

Biotechnology and Medicine of the Future

by Leroy Hood and Michael Hunkapiller

The 1980s appear to be the decade of biotechnology, and two of its aspects — recombinant DNA and microchemical instrumentation — will impinge greatly on medicine of the future.

THE 1980s APPEAR to be the decade of biotechnology. The striking new advances in biotechnology appear to offer virtually unlimited opportunities for attacking fundamental and applied problems in medicine, agriculture, energy, and chemical synthesis. We will discuss two aspects of biotechnology, recombinant DNA and microchemical instrumentation, and will consider how these advances impinge on medicine of the future.

DNA, GENES, AND DEVELOPMENT

The human organism starts its development as a single cell, the fertilized egg. This cell goes through a highly programmed and reproducible series of developmental stages. Initially the single cell doubles again and again to generate a ball of cells; then further division and cell migration lead to a hollow sphere, which elongates into a cylinder. The cylinder flattens and folds into three layers which in turn eventually give rise to the three fundamental cellular types from which the tissues and organs of the human organism are fashioned. Thus, from a single cell a human composed of 10^{14} cells with hundreds of specialized types develops. Two aspects of this developmental process raise intriguing questions that can be approached with the techniques of recombinant DNA. How do specialized cells turn on the particular genes they need for their differentiated functions? For example, how can red blood cells turn on the gene for making hemoglobin, the oxy-

gen-carrying molecule of human organisms? Second, what is the nature of the developmental program which directs every human organism to differentiate and mature in precisely the same way?

The information for expressing specialized proteins, such as hemoglobins, or controlling precisely the developmental pathway of the human organism is contained, in part, in the DNA language of its chromosomes. We can think of each human cell as having a library of encyclopedias with instructions on how to produce humans. The DNA language is made up of letters and words, just as the English language is. The DNA encyclopedias of each human cell contain 3 billion different letters of the DNA language. In DNA language the average word contains approximately 1000 letters. Hence human cells potentially have the information for 3 million different DNA words. Recombinant DNA technology permits scientists to isolate a single DNA word, place it in a test tube, and then study how this word contributes to the development of human organisms.

The 23 chromosomes contained in every human cell are the blueprint repositories containing the information for constructing human beings, and individual units of information or the words of the DNA language contained within the chromosomes are called genes (Figure 1). The gene or word is translated into the functional entities, called proteins, by a complex, computer-precise process denoted protein synthesis. The information contained in each gene is read from the so-called master computer tape, the genes of the human chromosomes, into a working computer tape called messenger RNA, which ultimately can be translated in special cellular machines into the proteins that give our body shape, form, and function (Figure 1).

The elegantly simple DNA molecule that makes up the informational backbone of the chromosomes has been beautifully fashioned for its primary function of information storage and transfer. A DNA molecule is made up of two oppositely oriented chains of nucleic acids, which are composed of four different letters or nucleotides — A, T, G, and C. (Figure 2). The two chains of the DNA molecule exhibit a molecular complementarity that confers upon DNA all of its fundamental properties. The letter A in one strand always pairs with the letter T in the second strand, as does the letter G always pair with the letter C (Figure 2). This letter pairing means that one DNA strand in a DNA molecule is the molecular complement of its partner, and, if separated, the two strands can readily find one another even in a complex mixture of DNA strands and

