kilometers below the magma reservoir. They, together with the accompanying harmonic tremor, were possibly caused by the sudden rise of new magma from depth into the shallow, highly swollen reservoir complex. Meanwhile, the tilt continued upward at its high rate. Within minutes, the reservoir could accept no more; magma forced its way upward, split north-south fissures across the floor of Halemaumau, and spilled onto the surface at 02:32 on 5 November, to begin the 1967–68 eruption (*12*).

References and Notes

1. T. A. Jagger and R. H. Finch, *Bull. Seismol. Soc. Amer.* **19**, 38 (1929).
2. R. M. Wilson, *Univ. Hawaii Res. Pub. 10* (1935).
3. J. P. Eaton, *Bull. Seismol. Soc. Amer.* **49**, 301 (1959).
4. R. W. Decker, D. P. Hill, T. L. Wright, *Bull. Volcanol.* **29**, 721 (1966).
5. This bench mark lies only 5 kilometers from the center of maximum uplift and doubtless is uplifted as the summit inflates. At this distance, however, the uplift is small—perhaps 5 or 10 percent of the maximum.
6. K. Mogi, *Bull. Earthquake Res. Inst. Tokyo Univ.* **36**, 99 (1958).
7. S. H. Laurila, *Rev. Geophys.* **1**, 553 (1963).
8. K. Kasahara, C. Okada, M. Shibano, K. Sasaki, S. Matsumoto, *Research on the Matsushiro Earthquakes*, pt. 3 (Earthquake Research Institute, Univ. of Tokyo, Tokyo, 1967), p. 93.
9. W. T. Kinoshita, H. L. Krivoy, D. R. Mabey, R. R. MacDonald, *U.S. Geol. Surv. Profess. Paper 475-C* (1963), pp. C114–C116.
10. J. P. Eaton, oral communication.
11. ———, *Amer. Geophys. Union Geophys. Monograph 6* (1962), p. 13.
12. Kilauea's 1967–68 eruption lasted from 5 November 1967 to 13 July 1968. All the activity was confined to the summit crater of Halemaumau, where more than 75 million cubic meters of basalt lava were erupted.
13. D. W. Peterson, *U.S. Geological Survey Map GO-667* (1967).

14. We thank the staff of the Hawaiian Volcano Observatory for invaluable aid and the staff of Hawaii Volcanoes National Park for logistical support. Publication of this article is authorized by the director of the U.S. Geological Survey.

Gene Regulation for Higher Cells: A Theory

New facts regarding the organization of the genome provide clues to the nature of gene regulation.

Roy J. Britten and Eric H. Davidson

Cell differentiation is based almost certainly on the regulation of gene activity, so that for each state of differentiation a certain set of genes is active in transcription and other genes are inactive. The establishment of this concept (*1*) has depended on evidence indicating that the cells of an organism generally contain identical genomes (*2*). Direct support for the idea that regulation of gene activity underlies cell differentiation comes from evidence that much of the genome in higher cell types is inactive (*3*) and that different ribonucleic acids (RNA) are synthesized in different cell types (*4*).

Little is known, however, of the molecular mechanisms by which gene expression is controlled in differentiated cells. As far as we are aware no theoretical concepts have been advanced which provide an interpretation of certain of the salient features of genomic structure and function in higher organisms. We consider here experimental evidence relating to these features. (i) Change in state of differentiation in higher cell types is often mediated by simple external signals, as, for example, in the action of hormones or embryonic inductive agents. (ii) A given state of differentiation tends to require the integrated activation of a very large number of noncontiguous genes. (iii) There exists a significant class of genomic sequences which are transcribed in the nuclei of higher cell types but appear to be absent from cytoplasmic RNA's. (iv) The genome present in higher cell types is extremely large, compared to that in bacteria. (v) This genome differs strikingly from the bacterial genome due to the presence of large fractions of repetitive nucleotide sequences which are scattered throughout the genome. (vi) Furthermore, these repetitive sequences are transcribed in differentiated cells according to cell type-specific patterns.

In this article we propose a new set of regulatory mechanisms for the cells of higher organisms such that multiple changes in gene activity can result from a single initiatory event. These proposals are presented in the form of a specific, relatively detailed model at the level of complexity which appears to us to be required for the genomic regulatory machinery of higher cells. We make no attempt to arrive at definitive statements regarding these proposed mechanisms; obviously evidence is not now available to support any model in detail. Our purpose in presenting an explicit theory is to describe the regulatory system proposed in terms of elements and processes which are capable of facing direct experimental test. It is hoped that our relatively detailed commitment will induce discussion and experiment, and it is expected that major modifications in concept will result.

Undoubtedly important regulatory processes occur at all levels of biological organization. We emphasize that this theory is restricted to processes of cell regulation at the level of genomic transcription.

We begin by describing our usage of certain terms and their role in the model, and then present the model itself. We then consider relevant experimental observations and certain testable implications of the model. Finally, some general implications of the model for evolutionary theory are mentioned.

Elements of the Model

The following definitions are intended only to clarify the usage of certain terms in our discussion of this model.

Gene: A region of the genome with a narrowly definable or elementary

Dr. Britten is on the staff of the Department of Terrestrial Magnetism, Carnegie Institution of Washington, Washington, D.C. 20015, and Dr. Davidson is on the staff of Rockefeller University, New York 10021.

function. It need not contain information for specifying the primary structure of a protein.

Producer gene: A region of the genome transcribed to yield a template RNA molecule or other species of RNA molecules, except those engaged directly in genomic regulation. We are using this term in a manner analogous to that in which the term "structural gene" has been used in the context of certain bacterial regulation systems (*5*). Products of the producer gene include all RNA's other than those exclusively performing genomic regulation by recognition of a specific sequence. Among producer genes, for example, are the genes on which the messenger RNA template for a hemoglobin subunit is synthesized, and also the genes on which transfer RNA molecules are synthesized.

Receptor gene: A DNA sequence linked to a producer gene which causes transcription of the producer gene to occur when a sequence-specific complex is formed between the receptor sequence and an RNA molecule called an activator RNA. We do not, in this model, wish to specify a mode of action for the receptor gene—that is, the nature of the molecular events occurring between the DNA, histones, polymerases, and so forth, present in the receptor complex. This model is concerned primarily with interrelations among the DNA sequences present in the genome.

Activator RNA: The RNA molecules which form a sequence-specific complex with receptor genes linked to producer genes. The complex suggested here is between native (double-stranded) DNA and a single-stranded RNA molecule (*6*). The role proposed for activator RNA could well be carried out by protein molecules coded by these RNA's without changing the formal structure of the model (*7*). Decisive evidence is lacking in higher cells, and we have chosen the simpler alternative (*8*). As the discussion of the evolutionary implications of this model will indicate, however, the probability of formation of new batteries of genes in evolution appears to differ greatly between these two alternatives.

