Buried in the cells of each newborn is a unique set of genetic instructions. These molecular blueprints not only shape how the child will grow and develop and whether it will have brown eyes or blue, but what sorts of medical problems it might encounter. The child’s genes forecast the likelihood of developing such disorders as cancer, heart disease, or Alzheimer’s disease later in life. How can doctors detect them in the morass of a person’s DNA to try to prevent their deadly effects? The following article, adapted from an account by scientists Stuart Orkin and Gary Felsenfeld, explores the trail of research that led scientists to answer those questions and open the door to gene testing, which is promising to transform medicine powerfully. It provides a dramatic example of how science works and how basic research leads to practical results that were virtually unimaginable when the research was done.

Beth M.’s father died of colon cancer, as did his mother. Then colon cancer was diagnosed in two of her brothers, both in their 40s. Beth, 37, felt that a curse was hanging over her family and worried about her future and that of her children. Could she have inherited from her father a tendency to develop colon cancer at an early age, just as she inherited his hazel eyes?

Fortunately for Beth, researchers have pinpointed the defective gene that has plagued her family and causes the odds of developing colon cancer to be 85% or greater. The gene is called MSH2, and researchers have developed an experimental genetic test for defects in this particular gene. Within a short time of having her blood drawn, Beth could stop worrying—the blood test showed that she did not inherit the MSH2 gene from her father. Although the test result did not free Beth from the possibility of ever developing colon cancer—the test predicts only the likelihood of developing colon cancer fostered by the MSH2 gene, and not from other causes—she now knew that the uncomfortable and expensive colon-cancer detecting procedures that she underwent each year were no longer necessary; nor was the excessive worry.

What Beth didn’t know was that the simple genetic test she had for the MSH2 gene would not have been possible without more than 50 years of research by many scientists who paved the way for pinpointing the genes that foster susceptibility to specific diseases. Most of the scientists had no idea that their quest for answers to such basic questions as how yeast cells detect and repair flaws in their genetic material would lead to practical genetic tests on people like the one for the MSH2 gene. These tests are also raising ethical, social, and legal questions as they take medicine into uncharted territory.
Unraveling the Nature of the Gene

A major milestone along the trail to gene testing was the discovery of the structure of deoxyribonucleic acid (DNA), the molecule that contains genes. It had been known since the middle of the 19th century, when the monk Gregor Mendel conducted his famous pea-breeding experiments, that physical traits such as height and color were passed from one generation to the next via units of inheritance that later came to be called genes. But the physical character of the gene had eluded scientists until 1944, when bacteriologist Oswald Avery of New York’s Rockefeller Institute showed that all that was needed to transform harmless bacteria into a type that can cause pneumonia was their uptake of DNA from a pneumonia-causing strain of bacteria. That experiment suggested that genes were made of DNA, and it launched many researchers on a quest to determine the exact structure of DNA as a means of unraveling how genes exert their influence on all living things.

Two of these researchers, Rosalind Franklin and Maurice Wilkins, of King’s College in London, studied the pattern generated when x-rays were scattered from DNA fibers. The photographic image immediately revealed that the DNA structure was regular and helical. With that information and knowledge of the chemistry of the DNA components, James Watson and Francis Crick, then at the Medical Research Council laboratories in Cambridge, England, began building molecular models that might account for the details in the photograph. The model that they ultimately proposed in 1953 contains two helically twisted strands connected to each other by a series of molecular “rungs.” They suggested that each rung was composed of one of two chemical “base pairs” called adenine (A)-thymine (T) or guanine (G)-cytosine (C). These young scientists correctly surmised that it was the order of those A, T, G, and C bases on the DNA strand that spelled out the genetic endowment of every living organism. They also recognized that the two strands could be separated for copying—a simple mechanism for passing on genetic information from one generation to the next.