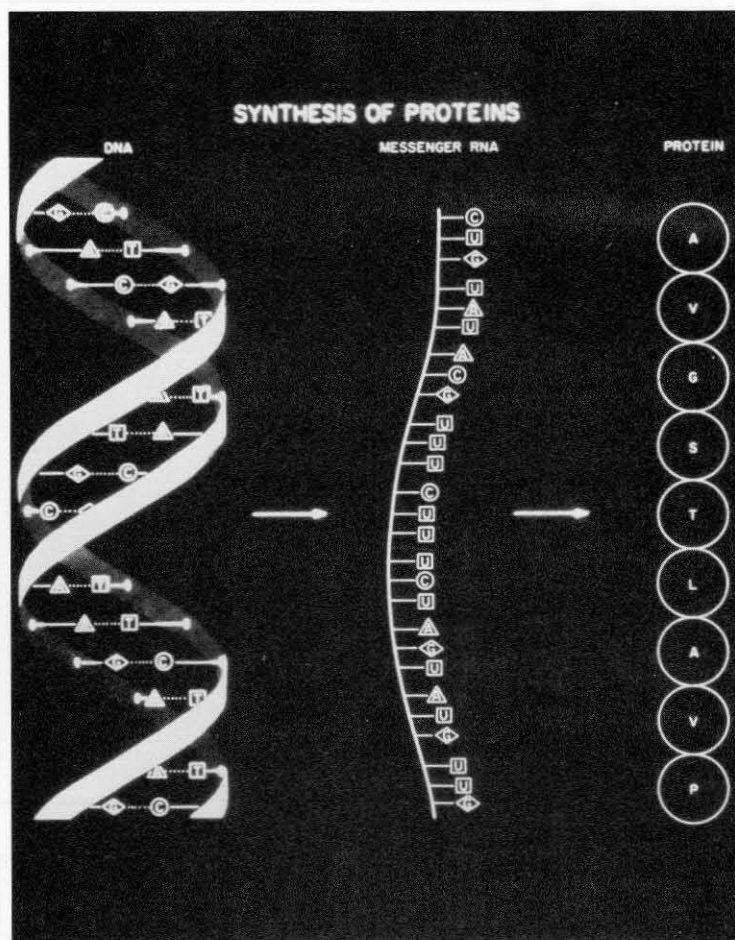


Figure 1 (above) illustrates the molecular relationship between DNA, RNA, and protein. Figure 2 (below) is a depiction of the Structure of DNA emphasizing the complementary base pairing between the two strands.

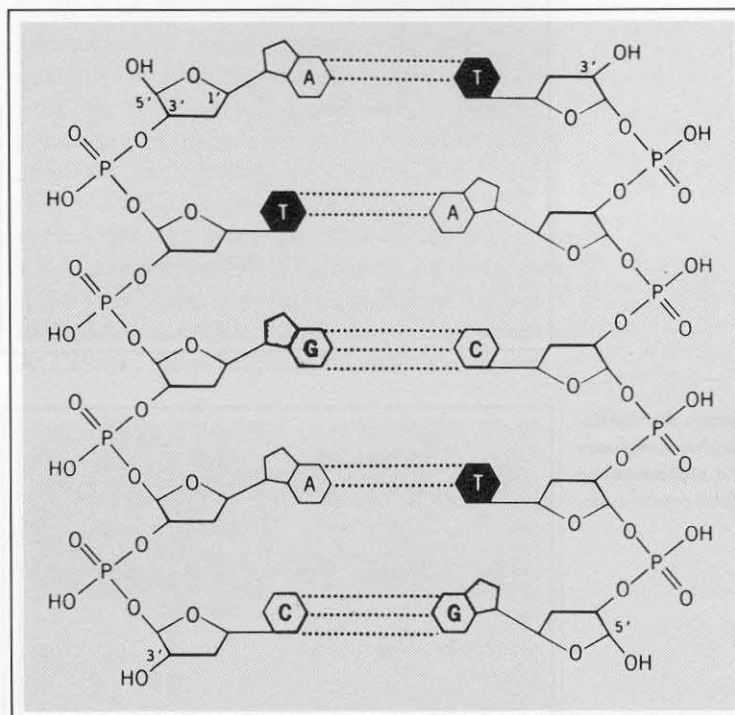
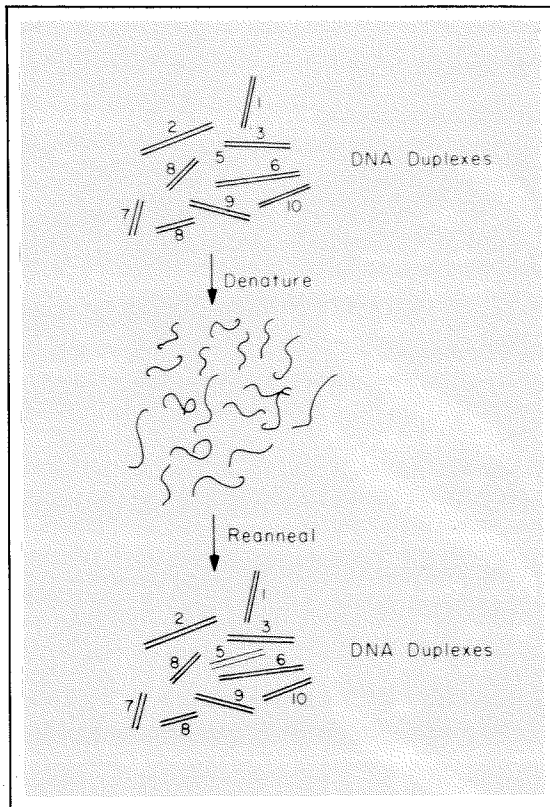
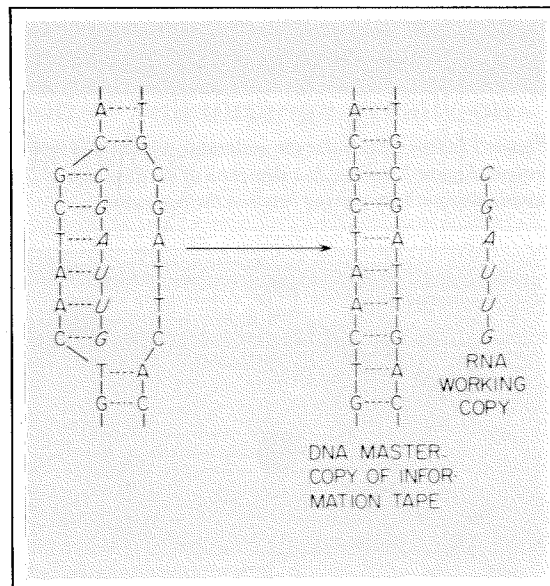


