

THE GORDON WILSON LECTURE

THE NEW GENETICS

DANIEL NATHANS, M.D.

BALTIMORE

These are exciting times in the field of genetics. As a result of recent, revolutionary advances in methodology, chromosomes of every organism, from viruses to man, are open to detailed chemical analysis. Genes can be altered at specific, pre-determined sites, and new genes can be constructed by assembly of units from diverse sources or by direct chemical synthesis. New information, new applications, and even new concepts are being reported at a rapid rate. In this lecture I would like to describe several aspects of the "new genetics", in particular some of its historical roots, the general strategy of the new methodology, some of the results and applications already in hand, and finally, a few directions for the future that seem especially promising.

HISTORICAL ROOTS

Science provides a continuous accumulation of knowledge and an ever deeper understanding of nature. What I am calling "the new genetics" did not arise *de novo*, but springs from discoveries concerning the chemical basis of heredity made over the past four decades and more. The study of heredity took a decisive chemical turn in the early 1940's when Oswald T. Avery and his colleagues at the Rockefeller Institute for Medical Research discovered that deoxyribonucleic acid—DNA—was the transforming principle that converted non-encapsulated, avirulent pneumococci into encapsulated, virulent organisms (1). Avery—whose portrait is shown in Figure 1, taken from René Dubos' charming book "The Professor, The Institute and DNA" (2)—spent most of his career with the pneumococcus. His advice to a young visitor—Barry Wood, my predecessor at Hopkins—epitomizes Avery's attitude toward research: don't be distracted by the "surface nuggets", he told Wood, but "dig a deep hole in one place, hoping to hit a vein" (2). The vein that Avery himself hit was no less than the discovery—to the disbelief of his contemporaries—that DNA is the chemical substance of heredity. Thereafter, attention was focused on the role of DNA in inheritance. The key to understanding was in the three dimensional structure of DNA, the

From the Department of Microbiology, Johns Hopkins University School of Medicine, Baltimore, Maryland.

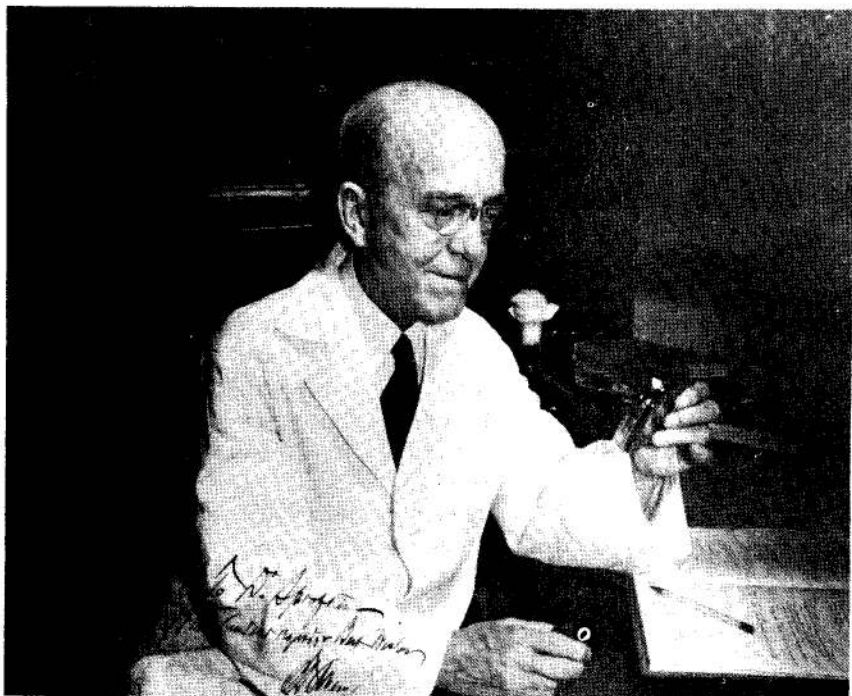


FIG. 1. Oswald T. Avery, in his laboratory at the Rockefeller Institute for Medical Research. (From reference 1, reprinted with permission of the Rockefeller University Archives.)

double-stranded, helical molecule proposed in 1953 by James Watson and Francis Crick in which the strands are held together by specific pairing of the four nucleotide bases that make up the DNA of all living organisms (3). As shown schematically in Figure 2, the nucleotide bases are adenine (A), guanine (G), thymine (T), and cytosine (C). Base-pairing occurs by hydrogen bonding between A and T or C and G. The DNA of all cells and viruses has this same chemical structure. What distinguishes the DNA of different organisms or different genes of a given organism is the order of nucleotide bases along the DNA chain. It is this order or nucleotide sequence that encodes the diverse genetic programs of living forms.

Contemporaneous with the discoveries I just described was a second root of the new genetics that began quite independently of chemistry, namely the formal genetics of bacterial viruses, initiated by Max Delbrück and Salvatore Luria (4). Viruses are the simplest known biological replicating units. They have a small number of genes clustered in a molecule of nucleic acid, analogous in many ways to a tiny piece of the

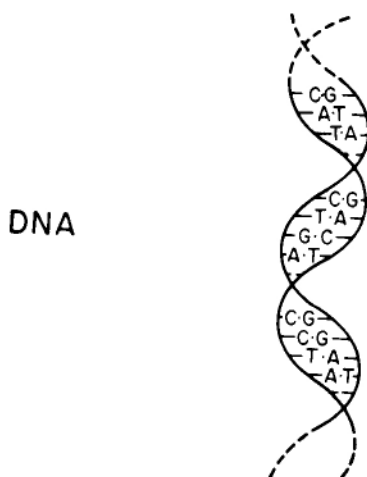


FIG. 2. The structure of DNA proposed by Watson and Crick (3).

DNA of cellular chromosomes. Because of their relative simplicity, viruses have served, and continue to serve, as model systems for studying genetic mechanisms found in living organisms generally. The combination of viral (and later, bacterial) genetics with the biochemistry of nucleic acids led to a golden age of discovery concerning the nature of genes and how they determine the phenotype of an organism. Among the major discoveries of this period were the protein code enciphered in the nucleotide sequence of DNA, the way this code is deciphered, and how gene activity is regulated in microorganisms. From these investigations a fundamental generalization emerged, backed by detailed biochemical evidence, as schematized in Figure 3. Genetic information encoded in the nucleotide sequence of DNA is expressed by transcription of the base sequence into messenger ribonucleic acid—mRNA—and then translated into the amino acids of the proteins. Three contiguous bases specify one amino acid of the protein, for example AUG specifies methionine (Met), GAC specifies aspartic acid (Asp), etc., as shown in Figure 3. Not shown in the figure are sequences in the DNA or RNA that serve as regulatory signals involved in controlling the rate of transcription or the rate of translation of particular genes.