A few years after Watson and Crick clarified the structure of DNA, several other researchers, notably Marshall Nirenberg, at the National Institutes of Health, and Har Gobind Khorana, at the University of British Columbia, deciphered the genetic code that all living cells use to translate the series of bases in their DNA into instructions for the production of the thousands of proteins that determine the cell’s structure and carry out all its functions, including determining such genetic traits as eye color and susceptibility to cancer. The researchers discovered that each triplet of bases (CTG, for example) codes for one amino acid (in this case leucine) or for a signal to start or stop building the long chain of amino acids that creates a protein.

Genetic Errors Cause Disease

The precise arrangement (sequence) of A, C, G, and T bases on a DNA strand is the recipe that encodes the exact sequence of a protein. If the recipes have extra bases or misspelled bases or if some are deleted, the cell can make a wrong protein or too much or too little of the right one. These mistakes often result in disease. In some cases, a single misplaced base is sufficient to cause a disease, such as sickle cell anemia.

Errors in our genes, our genetic material, are responsible for an estimated 3,000-4,000 hereditary diseases, including Huntington disease, cystic fibrosis, and Duchenne muscular dystrophy. What’s more, altered genes are now known to play a part in cancer, heart disease, diabetes and many other common diseases. Genetic flaws increase a person’s risk of developing these more common and complex disorders. The diseases themselves stem from interactions of such genetic predispositions and environmental factors, including diet and lifestyle. Some experts estimate that half of all people will develop a disease that has a genetic component.

Understanding the genetic code did not directly lead researchers to disease genes. Their ability to decipher the genetic messages encapsulated in DNA was stymied by the overwhelming number of such messages carried in the DNA of each cell. A human cell (except sex cells—sperm and egg cells—and some blood cells that have no nuclei) contains about 6 feet of DNA molecules tightly coiled and packed into 46 chromosomes—rod-like structures in the cell nucleus that are formed from DNA covered with proteins. This DNA is made up of 3 billion base pairs. If printed out, those base pairs would fill more than 1,000 Manhattan telephone directories. When researchers tried to break up DNA molecules into more manageable pieces, however, they ended up with a chaos of random fragments whose order in the original DNA was lost.
The Cutting Edge

In the late 1960s, a useful molecular tool came to the rescue of these frustrated researchers, thanks to a series of studies by Werner Arber, in Switzerland, and Hamilton Smith, at Johns Hopkins University. These investigators were studying what at first seemed to be an unrelated problem. They were interested in understanding how some bacteria resist invasion by viruses. When viral DNA enters these bacteria, it is cut into small pieces and inactivated by enzymes called endonucleases. Smith showed that one of these enzymes cut the DNA at a specific short DNA sequence. Smith’s colleague Daniel Nathans recognized that this provided a means of cutting a large DNA molecule into well-defined smaller fragments, and he used the method to generate the first physical map of a chromosome, that of the small monkey virus SV40. The map allowed Nathans to determine the arrangement of the individual genes within the DNA that forms the viral chromosome. With clairvoyance, Nathans speculated that larger chromosomes might be studied similarly. This heralded the mapping of chromosomes, an activity that forms the basis for the assignment of a disease gene to a specific region on a particular human chromosome.

The DNA cutting enzyme that Smith isolated was the first of over 1,000 “restriction enzymes” that have been discovered in just a few decades. Restriction enzymes not only allow chromosome mapping, they also enable researchers to generate large amounts of any specific DNA sequence of interest. These enzymes usually do not cut straight across the two strands of DNA, but cut in a staggered fashion. Consequently, their cuts create short, single-stranded tails on the ends of each fragment, called sticky ends. The sticky ends can be joined to other DNA strands with the aid of another type of enzyme, called ligase. By 1973, researchers were using restriction enzymes to cut specific DNA sequences of interest and join them to the DNA of bacteria. The bacteria then generated copies of the selected DNA with their own DNA each time they divided. Because a single bacterium grows rapidly, producing more than 1 billion copies of itself in 15 hours, large quantities of a specific DNA sequence can be produced in this manner—called cloning. This DNA can either be used for further study or to make DNA probes (see below).

