DOCUMENT RESUME

ED 291 550   SE 048 875

AUTHOR   Pines, Maya
TITLE   The New Human Genetics. How Gene Splicing Helps
                Researchers Fight Inherited Disease.
INSTITUTION   National Inst. of General Medical Sciences (NIH),
                Bethesda, Md.
REPORT NO   NIH-84-662
PUB DATE   Sep 84
CONTRACT   263-MD-332131-2
NOTE   68p.
PUB TYPE   Reports - Descriptive (141)

EDRS PRICE   MF01/PC03 Plus Postage.
DESCRIPTORS   Cancer; *Disease Control; Diseases; *DNA;
                *Embryology; *Genetic Engineering; *Genetics;
                Medicine; Reproduction (Biology); Science and
                Society; Scientific and Technical Information; Sickle
                Cell Anemia; *Technological Advancement

ABSTRACT
                The science of genetics is perceived to offer hope
that a large number of the 3,000 inherited diseases which afflict
human beings may be prevented or controlled. This document addresses
some of the advances that have been made in this field. It includes
an introduction and sections on: "The Beginning of Human Genetics";
"Unlocking the Secrets of DNA"; "The Development of Recombinant DNA
Techniques"; "A New Understanding of Sickle-Cell Disease and Other
Blood Disorders"; "How to Use DNA Fragments to Detect a Disease";
"How DNA Probes Find Their Match"; "Stalking the Most Elusive Genes";
"The Rapid Growth of Gene Mapping"; "Zeroing In on Cancer Genes"; "A
Bank of Living Human Cells"; "GenBank: A National Database of Nucleic
Acid Sequences"; and "The Promise of Genetic Therapy." A glossary of
terms is also included. (TW)

National
Institute
General
Medical
Sciences

ED291550

THE NEW HUMAN GENETICS
HOW GENE SPLICING HELPS RESEARCHERS FIGHT INHERITED DISEASE

2

ERIC
Full Text Provided by ERIC

What Is NIGMS?

The National Institute of General Medical Sciences (NIGMS) is unique among the components of the National Institutes of Health (NIH) in that its main mission is the advancement of the basic medical sciences. It supports selected research and research training programs in areas that underlie all medical investigation, such as cell and molecular biology and genetics. Ten of the scientists whom it funded have won Nobel Prizes since 1970 for their pioneering work on the structure and function of genetic material, the action of hormones, and the structure of the cell. Knowledge resulting from this work contributes directly to the progress of research on specific diseases in the other Institutes of NIH. NIGMS also develops and supports interdisciplinary studies in biophysics, pharmacology, biorelated chemistry, physiology, and trauma and burn research. Many of the researchers mentioned in this brochure worked with NIGMS support.

THE NEW HUMAN GENETICS

HOW GENE SPLICING HELPS RESEARCHERS FIGHT INHERITED DISEASE

ERIC
Full Text Provided by ERIC

CONTENTS

6

With the aid of new techniques, geneticists determined that this baby had inherited only one of his parents' genes for thalassemia and will be unaffected by this blood disorder.

7

INTRODUCTION

At first the young couple from Bethesda, Maryland, thought they would never dare to have children of their own. Though they were perfectly healthy, they knew that they both carried a gene for thalassemia, a life-threatening hereditary blood disease. Any child of theirs would have a 25 percent chance of inheriting a double dose of this gene and acquiring the disease—a chance they were not willing to take.

They changed their minds in the late 1970's, when more sensitive tests of red blood cells allowed scientists to detect this genetic disorder even in fetal blood. At about the same time, it became possible to withdraw blood samples from the fetus through a technique called fetoscopy.

The young woman became pregnant in 1979 and went to be tested at the Yale University Medical School's Genetics Center in New Haven, Connecticut, much of whose basic research is supported by the National Institute of General Medical Sciences (NIGMS). To her intense relief, she learned that the fetus was unaffected by the disease. The baby is now a healthy and active preschooler. However, the procedure is not without hazard and his parents were not too eager to risk it again.

By 1982, the science of human genetics had advanced so much that it was no longer necessary to take blood from the fetus for prenatal detection of thalassemia. Geneticists no longer needed to look for the abnormal product of an abnormal gene. Instead, they could look for the gene itself by examining the "master molecule of life"—the genetic material called deoxyribonucleic acid, or DNA—in the nuclei of cells. DNA could be extracted from any type of fetal cell, and since fetuses normally shed some of their cells into the amniotic fluid that surrounds them, these cells could be withdrawn with a hypodermic needle through the much safer technique of amniocentesis.

Early in her next pregnancy, the young woman went to The Johns Hopkins University School of Medicine in Baltimore, Maryland, to see Dr. Haig H. Kazazian, an expert in the detection of thalassemia by means of recently developed recombinant DNA (gene-splicing) techniques. The DNA tests showed that the fetus had inherited only her gene for thalassemia—not her husband's—and would escape the disease. And now the couple has another healthy baby boy.

In the future, fetal cells for DNA tests may be obtained even earlier in pregnancy through a new, still-experimental method called chorionic villus sampling, which provides results at 9 to 11 weeks of pregnancy (compared to 18 to 22 weeks for amniocentesis).

The science of genetics offers the first real hope that a large number of the 3,000 inherited diseases which afflict human beings may be
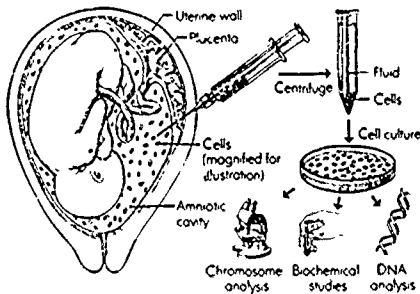
Amniocentesis—the most wide
spread technique of prenatal di
agnosis. Cells shed by the de
veloping fetus are extracted
from a sample of amniotic fluid
withdrawn from the expectant
mother's uterus by means of a
hypodermic needle. The cells
are cultured and then tested for
chromosomal defects, such as
Dc :n syndrome, and for certain
biochemical defects In addition,
scientists can now analyze the
DNA of these cells directly,
identifying specific genetic
errors

prevented or controlled. At present
this science is mainly diagnostic, al-
lowing some people to make in-
formed decisions about whether to
have children or not, and warning
others about their risks of develop-
ing certain illnesses. But it holds the
promise of many kinds of treatment.

Some genetic diseases already
can be treated effectively by chemi-
cal means. For example, in 1968
geneticists saved the life of an 8-
month-old boy who was rushed to
the Yale Hospital in a coma and
near death. While doctors gave him
emergency treatment, laboratory
tests revealed that the child's urine
contained strikingly large quantities
of methylmalonic acid, a chemical
that is usually present to this extent
only in cases of vitamin B-12 defi-
ciency; yet his blood tests showed
plenty of vitamin B-12.

Geneticists were called in. They
took a sliver of skin from the boy's
forearm, minced it, and put it in a
culture medium where the cells
could grow and divide. Then they
studied how these cells responded
to various chemicals. This pointed to
an unusual defect: The boy's cells
lacked a certain enzyme which nor-
mally activates vitamin B-12 and al-
lows it to break down the methylma-
lonic acid into simpler chemicals.
Because of this genetic defect, his
body needed 1,000 times the nor-
mal amount of vitamin B-12; no
lesser amount would do.

The doctors then injected an enor-
mous dose of vitamin B-12 into the

little boy's veins. As if by miracle, he recovered. The quantity of methyl-molonic acid in his urine decreased dramaticolly. The bosic defect re-moined, but from then on, with continuous high doses of vitamin B-12, he developed nearly normolly.

This was the first documented example of a recessive genetic disorder (which appears only when genes for it ore inherited from both parents) involving vitomin metabolism. Since its discovery, researchers have identified some 25 other inherited disorders thot respond to high doses of vitamins. Many of these conditions affect the central nervous system and produce mentol retardation, seizures, goit abnormalities, or behavioral disorders resembling schizophrenia. Some of the disorders can now be diagnosed prenotally, if the parents ore aware of their risk. Treatment can begin even before birth, with high doses of vitomins to the expectant mother for use by the developing baby.

All too frequently, however, the primary biochemical defect in genetic diseoses remains unknown, or else scientists do not yet have a way to deliver the missing chemical to the appropriate cells. Solutions to both these problems await the further development of humon genetics and the application of recombinant DNA ond other technologies.
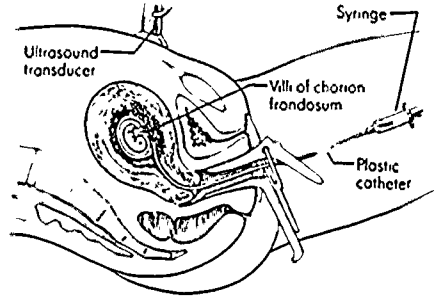
Chorionic villus sampling—a new and still-experimental method of prenotal diagnosis which provides results as early as the 9th week of pregnancy Fetal cells from the chorionic villi (protrusions of a membrane called the chorion which surrounds the fetus during its early development) are suctioned out through the uterine cervix and their DNA is analyzed. The preliminary results of this process can be obtained within a doy

THE BEGINNING OF HUMAN GENETICS

Genetic diseases used to be consid-ered quite rare. Today it is recog-nized that innumerable people suffer the consequences of disorders due wholly or in part to defective genes or chromosomes (the rod-shaped packages of genetic material inside the nucleus of a cell).

Genetics is now progressing so rapidly, on so many fronts, that it is revolutionizing medical research. It is producing a new understanding of how cancer develops, for instance. It is detecting the differences be-tween various forms of heart dis-ease. It is helping researchers de-sign more effective and less harmful drugs. It is providing precise infor-mation on who is most vulnerable to what kind of illness, and who should particularly avoid certain environ-mental agents. Perhaps most impor-tantly, it is bringing new insights into the function of regulatory genes which affect all human growth and development, from birth to death.

As recently as 1956, however, scientists were uncertain about the correct number of chromosomes in a human cell. Mammalian genetics still depended primarily on the slow-paced method of mating two ani-mals and studying their offspring. This approach worked quite well in mice. But since human beings have relatively few children, who take a long time to grow up and repro-duce, the study of human genetics was particularly difficult. It remained largely an observational science,
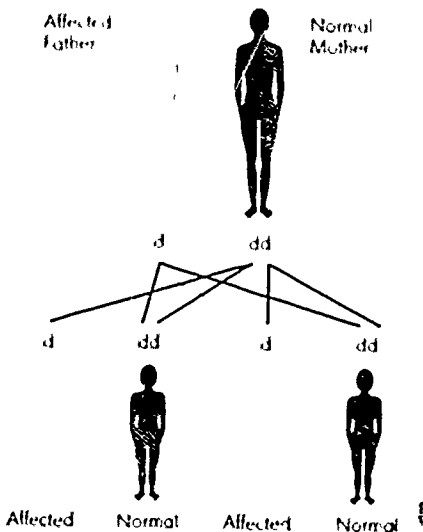

Nucleus
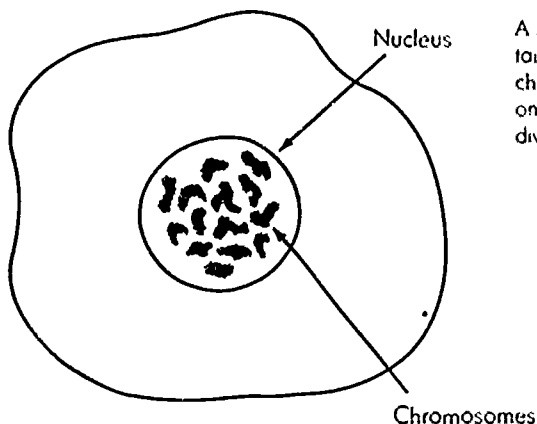
Chromosomes

A human cell The nucleus con-tains the genetic material, the chromosomes become visible only during certain stages of cell division

much as atomic physics was between the time of the Greeks and the 19th century. Early physicists had deduced the existence of atoms from the properties of matter, but they had no proof of it. Similarly, geneticists deduced the existence of genes from the properties of organisms and their progeny, but could neither analyze nor manipulate the particles about which they built elaborate theories. Their experiments dealt with the entire animal, rather than with chromosomes or genes.

The idea that human traits are under the control of distinct factors (later called genes), half coming from the father and half from the mother, goes back to the early 1860's and the experiments of the Austrian monk Gregor Mendel with different types of pea plants. He showed that in some cases an inherited trait will be expressed because of the presence of a single "dominant" factor, while in other cases two "recessive" factors are required for a trait to be expressed.

At about the same time, it became clear that, in animals, hereditary factors are transmitted through the egg and the sperm, and that each of these sex cells contains only half the normal number of chromosomes, which scientists had just learned to see under the microscope by means of special dyes ("chroma" comes from the Greek word for "color" and "some" from the Greek word for "body"). When a sperm

Affected Father   Normal Mother

d   dd

d   dd   d   dd

Affected   Normal   Affected   Normal

How Dominant Genetic Disorders are Inherited
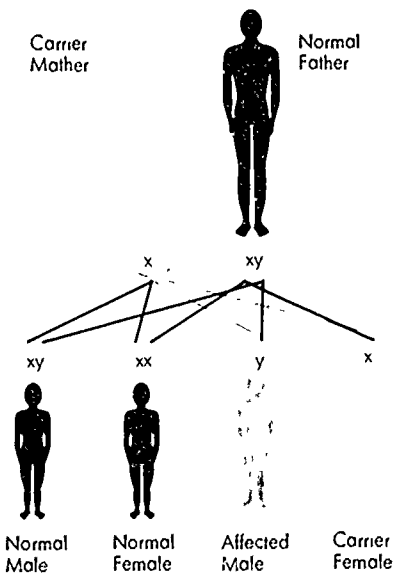One affected parent has a single faulty gene (D) which dominates its normal counterpart (d)

Each child has a 50 percent risk of inheriting the D—and the disorder—from the affected parent

12

9

enters an egg, these two haploid cells (cells with holf the usuol number of chromosomes) fuse to form a new cell with the full number of chromosomes.

One of the first things that researchers noticed about human chromosomes was a sex difference. In 1905, they discovered thot one particulor chromosome, which they colled "Y," could be found only in mole cells, together with on "X" chromosome. By contrast, female cells hove two copies of the X chromosome. Eventuolly it became cleor thot when sex cells are formed, each egg receives one X chromosome, while sperm corry either an X or o Y chromosome. Therefore the sex of offspring is determined by the sperm. If on egg is fertilized by a sperm that carries a Y chromosome, the fetus is mole; if the sperm carries on X chromosome, the fetus is female.