Integrator gene: A gene whose function is the synthesis of an activator-RNA. The term integrator is intended to emphasize the role of these genes in leading, by way of their activator RNA's, to the coordinated activity of a number of producer genes. A set of linked integrator genes is activated together in response to a specific initiating event, resulting in the concerted activity of a number of producer genes not sharing a given receptor gene sequence.

Sensor gene: A sequence serving as a binding site for agents which induce the occurrence of specific patterns of activity in the genome. Binding of these inducing agents is a sequence-specific phenomenon dependent on the sensor gene sequence, and it results in the activation of the integrator gene or genes linked to the sensor gene. Such agents include, for example, hormones and other molecules active in intercellular relations as well as in intracellular control. Most will not bind to sensor gene DNA, and an intermediary structure such as a specific protein molecule will be required. This structure must complex with the inducing agent and must bind to the sensor gene DNA in a sequence-specific way.

Battery of genes: The set of producer genes which is activated when a particular sensor gene activates its set of integrator genes. A particular cell state will usually require the operation of many batteries.

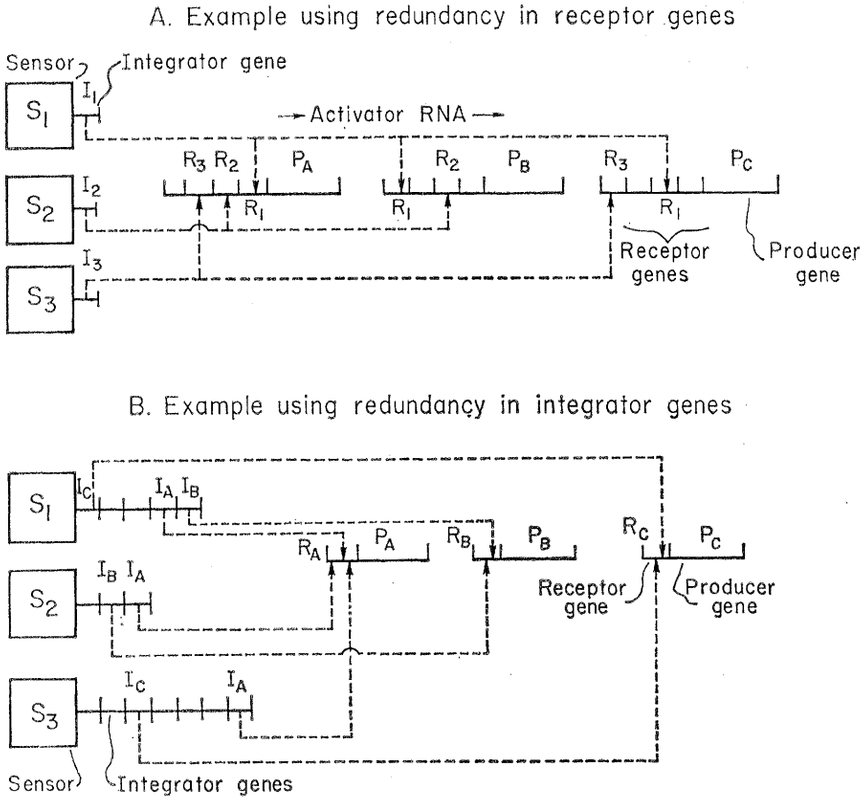

Fig. 1. Types of integrative system within the model. (A) Integrative system depending on redundancy among the regulator genes. (B) Integrative system depending on redundancy among the integrator genes. These diagrams schematize the events that occur after the three sensor genes have initiated transcription of their integrator genes. Activator RNA's diffuse (symbolized by dotted line) from their sites of synthesis—the integrator genes—to receptor genes. The formation of a complex between them leads to active transcription of the producer genes P_A, P_B, and P_C.

Integrative Function of the Model

The concerted activation of one or more batteries of producer genes is considered to underlie the existence of diverse states of differentiation. Examples of two basic aspects of the proposed integrative function appear in Fig. 1. In each case, the producer genes shown are integrated into three different, very small batteries. Sensor gene S_1 and its integrator specify the activation of producer genes P_A, P_B, and P_C; S_2 that of P_A and P_B; and S_3 that of P_A and P_C.

In Fig. 1A, the control pattern depends on the existence of redundant receptor sequences in the receptor gene sets of the three producer genes. Inclusion of a particular producer gene in each of the batteries calling on it depends on the presence of the appropriate receptor gene adjacent to the producer gene. Thus, in the case where there is only one integrator gene per sensor as in Fig. 1A, there will be as many copies of a given receptor gene

sequence as there are producer genes in a battery.

In the case shown in Fig. 1B, however, redundancy is present between the integrator genes of different integrator sets. A particular producer gene, in this example, is included in each of several batteries calling on it by virtue of the inclusion of the same integrator gene adjacent to each of the appropriate sensor genes. Here there will be as many copies of a given integrator-gene as there are batteries that call on its producer gene. For certain commonly required genes, for example those used in the fundamental biochemistry of each cell, this could be a very large number indeed.

Systems of the type portrayed in Fig. 1A might be most useful in the case where the producer genes to be integrated direct the synthesis of enzymes whose function is tightly coordinated physiologically, for example, the ten enzymes of the urea synthesis system. Where the system is needed, all the genes would be needed. The system portrayed in Fig. 1B is a more powerful integrative system since it can govern a larger diversity of producer genes. The number of receptor sequences governing each producer sequence is probably small since transcription of a producer gene sequence is not likely to be activated from a great distance along the DNA strand. There is no reason a priori, on the other hand, to restrict the number of integrator genes per integrator set, except for the requirement that the integrator genes not be so distant that there is a high probability of their being separated by translocation.

In this model, regulation is accomplished by sequence-specific binding of an activator RNA and not by sequence recognition on the part of histones. The latter seem clearly to be the general inhibitors of transcription in the genome, but evidently these general inhibitors do not possess sufficient diversity to be considered as sequence-specific regulatory elements themselves (*9, 10*). We have assumed that, unless otherwise specified, the state of the higher cell genome is histone-mediated repression and that regulation is accomplished by specific activation of otherwise repressed sites, rather than by repression of otherwise active sites.

Figure 2 combines the elements and systems we have thus far described. In the remainder of our discussion we consider various properties and consequences of the minimal model, as portrayed in this figure. The magnitude of the producer gene batteries is only suggested by the diagram in Fig. 2, and of course no attempt has been made to portray the actual complexity of the system, that is, to illustrate the number of elements whose function is likely to be integrated in a living cell. Obviously, the coordination of many batteries of genes is required in order to account for massive changes in differentiated state, such as the neogenesis of a tissue during development. We visualize such phenomena as being mediated by sensor genes sensitive to the products of integrator genes in other integrative sets. In other words, a single inducing agent could lead to the activation of a number of sensor-integrator sets, activating a vast number of producer genes.