Figure 3. The denatured, separated strands of DNA molecules base pair specifically to their complements and reform the double-stranded structure when the denaturing conditions are removed.



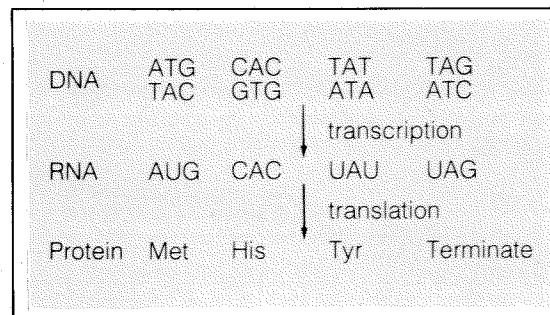
reassociate to form the duplex DNA molecule (Figure 3). The implications of this molecular complementarity give us fundamental insight into how DNA carries out its two most important functions — the replication of itself and protein synthesis. In order to replicate and create the DNA molecules for new cells, one DNA strand can serve as a template against which the second DNA strand is constructed by assembling together the letters in the new chain via the rules of molecular complementarity. Thus the two strands of a DNA molecule can separate and reassemble their complementary partners to generate two DNA molecules from one. Likewise, information from that functional unit of the DNA molecules, the gene, is transferred to the RNA computer tape, again by the process of molecular complementarity in which the RNA molecule is synthesized using one DNA strand as a template (Figure 4). This is possible because the RNA language also contains four different letters each of which uniquely pairs with a single DNA letter. The fundamental basis for recombinant DNA techniques resides in molecular complementarity because separated DNA chains can find and interact with their complementary partners.

Figure 4 shows the transcription of DNA into RNA, in which the molecule of RNA is formed with a base sequence that is complementary to one of the two original DNA strands.



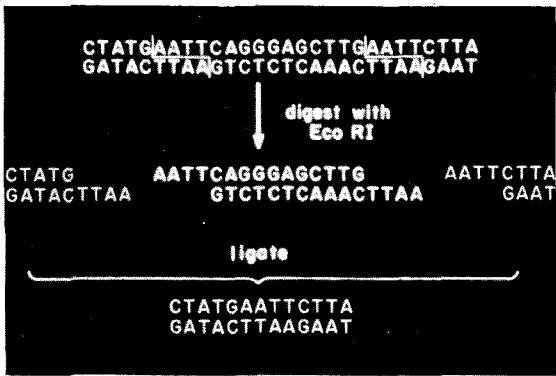
The molecular code by which information is transferred from the DNA master computer tape through the RNA working computer tape and ultimately to proteins is a fascinating process (Figure 5). The DNA molecule is double stranded and its information is stored via a language containing four letters. The RNA molecule is single stranded and also uses, as noted above, a language with four letters. Proteins, the functional units expressing gene information, are linear chains composed of 20 different letters called amino acids. Each amino acid is specified by a trio of letters in DNA and RNA language. Thus, triplets of DNA letters can be translated into single protein letters, and, conversely, protein letters can be reverse translated into DNA triplets. This ability to translate the genetic code dictionary in both the forward and reverse directions is fundamental to the use of the microchemical instrumentation we shall describe later.

Figure 5 illustrates the translation of nucleic acid sequence into amino acid sequence using the triplet genetic code.