CHEMICAL DISSECTION OF DNA—RESTRICTION ENDONUCLEASES

With rare exception, the investigations I have been describing were confined to microbes; the chromosomes of higher organisms were too complex for comparable analysis. Therefore genetic mechanisms in higher organisms such as mammals were largely inaccessible. The recent meth-

Expression of Genes

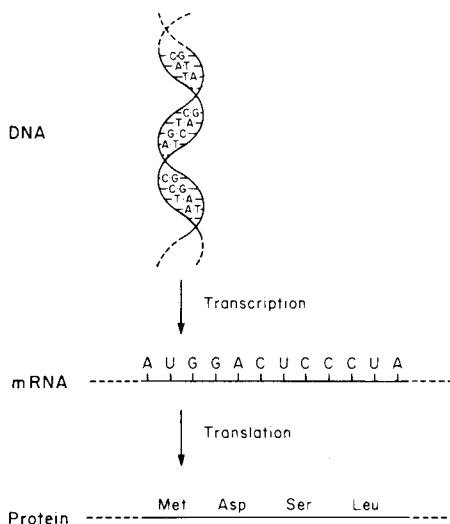


FIG. 3. The expression of the genetic code—from DNA to RNA to protein.

odological advances I referred to earlier have now overcome this limitation. In effect, they have made it possible to study very complex chromosomes, including those of man, segmentally, as if each segment were the chromosome of a virus, and to extend the analysis of any DNA molecule to the ultimate level, the sequence of nucleotides—the code words—along the DNA chain.

One of the seminal developments that made this extension possible was the discovery of enzymes that cut DNA at specific sites, the *restriction endonucleases*. In the early 1950's it was observed that bacteria have a primitive immune system, later identified at the molecular level by Werner Arber (5), namely, enzymes that recognize and break down foreign DNA. When that happens, we say the enzyme "restricts" the DNA. In 1968, my colleague at Hopkins, Hamilton Smith, discovered that restriction enzymes cut DNA at specific nucleotide sequences (6), at "code words" recognized by particular enzymes, as illustrated in Figure 4. Note that each enzyme listed recognizes a different sequence and that some cut the two DNA chains at directly opposite sites, whereas others produce staggered breaks in the DNA. The latter type of scission gives rise to so-called "sticky-ended" fragments that can rejoin by base-pairing of their single-stranded tails. At the present time about 150 restriction enzymes have been isolated from bacteria, recognizing in the aggregate about 50 different nucleotide sequences. Restriction endonucleases are

RESTRICTION ENDONUCLEASES

<u>ENZYME</u>	<u>SEQUENCE</u>
<u>Hin</u> d II	$ \begin{array}{ccccccc} & & & \downarrow & & & \\ G & T & P y & P u & A & C & \\ C & A & P u & P y & T & G & \\ & & & \downarrow & & & \end{array} $
<u>Hin</u> d III	$ \begin{array}{ccccccc} & \downarrow & & & & & \\ A & A & G & C & T & T & \\ T & T & C & G & A & A & \\ & & & & \downarrow & & \end{array} $
<u>Eco</u> R \pm	$ \begin{array}{ccccccc} & \downarrow & & & & & \\ G & A & A & T & T & C & \\ C & T & T & A & A & G & \\ & & & & \downarrow & & \end{array} $
<u>Hpa</u> II	$ \begin{array}{cccc} & \downarrow & & \\ C & C & G & G \\ G & G & C & C \\ & & \downarrow & \end{array} $

FIG. 4. Nucleotide sequences in DNA recognized by restriction endonucleases (6). The enzyme names are contractions of genus and species names of the bacteria from which the enzymes are prepared (e.g., *Hind*II is enzyme number 2 from *Hemophilus influenzae* strain d). The arrows indicate sites of cleavage of each DNA strand.

thus analogous to a set of trypsins and chymotrypsins, enzymes that cut protein molecules at specific amino acid residues. The value of such specific scalpels for DNA will become evident presently.

DISSECTION OF A MODEL CHROMOSOME

At the time Hamilton Smith was characterizing the first cleavage site-specific restriction enzyme, I was turning my attention to the analysis of a model mammalian chromosome, that of a small tumor virus, Simian Virus 40 or SV40 (7). As one of the simplest mammalian viruses, SV40 promised to be an attractive model system for investigating genetic mechanisms in the cells of higher organisms. As shown in Figure 5, the virus is a typical icosahedral particle within which is a mammalian type "minichromosome", consisting of DNA and histones clustered in aggregates called nucleosomes. When freed of histones the viral DNA is seen as a tiny ring of duplex DNA. SV40 DNA has about 5000 nucleotide pairs, enough DNA for only a few genes. However, in spite of its paucity of genetic information, SV40 can multiply by sequential expression of viral genes in the nucleus of simian or human cells, and it can cause transformation of cells to tumorigenicity in tissue culture or in a living animal. As diagrammed in Figure 6, when SV40 infects cells in which the virus can multiply, viral DNA finds its way to the nucleus, where an early gene

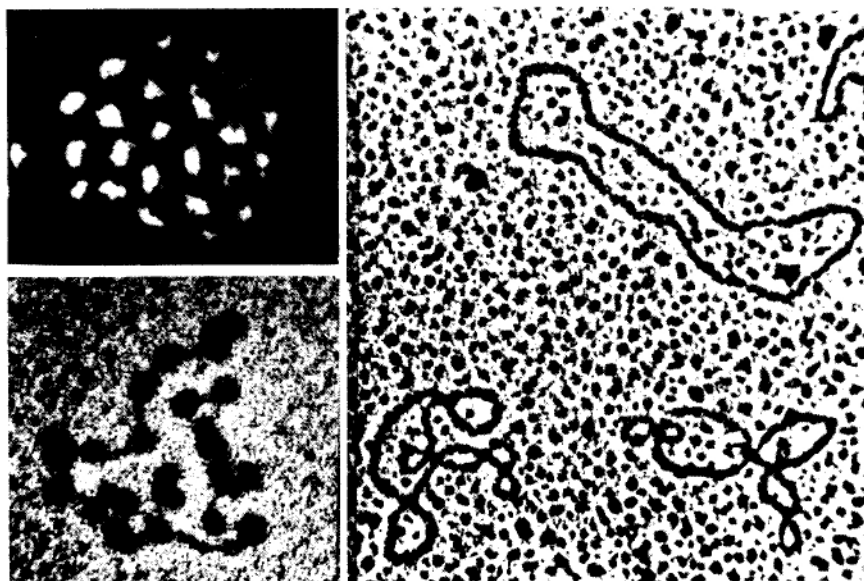


FIG. 5. Electron micrographs of Simian Virus 40 particles (upper left); its chromosome, consisting of DNA and nucleosomes (lower left); and naked viral DNA molecules (right).

