large numbers or all of the genes in an organism at the same time.

Haploid | Having one copy of each chromosome, as in a sperm or egg.

Haplotype | A set of closely linked genes or DNA polymorphisms inherited as a unit.

Histone | A type of protein found in chromosomes; histones attached to DNA resemble "beads on a string."

Homeobox | A DNA sequence found in genes involved in the regulation of the development of animals, fungi and plants.

Imprinting | The phenomenon in which a gene may be expressed differently in an offspring depending on whether it was inherited from the father or the mother.

Intron | A DNA sequence, or the RNA sequence transcribed from it, that interrupts the sequences coding for a gene product (exon).

Meiosis | The type of cell division that creates egg and sperm cells.

Microarray | Sometimes called a gene chip or a DNA chip. Microarrays consist of large numbers of molecules (often, but not always, DNA) distributed in rows in a very small space. Microarrays permit scientists to study gene expression by providing a snapshot of all the genes that are active in a cell at a particular time.

MicroRNA | A short piece of single-stranded RNA that does not encode a protein and controls the expression of genes.

Mitochondrion | The cell's power plant, supplying the energy to carry out all of the cell's jobs. Each cell contains up to 1,000 mitochondria. The structures contain their own small genomes, called mitochondrial DNA.

Mutation | A change in a DNA sequence.

Nucleotide | A building block of DNA or RNA. It includes one base, one phosphate molecule and one sugar molecule (deoxyribose in DNA, ribose in RNA).

Nucleus | The structure in the eukaryotic cell containing most of its genetic material.

Pharmacogenetics | The study of how people's genetic make-up affects their responses to medicines.

Protein | A molecule consisting of subunits called amino acids. Proteins are the cell's main building materials and do most of a cell's work.

Recombinant DNA | Hybrid DNA produced in the laboratory by joining pieces of DNA from different sources.

Replication | The process by which DNA copies itself in order to make a new genome to pass on to a daughter cell.

Ribosome | The cell structure in which proteins are manufactured. Most cells contain thousands of ribosomes.

RNA | Abbreviation for ribonucleic acid, the molecule that carries out DNA's instructions for making proteins. It consists of one long chain made up of nucleotides. Each nucleotide contains one base, one phosphate molecule and the sugar molecule ribose. The bases in RNA nucleotides are adenine, uracil, guanine and cytosine.

RNA interference (RNAi) | A gene-silencing process in which double-stranded RNAs trigger the destruction of specific RNAs.

RNA polymerase | An enzyme that transcribes a DNA sequence, creating mRNA.

RNA splicing | The process by which introns are removed and exons are joined together from an RNA transcript to produce an mRNA molecule.

Sequencing | Sometimes called DNA sequencing or gene sequencing. Discovering the exact order of the building blocks (see nucleotides) of a particular piece of DNA.

Stem Cell | A cell that can develop into many different cell types in the body.

Systems biology | A field that seeks to study the relationships and interactions between various parts of a biological system (metabolic pathways, organelles, cells and organisms) and to integrate this information to understand how biological systems function.

Telomere | A repeated DNA sequence that caps the ends of chromosomes.

Transcription | The first major step in gene expression, in which the information coded in DNA is copied into a molecule of RNA.

Translation | The second major step in gene expression, in which the instructions encoded in RNA are carried out by making a protein or starting or stopping protein synthesis.

Variant | A different version of a gene, one that has a slightly different sequence of nucleotides.

Discrimination Prohibited

Under provisions of applicable public laws enacted by Congress since 1964, no person in the United States shall, on the grounds of race, color, national origin, handicap, or age, be excluded from participation in, be denied the benefits of, or be subjected to discrimination under any program or activity (or, on the basis of sex, with respect to any education program or activity) receiving Federal financial assistance. In addition, Executive Order 11141 prohibits discrimination on the basis of age by contractors and subcontractors in the performance of Federal contracts, and Executive Order 11246 states that no federally funded contractor may discriminate against any employee or applicant for employment because of race, color, religion, sex, or national origin. Therefore, the programs of the National Institute of General Medical Sciences must be operated in compliance with these laws and Executive Orders.