Sequential patterns of gene activation, as in development, could result if certain sensors respond to the products of producer genes. In addition, the protein of a newly effective sensor assembly is, in the model, a product of a previously activated producer gene. Stabilization of a cell type in a given state of differentiation might also be explained in this way. Living systems continuously adjust their activities in accordance with their internal state, and it is evident that a requirement for sensors sensitive to feedback control by certain producer-gene products exists as well.

Fraction of the Genome Utilized for Regulation

Broadly speaking, genome size increases with the grade of organization of eukaryotes, as first pointed out by Mirsky and Ris in 1951 (*11, 12*). The wide range of genome sizes often observed among closely related creatures obscures the correlation. Organisms with large genomes presumably have a requirement for genomic information similar to that of their relatives with smaller genomes. This implies the evolutionary multiplication of the genome of ancestors possessing the minimum amount of DNA required to effect each grade of organization. It is thus useful to consider the minimum amount of DNA observable at each grade of organization. Figure 3 shows the minimum genome size (*13*) for some major steps in evolution between viruses and the higher chordates.

A reasonable explanation has not been advanced for the large genome sizes occurring at the higher organizational levels. Most of the known biosynthetic pathways are already represented in unicellular organisms. It is not possible to estimate the increase in number of producer genes required to specify structure and chemistry at the higher levels of organization. Nonetheless, it seems unlikely that the 30-fold

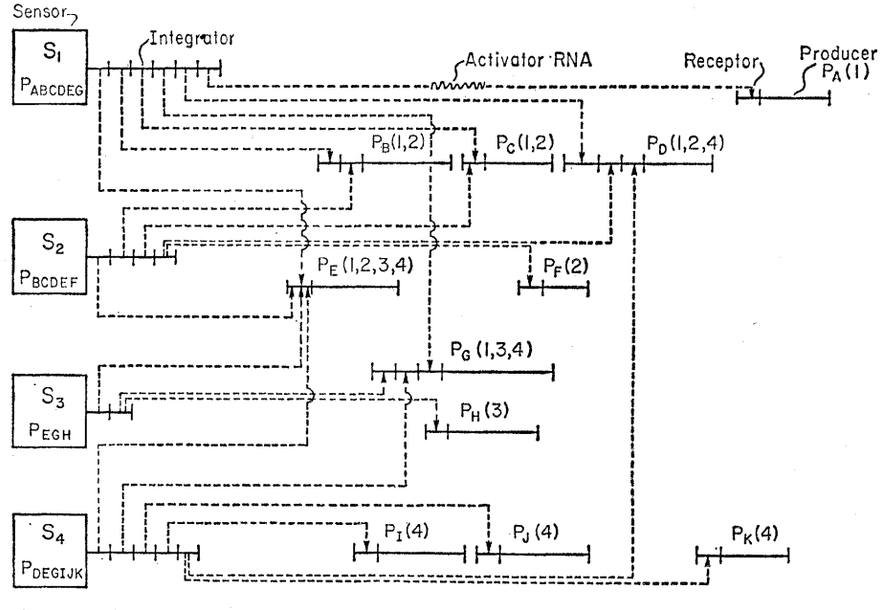

Fig. 2. This diagram is intended to suggest the existence of overlapping batteries of genes and to show how, according to the model, control of their transcription might occur. The dotted lines symbolize the diffusion of activator RNA from its sites of synthesis, the integrator genes, to the receptor genes. The numbers in parentheses show which sensor genes control the transcription of the producer genes. At each sensor the battery of producer genes activated by that sensor is listed. In reality many batteries will be much larger than those shown and some genes will be part of hundreds of batteries.

Table 1. Several of the functionally linked enzyme systems present in liver (*17*, chapter 12; *36*). Uridine monophosphate, UMP; adenosine monophosphate, AMP.

System	Number of enzymes
Glycogen synthesis	5
Galactose synthesis	6
Phosphogluconate oxidation	11
Glycolysis	12
Citric acid cycle	17
Lecithin synthesis	8
Fatty acid breakdown	5
Lanosterol synthesis	10
Phenylalanine oxidation	8
Methionine to cysteine	10
Methionine to aspartic acid	10
Urea formation	10
Coenzyme A synthesis	6
Heme synthesis	9
Pyrimidine synthesis (to UMP)	6
Purine synthesis (to AMP)	14

increase from poriferan to mammal can be attributed to a 30-fold increase in the number of producer genes. This problem cannot be escaped by attributing the large genome size to redundancy. Fifty-five percent of the DNA of the calf, for example, occurs in nonrepetitive sequences (*14*). This is enough DNA to provide almost 10^7 diverse producer-gene sequences the size of the gene coding for the beta chain of hemoglobin. A few other measurements have been made which indicate that such diversity in DNA sequence is general (*15*).

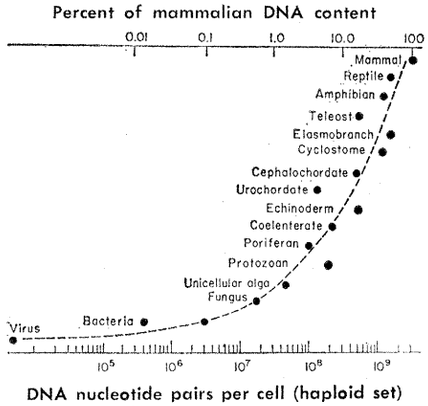

Fig. 3. The minimum amount of DNA that has been observed for species (*13*) at various grades of organization. Each point represents the measured DNA content per cell for a haploid set of chromosomes. In the cases of mammals, amphibians, teleosts, bacteria, and viruses enough measurements exist to give the minimum value meaning. However for the intermediate grades few measurements are available, and the values shown may not be truly minimal. No measurements were unearthed for acoela, pseudocoela and mesozoa. The ordinate is not a numerical scale, and the exact shape of the curve has little significance. The figure shows that a great increase in DNA content is a necessary concomitant to increased complexity of organization.

Quite possibly, the principal difference between a poriferan and a mammal could lie in the degree of integrated cellular activity, and thus in a vastly increased complexity of regulation rather than a vastly increased number of producer genes (*16*). Much of the DNA accumulating in the genomes toward the upper end of the curve in Fig. 3 might then have a regulative function. The model also suggests that a large amount of DNA could be devoted to regulatory function: consider integrator and receptor sequences which are not redundant. In this case a battery of producer genes would require a distinct integrator gene for each producer gene. Producer genes occurring in several batteries would require receptor genes corresponding to each battery. The resulting multiplicity of integrator and receptor genes might result in a much larger quantity of DNA in regulatory sequences than in producer sequences. It is likely that an ever-growing library of different combinations of groups of producer genes is needed as more complex organisms evolve. An effective way of storing the information specifying these combinations in the genome is to make use of sensors responsive to the activator RNA's of other integrative sets. Thus we propose that a higher level of integrator gene sets is accumulated. Each of these, when activated, could specify a very large program of producer gene activations by specifying the activity of a network of other sensor-integrator sets. Thereby many batteries of genes of the sort shown in Fig. 2 could be activated.