It took years of effort ond many technologicol odvonces before chromosomes could be seen cleorly enough to distinguish them from eoch other ond count them occurotely. Eoch species of plant or animol hos o chorocteristic number of chromosomes, these olso vary in size, length, ond other properties. In 1956, Drs. Joe-Hin Tjio, now ot the Notionol Institutes of Heolth (NIH),

Carrier Mother

Normal Father

x

xy

xy

xx

y

x

Normal Male

Normal Female

Affected Male

Carrier Female

How X-Linked Genetic Disorders are Inherited

The mother, who has o defective gene on one of her two sex chromosomes, is protected ogainst the defect because he; normal sex chromosome (x) compensotes for the defect on the other (X). The father has normal mole sex chromosomes (x and y)

Eoch mole child hos o 50 percent risk of inher;ing the foulty X and the disorder, on. 3 50 percent chance of irheriting the normal x chromosome

Eoch female child hos o 50 percent risk of inheriting the foulty X and becoming o carrier like her mother, and o 50 percent chance of inheriting two normal x chromosomes.

13

and Albert Levan of Sweden finally established the number of chromosomes in human cells. 46, or 23 pairs, with each pair containing one chromosome from the mother and one from the father. Twenty-two of these pairs are called autosomes, the two chromosomes of the 23rd pair are the sex chromosomes.

This led directly to cracking a major mystery. the cause of Down syndrome (mongolism). Geneticists soon found that this syndrome, which produces mental retardation and characteristic facial features in 1 out of 800 American children, is caused by a single extra chromosome number 21 in each cell. The

extra chromosome is believed to result most often from the incomplete separation of the chromosomes during the formation of the egg—an error which is more likely to occur with increasing maternal age. It can be detected before birth by examining the chromosomes in fetal cells collected from the amniotic fluid.

Several other forms of mental retardation and dozens of different disorders have now been traced to gross errors in the number or shapes of the chromosomes in each cell.

Going one step further and investigating the genes on the chromosomes proved far more difficult. It

Carrier
Father

Carrier
Mother

G   G

GG   G   G

Normal   Carrier   Carrier   Affected
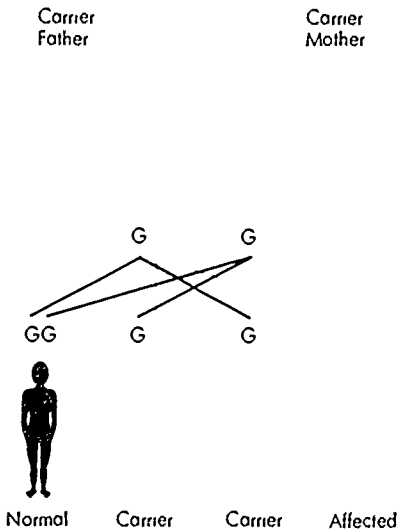
How Recessive Genetic Disorders are Inherited
Both parents, usually unaffected, carry a defective gene (g) but are protected by the presence of a normal gene (G) which is generally sufficient for normal function.

Each child has a 25 percent risk of inheriting a "double dose" of the g gene, which may cause a serious genetic defect; a 25 percent chance of inheriting two normal genes; and a 50 percent chance of being a carrier like both parents

With this historic photo, Drs. Joe-Hin Tjio and Albert Levan showed that there are 46 individual chromosomes in a normal human cell (magnification: ×2,300). Half of these chromosomes come from the mother, half from the father, and they are usually counted in pairs, except for the sex chromosomes (X and Y chromosomes).

1 2 3 4 5

6 7 8 9 10 11 12

13 14 15 16 17 18

19 20 21 22 X Y
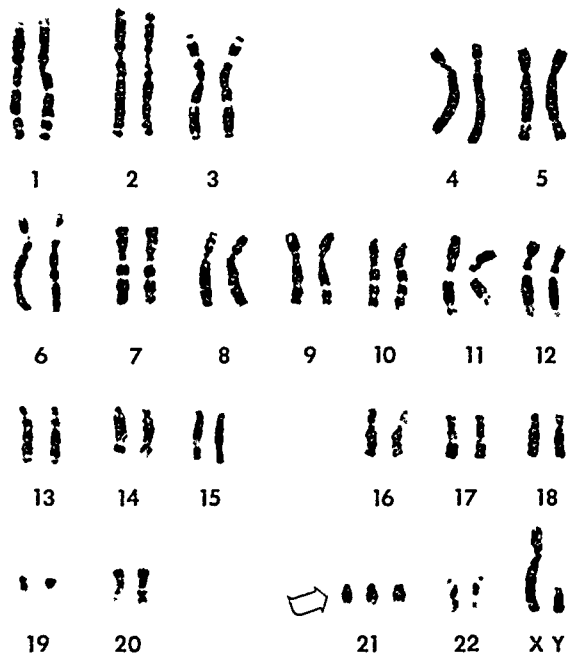
In Down syndrome (formerly called mongolism), there is an extra chromosome number 21 (three rather than the normal two). The extra chromosome most frequently results from incomplete separation of the chromosomes during the formation of an egg cell by the ovary. The condition thus represents a genetic accident rather than a strictly heritable disease. (The chromosomes in this photograph have been stained to reveal patterns of bands using more modern techniques than were available to Drs. Tjio and Levan in 1956.)

Common genetic disorders. Over 3,000 disorders due to a defect in a single gene or a chromosomal error have been identified. Some are evident at birth; others develop later.

Genetic Disease	Cause	Nature of Illness	Incidence	Inheritance
Down syndrome	autosomal chromosome abnormality	range of mental retardation	1 in 800	sporadic
Klinefelter's syndrome	sex chromosome abnormality	defect in sexual differentiation	1 in 2,000	sporadic
Cystic fibrosis	?	complications of excessively thick mucus secretion	1 in 2,000 Caucasians	autosomal recessive
Huntington's disease	?	progressive mental and neurolagical degeneration	1 in 2,500	autosomal dominant
Duchenne muscular dystrophy	?	muscular degeneration, weakness	1 in 7,000	X-linked
Sickle-cell disease	abnormal hemoglobin	impaired circulation, anemia, pain attacks	1 in 625 mostly black	autosomal recessive
Hemaphilia	defect in blood clotting factors	uncontrolled bleeding	1 in 10,000	X-linked
Phenylketonuria	enzyme deficiency	mental deficiency	1 in 12,000 mostly Caucasians and Orientals	autosomal recessive
Tay-Sachs disease	absence of an enzyme	buildup of fatty deposits in brain, leading to early death	1 in 3,000 Ashkenazic Jews	autosomal recessive
Lesch-Nyhan syndrame	enzyme deficiency	mental retardation, self-mutilation	1 in 100,000	X-linked

had been recognized since 1869 that a major constituent of cell nuclei was a substance which is now known as nucleic acid. Researchers later found that it really consists of two different substances, one containing the sugar ribose (this is ribo-nucleic acid, or RNA) and the other containing the sugar deoxyribose (this is DNA). In the 1920's it became clear that, unlike RNA, DNA was present only in the chromosomes. In 1944, Dr. Oswald Avery and his associates at the Rockefeller Institute in New York City discovered that DNA was directly involved in transferring hereditary characteristics from one strain of bacteria to another.

All living organisms, from bacteria to butterflies, from trees or fish to human beings, contain DNA that differs only in the number and arrangement of its components. But what is DNA?

UNLOCKING THE SECRETS OF DNA

DNA's three-dimensional structure was finally deciphered in 1953 by Drs. Francis H.C. Crick and James D. Watson, who were then working at the Medical Research Council laboratories in Cambridge, England. (Together with the English physicist Dr. Maurice Wilkins, they later won the Nobel Prize for this achievement.) Their now-famous double helix model of the DNA molecule explains both how DNA is built and how it replicates (makes more of itself).

According to this model, DNA is made up of two long and twisted strands. Each strand is composed of combinations of four smaller chemicals called nucleotides. Each nucleotide consists of one sugar mol-

ecule, one phosphate group, and one of four nitrogenous bases—adenine (A), thymine (T), guanine (G), or cytosine (C). These nucleotides line up next to each other like the two sides of a zipper, with the phosphate and sugar forming the outer ribbon, while the bases act like the interlocking teeth. Before cell division, new DNA must be synthesized and every gene must be replicated. To do this, the two strands separate and each one acts as a template for the formation of a mirror image, thereby producing two copies of the original DNA—two daughter DNA molecules whose sequences are identical to those of the parental double helix. When a cell divides, each daughter cell thus receives a
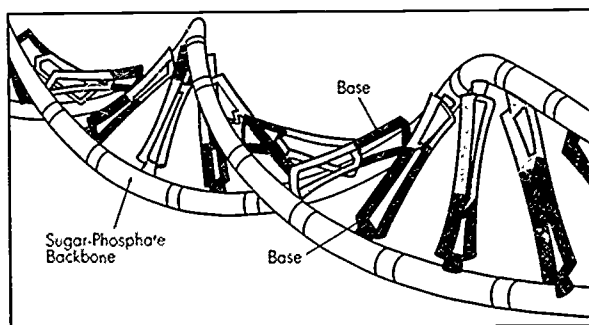
The double helix. model of the DNA molecule which embodies the code of heredity, as conceived by Drs. Francis H.C. Crick and James D. Watson in 1953 All living things reproduce themselves according to the genetic information in different sequences of DNA subunits (bases)

complete copy of the original cell's genetic information.

The two sides of the zipper can fit together only in one way. A pairs with T, and G with C. Because of this specific base pairing, if the sequence of one strand is known, that of its partner is also known. Each strand of the double helix thus specifies its complement, in the same sense that a photographic negative "complements" the positive image.

The four letters of this genetic alphabet can be used to write an infinite number of messages and instructions through which the genes direct the synthesis of thousands of enzymes and other proteins inside the cell. The language in which such instructions are written—the genetic code—was deciphered in the 1960's by Drs. Marshall W. Nirenberg of NIH, Severo Ochoa of New York University, and Har Gobind Khorana of the University of Wisconsin (work which also led to a Nobel Prize). The code consists of triplets of nucleotides (each triplet is called a codon) which are "read" in sequence along the DNA molecule.

Genes—which were mere abstractions for earlier scientists—can thus be studied and analyzed in terms of their constituent chemicals. Each gene is a series of codons which gives the instructions for building a specific protein. Each codon corresponds to one "word"—either one of the 20 amino acids which are the building blocks of

UUU	Phe	UCU	Ser	UAU	Tyr	UGU	Cys
UUC	Phe	UCC	Ser	UAC	Tyr	UGC	Cys
UUA	Leu	UCA	Ser	UAA	stop	UGA	stop
UUG	Leu	UCG	Ser	UAG	stop	UGG	Trp
CUU	Leu	CCU	Pro	CAU	His	CGU	Arg
CUC	Leu	CCC	Pro	CAC	His	CGC	Arg
CUA	Leu	CCA	Pro	CAA	Gln	CGA	Arg
CUG	Leu	CCG	Pro	CAG	Gln	CGG	Arg
AUU	Ile	ACU	Thr	AAU	Asn	AGU	Ser
AUC	Ile	ACC	Thr	AAC	Asn	AGC	Ser
AUA	Ile	ACA	Thr	AAA	Lys	AGA	Arg
AUG	Met (start)	ACG	Thr	AAG	Lys	AGG	Arg
GUU	Val	GCU	Ala	GAU	Asp	GGU	Gly
GUC	Val	GCC	Ala	GAC	Asp	GGC	Gly
GUA	Val	GCA	Ala	GAA	Glu	GGA	Gly
GUG	Val	GCG	Ala	GAG	Glu	GGG	Gly

The genetic code Each triplet of nucleotides codes for one amino acid, excepting three—the "stops"—which signify the end of a protein chain One amino acid, methionine, can also act as a signal to start protein production Each gene that codes for a protein is a series of triplets (codons) which gives the instructions for building that protein and thus influences a specific trait
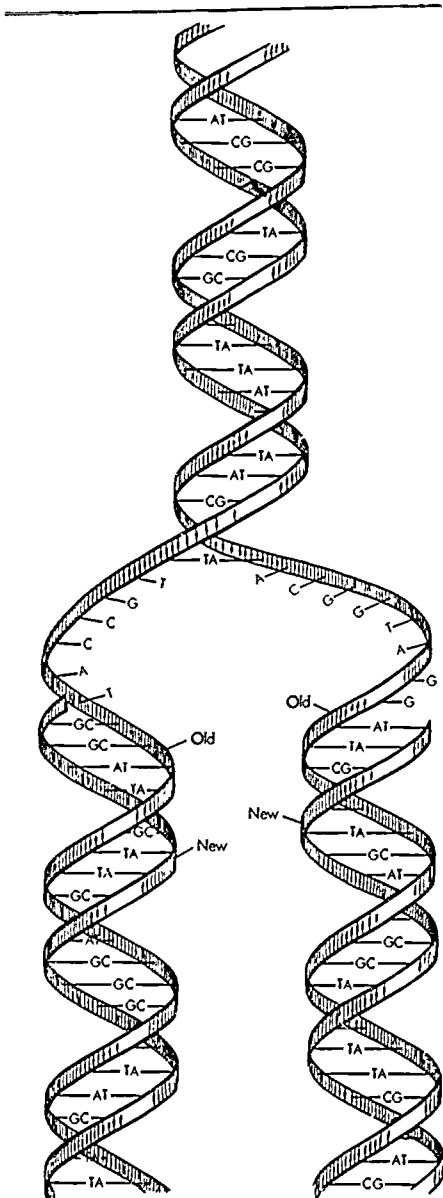
proteins, or a signal to start or stop constructing an amino acid chain. These instructions are transmitted not by the DNA itself, but by a copy, made of RNA, which acts as an intermediary. The original DNA remains safely in the nucleus, somewhat like the printing block in a printing press. The RNA copy is manufactured in the nucleus by transcribing just one strand of DNA, which codes the instructions.

In the late 1970's, scientists were surprised to find that in higher animals, including humans, the instructions are often interrupted by intervening sequences of DNA (introns) whose functions are not fully understood, but which clearly do not code for the proteins specified by the genes. The actual coding regions of genes are called exons. When information from the genes is transcribed into new strands of RNA, the introns are spliced out of the RNA. The resulting molecule, called messenger RNA, then moves out of the nucleus into the cytoplasm, where it can direct the production of protein.

DNA replication. To replicate before cell division, the DNA double helix separates and unwinds and each strand acts as a template for the formation of a mirror image according to the rules of base pairing: A with T, and G with C. This results in two daughter DNA molecules whose sequences are identical to those of the original DNA.

DNA transcription. The instructions contained in DNA are transmitted to other parts of the cell by an intermediary—a copy manufactured out of a different nucleic acid, RNA—while the original D^!^ remains safely in the nucleus. The RNA copy is made in the nucleus by transcribing (producing a new complement to) just one strand of DNA, which is sufficient to encode the instructions.