Infection by SV 40

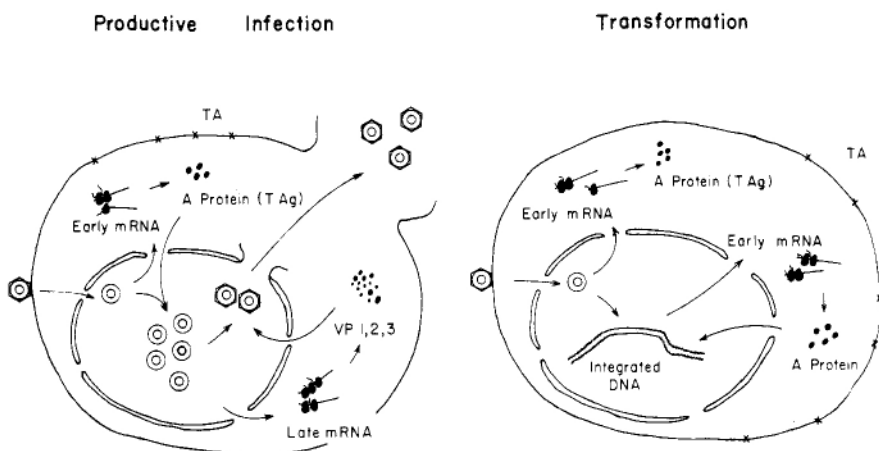


FIG. 6. Productive and transforming infection of cells by SV40. See text for an explanation of each figure.

encoding "T" (tumor) antigen is first expressed, viral DNA multiplies, and late genes encoding virion proteins (VP 1, 2, 3) are expressed. The final outcome is the formation of new virus particles and cell death. In contrast to such productive infection, when SV40 infects rodent cells, only the early gene is expressed, viral DNA does not replicate, but the viral chromosome becomes inserted into a cellular chromosome. Thereafter the viral DNA continues to express its T antigen gene and a cell surface antigen (TSTA), as a result of which the cell becomes tumorigenic. In sum, the interaction of this tiny viral genome with cells is a microcosm of regulatory phenomena related to virus multiplication and cell proliferation. My interest and that of other investigators was to use SV40 to analyze genetic regulation in normal and neoplastic cells. The first requirement was to determine the genetic organization—the molecular anatomy—of the viral chromosome. Restriction enzymes appeared to offer a direct approach to this end.

Our strategy was to cut SV40 DNA into specific fragments with a number of different restriction enzymes (Figure 7). The next step was to separate the fragments by electrophoresis in gels (Figure 8), and then to determine the size of each fragment and its relative position in the original viral DNA molecule. This provided a physical map of the genome, based on the positions of enzyme cleavage sites (Figure 9). We call this a *cleavage map* or *restriction map* of the genome. The availability of restriction fragments and cleavage maps opened the possibility of deter-

Specific Cleavage of SV40 DNA

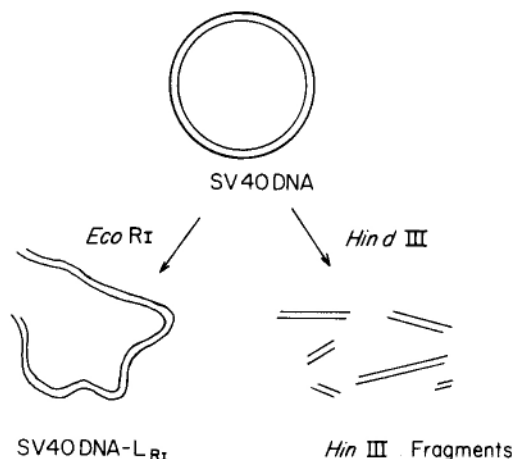


FIG. 7. Cleavage of SV40 DNA by a one-cut restriction enzyme (*EcoRI*) and by a multicut enzyme (*HindIII*).



FIG. 8. Electrophoresis of SV40 DNA fragments produced by digestion of viral DNA with *Hind*II plus *Hind*III restriction enzymes (8).

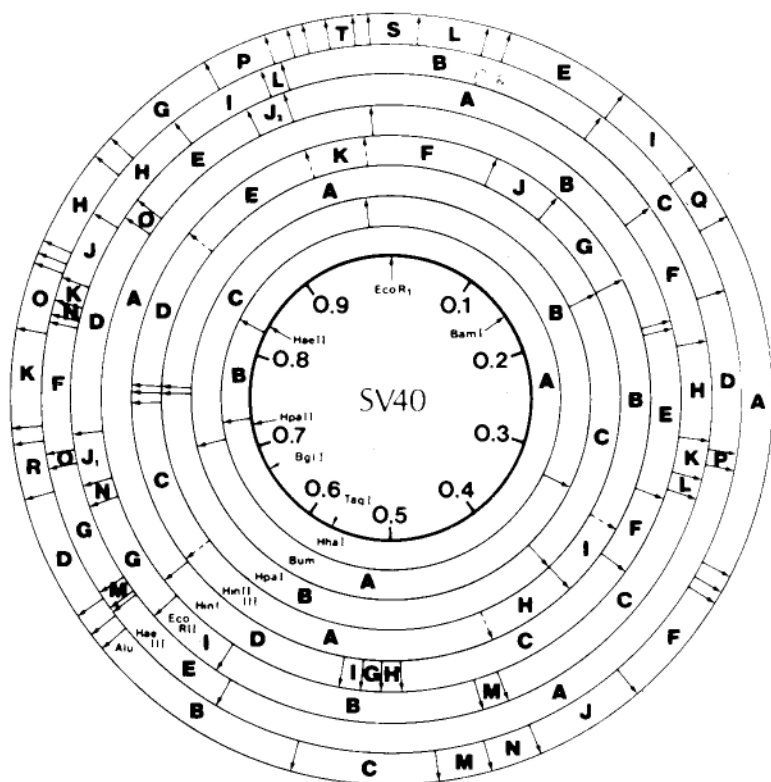


FIG. 9. A cleavage map of the SV40 genome (9). Inside the circle of DNA are map coordinates (0.1, 0.2, etc.) starting at the unique *EcoRI* site, and sites where other enzymes cut SV40 DNA once only (arrows). Each concentric ring shows the cleavage sites for a given multi-cut enzyme.

mining the entire *nucleotide sequence* of DNA molecules like that of SV40. Realization of this potentiality, however, was due to a second major methodological development, namely simple and rapid sequencing methods worked out in Fred Sanger's laboratory at Cambridge University (9) and Walter Gilbert's at Harvard (10). Gilbert's method is schematized in Figure 10a. A DNA fragment labeled at its ends with ^{32}P is separated into its component strands. Each strand is treated with reagents that result in random breakage at the site of one particular kind of nucleotide base (adenine-specific breakage is shown in Figure 10a), producing a set of ^{32}P -labeled chains of varying length, all beginning at the ^{32}P -containing end and ending at a position immediately preceding the attached base. Electrophoresis of this mixture allows one to determine the length of each such chain and hence the position of all A's (or G's or C's or T's)

Nucleotide Sequence Analysis (Maxam and Gilbert)