Experimental Justification of the Elements of the Model

There are five important classes of elements in this model: sensor genes, integrator genes, activator RNA, receptor genes, and producer genes. Is this degree of complexity really necessary? The particular set of elements we have postulated may of course not be the required ones. Five, however, is the minimum number of classes of elements which can carry out the following formally described process: (i) response to an external signal; (ii) production of a second signal; (iii) transmission of the second signal to a number of receptors unresponsive to the original signal; (iv) reception of the second signal; and (v) response to this event by activation of a producer gene and its transcription to provide the cell with the producer gene product. In the following sections we examine evidence that such a description is applicable to gene regulation in higher organisms, and explore evidence that suggests the existence of the elements of the model.

Integration of Physically Unlinked Producer-Gene Activity

We have assumed that a given state of differentiation depends on the coordinated activity of a number of biochemical systems. Each of these systems will probably contain a number of components. As an example, Table 1 lists some of the enzyme systems operating in one cell type, mammalian liver.

An underlying principle of this model is that producer genes active in any given tissue need not be physically linked in the genome. For physically adjacent producer genes, integration of activity could be based on the operation of gigantic polycistronic tissue-specific operons. There are good reasons for believing that this is not the case in eukaryotes. Some producer genes are called into activity in a number of different tissues, as illustrated in Table 2.

Table 2 shows the overlapping pattern of activity for 17 enzymes in 8 tissues. Direct contiguity of active producer genes could not produce this set of patterns if a single copy of each gene were present in the genome. Genetic evidence does not at present indicate the presence of multiple producer genes yielding identical products, except for ribosomal RNA and transfer RNA. An equally strong point can be made that control of the producer gene sets for the systems listed in Table 1 cannot be based on physical linkage of one set to the next in the liver genome. In other tissues, some but not all of these systems are functional (*17*). In other words, even where the producer genes within a physiologically coordinated enzyme system (Table 1) are linked, the same formal problem remains: a mechanism is required for coordinating the activity of the noncontiguous systems of producer genes characteristic of each state of differentiation. In at least some instances the

Table 2. Distribution of various enzymes in tissues of one organism, rat (*17*, Table XIII). + Means enzyme is present in amount 40 to 100 percent of that in the tissue where it is most plentiful; 0 means enzyme is essentially absent, that is, less than 8 percent the level of the tissue with the highest activity. If the level falls between 8 and 40 percent, or if data are lacking, space is left blank.

E.C. No.	Enzyme	Liver	Kidney	Spleen	Heart	Skeletal muscle	Small intestine	Pancreas	Brain
1.1.1.30	3-Hydroxybutyrate dehydrogenase	+		0	0	0		0	0
1.1.1.37	Malate dehydrogenase		0		+				
1.5.1.1	Pyrroline-2-carboxylate reductase		+	0	0				+
1.11.1.6	Catalase	+	+			0		0	0
1.11.1.7	Peroxidase	0	0	+	0	0	+		0
1.13.1.5	Homogentisate oxidase	+	+	0	0	0	0		0
2.1.1.6	Catechol methyltransferase	+		0	0	0	0		0
2.1.1.3	Dimethythetin-homocysteine methyltransferase	+		0		0			0
2.7.7.16	Ribonuclease	0	0		0	0		+	
3.1.1.1	Carboxylesterase		0	0		0		+	
3.1.1.5	Phospholipase	+		+	0	0	+		0
3.1.1.7	Acetylcholinesterase	0	0			0			+
3.1.1.8 } 3.1.1.9 }	Cholinesterases		0		+	0	+		
3.1.3.1	Alkaline phosphatase	0	+	0		0	+	0	0
3.1.3.2	Acid phosphatase			+	0	0			
3.1.3.9	Glucose-6-phosphatase	+		0	0	0			0
3.2.1.25	β-Mannosidase	+	+		0	0		+	0
3.2.1.30	β-Acetylamino deoxyglucosidase		+		0	0		0	0
3.2.1.31	β-Glucuronidase	+		+		0		0	
3.5.3.1	Arginase	+		0	0	0	0	0	0
3.5.4.3	Guanine deaminase	+	+	+		0		+	0
4.1.2.7	Aldolase	0	0	0	0	+			0
4.1.3.7	Citrate synthase	0	0		+	0			
4.2.1.3	Aconitate hydratase		+		+				0
6.3.1.2	Glutamine synthetase	+	0	0		0		0	+

integrated producer genes within each physiologically coordinated set are known to be noncontiguous in higher organisms. As an example, in the human the producer genes coding for the alpha and beta subunits of hemoglobin are unlinked (*18*). Another case concerns two of the enzymes of the phosphogluconic acid oxidation pathway (system No. 3 of Table 1) in *Drosophila melanogaster*. These are glucose-6-phosphate dehydrogenase (E.C.1.1.1.49) and 6-phosphogluconate dehydrogenase (E.C.1.1.1.43) whose genes are located on separate linkage groups (*19*). Evidently producer genes whose activity must be functionally integrated in the most intimate way can be located far apart in the genome. We conclude that within at least some functionally integrated producer-gene systems as well as among these systems, specification of particular patterns of activity requires a method of control other than one depending on contiguity of the producer loci.

The data considered so far provide instances of the type of pattern which our model is designed to interpret, but they do not indicate the extensiveness of the producer-gene batteries called forth in given conditions of differentiation. Table 3 lists some of the effects of estrogen on the uterus, an estrogen target tissue. Although we are ignorant of the diverse proteins involved in effecting these changes it is obvious that there must be many. Though in Table 3, of course, we present only a partial list, and the number of diverse producer genes required for each item on the list can only be guessed at present, this table provides a more realistic description of the magnitude of the problem of producer-gene integration than Tables 1 or 2 do. Analogous problems exist in explaining the integration of a multitude of producer genes in every cell lineage during development. As diverse cell lineages differentiate, a huge variety of qualitatively novel properties appear together. As in the case of the hormones, such processes of differentiation appear to require a mechanism for the simultaneous activation of many systems, such as proposed in this model.

Evidence for the Existence of Sensor Elements

There are many chemically defined agents that have the evident property of inducing large-scale changes in the producer-gene activity of specific target tissues. These agents now include steroid hormones, polypeptide hormones, several plant hormones, several vitamins, and several embryonic inductive agents (*20*). Frequently, the responsible agents also produce an alteration in the spectrum of RNA's being transcribed in the target tissues, as indicated by data obtained with RNA-DNA hybridization and studies in vitro of chromatin template activity (*20*); and these agents have been identified in the nuclear apparatus of the target cells (*20*). The most intensively studied system is perhaps estrogen response (Table 3). All of the above-mentioned forms of evidence exist for this hormone (*20*).