Complementary
DNA Strands

| | | | | | | | | | | | | | | | | | | |
T C C A A T G G C T T A T T T C C C T A

A G G T | | | | | | | G C T U A U U A G G G A T
A U G C G A A T A A | | | |
U C C

RNA Strand

Beginning of protein

This end translated first

Nearly completed protein

mRNA

Ribosome

RNA translation to protein. The single strand of messenger RNA (mRNA) carries the DNA's instructions to the ribosomes, tiny protein factories outside the nucleus. The ribosomes translate these instructions into growing chains of amino acids that become specific proteins.

THE DEVELOPMENT OF RECOMBINANT DNA TECHNIQUES

In order to zero in on individual genes, scientists needed a method that would consistently cut the long, ultra-thin strands of DNA at specific places. For nearly a decade they had no way to do this and progress in genetics research was slow. Considering that each human cell contains about 6 feet of DNA strands, coiled and packed into 46 tight bundles of chromosomes, and that this DNA is made up of 6 billion base pairs, the situation seemed almost hopeless. Even bacterial DNA was impossibly large, consisting of more than 4 million base pairs in the tiniest organisms, such as the intestinal bacterium Escherichia coli (abbreviated E. coli). Therefore, researchers focused on even smaller and simpler systems, such as viruses. When they tried to break up larger DNA molecules into more manageable pieces, they ended up with a chaos of random fragments whose order in the original DNA could not be established.

A happy finding in 1970 radically changed this picture. While studying Hemophilus influenzae, a bacterium which causes meningitis, a Johns Hopkins researcher, Dr. Hamilton Smith, noticed that it produced an enzyme which broke down the DNA of bacteriophages (viruses that attack bacteria). This self-defense enzyme did not harm the bacterium's own DNA, but it cut the DNA of a different bacterium, such as L. coli, into little pieces. Furthermore, all the cuts occurred at one specific base-pair sequence, which was later identified.

This was the first of a series of so-called "restriction enzymes," which have become the key to the manipulation and study of DNA. Soon afterward, Dr. Daniel Nathans, a col-

Nick
↓
-X-X-X-G A-A-T-T-C-X-X-X-
· · · ·   · · · ·
-X-X-X-C-T-T-A-A G-X-X-X-
↑
Nick
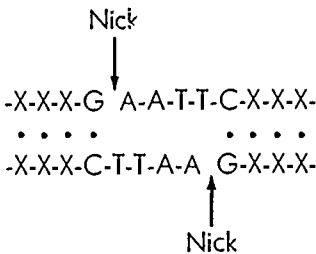
"Sticky ends" of DNA. The restriction enzyme EcoRI, produced by the E. coli bacterium, recognizes the DNA sequence GAATTC and cuts it between the G and A Since the two strands of the double helix are complementary, however, the restriction enzyme does not cut straight across both strands Instead, it produces a fragment with two exposed strips of single-stranded DNA at its ends. These exposed ends will "stick" ∩ other, similarly exposed single strands of DNA which have complementary sequences.

league of Smith's, coupled the ability of this and other restriction enzymes to cut DNA with a technique for separating the resulting fragments on the basis of size. This enabled him to create the first "restriction map" of a DNA molecule.

To date, over 200 restriction enzymes have been isolated from bacteria. Each one makes it possible to cut, or "digest," DNA at the point where particular sequences of nucleotides occur (a different sequence for each enzyme), producing fragments of different lengths. These enzymes do not usually cut straight across the two strands of DNA, however. The breaks often occur in a staggered fashion, creating short, single-stranded tails on the ends of each fragment when the fragments separate. These tails are called "sticky ends" 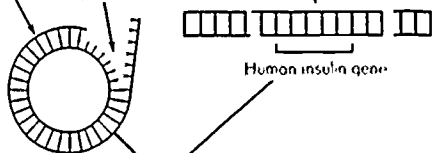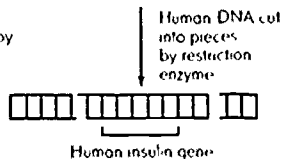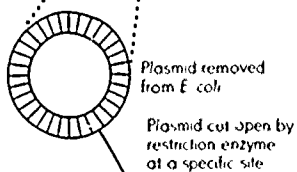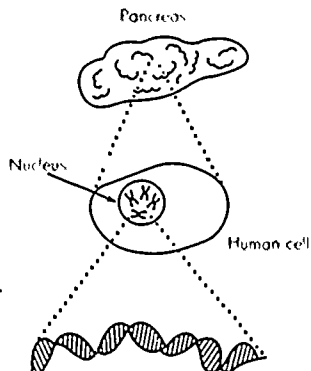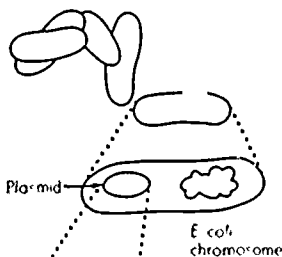because they tend to stick to complementary fragments by base pairing. Two fragments produced by the same enzyme can stick together in this way and then be permanently joined by adding another enzyme, DNA ligase. Scientists can thus splice together, or "recombine," various pieces of DNA.

The discovery of these recombinant DNA techniques produced extreme excitement among biologists. In 1973, Drs. Herbert Boyer and Stanley Cohen at the University of California, San Francisco, showed how these techniques might be used. They took advantage of the fact that inside many bacteria, in addition to a single large chromosome, there are tiny circular DNA molecules, called plasmids, each consisting of only a few thousand base pairs. The two researchers cut such E. coli plasmids with a restriction enzyme called EcoRI, for which each plasmid had only one recogni-

Bacterium   Circular Plasmids (several thousand base pairs each)   Main Circular Chromosome (4 million base pairs)
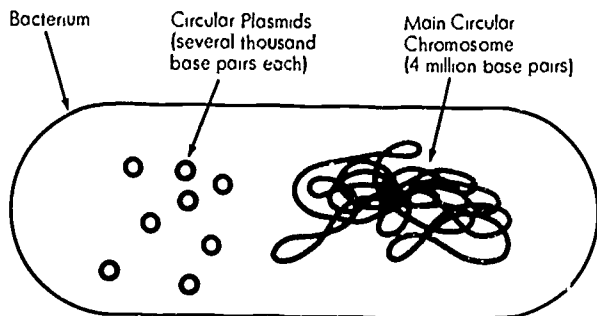
Plasmids. In addition to a main chromosome, bacteria (as well as yeast and some plants) contain plasmids, small circular molecules of DNA which replicate autonomously. Plasmids are easily moved from one cell to another or to the test tube. They can also be broken open with restriction enzymes. Scientists often insert foreign DNA into plasmids and then use them to transfer this DNA into other cells in recombinant DNA experiments.

E. coli bacteria, taken from human intestine

Pancreas

Plasmid

E. coli chromosome

Nucleus

Human cell

Plasmid removed from E. coli

Strand of DNA from human cell

Plasmid cut open by restriction enzyme at a specific site

Human DNA cut into pieces by restriction enzyme

Human insulin gene

Two pieces spliced together

Recombinant DNA (hybrid plasmid)

Human insulin gene

Hybrid plasmid inserted into E. coli cell
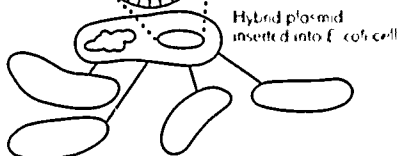
Bacteria with hybrid plasmid replicate, creating clone capable of producing insulin

Making a hybrid plasmid 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, break the plasmid open at a specific site by means of a restriction enzyme, and splice in insulin-making human DNA. The resulting hybrid plasmid can be inserted into another E. coli bacterium, where it replicates together with the bacterium, making it capable of producing large quantities of insulin.

23

tion site (the DNA sequence ot which o specific restriction enzyme mokes o cut); then they spliced in frogments of foreign DNA which hod olso been cut with EcoRI, ond seoled the plosmids bock together with ligose. The hybrid plosmids, eoch corrying one or more frog-ments of foreign DNA, were then tronsferred bock into E. coli bocterio, where they corried out the instruc-tions of the inserted DNA ond re-produced together with the bocte-rio's own genetic moteriol.

This opened up the possibility of isoloting ond studying olmost ony humon gene—o breothtoking vistu for geneticists. Humon DNA could be broken up with restriction en-zymes ond the frogments inserted ot rondom into bocteriol plosmids. The recombinont plosmids could then be introduced into o populotion of boc-terio. One or more of these bocterio would be likely to contoin o recom-binont plosmid corrying the specific humon gene thot wos wonted. The bocterio could then produce enor-mous quontities of this gene, since bocterio reproduce every 20 to 30 minutes. At thot rote, o single boc-terium could produce more thon o billion copies of itself in 15 hours.

The growth of gene-splicing tech-niques wos greotly helped olong by two reloted developments.

In 1975 ond 1976, research groups in Englond ond the U.S. devised new meons by which the nucleotide sequences of DNA seg-ments could be determined simply

ond ropidly. The opprooch token by the Americon scientists wos to use four different chemicols to breok o DNA strond into o series of frog-ments whose lengths depend on the distonce between the breok ond one rodiooctively lobeled end ot the strond. Becouse the chemicols cut ot specific nucleotides or poirs of nu-cleotides, reseorchers ore oble to identify the lost bose in eoch frog-ment. By studying the pottern of frogment lengths they con determine the order of the boses in ony DNA choin. The British method uchieves similor results by using enzymes ond modified nucleotides thot ore incor-porcted into the choin ond ollow identificotion of the terminol bose of eoch frogment. Both of these tech-niques hove been in extensive use since their development ond, os o result, the complete sequence of the DNA of severol smoller viruses is now known, ond the sequences of mony humon genes hove been determined.

Shortly ofter sequencing wos de-veloped, the previously formidoble tosk of putting together nucleotides to moke specific DNA sequences in the loborotory wos vostly simplified by new methods. Scientists con now produce sequences of 12 to 20 nu-cleotides in just o few doys.

By combining these techniques, reseorchers con now moss-produce o voriety of importont biologicol substonces, including insulin, humon growth hormone, ond the interferons.

A NEW UNDERSTANDING OF SICKLE-CELL DISEASE AND OTHER BLOOD DISORDERS

One of the victories of the science of human genetics has been its increasingly precise analysis of what goes wrong in various kinds of inherited blood disorders, particularly sickle-cell disease and the thalossemias (such as the illness for which the baby shown on page 4 was at risk).

Occasionally—perhaps because of an error when the cell replicates, or because of some outside influence such as a virus or radiation—the specific sequence of nucleotides in a DNA molecule is altered. Such changes are called mutations, and they can have either harmful, neutral, or, very rarely, beneficial effects. It is thought that every human being inherits about six or seven deleterious recessive mutations that, when transmitted to offspring who happen to receive the same deleterious gene from the other parent, can cause serious illness and even death.

Over 2 million Americans, mostly blacks, are carriers of the sickle-cell gene, for instance. This is an example of a mutation which generally does no harm to carriers, who also have one normal gene which compensates for the deleterious effects of the defective one; it even provides a selective advantage in certain parts of the world, since it offers some protection against malaria. However, when two carriers of this gene marry, each of their children has a 25 percent chance of inheriting a double dose of the sickle-cell gene and suffering from sickle-cell disease. This is a potentially lethal recessive disease which causes chronic anemia, jaundice, severe pain, and poor resistance to infection.

Carriers of the sickle-cell gene can be identified by a simple blood test. The disease itself was originally traced to a defect in hemoglobin, the oxygen-carrying protein of blood, back in 1949. With increasing precision, in 1956 it was traced to the substitution of just one amino acid (valine instead of glutamic acid) out of the nearly 150 amino acids that make up one subunit of the hemoglobin molecule. Recently it was tracked down to its fundamental source: the mutation of a single nucleotide in the DNA codon for glutamic acid (GAG), changing it to the codon for valine (GTG).

Because of this change in one nucleotide, the victim's hemoglobin is less soluble and under certain conditions it turns into a viscous, semi-

Normal Hemoglobin
- val - his - leu -thr - pro -   - glu -

Sickle-Cell Hemoglobin
- val - his - leu - thr - pro -   - glu -

Sequence of normal and sickle-cell hemoglobin differs by one amino acid

Normal red blood cells

In sickle-cell disease, many red blood cells are distorted from their normal round shape, often into the shape of a crescent or sickle. These distorted cells may obstruct the smaller blood vessels or be removed too rapidly by the spleen, causing anemia.

Sickle cells

solid gel. As a result, the normally round and very flexible red blood cells may become so inflexible that they can no longer squeeze through the small blood vessels. Some of these cells are also distorted into the shape of a crescent or sickle. When the abnormal cells obstruct the small blood vessels, oxygen can no longer reach the surrounding tissue, causing damage. Sometimes the abnormal cells are removed too rapidly by the spleen, causing anemia.

Several new methods of treatment are now being considered, including gene therapy (see page 53). But children with this disease often die young, after many medical crises.

"I am the mother of five children, three of whom have sickle-cell disease," says Ola Huntley, director of a sickle-cell self-help group in Los Angeles. "At the time we married, there were no tests for carriers. My second, third, and fourth children are presently suffering, the fourth was born before the third was diagnosed. It's rare that we can all be together at home, because they are constantly being hospitalized. For 25 years now, I've made those frequent trips to the hospital. There is only temporary relief. I see the pain in my children's eyes. I see them face major surgery. My daughter has experienced a stroke—a blood clot in the brain—and leg ulcers. My son has grand-mal seizures from blood clotting. In addition, they get ridicule from their peers. They have low self-esteem, and little motivation to live . . . ."

The disease is extremely variable, however, and some people with sickle-cell disease lead fairly normal lives. As yet there is no way to predict the severity of the disease with any reliability.

With the aid of recombinant DNA techniques, sickle-cell disease can now be detected prenatally. This kind of screening became possible when researchers found a way to tell the difference between normal and abnormal forms of the beta-globin gene, which codes for one of the two types of protein chains that make up hemoglobin.