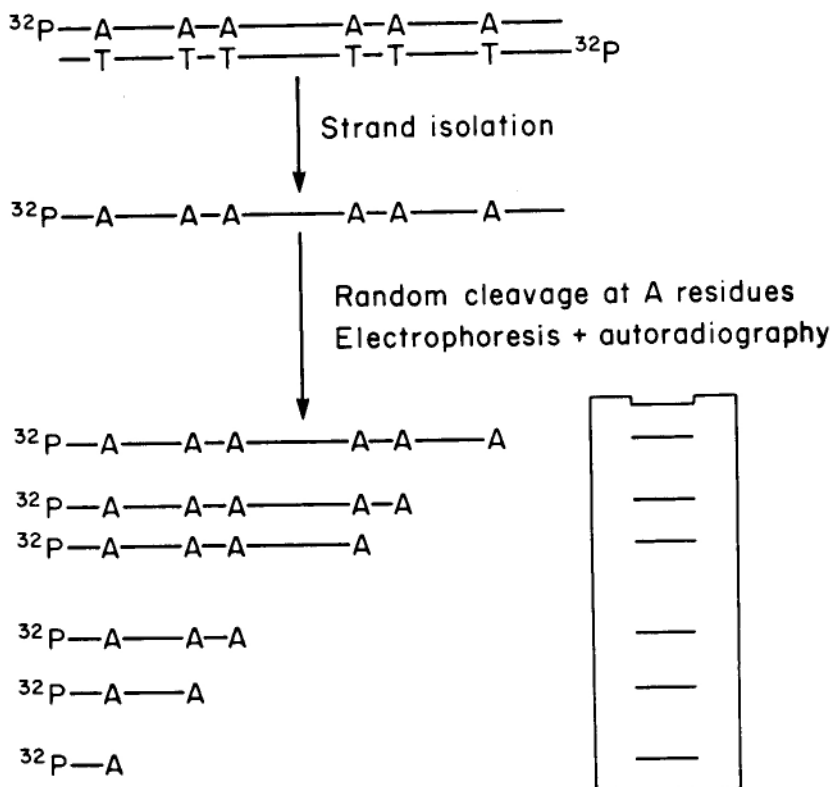


FIG. 10a. Outline of the Maxam-Gilbert method for determining the nucleotide sequence of a DNA fragment.

relative to the ^{32}P end. An actual analysis of this type is illustrated in Figure 10b, where it can be seen that the nucleotide base sequence can be read directly from the electropherogram. When applied to SV40 DNA by Sherman Weissman at Yale (11) and Walter Fiers at Ghent (12), the Gilbert-Maxam method yielded the entire sequence of 5243 nucleotide pairs.

Starting with the restriction map, and later with the nucleotide sequence map, it has been possible to locate in the SV40 DNA, viral genes and some of the regulatory signals very precisely, in fact at the sequence

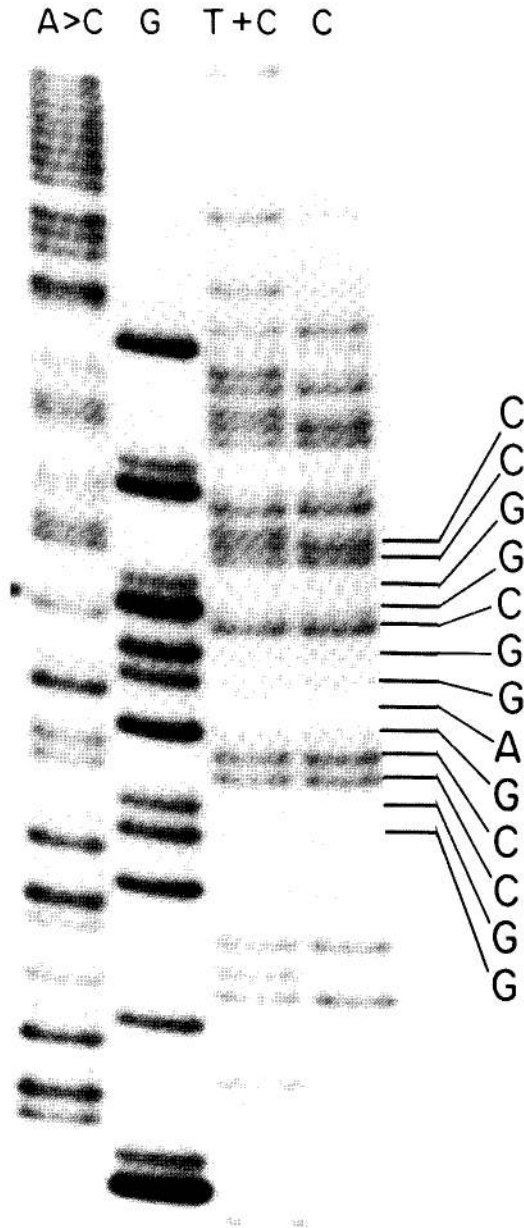


FIG. 10b. An actual sequencing gel from which the sequence can be read directly.

GENE EXPRESSION IN EUKARYOTES

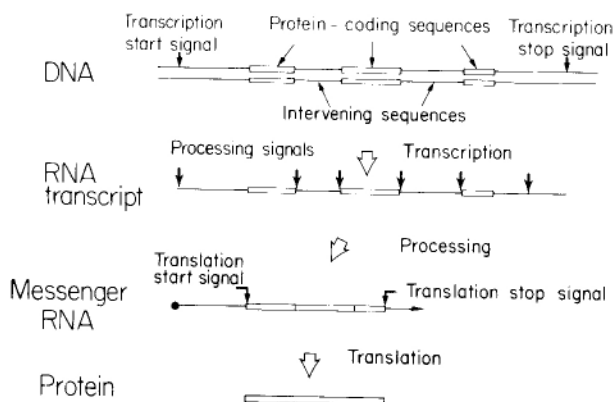


FIG. 12. Gene expression in eukaryotic cells.

level. A great deal of work from many laboratories is summarized in Figure 12. The inner circle shows the *Hind* cleavage sites and their map coordinates. Other circles indicate the positions of genes for T antigens (T and t), and for virion proteins VP 1, 2, 3. Shown at the top is the signal for the start of DNA replication (*ori*). Not shown explicitly in the map, but of considerable interest, is the transforming segment of the SV40 genome (14), i.e., the part of the molecule responsible for tumorigenesis, namely the segment coding for T antigen.

One of the most significant findings to emerge from the detailed analysis of the SV40 genome (and that of adenovirus) (15, 16) was the discovery that animal viruses have *split genes*, i.e., mRNA for a given protein is derived from discontinuous segments of DNA separated by intervening sequences (Figure 12). This completely unexpected circumstance turns out to reflect a fundamental property of genes from eukaryotic (nucleated) cells and their viruses, a property that adds a new dimension to evolution and to genetic regulation (17). As illustrated in Figure 12, regulation of gene expression in eukaryotic cells can occur not only at the level of transcription and translation, but also at the level of RNA processing whereby segments of RNA separated by intervening sequences are spliced together to form mRNA.

RESTRICTION ANALYSIS OF COMPLEX CHROMOSOMES

I pointed out earlier that the new genetic methods also allow detailed analysis of vastly more complicated chromosomes, such as those of mammalian cells. I turn now to some of these developments. As you know, mammalian nuclei have a number of individual chromosomes, as



FIG. 13. Spread of human chromosomes, visualized by light microscopy. The average chromosome contains about 20,000 times as much DNA as the SV40 genome.

illustrated in Figure 13, which shows the 23 pairs of human chromosomes. Each chromosome consists of DNA densely compacted by histones. Figure 13 is a light microscope picture in contrast to the electron micrograph of the SV40 virus chromosome shown in Figure 2. The viral chromosome would be invisible on this scale; its DNA content is around 1/20,000 or so of that of an average single human chromosome. In other words, each human cell has about 1 million times more DNA than does SV40, about 5×10^9 (5 billion) nucleotide pairs. How can one even begin a chemical dissection of so huge a genome? As you'll see, the same methods used to analyze the SV40 genome—with important additions—can be applied.