Sifting Out Telltale Genetic Sequences

The overwhelming number of genes in a human cell presented a major hurdle for researchers who wished to detect in a person’s DNA a specific disease-fostering gene. For example, researchers wanted to detect the gene for the type of hemoglobin—the oxygen-carrying blood protein—that is absent in a severe form of anemia, known as Cooley’s anemia. This disease can cause such debilitating illness or even death at an early age that scientists wanted to devise a way to detect it before birth.

Researchers had used the genetic code to determine the DNA sequence that codes for the series of amino acids that make up part of the hemoglobin protein. But they had no means of sifting the telltale DNA out of the DNA of the 100,000 other genes in

With the aid of new techniques, geneticists determined that this baby had inherited only one of his parents’ genes for thalassemia—an inherited disease characterized by mild to severe anemia—and thus will be unaffected by this blood disorder. (Division of Research Services, NIH)

The 46 chromosomes in the nucleus of a human cell pack about 6 feet of DNA strands, the genetic “blueprint” for each individual—in this case, a male as noted by the XY designation. This DNA is made up of 3 billion base pairs, enough information to fill over 1000 Manhattan telephone directories. (Copyright 1996 by Custom Medical Stock Photo)
a human cell, so that they could tell whether the gene was normal or not.

A way around the problem was discovered in 1975 when a Scottish scientist, Edward Southern, developed a powerful method to pinpoint a specific genetic sequence. Restriction enzymes were used to cut DNA into fragments, which were then separated by size by being sifted through a porous jelly-like substance through which an electric current is passed. The smaller fragments move faster through the gel than the larger ones, so that the DNA fragments from different genes end up at different positions. Then the separated fragments are blotted out of the gel onto a sheet of paper without changing their relative positions, and the paper copy of the gel is bathed in a solution containing radioactive DNA molecules, called probes. These probes are cloned DNA fragments with sequences that match a DNA sequence in the gene of interest (for example, the hemoglobin gene). Matching means that an A on one strand of the probe is matched by a T on a strand from the gene, a G is matched by a C, and so on for long distances along the two DNA molecules. This allows the DNA probe and the DNA fragment that is stuck to the paper to pair, making that region of the paper radioactive.

Because the probes are radioactive where they bind to the paper, they give off signals that can be made visible on x-ray film. This simple method, named Southern blotting after its creator, allowed researchers to detect a single DNA fragment from the hemoglobin gene among more than 100,000 other fragments in the same gel. Shortly after Southern blotting was developed, researchers used it to develop a prenatal test for Cooley’s anemia and other rare conditions for which telltale DNA sequences were known.

**Honing the Search for Disease Genes**

A serendipitous observation in 1978 by researchers at the University of California, San Francisco, made Southern blotting useful in the detection of several more common disorders. Yuet Wai Kan and Andrée-Marie Dozy were studying patients with sickle-cell anemia, a hereditary disease in which a single change in DNA gives rise to a defective form of hemoglobin that fosters painful and sometimes fatal blood clots. The researchers noticed, after they used a restriction enzyme to cut the DNA of patients with sickle-cell anemia, that most of the patients had a DNA fragment containing the beta-hemoglobin gene that was 13,000 base pairs long. People without sickle-cell anemia often lacked this particular DNA fragment after their DNA was cut by the same enzyme. Because the fragment produced was different in size from that normally seen, it was called a restriction-fragment-length polymorphism (RFLP).

RFLPs have been found in association with many common genetic disorders, including Huntington disease and some kinds of cancers. The RFLPs allow scientists to use Southern blotting to detect a disease or disease susceptibility in a person without knowing the precise DNA sequence of the gene that fosters it.

---

**Advances in Genetics Research That Led to Gene Testing**

This timeline shows the chain of research events that led to the gene testing that predicts the likelihood of developing various diseases. It is rich in examples of the contributions of basic research to unexpected outcomes of immense societal benefit.