In addition, Maurer and Chalkey (*21*) have isolated from calf endometrial chromatin a protein that binds 17β-estradiol. The binding is stereospecific, noncovalent, and strong (the Michaelis constant, K_m, for binding is $2 \times 10^{-8} M$); and the responsible protein appears not to be a histone. It does not bind steroids as closely related as 17 α-estradiol or diethylstilbestrol. Such a protein, in combination with the specific external agent for which it is the receptor, must interact with the genome in a sequence-specific way, since this interaction results in the activation of only a certain group of genes. Consider a system in which the genomic binding sites are simply adjacent to all the producer genes activated by the external chemical agent. Such a system would appear to possess a limited integrative function which might be utilized for certain small gene batteries. However, the binding of an

Table 3. Some effects of estrogen on uterine cells.

Effect	Ref.
Increase in total cell protein	(37)
Increase in transport of amino acids into cell	(37, 38)
Increase in protein synthesis activity per unit amount of polyribosomes	(37)
Increased synthesis of new ribosomes	(37)
Alteration of amounts of nuclear protein to nucleus	(37)
Increased amount of polyribosomes per cell	(37)
Increase in nucleolar mass and number	(39)
Increase in activity of two RNA polymerases	(37)
Increase in synthesis of contractile proteins	(40)
Imbibition of water	(41)
Increased synthesis of many phospholipids	(42)
Increased de novo synthesis of purines (dependent on new enzyme synthesis)	(43)
Alteration in membrane excitability	(40)
Alteration in glucose metabolism	(44)
Increase in synthesis of various mucopolysaccharides	(45)

external agent to a sequence-specific site on the genome could lead to the activation of a large number of distant producer genes. This is exactly the role the sensor elements of this model carry out. Implicit in the available data on hormone action are genomic elements performing some of the functions of the producer and integrator genes of the model.

Evidence Suggesting the Existence of Activator RNA's

Many of the properties attributed to the RNA in our model are actually those of a certain class of RNA molecules already described extensively; yet no known function has so far been attributed to this class of RNA's. The activator RNA molecules of the model have the following properties that can be tested. (i) They will, in the main, be confined to the nucleus, that is, they are not precursors of cytoplasmic polysomes. (ii) When observed in their functional role, they would be found in chromatin, bound to DNA in a sequence-specific manner. (iii) They are often the product of the redundant fraction of the genome. (iv) They include sequences not present in the polysomes carrying producer-gene templates, that is, most or all cytoplasmic polysomes. Table 4 summarizes some recent studies of RNA's which seem to fulfill condition (i). These RNA's have the suggestive properties of nuclear location, heterogeneity, and probable lack of "precursor" relation to cytoplasmic polysomal templates.

The hybridization experiments of McCarthy and Shearer (*22*) (Table 4) were performed at relatively low concentrations of nucleic acid and at short incubation time. Therefore, the RNA's they describe are the products of the redundant fraction of the genome. The presence of sequences specific to the nucleus and their absence from the cytoplasm is indicated by competition experiments. Furthermore, the nuclear RNA's contain sequences binding as much as five times more DNA than the cytoplasmic RNA at empirical saturation of the DNA with RNA. These and the other data of Table 4 show that RNA's are already known which might fulfill the functions we have assigned to activator RNA's.

At the heart of the model regulation system lies nucleus-confined RNA which determines the pattern of cellular gene activity, and this remains a key area of uncertainty. Bekhor, Kung, and Bonner (*10*), and also Huang and Huang (*10*), have presented evidence suggesting that sequence-specific binding between chromosomal RNA and genomic DNA determines the sites at which the transcription-inhibiting chromatin proteins bind to the DNA (*23*). Thus, according to these experiments, sequence recognition between the special chromosomal RNA's and the DNA specifies the pattern of gene activity (*10, 23*). Furthermore, the chromosomal RNA's have the property of binding, in what is apparently a sequence-specific way, to double-stranded native DNA (*7*). The significance of this line of investigation for experimental test of the idea of activator RNA's is obvious.

Large Changes in Transcription of Redundant Sequences

It is a striking fact that very large changes in the spectrum of RNA's deriving from repetitive sequences are observed when the state of differentiation alters. This knowledge is derived from RNA–DNA hybridization experiments carried out at relatively low concentrations of nucleic acid and short annealing times, so that reaction of RNA with any but the repetitive sequences in the genome is precluded. The spectrum of RNA's present or in the process of being synthesized in different tissues (*4, 24*), both in hormone response (*25*) and in embryonic development and differentiation (*26, 27*), has been investigated with competition procedures. In these experiments RNA from a cell type in one state of differentiation is used to compete with RNA from a cell type in another state of differentiation for binding sites in the repetitive fraction of the DNA. This type of analysis has shown that different families of repetitive genomic sequences are represented in the RNA of cells in diverse states of differentiation. Changes as large as 100 percent (apparent complete lack of homology) in the measured RNA's have been observed—for example, in successive stages of the embryogenesis of *Xenopus* (*27*). It is not particularly obvious why such changes should be detected, since the populations of producer genes active in each state of differentiation might be expected in general to be strongly overlapping. One possible explanation would be that much of the pulse-labeled RNA monitored in these studies is the rapidly turning over product of different regulatory genes such as the integrator genes of this model.

Regulatory Genes Known in Higher Organisms

The model suggests that a sizable portion of the functional genes in differentiated cell types may be regulatory genes (integrator and receptor genes). If this is so, it might be expected that, despite the difficulty of detecting such genes with classical genetic procedures, a certain number of apparent regulatory mutations would be known in higher organisms. The distinguishing characteristic of such regulatory loci would be pleiotropic effects on the activity of a number of producer genes, particularly with reference to a pattern of integration on the part of the latter. A number of good cases of this genre actually exist, particularly for drosophila and maize. A notable example is the Notch series of x-chromosome deficiencies (*28*), some of which are sharply localized. Notch mutants display a very large variety of developmental abnormalities—all affecting early embryonic organization—for example, failure to form a complete gut, failure of meso-

Table 4. Nucleus-confined, apparently heterogeneous RNA's of unknown function.

Source	Size of RNA	Compositional peculiarity if any	Turnover	Evidence against precursor relation with cytoplasmic RNA	Reference
HeLa cell nuclei	Heterodisperse 10*S* to 65*S*	29 to 32% uridine	Rapid	Composition; size; absence of association with nascent proteins	(*46*)
L-cell nuclei			Rapid	Sequences present in nuclear RNA absent in cytoplasmic RNA	(*22*)
Kidney and liver cell nuclei				Sequences present in nuclear RNA absent in cytoplasmic RNA	(*47*)
Reticulocyte nuclei	Heterodisperse 30*S* to 80*S*	29 to 31% uridine	Rapid	Kinetics (cytoplasmic mRNA does not turn over at all); size; base composition	(*48*)
HeLa cell nuclei	100 to 180 nucleotides	Extensive methylation	Extremely low	Composition; small size	(*49*)
Pea seedling nuclei	40 to 60 nucleotides	Presence of dihydropyrimidines		Sequences present in nuclear RNA absent from cytoplasmic RNA; composition	(*50*)

dermal differentiations to occur, overly large neural structure, and subnormal ectodermal skin production. Their effects are clearly pleiotropic. The multiplicity of the actual primary failures of these mutants is unknown. That is, no comparison can be made of the number of diverse producer genes affected simultaneously, as opposed to the array of sequential effects that follow the initial primary effects. Nonetheless, the effect of the Notch genes on the organization of the embryo is consistent with what would be expected of mutations in integrator gene sets. Many similar cases are known in which specific organizational lesions result from simple mutations affecting a small region of the genome (*29*). Studies with drosophila imaginal disk cell determination and transdetermination carried out by Hadorn and his associates (*30*) also demonstrate the existence of an apparatus in the genome for specifying integrated patterns of activity in the various cell types deriving from the disk cells. In experimental imaginal disk systems, highly exact specification of the patterns of producer gene activity is heritable through many cell divisions and is separated in time from producer gene function per se (that is, manifest differentiation).