RECOMBINANT DNA TECHNIQUES

The breakthrough in the development of recombinant DNA techniques came with the identification of enzymes that could cleave DNA molecules at precise points (restriction endonucleases) and enzymes that permit the joining together of distinct DNA fragments (ligases). The typical human chromosome contains approximately 10^8 nucleotide pairs. Restriction enzymes may cleave this large DNA molecule into regular gene-sized



fragments (Figure 6). The restriction endonucleases cleave DNA molecules at precise points determined by nucleotide recognition sites of four or six letters. For example, one restriction endonuclease cleaves all DNA molecules each time the base sequence GAATTC appears. Often the DNA strands are cleaved in an offset manner so that staggered single-stranded ends remain that exhibit a molecular complementarity for one another that allows any fragments produced by a particular restriction enzyme to be joined back together (Figure 6). Restriction endonucleases arose in bacteria as a defensive mechanism for destroying foreign DNAs. The bacterial DNAs are protected from cleavage by modifying nucleotides in the restriction enzyme recognition signal so that their own particular restriction enzymes cannot cleave their own DNA. This fortunate accident of nature has given molecular biologists hundreds of different restriction endonucleases to facilitate the cloning of human genes.

A second category of enzymes called ligases permits DNA fragments to be joined together, often using the molecular complementarity of their staggered ends (Figure 6). The ability to join together any two fragments of DNA permits human gene fragments to be spliced into small DNA circles called plasmids that can be grown up in large quantities in bacteria such as *E. coli* (Figure 7). The approach is to genetically modify the plasmid so that a particular restriction enzyme cleaves it at a single site. Then fragments of human genes cleaved by the same restriction enzyme can be inserted into plasmids, the human and plasmid DNAs can be ligated together, and the recombinant plasmids can be used to infect the bacteria *E. coli*. The recombinant plasmid is amplified to make hundreds of copies of the human gene in each bacterium (Figure 7). Thus large quantities of the particular human gene can be grown up in bacteria. The individual *E. coli* containing plasmids with different human genes can then be plated out on appropriate media to generate thousands of individual colonies, which

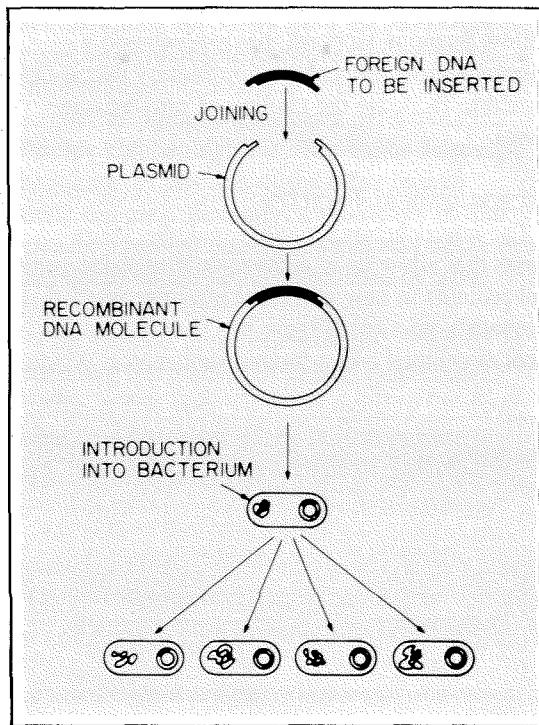


Figure 6 at the far left depicts the action of restriction enzymes and ligases. The restriction enzyme recognizes specific base sequences and clips the DNA strands at these points, leaving ends of uneven length on the original complementary strands. The DNA fragments can then base pair at these ends with any piece of DNA cut by the same enzyme, and hence possessing the appropriate complementary base sequences. The splicing action of the ligases can then be used to form an entirely new hybrid DNA sequence from the two fragments. Figure 7 (near left) is a schematic representation of the cloning of human gene fragments into plasmids and the plasmids into *E. coli*.

in turn can be screened to identify the bacterial colony containing the particular human gene of interest.

RARE-MESSAGE GENES

In principle the cloning procedure described above will lead to 3 million different clones of *E. coli*, each containing a different human gene. The critical question is how, from among the 3 million bacteria with their human genes, the single bacterium that contains the gene of interest — for example, an interferon gene — can be selected. If the gene in question synthesizes large quantities of the messenger RNA, such as do the genes for hemoglobin and antibodies, molecular probes can be made from the RNAs which readily allow the corresponding genes to be cloned (Figure 8). Conversely, if the gene of interest synthesizes very small quantities of messenger RNA, as is the case with the interferon gene, then it is termed a rare-message gene (Figure 8), and its isolation is a difficult and complicated process. At Caltech, the two authors in collaboration with

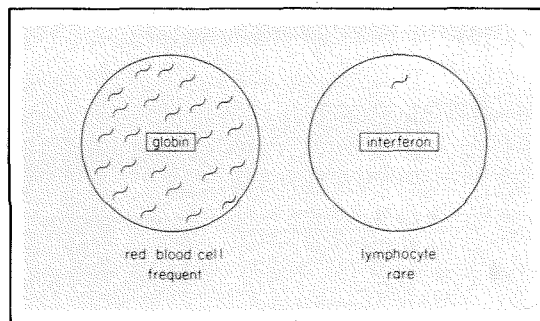


Figure 8. Examples of rarely and frequently expressed messenger RNAs.