- **1960-1966**
  Marshall Nirenberg, Har Gobind Khorana, and their colleagues decipher the genetic code that all living cells use to translate the series of bases in their DNA into instructions for the production of proteins.

- **1960-1965**
  Gregor Mendel conducts his pea-breeding experiments that show how physical traits, such as height and color, are passed from one generation to the next through genes.

- **1953**
  Using an x-ray pattern of DNA generated by Rosalind Franklin and Maurice Wilkins, James Watson and Francis Crick publish their double-helix model of DNA. This model accurately predicts that the order of four repeating molecules, known as bases, on the DNA strand spells out the genetic endowment of every living organism.

- **1944**
  Oswald Avery shows that the injection of DNA into bacteria causes genetic changes in them.

- **1970**
  Hamilton Smith serendipitously discovers the first restriction enzyme that cuts DNA at specific sites. Daniel Nathans uses such restriction enzymes to generate the first physical map of a chromosome.

This article was published in 1996 and has not been updated or revised.
RFLPs are used in a second way. An RFLP that is tightly linked to a specific disease (that is, where people who have the disease almost always have the specific RFLP) lies near the sequence of DNA that houses the disease-fostering gene. Therefore, by finding the cutting site on the DNA that created the RFLP, the disease gene itself can eventually be isolated.

The search for RFLPs linked to specific disorders can sometimes be narrowed to a specific region of a chromosome with the aid of chromosomal staining. Developed in the 1970s, chromosomal stains reveal a pattern of light and dark bands that reflects regional variations in the amounts of A and T bases versus G and C bases. Under a light microscope, differences in size and banding pattern distinguish the 23 chromosome pairs from each other and reveal major chromosomal abnormalities, including missing, added, or misplaced pieces of chromosomes. The eye cancer retinoblastoma, for example, often is associated with a missing band on chromosome 13; this finding led researchers to look for retinoblastoma-associated RFLPs within that region of the chromosome.

Researchers can also track a specific RFLP to a section of a particular chromosome by tagging it with an observable label (one that is fluorescent or radioactive). The location of the labeled RFLP can be detected after it binds to its complementary sequence of bases in an intact chromosome. This technique is known as in situ hybridization.
Spelling Out Disease Genes

To improve the precision of gene testing, scientists needed to spell out the actual DNA sequence of disease-fostering genes (the order of the A, C, G, and Ts). This endeavor was greatly aided by the development of a method for rapidly sequencing DNA by Harvard University researchers in 1977. One of the researchers, Walter Gilbert, had been trying to understand how particular genes in bacteria were turned off (prevented from generating the proteins for which they code) and he saw that he would not make much headway unless he could determine the sequence of particular segments of the bacterial DNA. He then worked with his colleague Allan Maxam to concoct a novel method that combines chemicals that cut DNA only at specific bases with radioactive labeling and Southern blotting to determine quickly the precise sequence of long DNA segments. A different but equally successful DNA sequencing method was developed at about the same time by Fred Sanger in Cambridge, England.

In the early 1980s, the Maxam-Gilbert and Sanger methods for DNA sequencing were improved and automated to speed the process. The pinpointing of disease-fostering genes could now proceed relatively quickly. Using a technique called positional cloning, researchers first zero in on the chromosome likely to house a disease gene by using chromosomal staining, in situ hybridization, or other techniques. Once the chromosomal home of the disease gene has been identified, they look for RFLPs or other genetic markers in that location that are tightly linked to the disease. They then determine the sequence of the DNA bases in the region of the telltale markers. They know that their search is over if they pinpoint a DNA sequence that is found only in people with the disease in question. Positional cloning has thus far been used to find over 50 disease genes, including the gene for cystic fibrosis, Duchenne muscular dystrophy, some types of Alzheimer’s disease, and early-onset breast cancer.

Revolutionary Copying Technique Developed

Genetic tests for the disorders caused by disease genes could not make it from the laboratory bench to the doctor's office until researchers developed an easy and inexpensive way to copy specific DNA molecules. That way, the small amount of DNA extracted from a person’s blood or tissue sample could be multiplied into the large quantities needed for DNA sequencing.