Genes are known in maize which display control over producer genes and are located in the genome at sites distant from the producer genes that they control (*31*). In addition, McClintock and others have demonstrated the presence of other control sites adjacent in the genome to the same producer genes as those controlled by the distant regulatory elements (*31*). Control of the expression of the producer gene is accomplished through the interaction of the distant regulatory gene with the contiguous regulatory gene. This point has been demonstrated by insertion of the contiguous regulatory genes at different sites in the genome, near known genes, which then respond to the same control system governed by the distant regulatory unit. An example is the system termed Ac-Ds. Here the distant regulatory element Ac (which behaves as an integrator gene of this model) can be made to govern producer genes in other chromosomes such as the gene series for synthesis of anthocyanin pigment. Establishment of Ac control over the pigment synthesis system is accomplished by transposing the contiguous regulatory element responsive to Ac (Ds) to the loci of the anthocyanin producer genes (Ds thus behaves like a receptor gene of this model). In several ways, these and other data presented by McClintock (*32*) would seem easily to fit a model such as that presented here.

DNA Sequence Repetition

The existence of repeated sequences in higher organisms led us independently to consider models of gene regulation of the type we describe here. This model depends in part on the general presence of repeated DNA sequences. The model suggests a present-day function for these repeated DNA sequences in addition to their possible evolutionary role as the raw material for creation of novel producer gene sequences. The apparently universal occurrence of large quantities of sequence repetition in the genomes of higher organisms (*14*) suggests strongly that they have an important current function.

The quantity of DNA in repeated sequences, the frequency of repetition (that is, number of times a given sequence is present per genome), and the precision of the repetition show great variation among species. Frequencies from 100 to 1,000,000 have been observed, and the quantities of DNA involved range from 15 to 80 percent of the total DNA. The usual relation between repeated sequences is not that of a perfect copy (*33*), but the sharing of most of the nucleotides in a sequence extending for at least a few hundred nucleotides. Repeated sequence families in the DNA are observed, with degrees of similarity varying from perfect matching to matching of perhaps only two-thirds of the nucleotides. Expression of families of repeated sequences by transcription into RNA shows tissue specificity (as mentioned above) in spite of the fact that the individual families contain these widely divergent sequences.

In the cases studied there is good evidence that the repeated sequences are scattered throughout the DNA. For example in bovine DNA, 75 percent of all fragments about 5000 nucleotides long contain a segment of repeated DNA (*34*). When the fragment size is reduced to about 500 nucleotides, only 45 percent contain repeated sequences. Therefore, the typical bovine DNA fragment of 5000 nucleotides is a composite of lengths of repeated sequence and nonrepeated sequence. For longer fragments (20,000 or so nucleotides), there is suggestive evidence (*14*) that more than 95 percent contain repetitive sequences. Therefore, for bovine DNA (and probably that of other organisms) repeated sequences are intimately interspersed with nonrepeated sequences, throughout the length of the genome. This is precisely the pattern required in our model if repeated sequences are usually or often regulatory in function.

Evolutionary Implications of the Model

Any evolutionary changes in the phenotype of an organism require, in addition to changes in the producer genes, consistent changes in the regula-

tory system. Not only must the changes be compatible with the interplay of regulatory processes in the adult, but also during the events of development and differentiation. At higher grades of organization, evolution might indeed be considered principally in terms of changes in the regulatory systems. It is therefore a requirement of a theory of genetic regulation that it supply a means of visualizing the process of evolution.

Inactivity of New Genetic Material

A characteristic of this model is that DNA sequences are inactive in transcription, unless specifically activated. Thus the genome of an organism can accommodate new and even useless or dangerous segments of DNA sequence such as might result from a saltatory replication (*35*). Initially these sequences would not be transcribed, and thus would not be subject to adverse selection. Only by inclusion in integrated producer gene batteries (through translocation of receptor genes) would their usefulness as producer genes be tested.

Formation of New Integrative Relations

A peculiar combination of conservatism and flexibility is supplied by the model system. Preexisting useful batteries of genes will tend to remain integrated in function. At the same time, there is the potentiality of formation of new integrative combinations of preexisting producer genes. These combinations would be the result of translocations, principally among the integrator gene sets. Less often, new producer gene batteries would result from events in which receptor genes are translocated into positions contiguous to other producer genes.

We visualize many of the integrator genes and receptor genes as being members of families of repeated DNA sequences. It is known that new repeated sequence families have originated periodically in the course of evolution (*35*). The new families of repeated sequences might well be utilized to form integrator and receptor gene sets specifying novel batteries of producer genes. Thus saltatory replications can be considered the source of new regulatory DNA. All that is required for regulatory function in this model is sequence complementarity (translocation of members of the same repetitive sequence family to integrator and receptor positions). Almost any set of nucleotide sequences would suffice. *The likelihood of utilization of new DNA for regulation is thus far greater than the likelihood of invention of a new and useful amino acid sequence*, since for the latter case great restrictions on the nucleotide sequence exist.

Changes in the integrator systems make possible the origin of new functions and possibly even of new tissues and organs. In other words, the model supplies an avenue for the appearance of novelty in evolution by combining into new systems the already functioning parts of preexisting systems.

Divergence within Repeated Sequence Families

Individual sequences may differ from others in a family as a result of many base changes. We presume that binding of activator RNA to the receptor genes will occur for a degree of sequence homology far short of perfect complementarity. However, at some degree of divergence, binding would be lost, and a producer gene would fail to be activated as a part of its previous battery. Eventually, the process of divergence might yield regulatory DNA in which the original patterns of repetition are no longer observable. In this way, nonrepeated (unique) regulatory DNA could arise, leading to the situation discussed earlier with respect to the fraction of the genome utilized for regulation.

The possibility of increasing sequence divergence among integrator and receptor genes suggests a novel evolutionary mechanism. The divergence of regulatory sequences can be expected to be reversible. If the degree of complementarity required for binding between activator RNA and receptor sequence is fairly low then a reasonably good probability would exist for a subsequent base change to restore the complementarity lost by an earlier change. Intermediate degrees of transcription of certain producer genes will probably result since sequences with a degree of complementarity near some critical value will bind only part of the time. Natural selection could then reversibly affect the integration of individual producer genes into batteries. The potentiality for smoothly changing patterns of integration among many sets of producer genes supplies a mechanism for direct adjustment by natural selection of the organization of systems of cellular activity. In other words, the model implies that selective factors can influence the integrative configurations in which an organism uses its genes.