Edward Southern in Edinburgh is responsible for an elegant, sensitive procedure for mapping cellular genes by means of restriction enzymes

(18). His procedure is schematized in Figure 14. Total cellular DNA (prepared, for example, from leukocytes or other cells) is digested with a given restriction enzyme and the resulting fragments are fractionated by electrophoresis. Since the DNA is so complex, a very large number of fragments of various lengths is produced. Therefore the electropherogram shows a continuous smear of DNA. This DNA is next denatured to separate individual strands and these are transferred to nitrocellulose sheets. To detect fragments derived from a particular gene one reacts the sheet with purified ^{32}P -mRNA (or a DNA copy thereof) derived from

Restriction Analysis of Cellular DNA

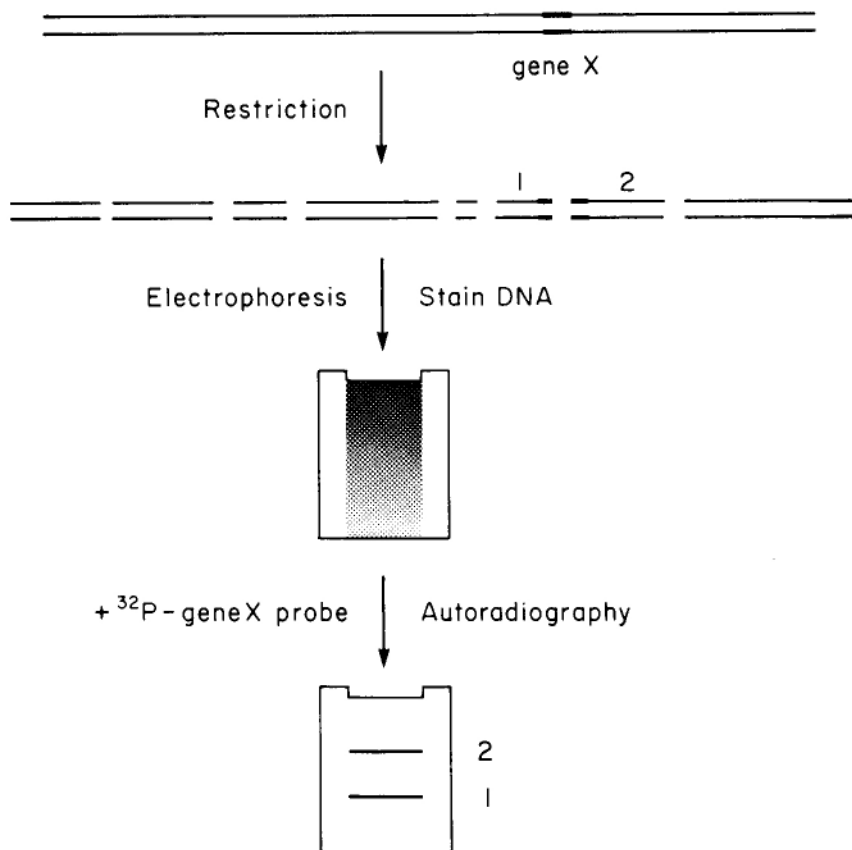


FIG. 14. Restriction analysis of cellular DNA by the fragment transfer and hybridization procedure of Southern (18).

that gene. For example, if one wishes to map the gene for globin, a ^{32}P copy of mRNA isolated from reticulocytes can be used as the probe. Since mRNA has a nucleotide sequence complementary to that of its gene, it will base-pair with that gene, allowing one to visualize those DNA fragments that are part of the gene of interest (Figure 14).

Figure 15 shows the results of restriction mapping of human globin genes by the Southern procedure (19). The cleavage map is analogous to that of SV40, showing the positions of globin coding sequences, intervening sequences, and surrounding DNA that probably contains regulatory signals. Also shown in Figure 15 are restriction sites in the intervening sequences that are present in DNA from some individuals but not others. Such *restriction polymorphisms* occur with high frequency (19) and will serve as valuable markers in studies of human populations. For example, Kan discovered that most patients with sickle cell anemia are missing a restriction site adjacent to the β globin gene (20). This marker has recently been used for pre-natal diagnosis of sickle cell disease or trait, as illustrated in Figure 16. Since random cloned DNA fragments can also be used as probes, we can expect rapid discovery of many restriction polymorphisms of this type.

MOLECULAR CLONING OF DNA

Restriction mapping using total cellular DNA has limitations. What is needed is a way of isolating from very complex mixtures of DNA frag-

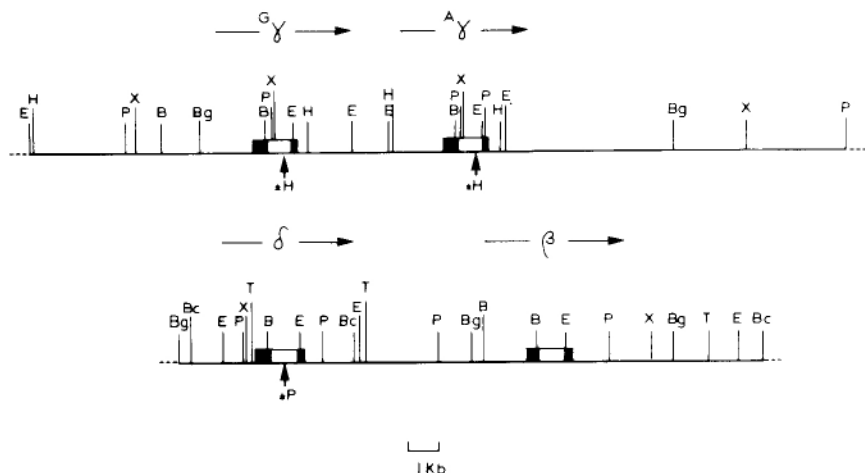


FIG. 15. Restriction map of human globin and surrounding genes. (Reprinted from reference 19, with permission of *Cell*, MIT press.) Capital letters refer to specific restriction enzyme sites. Black segments represent DNA sequences coding for globin and the white segments represent intervening sequences. Below the map are indicated extra restriction sites found in DNA of some individuals but not others (see text).