At first they used an indirect approach. If they could not recognize the defective gene itself, why not look for a "marker"—any identifiable variation in DNA that was close enough to the beta-globin gene to be likely to be inherited together with it? That is what Drs. Yuet Wai Kan and Andrées M. Dozy of the University of California, San Francisco, did in 1978, and their method is worth describing in some detail.
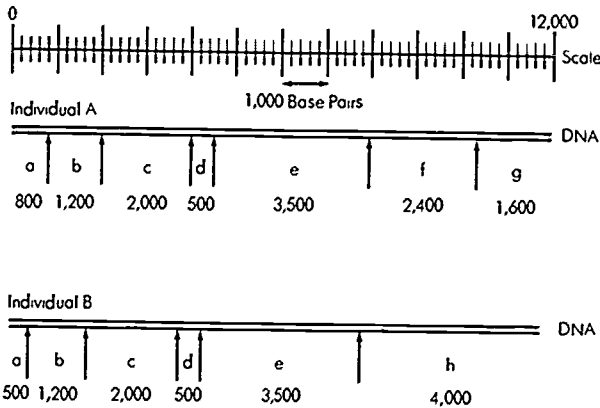
27

HOW TO USE DNA FRAGMENTS TO DETECT A DISEASE

As explained above, restriction enzymes cut DNA at specific nucleotide sequences which they recognize. The length of each resulting DNA fragment depends on the distance from one recognition site to the next.

At a point about 7,000 nucleotides away from the beta-globin gene on chromosome 11, Drs. Kan and Dozy noted a natural variation in DNA which appeared to have no effect on health. This was the first DNA polymorphism, or inherited variant, ever discovered. The researchers found .. ut the DNA of some people had a recognition site for the restriction enzyme Hpal at this point, while other DNA did not. If Hpal cut through the DNA at this point, one of the fragments it produced was a fairly short piece containing the beta-globin gene—a fragment only 7,600 base pairs in length. However, if the DNA lacked this sequence and Hpal could not make a cut at that point, the beta-globin-containing fragment was longer—13,000 base pairs in length.

The various fragments generated by such cuts can be separated by electrophoresis, a process in which an electric field makes the DNA fragments move through porous agarose gels. The smaller fragments move faster than the larger ones, so

0   12,000 Scale

1,000 Base Pairs

Individual A

DNA

| a | b | c | d | e | f | g |

800 1,200 2,000 500 3,500 2,400 1,600

Individual B

DNA

| a | b | c | d | e | h |

500 1,200 2,000 500 3,500 4,000

Restriction fragment length polymorphism (RFLP). When a specific restriction enzyme cuts DNA, it may produce fragments of different sizes in the DNA of different people. For example, EcoRI will cut DNA wherever the base sequence GAATTC occurs (represented here by arrows). But while individual A has six of these recognition sites, yielding fragments a through g, individual B has only five recognition sites. Thus the length of fragment h from individual B's DNA (4,000 base pairs) equals the sum of the lengths of fragments f and g produced by EcoRI in individual A's DNA. Such natural variations, or polymorphisms, in restriction fragment lengths are inherited.

28

their sizes can be determined by examining their positions in the gels.

The researchers then used this procedure to analyze DNA from the cells of many American black families and established that, in 70 percent of these families, the presence of longer DNA fragments after exposure to Hpal was associated with sickle-cell disease. This meant that in families where this link was shown, the occurrence of such fragments

after exposure to Hpal could be used as a predictor of disease.

Shortly afterward, several researchers went a step further and found direct evidence of the sickle-cell gene. It was a matter of trying until they found the most appropriate restriction enzyme—in this case, Mstll, which cuts DNA right in the middle of the beta-globin gene as well as in many other places. When the beta-globin gene is normal,

Larger

350
Base
Pairs

150
140

Direction
of movement
of fragments
through the gel

54

39
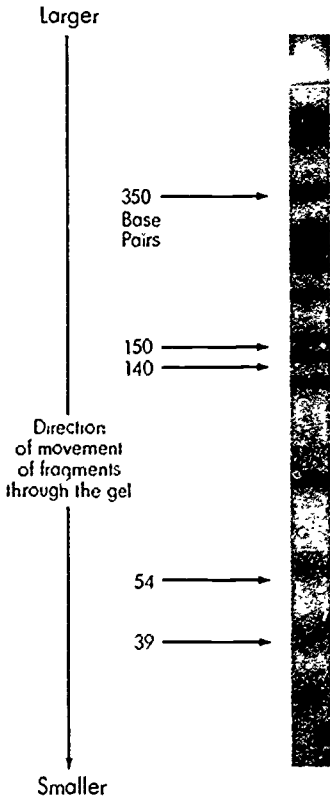
Smaller

Electrophoresis. To separate DNA fragments according to their size and detect RFLP's, scientists use electrophoresis, a process in which an electric field makes the DNA fragments migrate through a gel. Since the smaller fragments move through this gel faster than the larger ones, the length of any fragment can be determined by comparing its position in the gel to the position of DNA molecules of known sizes.

29

MstII cuts through it at the sequence CCTGAGG, producing two fragments of 1,150 and 200 base pairs in length. However, the sickle-cell mutation converts the DNA sequence at this point to CCTGTGG, thereby abolishing the MstII recognition site, this results in just one fragment of 1,350 base pairs after treatment with MstII.

To find out whether any given DNA contains the sickle-cell mutation, therefore, one need simply know what size fragments MstII produces at this point. But MstII and other restriction enzymes cut human DNA in many other places as well, and picking out these fragments from many others of similar sizes is somewhat like finding a needle in a haystack. Geneticists get around this problem by using specially prepared, radioactively labeled DNA probes.
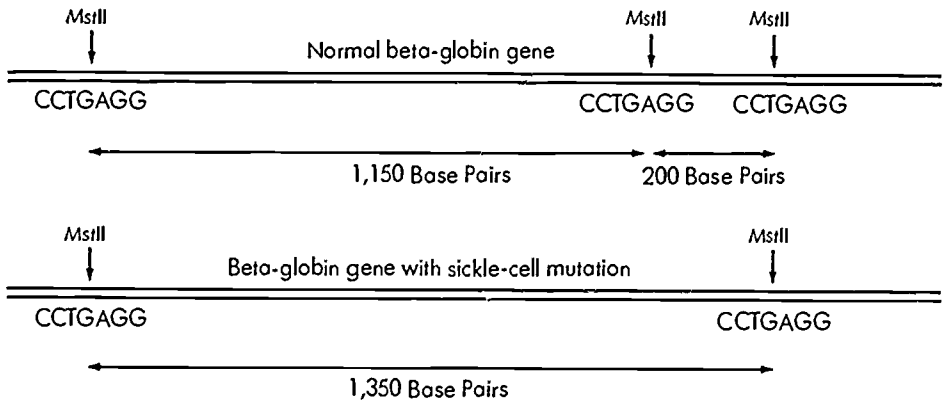
Direct evidence of the sickle-cell gene. It has become easy to detect the presence or absence of the sickle-cell gene through DNA analysis. If the restriction enzyme MstII cuts a person's DNA at the sequence CCTGAGG in the beta-globin gene, producing two fragments (1,150 and 200 base pairs in length), the gene is normal. But if MstII cannot make this cut because the recognition site has been abolished by a mutation that changed the sequence to CCTGTGG, there will be only one fragment, 1,350 base pairs in length—evidence that the person from whom this DNA was taken carries a gene for sickle-cell disease.

HOW DNA PROBES
FIND THEIR MATCH

A probe is a short sequence of single-stranded DNA that is complementary to the DNA sequence one seeks. Such probes take advantage of the fact that single strands of DNA automatically seek out complementary strands whose bases pair up with them, G with C and A with T. In order to be recognized by a probe, however, the DNA that is exposed to it must first be separated ("denatured") into single strands. The two strands of the DNA double helix are easily denatured by exposing them to near-boiling temperatures or to extremely alkaline conditions.

Probes are often made out of DNA segments that have been cloned, or individually reproduced, inside *E. coli* bacteria. For example, the probe used to detect the sickle-cell mutation is a fragment of a cloned beta-globin gene. Before being used, it is treated to make it radioactive so that it can be detected in the midst of large amounts of other DNA.
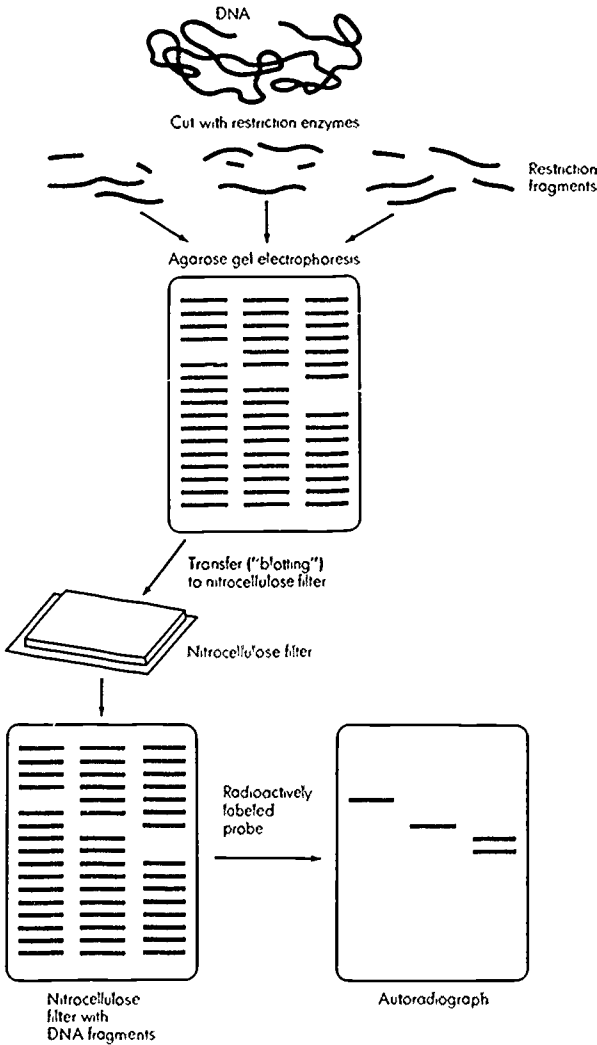
Any DNA that is suspected of having the sickle-cell mutation is first cut by MstII. The resulting fragments are then separated by agarose gel electrophoresis and treated chemically to denature them, producing single strands of DNA. Next these fragments are transferred ("blotted") onto a nitrocellulose filter to which they become bound, still in the same pattern as they were in the agarose gel. (This technique is called Southern blotting.) At this

point, the filter is exposed to the radioactively labeled probe, which will stick, or hybridize, only to the specific DNA sequences that are complementary to it. These hybridized sequences will then give off a radioactive signal that can be visualized by exposing the filter to X-ray film, a procedure known as autoradiography.

If a band corresponding to a DNA fragment of 1,350 base pairs shows up on the autoradiograph, it represents the sickle-cell gene. If two bands corresponding to the shorter fragments (1,150 and 200 base pairs) show up, this represents a normal gene. If both the longer band and the two shorter ones appear, the person has inherited one copy of the defective gene and one copy of the normal gene, and will be a symptomless carrier. The sickle-cell mutation can thus be diagnosed directly.

Not many genetic diseases have their DNA errors right in the middle of a restriction enzyme's recognition site, as is the case for the sickle-cell gene. Therefore, different strategies must be used to detect them.

One strategy is to manufacture probes out of laboratory chemicals by synthesizing strings of nucleotides called oligonucleotides. Such oligonucleotides may be engineered to match normal genes perfectly and will hybridize to them. If the match is not perfect, however—if a single nucleotide does not match—the hybrid will be unstable and fall apart under

DNA

Cut with restriction enzymes

Restriction fragments

Agarose gel electrophoresis

Transfer ("blotting") to nitrocellulose filter

Nitrocellulose filter

Nitrocellulose filter with DNA fragments

Radioactively labeled probe

Autoradiograph

Southern blotting This technique for transferring DNA fragments from gel electrophoresis (in which the fragments are separated according to their size) to a nitrocellulose filter (to which they become bound, still in the same pattern) makes it possible for scientists to pick out specific fragments of DNA. Once the fragments (represented here by lines, but more often in the form of a continuous smear of DNA) are on the nitrocellulose filter, they can be exposed to radioactive probes that will hybridize (stick) to any complementary sequence The hybridized fragments will then give off radioactive signals that can be made visible on X-ray film in a procedure known as autoradiography.

32

certain conditions. Mutant genes that are responsible for particular diseases and have a single "point mutation" (a substituted, missing, or extra nucleotide) can now be detected in this way, assuming one knows exactly where to look.

The most common form of thalassemia among people who come from Mediterranean areas, for instance, has been traced to the substitution of one nucleotide (A for G) in the beta-globin gene, which results in a reduced production of hemoglobin and severe anemia. However, this substitution does not alter any known restriction enzyme site. Recently researchers have prepared synthetic DNA probes that are replicas of the normal and mutant nucleotide sequences. These probes are beginning to be used in prenatal diagnosis of thalassemia among families of Greek or Italian ancestry, in whom the disease is most prevalent.

Similar methods have made it possible to detect alpha 1-antitrypsin deficiency, an enzyme deficiency which predisposes those who harbor it to severe and often fatal cirrhosis of the liver in infancy, and to obstructive emphysema in adulthood. Parents whose children have died of cirrhosis often seek prenatal diagnosis before having other children. But until recently such diagnosis required tests of fetal blood obtained through fetoscopy.

In 1983, researchers succeeded in preparing a 19-nucleotide probe which was able to detect the cause of this deficiency—the substitution of one nucleotide in the alpha 1-antitrypsin gene—in cells taken from amniotic fluid. As a result, parents can now know early in pregnancy whether a fetus will be affected.

Since the enzyme deficiency brings permanent damage to the lungs, children with this condition must stay away from industrial pollutants as well as cigarette smoke. "If they do smoke, they will have emphysema by their late 30's and die within a few years," says Dr. Savio L.C. Woo of the Baylor College of Medicine in Houston, Texas, one of the developers of this test.

In the future, children who are found to have this deficiency may also be aided by injections of the missing protein, much as diabetics are aided by insulin.

STALKING THE MOST ELUSIVE GENES

In hundreds of other genetic diseases, the biochemical fault is still unknown and there is as yet no clue to where the error may lie among the 6 billion base pairs in human DNA.

About 10 million Americans are carriers of cystic fibrosis, for instance—the most widespread recessively inherited disease among Caucasians in the U.S. If carriers marry, their offspring may inherit this life-threatening illness. Cystic fibrosis produces thick mucus in the pancreas and lungs which prevents normal digestion and breathing. It used to kill its victims before the age of 6, they starved despite an adequate diet or succumbed to severe infections. Today, enzyme therapy and antibiotics help many of these children survive to adulthood, but the majority still die young, and they all require constant care, including regular drainage of their lungs and other special measures to get rid of the mucus.

Although the damage caused by this disease is obvious, its biochemical basis is not. At present, carriers of the cystic fibrosis gene cannot be detected until they have produced a child with the disease. Nor can the disease be identified prenatally. However, a new approach is giving scientists a chance to locate some of the most elusive genes—and cystic fibrosis is a prime target for this kind of research.