The families of repeated sequences that appear and remain in the genome of a species affect the rate at which newly integrated systems of producer genes will arise. Thereby, the rate of evolution is affected. It follows that the rate of evolution will be acted on by natural selection.

The issues raised in considering the evolution of the regulatory systems themselves are of a magnitude which is really out of reach in this brief discussion. However, the model offers interesting and surprising predictions. The properties of the model regulatory system suggest that *both the rate and the direction of evolution (for example, toward greater or lesser complexity) may be subject to control by natural selection.*

Summary

A theory for the genomic regulation systems of higher organisms is described. Batteries of producer genes are regulated by activator RNA molecules synthesized on integrator genes. The effect of the integrator genes is to induce transcription of many producer genes in response to a single molecular event. Current evidence suggesting the existence of elements of this model is summarized. Some evolutionary implications are indicated.

References and Notes

1. The variable gene activity theory of cell differentiation was explicitly proposed in the early 1950's, by A. E. Mirsky [in *Genetics in the Twentieth Century*, L. C. Dunn, Ed. (Macmillan, New York, 1951), p. 127] and by E. Stedman and E. Stedman [*Nature* **166**, 780 (1950)]. T. H. Morgan, among several others, had earlier considered this idea [*Embryology and Genetics* (Columbia Univ. Press, New York, 1934)].
2. Evidence for the equivalence of differentiated cell genomes comes from a variety of sources, including regeneration experiments, nuclear transplantation experiments, early embryological studies in which nuclei were positioned in cells other than those normally receiving them, measurements of DNA content per cell, and so forth. [See E. H. Davidson, *Gene Activity in Early Development* (Academic Press, New York, 1968), pp. 3–9.] Critical recent evidence has been provided by Gurdon's demonstration that differentiated intestinal cell nuclei from *Xenopus* tadpoles can direct the development of whole frogs when reimplanted into enucleate eggs [J. B. Gurdon, *Develop. Biol.* 4, 256 (1962)], and by DNA reassociation studies in which the DNA of different tissues was shown to be indistinguishable in sequence conte➡[B. J.