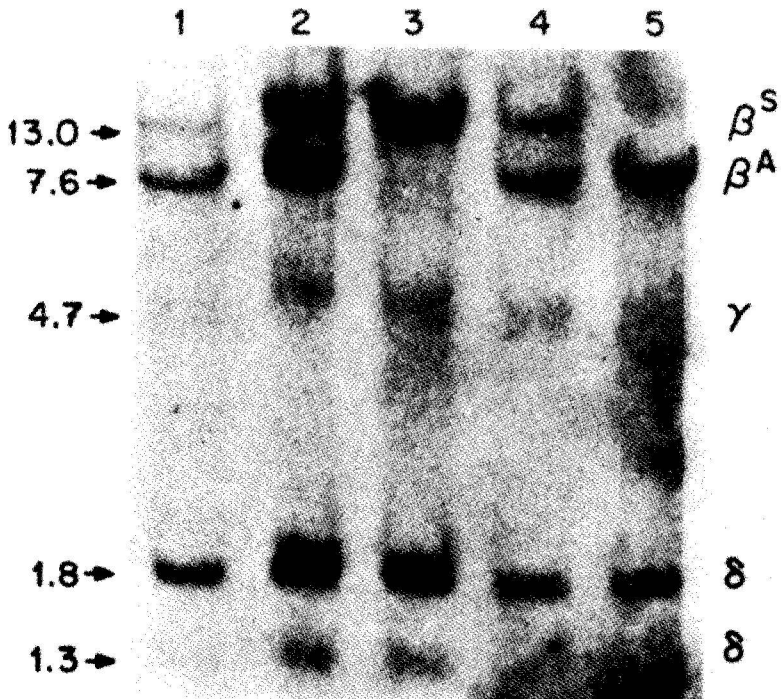


FIG. 16. Pre-natal diagnosis of sickle cell trait by restriction mapping of DNA from amniotic cells. (From reference 21, with permission.) DNA was digested with restriction endonuclease *Hpa* I, electrophoresed, transferred to nitrocellulose, and probed with ^{32}P globin DNA. The figures on the left indicate fragment sizes in kilodaltons; the symbols on the right indicate DNA fragments derived from sickle β globin (β^S), normal β globin (β^A), γ and δ globin, respectively. Column 1, DNA from the father of the fetus; column 2, mother; column 3, a sib with sickle cell disease; column 4, the fetus; column 5, a normal control. Note that father, mother and fetus are heterozygous (have both β^S , and β^A bands), the sib has only β^S , and the normal only β^A .

ments individual genes and their surrounding DNA in homogeneous form and in sufficient quantity to analyze in chemical detail, as was done for SV40 genes. Development of methods to do exactly that was another major advance, namely, the *molecular cloning of DNA* in microorganisms, as worked out by Stanley Cohen, Herbert Boyer, and their colleagues (22). "Molecular cloning" denotes the propagation of DNA molecules, all of which are derived from the same parental molecule. The procedure developed by Cohen and Boyer takes advantage of replication of virus-like genomes called "plasmids" present in certain bacteria. Figure 17 shows an electron micrograph of a ruptured *Escherichia coli* cell whose DNA has spread out on the microscope grid. At the bottom of the picture is a tiny ring of DNA separate from the rest—a bacterial plasmid.

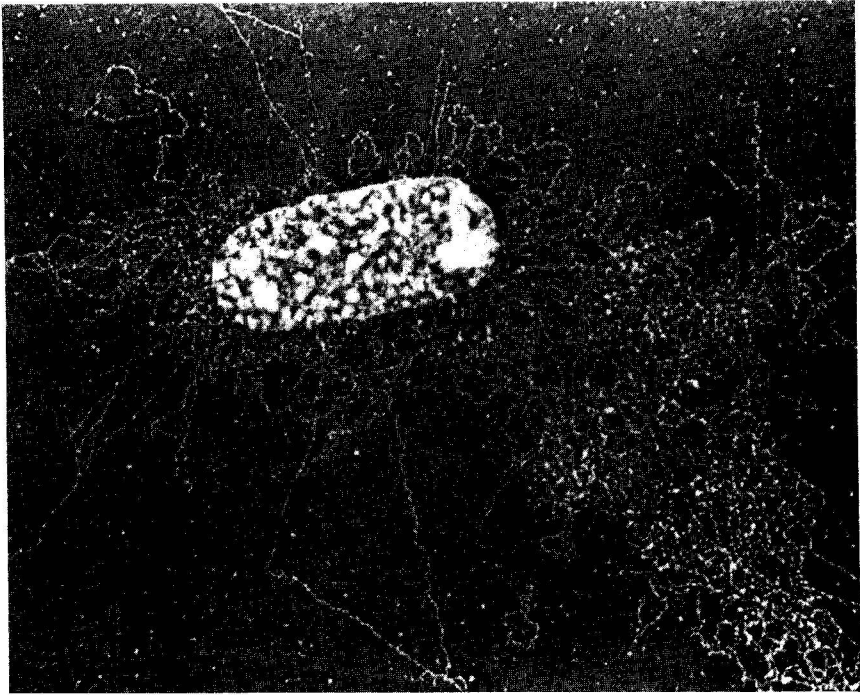


FIG. 17. Electron micrograph of a ruptured *E. coli* cell, showing cellular DNA and a plasmid (bottom center). (From reference 23, with permission of Dr. J. Griffith.)

Plasmids multiply inside growing cells, often comprising a substantial fraction of the total DNA. The usefulness of plasmids for molecular cloning is due to the fact that DNA from any source can be inserted *in vitro* into the plasmid, and the recombinant plasmid, when taken up by bacteria, will multiply inside the cell, thus propagating the DNA insert (Figure 18). If the plasmid has an antibiotic-resistance gene, bacteria containing a recombinant plasmid can be cloned readily simply by spreading them out on agar medium with antibiotic, and allowing them to grow into colonies. Those colonies with the recombinants of interest are detectable by hybridization of their DNA with suitable radioactive RNA or DNA probes. This procedure, or some variation of it, is being widely applied to prepare large amounts of cellular or viral genes in homogeneous form suitable for chemical analysis. Let me cite a few examples of particular interest to this audience.

Philip Leder in Bethesda and Susumu Tonegawa in Basel have used recombinant DNA methods to prepare mouse *immunoglobulin genes* (25, 26). Having cloned genes of light chains of immunoglobulins from DNA of plasmacytoma (myeloma) cells, they constructed detailed restric-