Once again, scientists doing basic research overcame a stumbling block. In a small California biotechnology company, the Cetus Corporation, a young scientist, Kary Mullis, was employed to generate new ideas instead of doing bench experiments. He recognized that rather than relying on bacteria to duplicate selected DNA in a cloning process, he could use just the enzymes—called DNA polymerases—that bacteria themselves use to copy DNA. He developed a method, called the polymerase chain reaction and abbreviated as PCR, that allows the enzymes to be used to amplify any specific DNA sequence in a test tube.

There was only one catch—the method worked only on single-stranded DNA and the heating that is needed to unzip the two strands of DNA kills the polymerases. Fortunately, researchers several years earlier had isolated bacteria that had the amazing ability to thrive at temperatures near that of boiling water in
hot springs. Scientists discovered that the polymerase isolated from these bacteria could survive the high temperatures needed for PCR. By the late 1980s, the PCR technique had spawned a number of practical developments, of which gene testing is only one.

**Tracking a Colon-Cancer Gene**

By 1990, researchers had the techniques that they needed to search for the gene that causes an important form of colon cancer known as hereditary nonpolyposis colorectal cancer (HNPCC). People who inherit the HNPCC gene have an 80% or greater chance of developing colon cancer and other cancers, usually at an early age. Women with the gene also face a markedly increased risk of uterine and ovarian cancer.

The search for the HNPCC gene had actually begun more than 30 years earlier in the laboratories of several scientists who were working on bacteria and yeast to learn what happens to the occasional genetic errors that occur during sexual reproduction or when DNA is damaged by particular chemicals. These errors stem from DNA that has incorrectly paired bases (an A paired with a G, for example, instead of with a T). Studies showed that bacterial and yeast cells had a way of snipping out the mismatched bases in a process called mismatch repair. Through genetic analysis in both bacteria and yeast, several genes critical to the mismatch pathway were identified and their protein products were characterized.

Paul Modrich had devoted much of his academic career to working out the details of this repair mechanism in bacteria. By 1992, Richard Kolodner and his colleagues at the Dana Farber Cancer Institute in Boston had isolated a gene called Mut S homolog 2 (MSH2), which was needed for mismatch repair in yeast. Kolodner speculated that people probably also had a similar gene that governed mismatch repair. Because the process is vital to the functioning of cells, Kolodner assumed that errors in a human version of MSH2 and other mismatch-repair genes would cause some human diseases. With that in mind, he and his collaborators used PCR to detect the human MSH2 gene at the end of 1993.

Meanwhile, Bert Vogelstein at Johns Hopkins University and his colleagues in Finland were studying the families of people afflicted with HNPCC. They had used positional cloning to pinpoint the gene for the condition and published its sequence two weeks after Kolodner and his collaborators had published the sequence of their human MSH2 gene. The two sequences were identical: Vogelstein’s disease gene for HNPCC was the same as Kolodner’s MSH2 gene. Shortly after, a second mismatch-repair gene was found also to cause HNPCC. Vogelstein, Kolodner, and other researchers have since developed genetic tests for these two genes. The tests can tell people, in families prone to HNPCC, whether they have one of the genes that foster it. As many as 1 in 200, or 1.25 million Americans, may carry one or the other of these altered genes. People found to carry an altered gene can be counseled to adopt a high-fiber, low-fat diet in the hope of preventing cancer. They can also be advised to start yearly colon examinations at about age 30. Such examinations should help physicians to detect any precancerous growths on the colon so that they can remove them before the growths turn malignant. For those people, like Beth, who turn out not to carry the altered genes, the diagnostic test can provide a huge relief, removing the fear they have lived under as well as the need for frequent colon examinations.
Gene Testing Poses Social Dilemmas

Unfortunately, people with many diseases, such as Huntington disease, will initially find that their disease is predictable through gene testing, but not yet preventable or curable. Gene testing in those cases poses a number of ethical, psychological, and legal dilemmas. Recent federal legislation reduces the concern that people might be denied insurance or employment solely because they test positive for some condition. Several states have passed laws that prohibit insurance discrimination based on the results of gene tests. But the emotional distress that testing positive can promote is also an important concern.