Professor of Biology William Dreyer and Staff Fellow Suzanna Horvath, have developed a microchemical instrumentation facility that is designed to facilitate the cloning of rare-message genes.

MICROCHEMICAL FACILITY

The microchemical facility includes instruments that allow us to analyze or synthesize genes and proteins in new and more effective ways than previously possible (Figure 9). The first instrument developed was a protein sequenator, which

Figure 9. A capsule description of the instruments that comprise the Caltech microchemical facility and the primary function of each.

1. Protein Microsequenator	- determines order of amino acids in proteins
2. Peptide Synthesizer	- synthesizes small peptides and fragments of proteins
3. DNA Synthesizer	- synthesizes oligonucleotides (genes)
4. DNA Sequenator	- determines order of nucleotides in genes
5. Gene Analyzer	- fingerprints genes

Figure 10 is a diagram of the separation of complex protein mixtures by two-dimensional polyacrylamide gel electrophoresis. A: The proteins are separated in the first dimension on the basis of their molecular size. B: They are then separated in a second dimension on the basis of electronic charge on their surface. This technique allows separation of several thousand proteins in a single experiment.

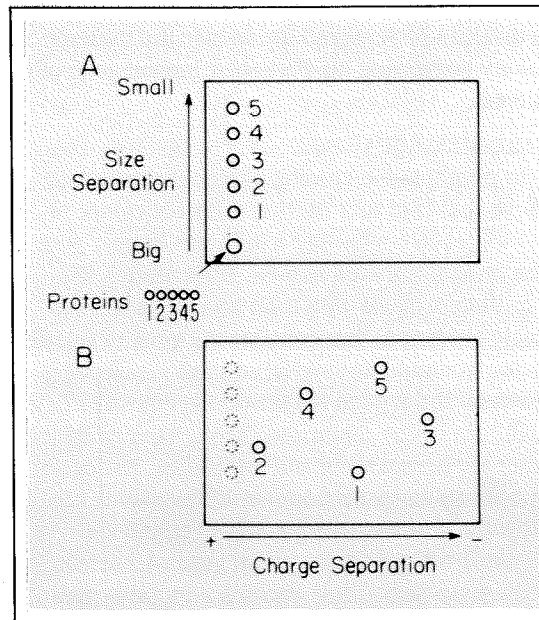
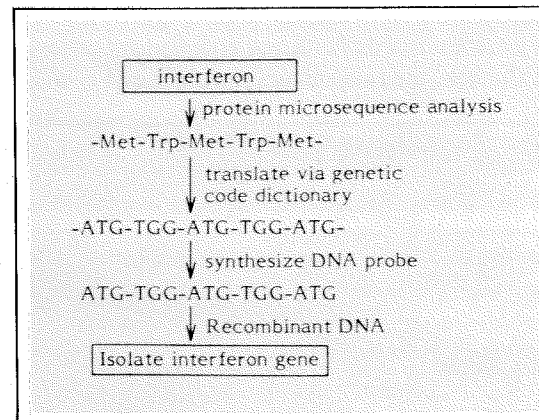


Figure 11 gives a strategy for cloning of rare-message genes based on synthesis of nucleic acid probes using protein structure obtained by micro-sequencing.



determines the linear order of amino acids in a protein by acting as a pair of molecular scissors to remove amino acid subunits from the proteins one at a time. A second analytic instrument must be used to identify each amino acid. The main virtue of the Caltech microsequenator is that it is approximately 100 times more sensitive than other sequenators. Indeed, the Caltech sequenator is so effective that it could be ten times more sensitive if we had an analytic instrument of sufficient sensitivity to characterize the individual amino acids at these low levels.

To circumvent this difficulty, we have, in collaboration with Professor Dreyer, spent the last six years collaborating with engineers at the Jet Propulsion Laboratory to design an ultrasensitive mass spectrometer that is capable of analyzing amino acid residues at levels far below those previously possible. This instrument is nearing completion and is sufficiently sensitive that we can use the highest resolution analytic separation tool in protein biochemistry, two-dimensional gel electrophoresis, as a tool for preparing sufficient quantities of protein for sequence analysis with the microsequenator (Figure 10).

The importance of obtaining this very high level of sensitivity for the analysis of proteins can be understood in the context of the third instrument, whose development we have just finished. The DNA synthesizer, or so-called gene machine, has the capacity to synthesize fragments of genes relatively rapidly. It is interesting to note that although the chemistries for protein-sequence analysis and DNA synthesis are very different, the instrumentation and programming requirements for these two chemistries are remarkably similar, and indeed, these two machines — the protein sequenator and DNA synthesizer — are strikingly similar to one another.