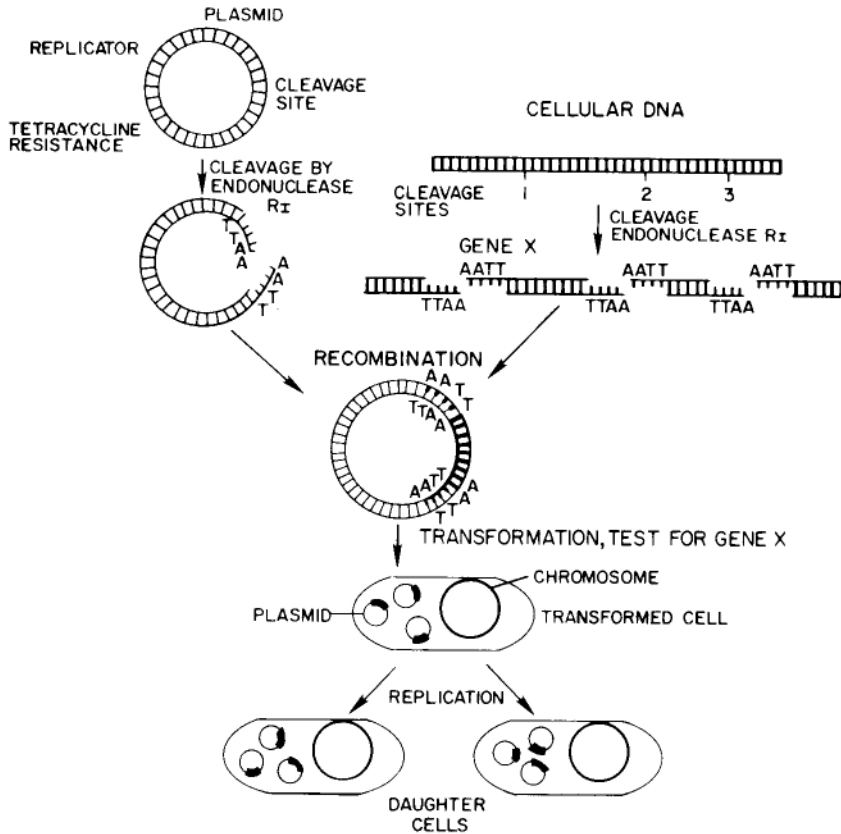


FIG. 18. Molecular cloning of DNA in a bacterial plasmid (24). A recombinant plasmid is constructed *in vitro* by joining linearized plasmid DNA carrying an antibiotic resistance gene and a DNA segment to be cloned. The recombinant is used to transform bacteria to antibiotic resistance. Resistant bacteria are then cloned. Each clone can then be tested for the presence of the inserted DNA by hybridization to a radioactive RNA or DNA probe. (Re-drawn from reference 24.)

tion maps and nucleotide sequence maps of the genes and surrounding DNA. Figure 19 is a summary diagram of some of their results. First note that the light chain of an antibody molecule consists of four parts: a leader or L segment, a variable or V segment, a constant or C segment, and a junctional or J segment (recognized only after the DNA analyses to be described). In immunoglobulin-producing cells of a particular plasmacytoma the light chain gene has the structure shown in Figure 19. The V-coding segment is adjacent to the J-coding segment, but far from the C-coding segment. All the alternative J-coding segments are part of a large intervening sequence, spliced out during formation of the mRNA.

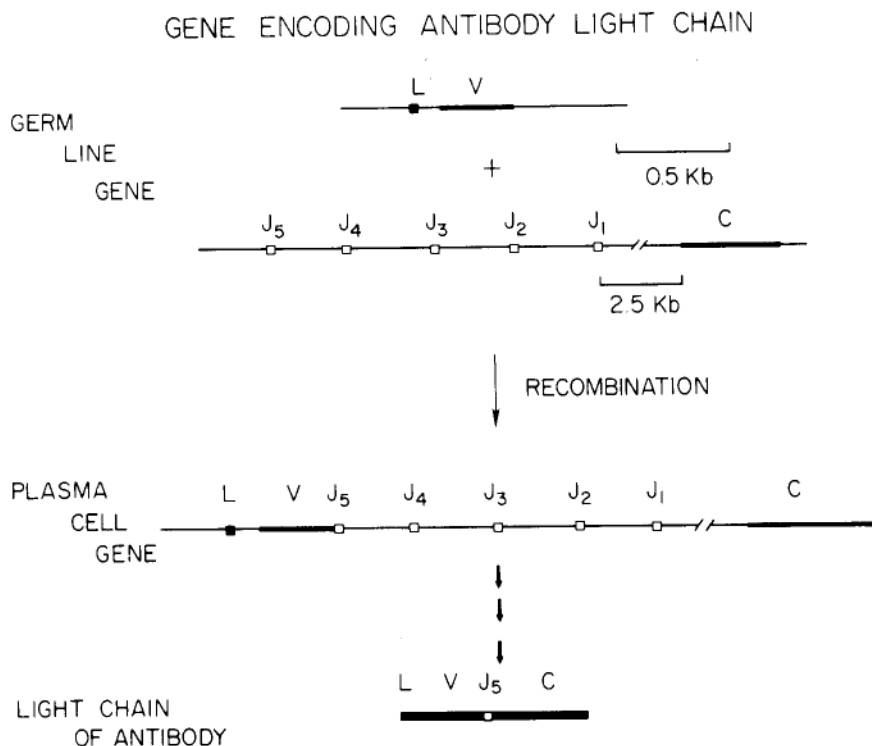


FIG. 19. The structure of the gene for a light chain of immunoglobulin. See text for details. (Re-drawn from reference 25.)

Now note that in germ line (embryonic) DNA the V and J segments are further separated; in fact their relative positions in chromosomal DNA is not yet accurately known. Evidently, during the differentiation of antibody-producing cells, V segment DNA recombines with J segment DNA to produce the gene structure found in plasma cells! This recombination event is thought to activate the gene.

A second example of the application of molecular cloning has as its objective the production of a *vaccine against hepatitis B virus (HBV)*. As you know, a substantial percentage of the human population is latently infected with HBV. HBV is associated with serious disease—acute and chronic hepatitis and hepatocellular carcinoma. Holding back understanding of this virus and preparation of a vaccine is the inability to grow HBV in tissue culture. Very recently the viral genome has been cloned in bacteria and its entire nucleotide sequence determined (27). The gene coding for HBV surface antigen, antibodies to which appear to be protective, has been identified, and several research groups are now attempting to construct recombinant plasmids that will direct the production of the

surface antigen in bacteria. From the bacterial product it may be possible to prepare a useful vaccine.

Finally I want to cite an example of the production of useful *polypeptide hormones* in bacteria by recombinant DNA methods, an application that has gotten much attention in the popular press. As a result of decades of research with bacteria and bacterial viruses, we know enough about the genetic elements that control gene expression to construct recombinant plasmids with very high levels of transcription and translation of gene inserts. Essentially any gene can be activated by putting it next to the appropriate signals. In this way genes for insulin, growth hormone, and somatostatin have been made to function in bacteria. The somatostatin case is instructive for another reason: the gene was totally synthesized from individual nucleotides, the order of nucleotides being deduced from the known amino acid sequence of somatostatin and the genetic code (Figure 20) (28). When inserted into an appropriate plasmid, the synthetic gene was translated into the amino acid sequence of somatostatin.

FUTURE DIRECTIONS

What new developments appear likely in the future? First of all, and most important in the long run, the new genetics is likely to provide fresh insights into genetic mechanisms of higher organisms: the structure of gene clusters and regulatory elements; what molecules interact with regulatory elements to control gene action; how hormones work at the gene level; and in time the nature of complex genetic programs that govern the growth, development and specialized functions of higher organisms. Many of these insights are likely to have considerable impact on our understanding and control of diseases ranging from inborn errors of metabolism through autoimmune disease to cancer. At the applied level, we are already seeing promising starts in the microbial production of medically useful products, such as polypeptide hormones, viral proteins, enzymes, and antibiotics. One can foresee also the production of polypeptide analogues—by *in vitro* mutation or chemical synthesis of genes—some of which may be therapeutically useful. We will see increasing use of restriction analysis to diagnose disease or genetic predisposition to disease, based on extensive polymorphism for restriction sites in the

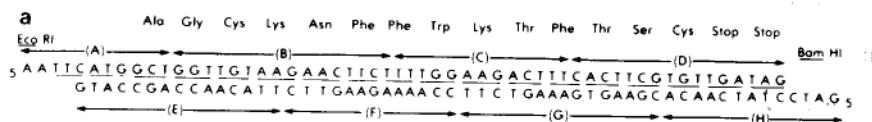


FIG. 20. Synthetic gene for somatostatin. The amino acid sequence of somatostatin is shown above, and the blocks of nucleotides that were synthesized and joined are indicated below. (From reference 28.)

human population. And perhaps also we may see the emergence of gene therapy for certain disorders. Outside of medicine, one can anticipate the production of many industrially useful compounds by recombinant methods, increasing exploration of microbial energy generation, and possibly important agricultural applications.