Scientists who studied genetic diseases in the past tried to identify the defective gene product, such as a faulty protein, and then worked back to the molecular basis of the disorder, such as the replacement of one amino acid by another. Studies of the DNA error that produced the amino acid replacement came later, if at all. "But we now have the ironic situation of being able to jump right to the bottom line without reading the rest of the page, that is, [to identify the gene] without needing to identify the primary gene product or the basic biochemical mechanism of the disease," Dr. David E. Comings of the City of Hope National Medical Center wrote in an *American Journal of Human Genetics* editorial in 1980. He called this approach the "New Genetics."

The New Genetics makes it theoretically possible to diagnose any disorder which can be linked to a DNA polymorphism in a family. It really began in 1978, when Drs. Kan and Dozy detected sickle-cell disease by taking advantage of DNA variations that were inherited together with the presence or absence of the sickle-cell gene within families. Researchers are now discovering so many other restriction fragment length polymorphisms (RFLP's) that, in any family, a given gene defect is likely to be linked to one. Perhaps the most dramatic achievement in this area to date has been the discovery of a DNA polymorphism linked to the dreaded Huntington's disease.

Unlike other genetic disorders

which show up soon after birth, Huntington's disease may not appear until the victim is 35 or 40 years old. Then it produces progressive and irrevocable deterioration of the brain. Its first signs, such as personality changes or abnormal movements, are often misdiagnosed. (The folksinger Woody Guthrie, who died of the disease, was originally thought to be an alcoholic, then schizophrenic.) Eventually facial expressions become distorted, head, hands, and shoulders seem to writhe in constant motion, speech blurs. As the disease advances, the victim becomes totally disabled and irrational. It is a slow process, leading to death within 10 to 15 years. There is no effective treatment or cure.

Huntington's disease is caused by a dominant gene, so each child of an affected parent stands a 50 percent chance of inheriting the gene and therefore developing the disease later in life. Generally the illness is not recognized until the victims have had children of their own, their offspring must then live in fear of suffering the same fate.

Until 1983 there was no way to
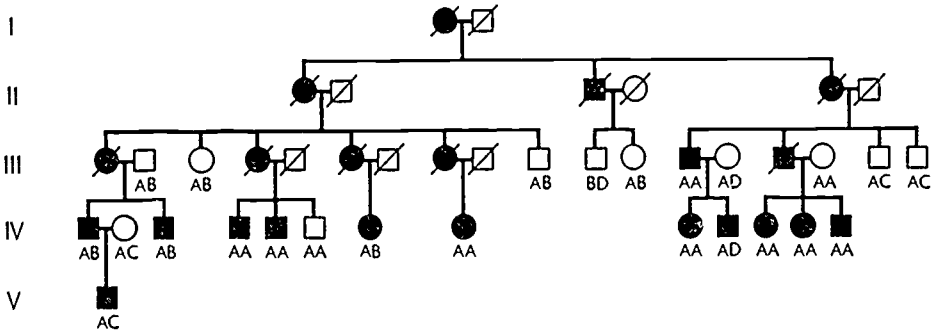
Pedigree of an American family with many victims of Huntington's disease. The victims are indicated in black. Circles represent females and squares represent males. The original victim (top circle, with slash indicating that she is deceased) transmitted the disease to all three of her children. Six of these children's children were spared, but six others inherited the gene for the disease and 12 living members of this family have disease symptoms. Based on restriction fragment length polymorphism studies, one other family member is expected to develop the disease.

identify carriers of the gene before their symptoms developed. Nor was there any clue to what caused the disease or where to seek remedies. Since scientists had no idea where the Huntington's disease gene was located, they could not even look for a restriction enzyme site that was close to it, as Drs. Kan and Dozy had done for sickle-cell disease a few years earlier.

At Massachusetts General Hospital in Boston, molecular biologist Dr. James F. Gusella decided to cut up the DNA of Huntington's disease victims with restriction enzymes and compare the resulting fragments with fragments of similarly treated DNA from unaffected members of the same families. He wanted to see whether the lengths of these fragments varied in any way that could be linked to the appearance of Huntington's disease.

This required, first of all, finding a very large family with many living members who were willing to be tested and were old enough either to show signs of Huntington's disease or to be likely to have escaped it. (Past the age of 60, the risk of developing Huntington's drops to 1 percent.) Through Dr. P. Michael Conneally, a professor of medical genetics at Indiana University in Indianapolis who maintains a Huntington's disease roster, such a family was found.

Next Dr. Gusella began to look for patterns in the millions of fragments which he obtained after cut-

ting up DNA taken from the cells of both Huntington's disease victims and their healthy relatives. He intended to expose these fragments to a whole "library" of hundreds of radioactively labeled probes, one by one. Any variations in the resulting autoradiographs would then be checked for possible linkage to Huntington's disease, the statistical analysis involved would be done by Dr. Conneally with the aid of a special computer program.

The very first batch of 12 probes which Dr. Gusella used produced a promising lead. One of the probes, a DNA sequence arbitrarily called G8 and whose function was unknown, stuck to fragments of different lengths in DNA from various members of the family, these fragments could form any one of four patterns, designated with the letters A, B, C, and D; and it appeared that the A pattern was inherited along with Huntington's disease.

In any study of linkage, scientists need a large number of cases in which a disease correlates with some identifiable marker, the greater the number of such cases, the higher the odds that the two factors are linked. The American family was large by American standards, but had only 25 living members of appropriate ages. While this was enough to provide a hint of a link between the DNA fragment lengths and the disease, it was not enough to prove it.
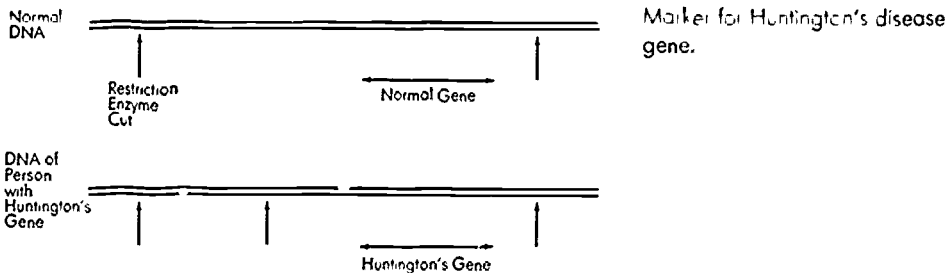
Fortunately, an NIH team had

been studying an extraordinary group of interrelated families who live along the shores of Lake Maracaibo in Venezuela and have the largest concentration of Huntington's disease in the world—nearly 100 living victims and 1,100 persons at risk for the disease. For 3 years in a row, this team took blood and skin samples from people who clearly had the disease, as well as from their healthy relatives. One complete set of blood samples was carried to Dr. Gusella. Another set, together with the skin samples, was taken to the NIGMS Human Genetic Mutant Cell Repository in Camden, New Jersey, where they were established as cell cultures, frozen, and stored. These cells are now available to researchers.

At Massachusetts General, Dr. Gusella extracted DNA from the Venezuelan families' white blood cells (red blood cells have no nuclei, and therefore no DNA) and set to work. He cut up the DNA with a restriction enzyme called HindIII, as he had done with the DNA from the American family, then exposed the fragments to the G8 probe.

Next Dr. Conneally analyzed Dr. Gusella's data on the families' DNA fragments. Although he found that in this family a different pattern of fragment lengths—the C pattern—was inherited together with Huntington's disease, the link between this pattern and the disease held up very well. As the information poured in, he became exuberant. "The chances were 600 million to 1 that we had the marker," he said. "And when we studied some more of the Venezuelan samples, the chances increased to over a billion to one—about as certain as you can get."

If G8 continues to be a suitable

Marker for Huntington's disease gene.

Normal DNA

Restriction Enzyme Cut

Normal Gene

DNA of Person with Huntington's Gene

Huntington's Gene

probe, it may soon be possible to diagnose who in a particular family carries the Huntington's gene and who is free of it—assuming that there are enough key family members to determine which pattern of DNA fragments travels with the Huntington's disease gene in that family.

The 100,000 Americans who are at risk for Huntington's may thus be able to learn their fate, if they wish to do so. The burden of Huntington's disease could then be drastically reduced within a few generations, if couples who know that one of them has the gene refrain from bearing children. Prenatal diagnosis would be another potential option for such couples. At the same time, prospective parents who might hesitate to have children for fear of transmitting the disease would be freed to do so if they knew that they did not carry the Huntington's gene.

Another happy feature of Dr. Gusella's and Dr. Canneally's finding was that Dr. Susan L. Naylor of the Roswell Park Memorial Institute in Buffalo, New York, who had been working with the G8 sequence for unrelated reasons, had already located this sequence on human chromosome number 4. This indicates that the Huntington's gene is also on chromosome 4. Knowing this, scientists can begin to zero in on the Huntington's gene itself. Ultimately, the gene can be isolated and sequenced. Once this happens, diagnosis should be much easier and applicable to anyone, even without analyzing many family members. Even more importantly, scientists may then be able to decipher the biochemical defect that is involved in Huntington's disease and perhaps find ways to counteract it.

THE RAPID GROWTH OF GENE MAPPING

In 1911, scientists noticed that only men had a certain kind of color blindness. The trait seemed to be passed down from color-blind fathers to their color-blind grandsons, through daughters who had normal vision. The scientists deduced that the gene for color blindness must be on the X chromosome. They reasoned that a female (who has two X chromosomes) must be protected against color blindness by inheriting one X chromosome carrying a normal gene and another carrying a defective gene, and that the normal gene compensates for the defective one, however, a male (who has one X chromosome from his mother and one Y chromosome from his father) lacks this protection. Therefore each son of a woman who carries a defective gene on an X chromosome has a 50 percent chance of inheriting the defect.

This was the first recorded instance of gene mapping, the attempt to locate genes on specific chromosomes—a field which is now undergoing explosive growth. Gene mapping is proving extremely important both in the diagnosis of genetic diseases and in research on the controlling regions that turn certain genes on and off.

Although it was relatively easy to assign certain genes to the X chromosome, scientists had to wait half a century before they could assign genes to particular human autosomal chromosomes, and even longer to map the position of any gene on a chromosome.

A first breakthrough come in the 1960's, when researchers learned how to fuse human cells with mouse cells in a test tube. This created hybrid cells in which human chromosomes, or chromosomal fragments, could be isolated and recognized. The human cells used for such hybrids are mixed with cells derived from mouse tumors and treated with a variety of agents that promote the formation of bridges between them, new fused or hybrid cells with a large number of both human and mouse chromosomes soon appear. Generally all the mouse chromosomes in these hybrid cells are retained, while the human chromosomes are gradually lost as the cells multiply. With luck, the hybrid cells end up with only one or a few human chromosomes remaining in each cell, or better yet, a fragment of one human chromosome. Any human protein produced by such cells is evidence that the gene for this protein is located on the particular human chromosome or fragment remaining in the hybrid cell. With the aid of this procedure, more than 100 human genes were localized to specific chromosomes.

An even newer technique is in situ (the Latin term for "in place") hybridization, in which scientists stop the division of human cells at a stage when each chromosome is highly condensed and clearly visible under a light microscope. Then they bind or hybridize the chromosomes
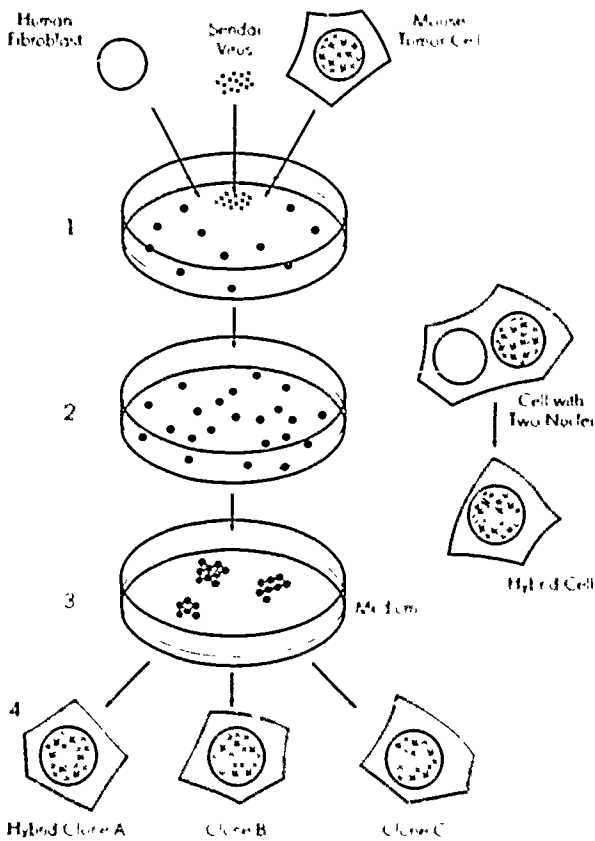
Human-mouse hybrid cells To study individual human chromosomes, scientists create human-mouse hybrid cells that contain one or a few human chromosomes These hybrids are formed by mixing human fibroblasts (connective tissue cells) with mouse tumor cells and adding a fusion-enhancing agent such as inactivated Sendai virus (1). Some of the cells fuse (2), first forming cells with two nuclei and then forming new hybrid cells with a single nucleus When placed in a special medium, the hybrid cells proliferate and form colonies (3) The hybrid cells retain all the mouse chromosomes but, with continued growth, lose most of their human chromosomes in a random fashion (4) Thus, each hybrid clone (descendant of a single hybrid cell) has a different complement of human chromosomes which can be studied separately

40

Chromosomes of a normal human male prepared using high-resolution banding techniques that reveal approximately 850 bands on 23 chromosomes.