These three machines permit us to approach the routine and rapid cloning of rare-message genes in the following manner (Figure 11). Often small amounts of a rare-message gene product such as human interferon can readily be obtained for protein sequence analysis. After a portion of the protein has been analyzed by the protein sequenator, this sequence is reverse translated into DNA language. A small DNA fragment 12-17 nucleotides in length is then synthesized with the DNA synthesizer and used as a molecular probe to clone the rare-message gene. Thus, a task that took in the past two or even three years to achieve, in the future can be accomplished in less than one month. This use of the microchemical facility for the cloning of rare-message genes accordingly opens up the potential for the cloning of an enormous variety of interesting and biomedically rele-

vant human genes. Indeed, we hope to be able to clone in the future virtually any gene whose protein product can be visualized on a two-dimensional gel (Figure 10). Thus it will be possible to clone thousands of human genes that were heretofore inaccessible to the techniques of recombinant DNA.

We also are at work designing a machine that we hope will eventually allow us to fingerprint human genes. If we are successful in this endeavor, gene fingerprinting will open up some very exciting new areas for human medical genetics. These two interrelated biotechnologies — recombinant DNA and microchemical instrumentation — will have a great impact on medicine of the future.

From "Prenatal Diagnosis of Genetic Diseases" by Theodore Friedmann. Copyright November 1971 by Scientific American, Inc. All rights reserved.

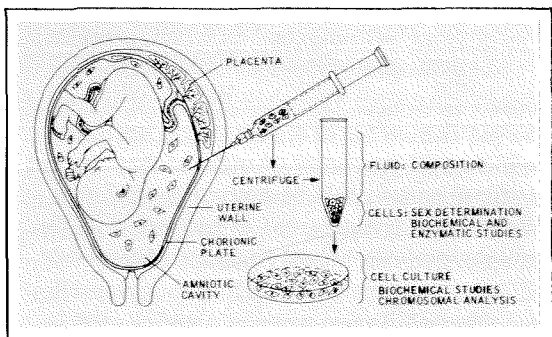


Figure 12 is a schematic representation of analysis of fetal chromosomes by amniocentesis.

HUMAN GENETIC DISEASES

Approximately 2500 different human genetic diseases have been described. Some of these can be diagnosed prenatally by a procedure called amniocentesis (Figure 12). In this procedure a needle is placed in the amniotic sac of the unborn fetus and used to withdraw fluid and cells that are derived primarily from the lungs of the fetus. The fluid and cells can be examined to determine whether they have biochemical abnormalities indicative of any one of approximately 200 human genetic diseases. Mongolism, for example, can readily be detected by the presence of an extra human chromosome 21 in the cells of the fetus. Tay-Sachs disease can be readily diagnosed because it has a protein abnormality that is present in the amniotic fluid. The techniques of recombinant DNA, at least in theory, offer us the potential to diagnose directly virtually every one of the 2500 different human genetic diseases. The approach would be to fingerprint human genes so as to distinguish normal and abnormal genes. The fingerprinting would compare the sizes of DNA fragments from the gene of interest produced by restriction enzymes (Figure 13). DNA from nor-

mal individuals and DNA from the fetal cells derived by amniocentesis could thus be fingerprinted to unambiguously determine whether, for example, the unborn fetus has a defective gene for sickle cell hemoglobin, Tay-Sachs disease, mongolism, or any one of the other human genetic diseases. Recently, a laboratory in San Francisco reported that the normal and sickle cell genes could be readily distinguished by gene fingerprinting.

If we succeed in the microchemical facility in automating the process of human gene fingerprinting, it will open up the possibility for diagnosing *in utero* many of the human genetic diseases. This type of analysis will be particularly important with diseases such as Huntington's chorea, which do not manifest themselves until the individual is 30 or 40 years old. At this time the afflicted person undergoes rapid deterioration in a particular region of his brain. There is no therapy for these patients. Since this disease can be caused by just a single copy of the defective gene, the tragedy is that the carriers of this gene defect often have children before they manifest the disease in their middle adult life. Their children have a 50 percent probability of acquiring the disease. It is psychologically devastating for those children to know that they have to live 30 or 40 years before learning whether they have this inevitably fatal genetic disease. Thus the fingerprinting of human genes will afford a revolutionary and unique opportunity for diagnosing human genetic disease *in utero* in time for a therapeutic abortion or other appropriate measures.