I began by noting that these are exciting times in genetics, and I hope I have succeeded in conveying some of the excitement felt by those working in this field. Since areas of inquiry opened up by the new genetics concern fundamental and complex phenomena of higher organisms as well as a broad range of applied biology, the excitement is likely to continue for some time to come.

REFERENCES

1. AVERY, O. T., MACLEOD, C. M., AND McCARTY, M.: Studies on the chemical nature of the substance inducing transformation of pneumococcal types. Induction of transformation by a desoxyribonucleic acid fraction isolated from pneumococcus type III. *J. Exp. Med.* **79**: 137, 1944.
2. DUBOS, R. J.: *The Professor, the Institute and DNA*, The Rockefeller Univ. Press, N.Y., 1976.
3. WATSON, J. D. AND CRICK, F. H. C.: A structure for deoxyribose nucleic acids. *Nature* **171**: 737, 1953.
4. LURIA, S. E. AND DELBRÜCK, M.: Mutations of bacteria from virus sensitivity to virus resistance. *Genetics* **28**: 491, 1943.
5. ARBER, W.: Host-controlled modification of bacteriophage. *Ann. Rev. Micro.* **19**: 365, 1965.
6. SMITH, H. O.: Nucleotide sequence specificity of restriction endonucleases. *Science* **205**: 455, 1979.
7. NATHANS, D.: Restriction Endonucleases, Simian Virus 40, and the New Genetics. *Science* **206**: 903, 1979.
8. DANNA, K. J. AND NATHANS, D.: Specific cleavage of Simian Virus 40 DNA by restriction endonuclease of *Hemophilus influenzae*. *Proc. Natl. Acad. Sci. (USA)* **68**: 2913, 1971.
9. SANGER, F. AND COULSON, A. R.: A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *J. Mol. Biol.* **94**: 441, 1975.
10. MAXAM, A. M. AND GILBERT, W.: A new method for sequencing DNA. *Proc. Natl. Acad. Sci. (USA)* **74**: 560, 1977.
11. REDDY, V. B., THIMAPPAYA, B., DHAR, R., SUBRAMANIAN, K. N., ZAIN, B. S., PAN, J., GHOSH, P. K., CELMA, M. L., AND WEISSMAN, S. M.: The genome of Simian Virus 40. *Science* **200**: 494, 1978.
12. FIERS, W., CONTRERAS, R., HAEGEMAN, G., ROGIERS, R., VAN DE VOORDE, A., VAN HEUVERSWEYN, H., VAN HERREWEGHE, J., VOLCKGERT, G., AND YSEBAERT, M.: Complete nucleotide sequence of SV40. *Nature* **273**: 113, 1978.
13. GHOSH, P. K., REDDY, V. B., SWINSCOE, J., LEBOWITZ, P., AND WEISSMAN, S. M.: Heterogeneity and 5'-terminal structures of the late RNAs of Simian Virus 40. *J. Mol. Biol.* **126**: 813, 1978.
14. GRAHAM, F. L., ABRAHAMS, P. J., MULDER, C., HEIJNEKER, H. L., WARNAAR, S. O., DE VRIES, F. A. J., FIERS, W., AND VAN DER EB, A. J.: Studies on *in vitro* transformation by DNA and DNA fragments of human adenoviruses and Simian Virus 40. *Cold Spr. Harbor Symp.* **39**: 637, 1974.

15. BERK, A. J. AND SHARP, P. A.: Sizing and mapping of early adenovirus mRNAs by gel electrophoresis of S1 endonuclease-digested hybrids. *Cell* **12**: 721, 1977.
16. BROKER, T. R., CHOW, L. T., DUNN, A. R., GELINAS, R. E., HASSELL, J. A., KLESSIG, D. F., LEWIS, J. B., ROBERTS, R. J., AND ZAIN, B. S.: Adenovirus—2 messengers—an example of baroque molecular architecture. *Cold Spr. Harbor Symp.* **62**: 531, 1977.
17. DARNELL, J. E., JR.: Implication of RNA-RNA splicing in evolution of eukaryotic cells. *Science* **202**: 1257, 1978.
18. SOUTHERN, E. M.: Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* **98**: 503, 1975.
19. JEFFREYS, A. J.: DNA sequence variants in the $\epsilon\gamma$ -, $\alpha\gamma$ -, δ - and β -globin genes of man. *Cell* **18**: 1, 1979.
20. KAN, Y. W. AND DOZY, A. M.: Antenatal diagnosis of Sickle-Cell anaemia by DNA analysis of amniotic-fluid cells. *Lancet* **2**: 910, 1978.
21. PHILLIPS, J. A., III, SCOTT, A. F., KAZAZIAN, H. H., JR., SMITH, K. D., STETTEN, G., AND THOMAS, G. H.: Prenatal diagnosis of hemoglobinopathies by restriction endonuclease analysis: Pregnancies at risk for Sickle Cell anemia and S-Q^{Ara}h disease. *Johns Hopkins Med. J.* **145**: 57, 1979.
22. COHEN, S. N., CHANG, A. C. Y., BOYER, H. W., AND HELLING, R. B.: Construction of biologically functional bacteria plasmids *in vitro*. *Proc. Natl. Acad. Sci. (USA)* **70**: 3240, 1973.
23. GROBSTEIN, C.: The recombinant-DNA debate. *Scientific Amer.* **237**: 22, 1977.
24. COHEN, S. N.: The manipulation of genes. *Scientific Amer.* **233**: 24, 1975.
25. SEIDMAN, J. G., MAX, E. E., AND LEDER, P.: A κ -immunoglobulin gene is formed by site-specific recombination without further somatic mutation. *Nature* **280**: 370, 1979.
26. SAKANO, H., HÜPPI, K., HEINRICH, G., AND TONEGAWA, S.: Sequences at the somatic recombination sites of immunoglobulin lightchain genes. *Nature* **280**: 288, 1979.
27. GALIBERT, F., MANDART, E., FITOUSSI, F., TIOLLAIS, P., AND CHARNAY, P.: Nucleotide sequence of the hepatitis B virus genome (subtype ayw) cloned in *E. coli*. *Nature* **281**: 646, 1979.
28. ITAKURA, K., HIROSE, T., CREA, R., RIGGS, A. D., HEYNEKER, H. L., BOLIVAR, F., AND BOYER, H. W.: Expression in *Escherichia coli* of a chemically synthesized gene for the hormone somatostatin. *Science* **198**: 1056, 1977.