1   2   3   4   5

6   7   8   9   10   11   12

13   14   15   16   17   18

19   20   21   22   X   Y

41
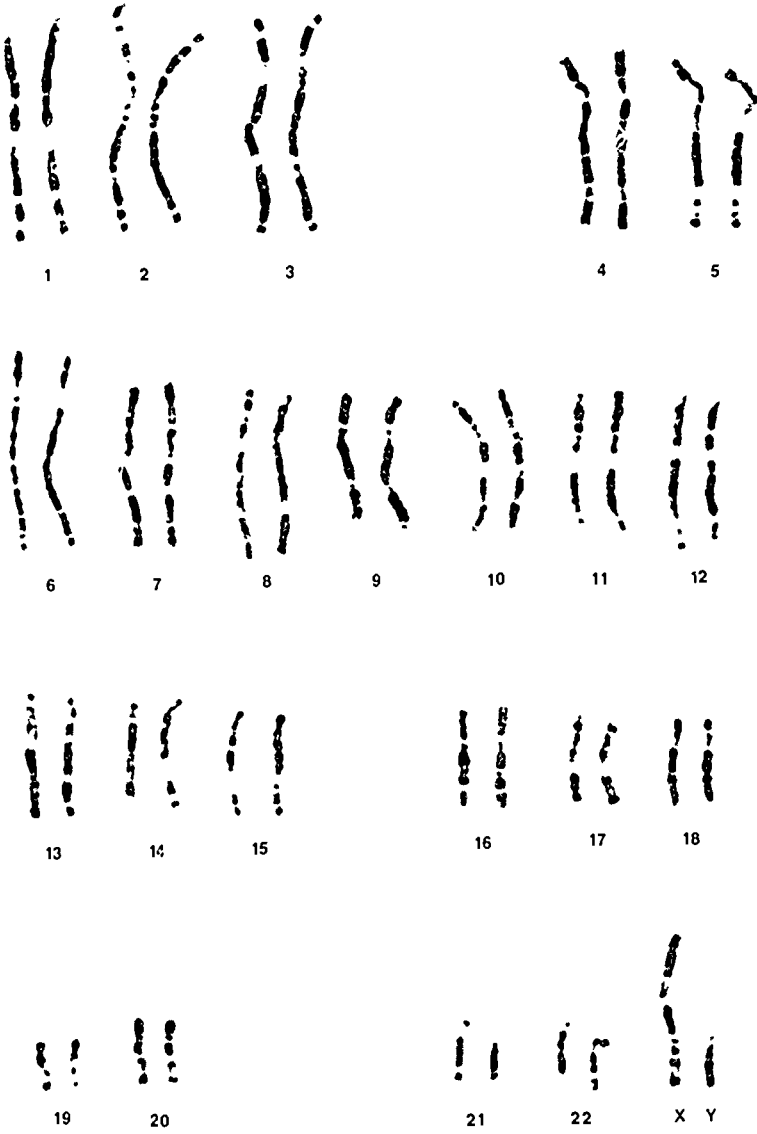
in these cells to a radioactively labeled copy of the piece of DNA that they are trying to map. When scientists find out, through autoradiography, which chromosome this piece of DNA sticks to, they can deduce the location of the DNA.

Techniques such as in situ hybridization and cell fusion are useful for mapping genes whose products are known. But in many cases, as with the genes responsible for Huntington's disease or cystic fibrosis, scientists know absolutely nothing about the nature of the gene's products. If they wish to map such genes, they must rely on the fact that any identifiable genetic variation, such as an RFLP, which is inherited along with a disease must be relatively close to the gene for this disease, on the same chromosome. Other techniques for mapping genes by direct chromosome analysis are currently being refined.

Once a gene is known to be on a particular chromosome and relatively close to a specific marker, scientists can begin to "walk" along the chromosome, using different DNA probes to come closer and closer to the gene. The hope is ultimately to isolate the gene, to understand what it makes, and to decipher how its malfunction can cause disease.

Gene mapping relies heavily on the finding that certain chemical stains produce characteristic patterns of light and dark bands on chromosomes. After chemical stain-

ing, each human chromosome or fragment of one has its own identifiable pattern when viewed under the microscope, and the individual bands serve as landmarks for the positions of specific genes.

In 1976, Dr. Jorge J. Yunis of the University of Minnesota developed high-resolution chromosome banding techniques which make it possible to see many more bands than before—up to 5,000 bands on a set of 23 human chromosomes. At this high degree of resolution, each band may hold as few as 10 genes.

By now over 3,500 human traits have been associated with specific genes, according to Dr. Victor A. McKusick of The Johns Hopkins University School of Medicine, who publishes a continually updated catalog of dominant, recessive, and X-linked traits. While a few of these traits are normal variations (such as the ABO blood types), 90 percent of the genes listed in his catalog are associated with diseases.

Over 600 of these genes have been assigned to specific chromosomes. For about half of them, the general area of the chromosome on which they are located is also known. In many cases their location has been narrowed down to just a few bands on the chromosome or to individual bands. Nearly all of this mapping has occurred in the past decade.

Eventually the entire set of genes in human cells may be mapped. Drs. David Botstein of the Massa-

Gene map Over 600 genes have been assigned to specific chromosomes, and about half of them have been mapped to certain locations on these chromosomes, as can be seen from this sampling. Many more genes (such as those for hemophilia and Lesch-Nyhan syndrome) have been mapped to the X chromosome than to the Y chromosome, which is carried only by males. Eventually the entire set of genes in human cells may be mapped.

A human gene map — chromosome ideograms with gene loci assignments, February 1 1984.
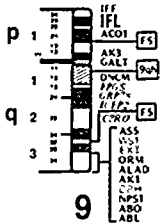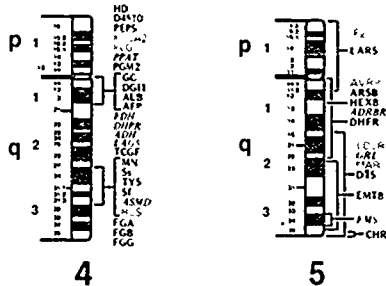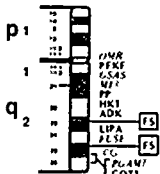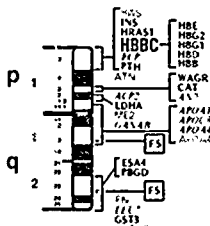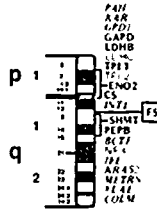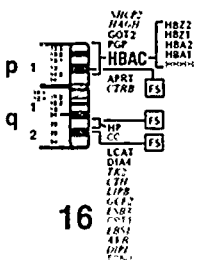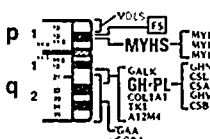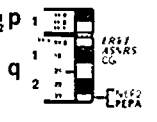
chusetts Institute of Technology, Raymond L. White and Mark Skolnick of the University of Utah, and Ronald Davis of Stanford University suggest using a variety of restriction enzymes to cut up DNA (especially that of people who belong to large families whose traits over several generations are well known), using DNA probes to study the resultin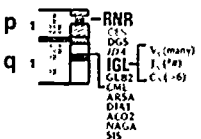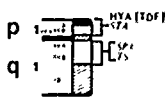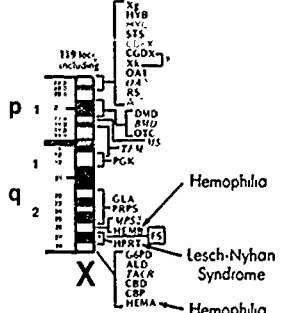g fragments, and systematically looking for restriction fragment length polymorphisms on each chromosome. They estimate that if scientists find just 150 polymorphisms spaced throughout the genetic material at reasonably regular intervals, "then all genes will be linked to one or another of the regions containing RFLP's and can thereby be mapped."

It might then become possible to understand better each person's genetic composition, or genotype—and also to unravel the mechanisms of many complex diseases, such as diabetes or breast cancer, which may be different in separate families. Diabetes may actually consist of several distinct disorders, each caused by one or more separate genes. These disorders only look alike because we do not understand them well enough. Over the past few years, researchers have discovered that more than 20 separate mutations can produce thalassemia. In order to prevent such illnesses or treat them most effectively, it is sometimes important to know which mechanism is involved in a given patient or family.

"Part of our goal is to sort out who is genetically predisposed to what disease," says Dr. Skolnick. If people know that they are particularly at risk for colon cancer, for instance, they can be examined more frequently and make sure that any benign growths which are found are removed, to prevent their becoming malignant.

Gene mapping is also helping researchers investigate some of the controlling elements in DNA that influence the expression of adjacent genes. These small sites, only about 100 base pairs long, act as "promoters," "terminators," "enhancers," "attenuators," or other kinds of regulators of genes. They hold the key to many of the changes that occur during growth and development, and they play a vital role in certain diseases.

For example, several related hemoglobin genes—those for fetal hemoglobin as well as those that are expressed in adulthood—are located next to each other on the same chromosome, but at various times some are switched on and others are switched off. In people with a double dose of the sickle-cell gene, the switch from fetal to adult hemoglobin usually leads to symptoms of the disease. Scientists would like to find out what controls this turning on or off in a specific region of DNA. Once they know where a gene is located, they can begin to look for its controlling elements,

which often ore not too for owoy.

It is generolly believed thot eoch humon cell (except the egg ond sperm cells) contoins oll the genes thot ore ovoiloble onywhere in the body. Only some of these genes ore turned on ot ony porticulor time in o given cell, however. Leorning how to turn on specific genes in ony humon cell might moke it possible to octi vote certoin functions so os to correct genetic defects. Leorning how to turn off other genes migh, prevent some diseoses, such os concer.

46

ZEROING IN ON CANCER GENES

One of the most spectacular results of the new human genetics has been the discovery that certain genes, called oncogenes, play a key role in the development of cancer. These findings are revolutionizing the study of cancer and may lead to entirely new methods of diagnosing, treating, or preventing many kinds of malignancies.

More than 20 human oncogenes have been identified as of mid-1984, and laboratories around the world are competing with each other to find more. Individual oncogenes appear to be necessary, but not sufficient, to trigger cancer, a process which is believed to require several steps. In some cases, chromosomal breaks, deletions, translocations, or insertions of foreign DNA place a potential oncogene near a regulatory element that activates it, in other cases, single point mutations in the coding sequences of certain genes may turn them into oncogenes. In addition, it is now believed that two or more oncogenes of separate classes need to cooperate with each other to make normal cells become malignant.

Most of the oncogenes were found after scientists discovered that certain viruses constantly move in and out of cells, picking up bits of DNA as they go. "Out of the millions of times this happens, an occasional virus picks up genes that make a recipient cell go crazy," explains Dr. John Cairns, professor of microbiology at the Harvard School of Public Health. "If you can find this virus, it hands you the cancer gene on a plate."

Thus a virus that infects chickens, called the Rous sarcoma virus, contains a powerful oncogene—the src (for sarcoma) gene—which stems from a closely related, normal gene of chickens. All mammals, fish, and even fruit flies have normal genes that are almost identical to src, which shows that these genes have persisted over millions of years of evolution. They are believed to code for same key proteins involved in cell growth and regulation, becoming activated as oncogenes by association with a retrovirus (an RNA-containing virus) such as the Rous sarcoma virus.

An entirely separate line of research which has also uncovered same oncogenes started with the question of how various chemicals and other non-viral mutagens cause tumors. Many kinds of chemicals, as well as ultraviolet light and radiation, can cause chromosomal breaks, translocations, or deletions, or produce point mutations in DNA. In most cases these changes are not significant or are taken care of by the cells' self-repair mechanisms. But sometimes they result in placing a gene that has the potential to become an oncogene next to a powerful activating signal which provokes it to malignant activity.

It had long been known that various leukemias (cancers of the blood) and lymphomas (cancers of

Normal
Chromosome 8

Normal
Chromosome 14

Translocation
Chromosome 8q −

Translocation
Chromosome 14q +
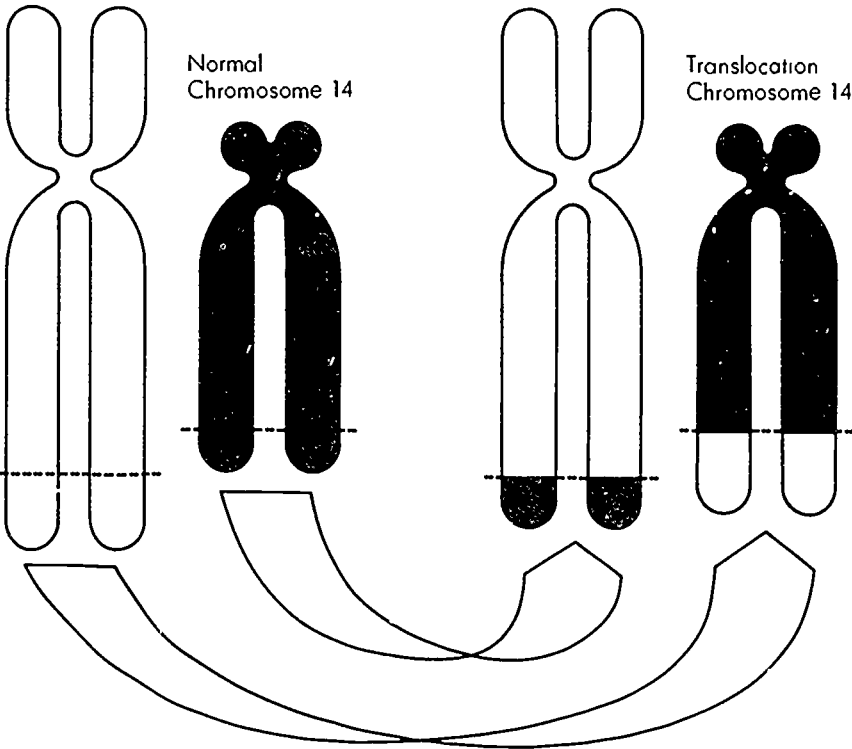
Chromosomal translocation. As cells replicate and divide, chromosomal errors may occur in which pieces of chromosomes break off, attaching to other chromosomes. One reciprocal translocation, in which two chromosomes exchange pieces, has been linked to a rare form of cancer called Burkitt's lymphoma. A piece of chromosome 8 containing a gene called myc breaks off and moves to chromosome 14, while a piece of chromosome 14 moves to chromosome 8. As a result, the myc gene seems to lose some of its regulatory sequences and becomes an activated oncogene.

the lymphoid tissues that play a key role in the body's defense against infection) are correlated with specific chromosomal translocations. Now researchers have found that in Burkitt's lymphoma a gene called myc is moved from its normal location on chromosome 8 to chromosome 14, where it lands near a gene for immunoglobulin. This juxtaposition appears to increase transcription of the myc gene. Through the chromosomal rearrangement process, the myc gene seems to lose some of its own regulatory sequences, acquires more active sequences which are normally involved in the production of antibodies, and becomes an activated oncogene.

Recently these two lines of research have converged, as some of the oncogenes that were found in chemically induced tumors turned out to be the same as the onco-

genes previously found in retroviruses. The myc gene, for instance, is associated with four distinct strains of viruses known to cause cancer in chickens.

Thus a common set of genes may hold the key to many kinds of cancer. Apparently these genes can become activated in various ways. by association with a retrovirus, by juxtaposition with activating signals from other genes, by point mutations, or by other means. Each activation is a step in the conversion of a normal cell into a tumor cell, but several such steps may be necessary to produce cancer.

Once the protein products of these cancer genes are identified, it may be possible to interfere with the action of either the genes or their products and thus halt or reverse the progress of cancer.

A BANK OF LIVING HUMAN CELLS

In order to study the role of mutations or other genetic errors in disease, scientists need to have access to living cells which contain these genetic changes. Since 1972, NIGMS has maintained a special facility for this purpose, the Human Genetic Mutant Cell Repository, generally known as the Cell Bank, at the Institute for Medical Research in Camden, New Jersey. It is the largest cell bank of this type in the world.