CANCER GENES

One of the most fundamental advances in cancer biology in the last ten years has been the isolation of the so-called cancer genes using recombinant DNA technology. These are genes that have the ability, when introduced into cells, to convert a normal cell into a cancerous cell

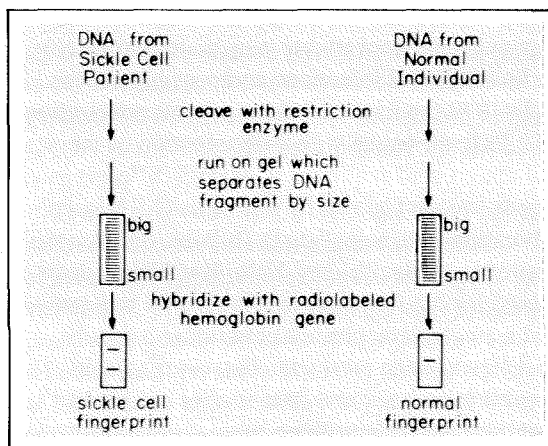
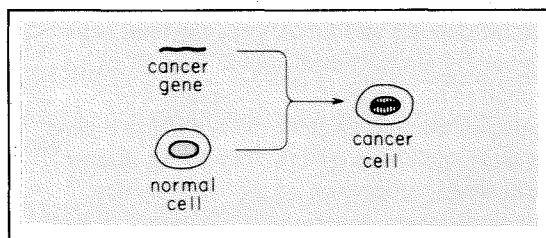


Figure 13 illustrates analysis of human genes by fingerprinting, using one-dimensional polyacrylamide gel electrophoresis and specific DNA probes.

Figure 14. Some types of foreign genes are oncogenic. When they become integrated into a host cell's chromosome, they may help trigger the cancerous state in that cell.



(Figure 14). The isolation of cancer genes presents revolutionary opportunities for understanding the molecular basis for cancer in that these genes can be studied individually to determine how they cause cancer. One of the striking features about cancer genes is that every human has in his chromosomes many normal genes that are remarkably similar to the cancer genes. Indeed, one scientist at MIT has been able to demonstrate that the cancer gene associated with human bladder tumors differs from its normal counterpart by just a single nucleotide in the entire gene. This is a revolutionary observation and raises many fascinating questions about what causes cancer genes to become different from their normal counterparts and what kinds of functions they carry out to actually cause the tumor process. Can environmental agents, for example, convert normal genes into cancer genes? These are questions that the techniques of modern recombinant DNA technology will readily allow us to begin to answer in the future.

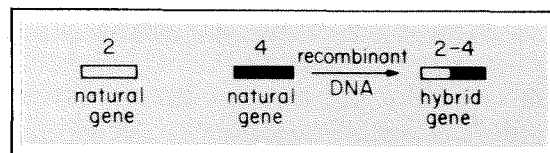
INTERFERON

The cloning of the gene for interferon was a scientific advance that fascinated the general public, whose interest in interferon arose from its potential use as an antitumor agent. The history of interferon is interesting though remarkably brief. Interferon was first identified in 1957, and until recently it was possible to isolate this protein only in extremely small quantities. The entire world's output in 1979 was 0.1 oz., and it cost approximately \$20 billion per pound. Even the small amounts of interferon available in the 1970s led to a great deal of clinical excitement because this molecule appeared to play an important role in fighting viral infections and possibly cancer.

In 1979 the first real chemical characterization of interferon began with determination of the partial amino acid sequence of several different interferons here at Caltech by one of us (MH) in collaboration with scientists at DuPont, the National Institutes of Health, and Yale. It was then determined for the first time that interferon was not just a single gene product, but rather a family of different gene products. These amino acid sequence analyses and later studies established that there are at least three basic classes of inter-

ferons, denoted alpha interferon, beta interferon, and gamma interferon. In 1980 the alpha and beta interferon genes were cloned. Interestingly, alpha interferon turned out to be encoded by a family of at least 15 different interferon genes. These alpha interferon genes were very similar to one another, and provocative questions were raised as to whether these different genes carried out distinct functions.

The existence of multiple copies of the alpha interferon gene family then permitted molecular biologists to perform a spectacular series of experiments that led to the synthesis of hybrid interferon genes (below). For example, the front half of one interferon gene could be fused to the back half of a second interferon gene to generate a hybrid gene. The interesting observation was that



the biological properties of some of these hybrid interferon genes could be profoundly different from those of either parent. It was reported, for example, that one hybrid alpha interferon gene had enormously enhanced cell-killing properties over either of the two parental genes used to construct the hybrid. Thus the existence of multigene families for interferons presents the molecular biologist with two kinds of opportunities. First, he can study the normal genes of this family to define their individual functions; perhaps some will be effective antiviral agents and others effective anticancer agents. Second, hybrid genes may be constructed to optimize certain particularly favorable properties in a way that could never have been done in natural genes. The early clinical studies of the various interferon genes suggest that they will indeed be effective agents for dealing with viral infections. Their ability to serve as therapeutic agents for cancer appears less certain.