Accessibility

This publication can be made available in formats that are more accessible to people with disabilities. To request this material in a different format, contact the NIGMS Office of Communications and Public Liaison at 301-496-7301; send e-mail to info@nigms.nih.gov; or write to the office at the following address: 45 Center Drive MSC 6200, Bethesda, MD 20892-6200. If you have questions or comments about this publication, you can use the same contact information to reach the office.

Additional Copies and Web Links

To order additional copies of *The New Genetics* or other free NIGMS publications, go to http://publications.nigms.nih.gov/order or use the contact information above.

The New Genetics is available online at http://publications.nigms.nih.gov/thenewgenetics.

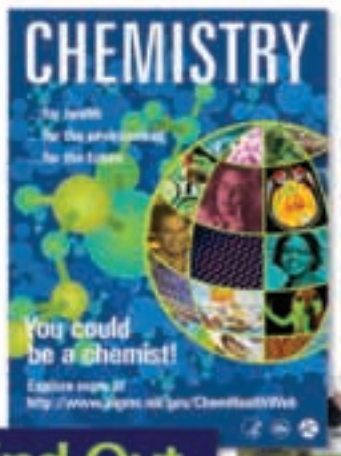
National Institute of
General Medical Sciences

NIH Publication No. 10-662
Revised April 2010
http://www.nigms.nih.gov

Read, Share, Recycle.

Findings Magazine

NIGMS Science Education Publications



Posters

CD-ROM Containing NIGMS Science Education Materials

ATTENTION READERS

We would like your comments on *The New Genetics*.
Please give us your feedback by filling out this postage-paid card.

1. How do you use this publication? _____

_____

2. What do you like best about this publication? _____

_____

3. How could we improve this publication? _____

_____

4. Do you have other suggestions or comments? _____

_____

Name _____

School/Organization* _____

Grade Level* _____

Address _____

_____

City _____

State _____ ZIP Code _____

E-mail (optional) _____

Phone (optional) _____

*required for multiple-copy orders

Browse and order free NIGMS educational publications at http://publications.nigms.nih.gov/order

Send me free poster(s) on:

- ❑ Living Laboratories
- ❑ Chemistry
- ❑ Find Out (what it's like to be a scientist)

❑ Send me a free CD-ROM containing all NIGMS science education materials for high school and early college students.

Send me a free copy of the NIGMS science education booklet(s) on:

- ❑ cell biology
- ❑ chemistry and biochemistry
- ❑ computers and biology
- ❑ genetics
- ❑ pharmacology
- ❑ structural biology/proteins

❑ Send me a copy of *Findings*, a free magazine with profiles of scientists and more.

❑ Add my name to the mailing list for free future issues of *Findings*.

- ❑ single copy
- ❑ class set ____ copies (limit 30)

DEPT OF HEALTH AND HUMAN SERVICES
NATIONAL INSTITUTES OF HEALTH
NATIONAL INSTITUTE OF GENERAL MEDICAL SCIENCES
45 CENTER DR RM 3AN.32 MSC 6200
BETHESDA MD 20892-6200

OFFICIAL BUSINESS
PENALTY FOR PRIVATE USE $300

NO POSTAGE
NECESSARY
IF MAILED
IN THE
UNITED STATES

BUSINESS REPLY MAIL
FIRST-CLASS MAIL PERMIT NO. 13155 WASHINGTON DC

POSTAGE WILL BE PAID BY NIH DHHS

NATIONAL INSTITUTES OF HEALTH
NATIONAL INSTITUTE OF GENERAL MEDICAL SCIENCES
OFFICE OF COMMUNICATIONS AND PUBLIC LIAISON
45 CENTER DR RM 3AN.32 MSC 6200
BETHESDA MD 20814-9692