Here, stored in liquid nitrogen at 316 degrees below zero Fahrenheit, are over 6,000 cell lines representing more than 300 genetic diseases. There are cells with every known type of chromosomal abnormality, including deletions, additions, inversions, translocations, and "fragile sites" where breaks are likely to occur. There are cells with identified biochemical defects, such as those from people with Tay-Sachs disease, in which a fatty material accumulates in the brain because of an enzyme deficiency, and cells in which the biochemical defect is still a mystery, as in cystic fibrosis. There are cells from individuals and cells from large families, cells taken from amniotic fluid; cells from victims of various hereditary diseases, as well as cells from healthy carriers of these diseases, and also cells from normal volunteers, used for comparison.

All cell cultures in the bank are available at modest cost to anyone doing research on genetic diseases.

A sliver of skin is removed from a young woman's forearm. After incubation in a nutrient solution, some of the cells in this tissue will grow and divide, forming a line of cultured cells.

Each year the Cell Bank ships out some 3,000 cell cultures to researchers in the U.S. and abroad. Until recently, the demand has been mostly for fibroblasts—connective tissue cells obtained from samples of human skin—which are particularly useful for studies of cell biochemistry and chromosomes. Many diseases which seem, in a patient, to be localized to a particular organ are actually caused by biochemical defects that show up even in cells taken from the skin. Now the demand is shifting toward lymphoblasts, the immortalized progeny of white blood cells that have been exposed to the Epstein-Barr virus (a virus associated with Burkitt's lymphoma and with infectious mononucleosis) and can grow forever. Lymphoblasts are proving more practical for studies of gene linkage, as in Huntington's disease.

Establishing such cell lines from skin or blood takes considerable expertise, as well as time. The cell cultures must be sterile, guaranteed to have specific abnormalities, and able to grow.

Fibroblasts can reproduce only a limited number of times—usually not more than 40 to 50—in tissue culture, and well before the end of this lifespan they slow down, taking much longer to divide. Lymphoblasts are immortal, but after being kept in continuous culture for a long time they often undergo chromosomal changes which may interfere with research. Researchers therefore need cells that are still in their prime.

"Before the repository, if you wanted to work on a genetic disorder like galactosemia, which is expressed in fibroblasts, you usually got a cell line from a friend," says Dr. Fred Bergmann, director of the Genetics Program at NIGMS. "But the cells might be contaminated, or half dead! So the repository does not usually begin with a cell line already in tissue culture, but with a primary source, such as a piece of skin from a patient."

A small circle of skin, 3 millimeters in diameter, is punched out—usually from the donor's forearm—under a local anesthetic. At the Cell Bank, this sample is cut into tiny pieces and "seeded" in a plastic flask, where it is covered with culture medium and left to incubate at body temperature for a few weeks. The cells divide, and when the surface of the flask is covered with cells, the population is split into other flasks. This process of tissue culture is repeated several times until there are enough cells for up to 100 glass ampules of 1 million cells each. But meanwhile the cells go through dozens of tests to make sure they are not contaminated by microorganisms such as bacteria, yeast, fungi, or mycoplasma. Finally—3 to 6 months after the skin sample's arrival—the cells are frozen inside their ampules and stored for future use.

Research by Drs. Joseph Goldstein and Michael Brown of the University of Texas Health Science
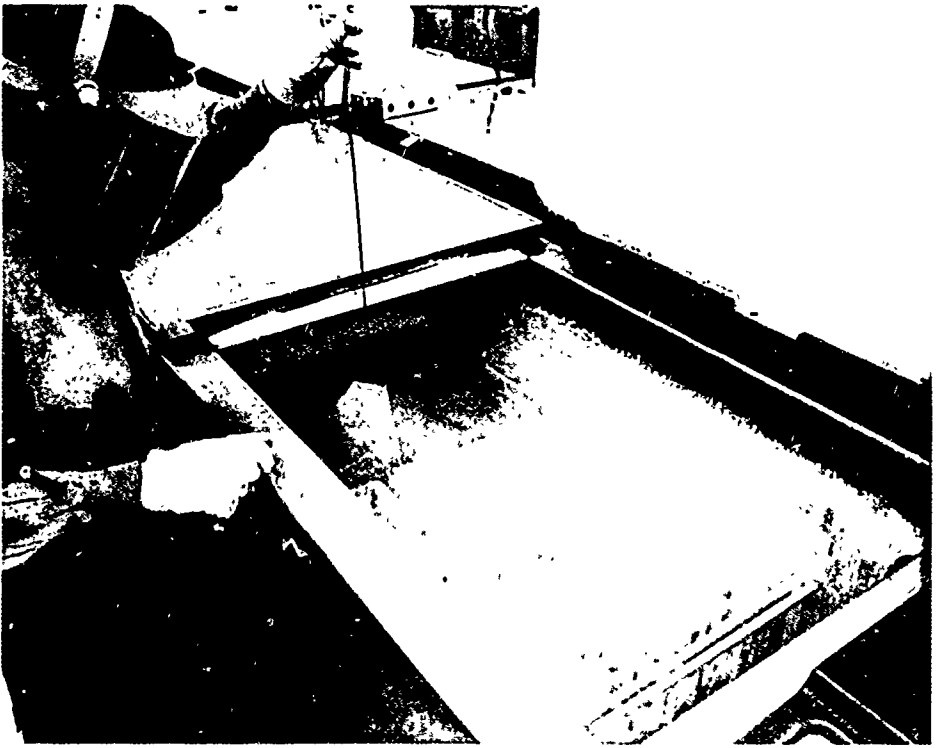
Vials containing cultured cells
are heat-sealed before being
placed for storage in tanks of
liquid nitrogen at −316 °F
(−196 °C) in the NIGMS Cell
Bank.

Center of Dallas provides a good example of how such cells are used. Working with fibroblasts, they discovered various defects in a specific type of receptor on the surface of some cells (receptors are proteins which recognize specific chemicals from outside the cell and send appropriate chemical signals into the cell's interior). These defects are determined by various forms of a dominant gene and greatly increase the risk of heart disease and early death. Among people who have suffered heart attacks before the age of 60, 1 out of 20 has been found to carry a gene for familial hypercholesterolemia, a disorder in which too much cholesterol is produced. Such genes occur in roughly 1 in 500 persons in the general population, they are thus among the most prevalent genes which, in single dose, lead to a specific and sometimes fatal disease. Very infrequently, two persons carrying these mutant genes marry. Each of their offspring then stands a 25 percent chance of inheriting a double dose of the gene for hypercholesterolemia. Such extremely rare individuals have a high likelihood of dying from heart disease before the age of 25.

All mammalian cells need a small supply of cholesterol. Normal cells either take up some cholesterol from the blood or produce it themselves, in response to an intricate feedback mechanism. Special receptors on the cell surface bind to LDL (low-density lipoproteins, which carry cholesterol in the blood) and bring cholesterol into the cell, this acts as a signal to the cell to reduce its own production of cholesterol. The system works pretty much like a thermostat to ensure that cells always have enough cholesterol for life and growth without producing too much of it. But Drs. Goldstein and Brown discovered that cells from people with familial hypercholesterolemia do not have a sufficient number of working LDL receptors on their surfaces, breaking the normal chain of control. These cells then go right on producing cholesterol, which spills out of the cells, accumulates in the victims' blood, and clogs their arteries.

"We have identified seven different mutant alleles—seven alternative forms of the gene for LDL receptors —which can produce familial hypercholesterolemia," says Dr. Brown. "Some patients can't make any LDL receptors at all. Others can make these receptors but can't transport them to the cell surface, so the receptors don't function. Or else the receptors get to the surface but are relatively less efficient."

Now that the researchers have isolated and sequenced most of the LDL receptor gene, they expect to develop techniques to diagnose these variant forms of hypercholesterolemia. Meanwhile, the insights produced by this sort of research have led to progress in treating the disorder. Instead of using drugs to reduce the amount of cholesterol in

the blood of such patients (and thus forcing their cells to produce even more cholesterol to make up for the loss), scientists are now developing drugs that appear to stimulate the production of normal LDL receptors. Genetic research pinpointing the most appropriate treatment for each type of hypercholesterolemia could lead to a dramatic reduction in the rate of circulatory disorders and heart attacks.

The Cell Bank now contains 22 different lines of fibroblasts from people with familial hypercholesterolemia, including the cells on which Drs. Goldstein and Brown based their discoveries.

The bank also has lymphoblasts grown from the white blood cells of members of some very large and special families. These cells are used primarily for linkage studies, in efforts to find reliable markers for various disorders. For instance, there are cell lines from 49 members of an Amish family with a high incidence of depressive and manic depressive disorders, in whom the psychiatric status of each person is known. And there is the collection of cells from 341 interrelated members of the Venezuelan Huntington's disease community, which scientists can use to study a variety of genetic disorders in addition to F   igton's disease.

The Cell Bank's collection is considered so valuable that elaborate steps have been taken to prevent its loss in some disastrous fire or other local accident. A complete duplicate of this collection is stored at the Cooper Hospital Medical Center in Camden, New Jersey, and both sets are protected by various automatic sensors and security systems.

GENBANK:
A NATIONAL DATABASE OF NUCLEIC ACID SEQUENCES

There are no cell cultures and not even a scrap of DNA or RNA in storage at GenBank®, the genetic sequence data bank created by NIGMS in 1982—just computers and lists of nucleic acid sequences from organisms as varied as baboons, bacteria, yeast, or E. coli, plus over 500 sequences of human and mouse DNA and RNA.

Now that stretches of nucleic acids of various kinds are being sequenced by researchers all over the world with increasing speed, a central library of sequences has become essential to prevent duplication of effort and to enable scientists to compare what they are looking at with all other known sequences. It would be impossible for researchers to do their own work and also keep up with the flow of information in this field.

GenBank is an international repository of all published nucleic acid sequences greater than 50 nucleotides in length. The sequences are catalogued, checked for accuracy, and annotated for sites of biological interest at the Los Alamos National Laboratory in Los Alamos, New Mexico. By the end of 1983, GenBank contained more than 4,000 separate sequences comprising over 2.8 million base pairs of DNA.

Users can call up on their own computers or request a printout of any sequence in which they are interested and find out many of the significant features within this sequence—for example, which regions code for proteins. The database allows researchers to search for similarities between a new sequence and all existing sequences, and to do sophisticated analyses that might reveal other important regions. It has proved particularly useful to scientists who are working on oncogenes, as well as to those who are looking for the DNA sequences that control gene replication or expression.

GenBank is supported by NIGMS as well as by several other Institutes or Divisions of NIH and some branches of government. It is administered by the Computer Systems Division of Bolt Beranek and Newman Inc. in Cambridge, Massachusetts, and shares information and collaborates with the Nucleotide Sequence Data Library of the European Molecular Biology Laboratory in Heidelberg, West Germany.

THE PROMISE OF GENETIC THERAPY

Many genetic diseases are lethal or extremely debilitating, and their treatment still leaves much to be desired. A few of these diseases can be treated by providing substances that are deficient in the body, or by removing other chemicals that accumulate to dangerous levels. In the disorders of vitamin metabolism mentioned earlier, for instance, large doses of specific vitamins can save lives. Some hormonal defects can also be remedied. Thus, inherited growth hormone deficiency is treated with growth hormone, this can either be taken from the pituitary glands of cadavers or—now that the gene for this hormone has been identified and cloned—mass-produced by bacteria.

Recently, scientists have found the gene for factor VIII, the rare and expensive blood-clotting factor that hemophiliacs need so they will not bleed to death (hemophilia is an X-linked bleeding disorder). Before this discovery, factor VIII had to be extracted in minute quantities from many donors' blood and thus carried a high risk of transmitting hepatitis, as well as some risk of passing on acquired immune deficiency syndrome (AIDS). Furthermore, many hemophiliacs could not afford to pay $10,000 a year or more for factor VIII therapy. But now that the gene has been isolated, it can be inserted into bacterial cells and factor VIII can be produced in large quantities, at low cost, without any risk of these complications.

Another highly treatable genetic defect is phenylketanuria (PKU). Children with this condition suffer from a metabolic disorder which prevents them from breaking down the amino acid phenylalanine, a normal constituent of most foods. As a result, phenylpyruvate, a toxic chemical, builds up in their blood, causing severe mental retardation. Fortunately, the damage can be reduced or prevented by a stringently restricted diet started right after birth, and screening of newborns for PKU is now mandatory in many states.

Such treatments help, but they do not cure the diseases involved. Moreover, most genetic disorders cannot be treated at all by this sort of environmental manipulation. Therefore researchers are now working on entirely new ways to correct genetic defects—by acting directly on the DNA in people's cells.

As mentioned earlier, one approach might be to switch on certain genes which would otherwise be inactive, so that they can take over the job of defective genes. This has already been tried on some victims of thalassemia and sickle-cell disease. However, the overall usefulness of this therapy remains to be determined.

Another possible approach is gene therapy—the introduction of normal genes into the chromosomes of cells that contain defective genes in the hope that the manipulated

53

cells will ultimately replace the defective ones, thereby curing the patient. The technical problems that must be solved before gene therapy can be used in medical practice are overwhelming, but according to Dr. W. French Anderson of NIH, "it is a procedure with enormous potential."

"Gene therapy. . . is conceptually no different from any therapy in medicine that attempts to improve the health of a sick patient," explains Dr. Arno G. Motulsky, director of the Center for Inherited Diseases at the University of Washington, Seattle. "The only difference is that DNA, rather than other biologicals, drugs, or surgery, is used as the therapeutic modality. This point is important because some critics claim that gene replacement represents a revolutionary departure in medical treatment. In fact, gene therapy for diseased tissues is no different from any other therapy. No change in the genes of the reproductive organs is attempted."

Only the first prerequisite for such therapy—the ability to clone a normal counterpart of a defective gene—is available so far. Three major hurdles remain. learning how to deliver the normal gene to the proper target cell and having it stay there, getting the normal gene to produce what it should and regulating it properly, and ensuring that the new gene does no harm.

Many new techniques of gene delivery are being developed. For example, some researchers insert genes into viruses and then let the viruses infect cultured cells. Others inject DNA directly into the nuclei of cells.

"In animals, we can sometimes deliver genes to target cells," summarizes Dr. Anderson. "Once in place, a few of these genes can be made to express themselves in the desired manner.

"For humans, we cannot carry out the first two steps in the transfer process [delivery and regulation] and therefore know nothing about the safety."