REPLACEMENT THERAPY

The techniques of recombinant DNA have already led to the isolation of human genes that encode insulin, growth hormone, interferons, and a variety of other biologically interesting molecules. These genes have been placed in bacteria or yeast in such a manner that large quantities of the corresponding human protein can be synthesized and isolated. Indeed, insulin and interferon derived from these recombinant DNA techniques will soon be commercially available. Thus the recombinant DNA techniques are going to play a

very important role in the isolation of genes that can be used to synthesize products for use in replacement therapy for individuals who have particular genetic deficiencies. As an example, synthetic human insulin will be useful in treating certain diabetics who cannot tolerate the animal insulin given successfully to most diabetics.

GENETIC SURGERY

With the ability of scientists to clone human genes, the natural question is, Can these genes be transferred directly into human beings who are deficient in these gene functions? For example, can a diabetic patient whose disease arises from a deficiency in the insulin gene be cured by having genetic surgery place good insulin genes in the relevant cells that synthesize insulin? Genetic surgery appears to be a long time in the future. Scientists do not yet know how to position a gene at a precise and appropriate position on the human chromosome. Moreover, they do not know how to place genes in foreign cells such that they can respond appropriately to the normal physiologic signals of their environment. The production of insulin, for example, should be increased when the levels of blood sugar rise. Indeed, with many genetic diseases, the deficiency occurs only in one tissue or population of cells in the organism. It is not clear how to insert functional genes into such a particular cell population. Thus there are a variety of technical difficulties that limit any immediate use of recombinant DNA techniques to carry out genetic engineering on human beings. However, the recent spectacular transfer of the rat growth hormone gene into eight mouse embryos resulted in the production of seven giant mice, presumably generated by the action of the rat growth hormone on the growing mouse embryos. Evidently, it will be possible to genetically engineer humans in the future, but genetic surgery raises interesting social and ethical questions.

BIOTECHNOLOGY AND MAN

The two biotechnologies that we have discussed here, recombinant DNA and microchemical instrumentation, will not only have a profound impact on the area of biomedical research, but as well on agriculture, energy, the synthesis of carbon-based chemicals, and perhaps even mining. In the long run, the greatest impact of genetic engineering probably will be in agriculture where one will acquire the ability to design plants that fix their own nitrogen, are incredibly resistant to drought and other harsh environmental conditions, are resistant to microorganisms and parasites, and may have enormously enhanced potential for protein and carbohydrate synthesis. Clear-

ly, biotechnology will have a profound impact on modern society that will continue into the indefinite future.

SOCIAL AND ETHICAL QUESTIONS

The potentialities of these biotechnologies have raised interesting social and ethical questions. What should be the criteria for determining whether to abort an unborn fetus with genetic defects? Clearly most individuals would agree that a fetus with Tay-Sachs disease, a disorder that leads to rapid degeneration of the brain in a newborn child, is a candidate for therapeutic abortion. In contrast, using amniocentesis to determine whether an unborn child is a boy or a girl and then employing a therapeutic abortion to obtain a child with the sex of your choice would be universally condemned. Most genetic diseases or other indications for therapeutic abortions lie between these two extremes, and clearly, society must struggle with these difficult decisions.

Man is now coming to a position where the engineer may engineer himself. Who should decide when and how human engineering can be carried out? What are appropriate experimental models for determining whether the techniques of human engineering are appropriately safe for human beings?

Societal questions are perplexing as well. The interrelationship between the genetic engineering companies and academia poses some serious questions regarding these relationships. On the one hand, the industrial interactions offer potentially enormous resources for academics to carry out their scientific programs. This is not a trivial point in these times when government support for science is waning significantly. At the same time, serious questions must be raised about how much time a professor may spend on outside commercial activities, whether he will employ his laboratory and students in the pursuit of applied research directly for industry, and perhaps most difficult of all, whether he will be forced into a secrecy which is dangerous for the free-exchange style of science that we have known in the past. It is our conviction that the opportunities and dangers can be balanced against one another so that academia and industry each provide for the other something it is missing.

The recombinant DNA techniques and microchemical instrumentation present man with unparalleled opportunities for modifying his environment, even himself. Perhaps the key question that is facing all of us at this juncture is whether man has the wisdom to take advantage of all of these opportunities and use them in an appropriate manner to their fullest potential. □