Law-makers, policy-makers, and scientists are exploring ways to address some of the societal implications of genetic testing. A 1994 Institute of Medicine report suggested a number of guidelines for genetic screening. These include extensive education and counseling of people receiving gene tests as well as the reservation of widespread genetic testing for treatable or preventable conditions of relatively high frequency. The Human Genome Project, whose goal is to sequence all the genes in human DNA, supports research aimed at developing policies or programs that maximize the benefits of genetic research while minimizing the potential for social, economic, or psychological harm. As genetic forecasting takes medicine to new heights, it forces the legal, ethical, and social policies that guide its usefulness to advance to new heights of sophistication as well.

Medicine Transformed

The HNPCC gene-test story is just one of many that illustrate how gene testing is dramatically transforming the practice of medicine. This new ability to probe genes, however, can be a double-edged sword. For some diseases, such as Alzheimer’s disease, the ability to detect a faulty gene has outpaced medical science’s ability to do anything about the disease that it causes. The possibility of genetic testing for these diseases raises a number of controversial social issues that scientists alone cannot resolve. Policy-makers, lawyers, and scientists are working hard to come up with answers to the many novel social questions that gene testing poses. (See box above.)

But for many disorders, by genetically forecasting a person’s likelihood of developing a disease long before symptoms appear, doctors have a much better chance to prevent the disorder or to detect it at its earliest stages when treatment is more likely to be successful. A newborn who tested positive in a genetic test for retinoblastoma, for example, was screened for eye tumors just a few weeks after he was born. His doctor discovered that he had two tumors in each eye, and was able to save the infant’s sight by treating him with radiation.

In addition, prenatal genetic testing for fatal or extremely debilitating conditions is reducing their incidence in the general population. A prenatal screening program for the severe form of anemia known as beta-thalassemia, for example, led to the near eradication of this deadly disease on the Italian island of Sardinia. A twenty-fold reduction in the incidence of the fatal neurologic disorder Tay-Sachs disease occurred in the Baltimore area following the development and implementation of a prenatal screening program for the Tay-Sachs gene. For parents who choose to carry to full-term a fetus with a debilitating genetic disorder, prenatal testing allows preparation for the medical care that their children will need.

In the long run, the deciphering of the genetics of human disorders is likely to lead to better treatments. Shortly after the genes for cystic fibrosis and Duchenne muscular dystrophy were discovered, for example, researchers developed experimental treatments for these disorders that aim to correct or replace the faulty genes responsible for them. Drugs that can alter or replace the protein gene products that are missing or defective in some cancers and other disorders also are being pursued.

None of those medical feats would have been possible if researchers had not been able to pursue their curiosity about such basic riddles as what governs a pea plant’s height, or how genes are passed from one bacterium to another. Many of the scientific breakthroughs that led to gene testing stemmed from serendipitous findings in totally unrelated studies and from basic research, much of it publicly funded, whose far-flung practical implications were completely unexpected by their discoverers—curious people who just wanted to understand how nature works.

This article, which was published in 1996 and has not been updated or revised, was adapted by science writer Margie Patlak from an article written by Howard Hughes Medical Investigator at Harvard University Stuart Orkin and National Institutes of Health scientist Gary Felsenfeld for Beyond Discovery™: The Path from Research to Human Benefit, a project of the National Academy of Sciences. The Academy, located in Washington, D.C., is a society of distinguished scholars engaged in scientific and engineering research and dedicated to the use of science and technology for the public welfare. For more than a century, it has provided independent, objective scientific advice to the nation.

© 1996 by the National Academy of Sciences

October 1996