Nevertheless, the recent progress in work with animals implies that the time is approaching when gene therapy on somatic (non-reproductive) cells could be done in humans. Such therapy would affect only the patient, not the patient's offspring, and would in some ways resemble an organ transplant. It might allow some single-gene disorders such as Lesch-Nyhan syndrome, which causes severe mental retardation and self-mutilation. to be treated for the first time.

Meanwhile such major illnesses as heart disease, concer, ond schizophrenia, each of which probably involves many different genes, seem far too complex for any kind of gene therapy. At present the manipulation of several genes at one time is not possible. However, some multigene diseases may be prevented or mitigated by means of drugs, or by manipulating lifestyle or other environmental variables. As scientists

discover the genetic risk factors for these diseases, people who find out that they are at particular risk may be able to reduce their chances of illness by changing specific aspects of their environment, thereby keeping the diseases at bay.

In these various ways, many medical mysteries are beginning to yield to the genetic approach. The discoveries sketched above represent only a fraction of current research on chromosomes, genes, recombinant DNA, gene mapping, oncogenes, genetic diseases, and treatment. As this research advances, it is enabling doctors to help millions of people prevent or control inherited diseases. At the same time, this research is laying the groundwork for a new medical science for the 21st century.

GLOSSARY

agarose gel—a porous, semi-solid material used for many research purposes, among them electrophoresis.

alleles—alternative forms of a gene, occupying a specific site on a chromosome, which determine alternative characteristics in inheritance.

amino acid—a building block of protein. Each protein consists of a specific sequence of amino acids.

amniocentesis—a method of prenatal diagnosis that involves withdrawal of a small amount of fluid from the amniotic sac that surrounds the fetus; the fluid contains cells shed by the fetus which can be analyzed.

amniotic fluid—the fluid that cushions the fetus inside the amniotic sac.

autoradiograph—image produced on an X-ray film by a radioactively labeled substance.

autosome—any of the non-sex chromosomes; in the case of normal humans, there are 22 pairs.

base—one of the five molecules that make up the informational content of DNA and RNA. In DNA, bases pair across the two chains of the double helix. adenine with thymine, and guanine with cytosine RNA is single-stranded and contains uracil instead of thymine.

base pairs—pairs of complementary nucleotides forming the DNA double helix.

carbohydrate—a class of compounds such as starch or cellulose that is broken down to sugars such as glucose and fructose.

carrier of genetic disease—a person who possesses a defective recessive gene together with its normal allele. Although the defective gene's product is not detectable, the gene can be transmitted to progeny who will have the genetic disease if another copy of the same recessive gene is inherited from the other parent.

cell—the basic subunit of any living organism; the simplest unit that can exist as an independent living system.

cell culture—the propagation of cells or tissues outside an organism, using special nutrients conducive to their growth; also known as tissue culture.

cell-surface receptor—a protein in the cell surface which selectively recognizes a specific chemical from outside the cell. This chemical must fit the receptor like a key in a lock.

cholesterol—a fatty substance found in all animal tissue.

chorionic villus sampling—a still-experimental method of prenatal diagnosis that involves withdrawing a sample of tissue from protrusions of a membrane called the chorion, which surrounds the developing fetus. The sampling is done directly through the uterine cervix and allows the detection of many fetal defects as early as the 9th week of pregnancy.

chromosome—a rod-like structure found in the cell nucleus and containing the genes. Chromosomes are composed of DNA and proteins. They can be seen in the light microscope during certain stages of cell division.

chromosome bands—patterns of light and dark bands produced by chemical staining of the chromosomes. Each chromosome or fragment of one has its own identifiable pattern of bands seen under the microscope, and the individual bands serve as landmarks for the positions of specific genes.

cloning—asexually producing multiple copies of genetically identical cells or organisms descended from a common ancestor. Compare with gene cloning.

codon—a triplet of nucleotides in the DNA or RNA molecule that codes for 1 of the 20 amino acids in proteins or for a signal to start or stop protein production. Each gene that codes for a protein is a series of codons which gives the instructions for building that protein.

cytoplasm—all the substance inside a cell, excluding the nucleus.

differentiation— the series of biochemical and structural changes that groups of cells undergo in order to form a specialized tissue.

DNA (deoxyribonucleic acid)—the substance of heredity, a large molecule which carries the genetic information necessary for the replication of cells and for the production of proteins. DNA is composed of the sugar deoxyribose, phosphate, and the bases adenine, thymine, guanine, and cytosine.

DNA denaturation—the separation of DNA into its two strands of nucleotides, for example by exposing it to near-boiling temperatures or to extremely alkaline conditions.

DNA probe—a specific sequence of single-stranded DNA used to seek out a complementary sequence in other single strands. The probe is usually made radioactive so that it can be detected.

DNA sequencing—determining the nucleotide sequences of DNA.

dominant—refers to a characteristic that is apparent even when the gene for it is inherited from only one parent.

dominant gene—a gene that is expressed even when its allele on the paired chromosome is different.

electrophoresis—a method of separating substances, such as DNA fragments, by using an electric field to make them move through a medium at rates that correspond to their electric charge and size.

enzyme—a protein which speeds up or catalyzes a specific chemical reaction.

Epstein-Barr virus—a virus associated with Burkitt's lymphoma and infectious mononucleosis.

Escherichia coli (E. coli)—a common intestinal bacterium which geneticists have used for many studies.

exons—coding sequences of genes which are retained (after excising introns) when mature messenger RNA is made.

factor VIII—a clotting factor which is lacking in the blood of people who suffer from hemophilia, exposing them to the danger of uncontrolled bleeding.

fetoscopy—a technique for withdrawing blood samples from a fetus with the aid of a device that permits direct visualization of the fetus and placenta.

60

fibroblast—a connective tissue cell, found in skin and other tissue.

fragile sites—chromosomal areas that are particularly vulnerable to breakage.

gene—a unit of heredity; a segment of the DNA molecule containing the code for a specific function.

gene cloning—isolating a gene and making many copies of it by inserting it into cells and allowing it to multiply

gene expression—the manifestation of the genetic material of an organism as specific traits.

gene library—a collection of DNA fragments from a cell type or organism which have been introduced into viruses or plasmids and which, taken together, represent the total DNA of that cell type or organism.

gene mapping—determining the relative locations of different genes on chromosomes.

gene splicing—joining pieces of DNA from different sources by using recombinant DNA technology

gene therapy—the introduction of a normal, functioning gene into a cell in which that gene is defective

genetic code—the language in which DNA's instructions are written It consists of triplets of nucleotides (codons), with each triplet corresponding to one amino acid in a protein structure or to a signal to start or stop protein production.

genetic engineering—altering genetic material to study genetic processes and potentially to correct genetic de-

fects. See recombinant DNA technology.

genetics—the scientific study of heredity—of how particular qualities or traits are transmitted from parents to offspring.

genome—the total genetic endowment packaged in the chromosomes. The normal human genome consists of 46 chromosomes.

genotype—the full set of genes carried by an individual, including alleles that are not expressed

germ cell—a sex cell (sperm or egg). It differs from other cells in that it contains only half the usual number of chromosomes. Male and female germ cells fuse during fertilization.

haploid cell—a cell with half the usual number of chromosomes, such as a sperm or egg cell.

hemoglobin—the oxygen-carrying protein found in red blood cells.

hormone—a "messenger" molecule of the body that helps coordinate the actions of various tissues, it is made in one part of the body and transported, via the bloodstream, to other parts, where it has a specific effect on cells.

Huntington's disease—a disease that generally appears in adulthood, producing progressive mental and physical deterioration, it is caused by a dominant gene.

hybrid cells—fused cells, usually of different organisms, which contain chromosomes from each organism.

hybridization—the placement of complementary single strands of nucleic acids together so that they will stick and

form a double strand. The technique of hybridization is used in conjunction with probes to detect the presence or absence of specific complementary nucleic acid sequences.

intron—a DNA sequence that interrupts the sequences coding for a gene product (exons). After information from the genes is transcribed into new strands of RNA, the introns are cut out of the RNA. The function of introns is still being explored.

linkage—the relationship between two genes, or between an identifiable trait and a genetic disorder. Genes that are located relatively close to each other on the same chromosome are said to be linked and generally are inherited together.

lipoproteins—compounds consisting of lipids (fatty substances such as cholesterol) and proteins.

lymphoblast—an immature lymphocyte that is immortalized when grown in cell culture.

lymphocyte—a white blood cell which is part of the immune system.

marker—a detectable genetic variant, such as one of the ABO blood types Some markers are found only among the victims of certain diseases and can be used to determine the presence of these diseases.

medical genetics—the study of the causes, symptoms, treatment, and prevention of genetic disorders.

messenger RNA—the ribonucleic acid molecule that transmits the genetic information from the nucleus to the cytoplasm, where it directs protein synthesis.

molecular genetics—the study of genetic mechanisms at the level of the molecules DNA and RNA and their components.

mutation—a change in the number, arrangement, or molecular sequence of a gene.

mycoplasma—tiny organisms, smaller than bacteria but larger than viruses.

nucleic acids—DNA and RNA, the molecules that carry genetic information.

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 cell containing the genetic material.

oligonucleotide—short string of nucleotides.

oncogenes—genes that may play a key role in the development of cancer.

phenotype—the entire expressed physical, biochemical, and physiological constitution of an individual, resulting from the interaction of the genetic endowment with the environment.

plasmid—a small, self-replicating molecule of DNA that is separate from the main chromosome in bacteria, yeast, and some plants. Because plasmids are easily moved from one cell to another or to the test tube, scientists often insert foreign DNA into them and use them to transfer this DNA into other cells in recombinant DNA experiments.

point mutation—a change in a single base pair in an organism' DNA.

polymorphism—an inherited variation, such as the ABO blood groups

probe—see DNA probe.

protein—a molecule composed of amino acids arranged in a special order determined by the genetic code, proteins are required for life processes.

receptor—see cell-surface receptor.

recessive—refers to a characteristic that is apparent only when genes for it are inherited from both parents.

recessive gene—a gene whose product is detectable only when its allele on the paired chromosome is the same.

recognition site—see restriction enzyme recognition site.

recombinant DNA—the hybrid DNA produced in the laboratory by joining piece of DNA from different sources.

recombinant DNA technology—techniques for cutting apart and splicing together pieces of DNA from different sources.

replication —formation of an exact copy. DNA replication occurs when each strand of DNA acts as a template for a new, complementary strand formed according to base-pairing rules.

restriction enzyme—an enzyme that recognizes a specific base sequence (usually four to six base pairs in length) in a double-stranded DNA molecule and cuts both strands of the DNA molecule at every place where this sequence appears.

restriction enzyme recognition site—the DNA site where a specific restriction enzyme cuts the DNA molecule.

restriction fragment length polymorphism (RFLP)—the presence of two or more variants in the size of DNA fragments from a specific region of DNA that has been exposed to a particular restriction enzyme. These fragments differ in length because of an inherited variation in a restriction enzyme recognition site.

restriction fragments—fragments of DNA produced by cuts made with restriction enzymes.

retrovirus—an RNA-containing virus that replicates by means of an enzyme (reverse transcriptase) which, upon infection of a host cell, makes a strand of DNA that complements the infecting virus strand, the double-stranded DNA produced in this way then becomes part of the host cell's chromosomal DNA and reproduces along with it, eventually also producing an RNA strand identical to the original virus.

RNA (ribonucleic acid)—a single-stranded nucleic acid which contains the sugar ribose. There are many forms of RNA, including messenger RNA, transfer RNA, and ribosomal RNA (all involved in protein synthesis), as well as several small RNA's whose functions are unclear.

sex cell—a reproductive or germ cell (egg or sperm).

sex chromosome—one of the chromosomes (X or Y) involved in sex determination. Normal human females have two X chromosomes in each cell, while normal males have one X and one Y.

sickle-cell disease—a potentially lethal recessive blood disorder caused by

the mutation of a single nucleotide in the gene for beta-globin, one of the protein chains that make up adult hemoglobin.

somatic cell—a non-reproductive cell. One of the cells composing all the parts of the body (e.g., tissues, organs) other than the germ cells

Southern blotting—a procedure for transferring DNA fragments from an agarose gel to a filter paper without changing their relative positions.

thalassemias—recessively inherited blood disorders caused by various mutations which reduce the synthesis of one of the protein chains that make up hemoglobin. The victims of severe thalassemia require frequent blood transfusions and often die in their teens or early twenties.

tissue culture—see *cell culture*.

transcription—the transfer of information from various parts of the DNA molecule to new strands of messenger

RNA, which then carry this information from the nucleus to the cytoplasm.

translation—the process of turning instructions from messenger RNA into protein in the cytoplasm.

translocation—an error occurring during chromosomal replication, whereby a fragment of one chromosome becomes attached to another chromosome.

X chromosome—a sex chromosome. Normal human females have two X chromosomes in each cell, while normal males have one X and one Y chromosome in each cell.

X-linked—refers to any gene found on the X chromosome or traits determined by such genes. Refers also to the specific mode of inheritance of such genes.

Y chromosome—a sex chromosome. Normal human males carry one X chromosome and one Y chromosome in each cell.

64

PHOTOGRAPH CREDITS

Crawford, J.T., Division of Research Services, NIH, page 4.

Tjio, J.-H., National Institute of Arthritis, Diabetes, and Digestive and Kidney Diseases, NIH, page 12.

Yunis, J.J., University of Minnesota, Minneapolis, MN, page 12.

Murayama, M., National Institute of Arthritis, Diabetes, and Digestive and Kidney Diseases, NIH, page 23.

Yunis, J.J., *Human Pathology* 12:494, 1981. W.B. Saunders Company, Philadelphia, page 38.

Institute for Medical Research, Camden, NJ, pages 47, 49.

ILLUSTRATION CREDITS

Watson, J.D., Tooze, J., and Kurtz, D T , Recombinant DNA: A Short Course W H Freeman and Company, New York, 1983, pages 16, 17, 19, 26, 29.

Splicing Life, President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research, 1982, pages 18, 20.

Gusella, J.F. and Housman, D., in Gershon, E.S., Matthysse, S., Breakefield, X.O., and Ciaranella, R.D. (eds.), Genetic Research Strategies for Psychobiology and Psychiatry. The Boxwood Press, Pacific Grove, CA, 1981, page 25.

Fiddes, J.C., California Biotechnology Inc., Palo Alto, CA, page 26.

Gusella, J.F., et al., Nature 306·235, 1983. Macmillan Journals Limited, London, page 32

Ruddle, F.H. and Kucherlapati, R S., Scientific American 231:1, 1974. W.H. Freeman and Company, New York, page 37.

McKusick, V A , The Johns Hopkins Hospital, Baltimore, MD, pages 40, 41

Marx, J.L., Science 218.985, 1982 American Association for the Advancement of Science, Washington, DC, page 45.