McCarthy and B. H. Hoyer, *Proc. Nat. Acad. Sci. U.S.* **52**, 915 (1964)].
→ V. G. Allfrey and A. E. Mirsky, *Proc. Nat. Acad. Sci. U.S.* **44**, 981 (1958); *ibid.* **48**, 1590 (196→ V. C. Littau, V. G. Allfrey, A. E. Mirsky, *ibid.* **52**, 93 (1964); J. Bonner, M. E. Dahmus, D. Fambrough, R. C. Huang, K. Marushige, Y. H. Yuan, *Science* **159**, 47 (1968).
4. J. Paul and R. S. Gilmour, *Nature* **210**, 992 (196→ B. J. McCarthy and B. H. Hoyer, *Proc. Nat. Acad. Sci. U.S.* **52**, 915 (1964).
5. F. Jacob and J. Monod, *J. Mol. Biol.* **3**, 318 (1961).
→ I. Bekhor, J. Bonner, G. K. Dahmus, *Proc. Nat. Acad. Sci. U.S.* **62**, 271 (1969).
7. In this case it would be desirable to use a distinctive term other than producer gene for the DNA sequences coding for such regulative protein molecules. The producer genes would then be defined as those sequences coding for RNA's other than those translated to produce proteins which recognize the receptor sequences.
8. J. H. Frenster [*Nature* **206**, 1269 (1965)] has proposed that RNA could act as a "derepressor." There is only a formal similarity to one part of this model—the activator RNA. An RNA molecule, fully complementary to the "unread" strand was supposed to maintain the DNA of a gene in a strand-separated condition and thus permit its transcription.
9. E. W. Johns and J. A. V. Butler, *Nature* **204**, 853 (1964); K. Murray, *Ann. Rev. Biochem.* **34**, 209 (1965); L. S. Hnilica, *Biochem. Biophys. Acta* **117**, 163 (1966); and H. A. Kappler, *Fed. Proc.* **24**, Pt. I (2), 601 (1965); J. L. Beeson and E. L. Triplett, *Exp. Cell. Res.* **48**, 61 (1967); D. M. Fambrough and J. Bonner, *J. Biol. Chem.* **243**, 4434 (1968); R. J. deLange, D. M. Fambrough, E. L. Smith, J. Bonner, *J. Biol. Chem.* **243**, 5906 (1968). In addition to lack of diversity of tissue specificity of histones shown by the above references, information also exists which indicates directly that histone-DNA recognition does not control the sequence-specific pattern of transcription in chromatin in vitro. [See J. Paul and R. S. Gilmour, *J. Mol. Biol.* **34**, 305 (1968)].
10. I. Bekhor, G. M. Kung, J. Bonner, *J. Mol. Biol.* **39**, 351 (1969); R. C. Huang and P. C. Huang, *ibid.*, p. 365.
→ A. E. Mirsky and H. Ris, *J. Gen. Physiol.* **34**, 451 (1951).
12. A. E. Mirsky, *Harvey Lec.* **Ser. 46**, 98 (1950–51).
13. Species whose genome sizes appear in Fig. 3 are as follows (from top to bottom): *Bos bos* [C. Leuchtenberger, R. Leuchtenberger, C. Vendrely, R. Vendrely, *Exp. Cell Res.* **3**, 240 (1952)]; *Chelonia mydas* [H. Ris and A. E. Mirsky, *J. Gen. Physiol.* **33**, 125 (1949)]; *Scaphiopus couchi* [E. Sexsmith, thesis, Univ. of Toronto (1968)]; *Tetraodon fluviatilis* [R. Hinegardner, *Amer. Natur.* **102**, 517 (1968)]; *Carcharias obscurus* (*11*); *Lampetra planeri* and *Amphioxus lanceolat*→[N. B. Atkin and S. Ohno, *Chromosoma* **23**, 10 (1967)]; *Acidea atra* (*11*); *Paracentrotus lividus* [R. Vendrely, *Compt. Rend. Soc. Biol.* **143**, 1386 (1949)]; *Cassiopea* and *Dysidea crawshagi* (*11*); *Amoeba hystolytica* (A. Gelderman, unpublished data); *Sacharomyces* [M. Ogur, S. Minckler, G. Lindegren, C. G. Lindergren, *Arch. Biochem. Biophys.* **40**, 175 (1952)]; *Escherichia coli* [J. Cairns, *J. Mol. Biol.* **4**, 407 (1962)]; *Mycoplasma* [H. Bode and H. J. Morowitz, *ibid.* **23**, 191 (1967)]; *Simian Virus* 40 [T. Ben-Porat and A. S. Kaplan, *Virology* **16**, 261 (1962)].
14. R. J. Britten and D. E. Kohne, *Science* **161**, 529 (1968).
15. In addition to calf and mouse (*14*), the following species have been shown to contain relatively large quantities of nonrepetitive DNA: *Xenopus laevis* (E. H. Davidson and B. R. Hough, *Proc. Nat. Acad. Sci. U.S.*, in press); Sea urchin and chicken (D. E. Kohne, unpublished data); *Drosophila*, 3 species (C. D. Laird and B. J. McCarthy, in press); Human (R. J. Britten, unpublished data). Even *Amphiuma* with a genome 30 times the size of the mammalian appears to have about 20 percent nonrepeated DNA [R. J. Britten, *Carnegie Inst. Wash. Year B.* **66**, 75 (1968)].
16. The choice of a mammal in this comparison could be objected to on the grounds that an enormous number of producer genes could be required for the construction of the super-complex central nervous system and immunoresponse systems of higher chordates. However, it seems unlikely to us that the enormous increase in genome size which has occurred in evolution could be explained in this way.
17. M. Dixon and E. C. Webb, *Enzymes* (Academic Press, New York, 1964).
18. C. J. Epstein and A. G. Motulsky, in *Progr. Med. Genet.* **4**, 97 (1965).
19. W. J. Young, *J. Hered.* **57**, 58 (1966).
20. Action of the steroid hormones aldosterone, cortisone, testosterone, and estrogen are known to depend on gene activity for their effects in target tissues: for testosterone, S. Liao, R. W. Barton, A. H. Lin [*Proc. Nat. Acad. Sci. U.S.* **55**, 1593 (1966)] and S. Liao [*J. Biol. Chem.* **240**, 1236 (1965)] demonstrate increase in prostate template-active RNA, and actinomycin-sensitive increase in RNA synthesis, with change in base composition. Template activity of muscle chromatin is also increased [C. B. Brewer and J. R. Florini, *Biochemistry* **5**, 3857 (1966)] and testosterone binds to a protein or proteins of prostate chromatin in vivo [N. Bruchovsky and J. D. Wilson, *J. Biol. Chem.* **243**, 5953 (1968)]. A similar range of evidence exists for the effect of estrogen on uterus [see T. H. Hamilton, *Science* **161**, 649 (1968)]. Striking increase in RNA synthesis in uterus occurs almost instantly with estrog→[A. R. Means and T. H. Hamilton, *Proc. Nat. Acad. Sci. U.S.* **56**, 1594 (1966)] and the spectrum of gene products made is different [B. O'Malley, W. L. McGuire, P. A. Middleton, *Nature* **218**, 1249 (1968)]. Corticosteroids cause many liver enzymes to be synthesized [G. Weber, S. K. Srivastava, L. Radhey, *J. Biol. Chem.* **240**, 750 (1965); D. Greengard and G. Acs, *Biochim. Biophys. Acta* **61**, 652 (1962)] and cause sharp increases in liver nuclear RNA synthesis [G. T. Kenny and F. J. Kull, *Proc. Nat. Acad. Sci. U.S.* **50**, 493 (1963)], resulting from both increased polymerase [O. Barnabei, B. Romano, G. Bitono, V. Tomasi, F. Sereni, *Biochim. Biophys. Acta* **113**, 478 (1966); W. Schmid, D. Gallwitz, C. E. Sekeris, *ibid.* **134**, 85 (1967)] and from increased template activity in isolated chroma→[M. E. Dahmus and J. Bonner, *Proc. Nat. Acad. Sci. U.S.* **54**, 1370 (1965)]. The specific effect of aldosterone on isolated bladder is abolished by actinomycin, and this hormone localizes in the bladder cell nucl→G. A. Porter, R. Bogorach, I. S. Edelman, *Proc. Nat. Acad. Sci. U.S.* **52**, 1326 (1964); → I. S. Edelman, R. Bogorach, G. A. Porter, *ibid.* **50**, 1169 (1963)]. The polypeptide gonadotropins follicular stimulating hormone and lutenizing hormone stimulate protein synthesis in ovary in an actinomycin-sensitive way [M. Civen, C. B. Brown, J. Hilliard, *Biochim. Biophys. Acta* **114**, 127 (1966)], and adrenocorticotropin (ACTH) promotes actinomycin-sensitive increase in adrenal RNA synthesis and steroidogenesis [R. B. Farese, *Fed. Proc.* **24** Pt. I (2), 306 (1965)]. Mobilization of Ca^{++} from bone, promoted by the polypeptide parathyroid hormone, is prevented by actinomycin [H. Rasmussen, D. Arnaud, C. Hawker, *Science* **144**, 1019 (1969)]. Several specific plant hormones, including giberellins, auxin, and cytokinins [reviewed by A. Trewas, in *Progr. Phytochem.* **1**, 114 (1968)] as well as antheridogens (B. R. Voeller, personal communication) also appear to act at the genomic level in tissues affected by them. Transport of Ca^{++} mediated by vitamin D in intestinal mucosa is blocked by actinomycin [J. E. Zull, E. Czarnowska-Misztall, H. F. deLuca, *Science* **149**, 183 (1965)]. Various embryonic processes known to occur as a result of specific inducing agents [see review by T. Yamada, in *Compr. Biochem.* **28**, 113 (1967)] are known to be sensitive to actinomycin treatment of the competent tissues, for example, in mouse pancreas induction [N. K. Wessels and F. H. Wilt, *J. Mol. Biol.* **13**, 767 (1965)] and neurulation in the frog [J. Brachet, H. Denis, F. deVitry, *Develop. Biol.* **9**, 398 (1964)].
21. H. R. Maurer and G. R. Chalkey, *J. Mol. Biol.* **27**, 431 (1967).
22. R. W. Shearer and B. J. McCarthy, *Biochemistry* **6**, 283 (1967).
23. These experiments are carried out by dissociating the chromatin in high salt and allowing the DNA and the proteins to reassociate by gradually lowering salt concentration in the presence of 5*M* urea. If the chromosomal RNA (RNA found initially bound to chromatin) is destroyed by ribonuclease or zinc hydrolysis, or if conditions are such as to prevent sequence-specific binding of the RNA to the DNA during the reconstitution procedure, the proteins reassociate in random positions and the reconstituted chromatin makes all the species of gene products made by totally deproteinized DNA. If the chromosomal RNA is allowed to associate with the DNA during the reconstitution process, on the other hand, the proteins of the chromatin evidently return to positions on the DNA close to those in the starting material, so that the reconstituted chromatin now synthesizes a spectrum of RNA's homologous to that of the original chromatin.
24. N. Miyagi, D. Kohl, R. A. Flickinger, *J. Exp. Zool.* **165**, 147 (1967); H. Ursprung, K. D. Smith, W. H. Sofer, D. T. Sullivan, *Science* **160**, 1075 (1968).
25. B. O'Malley, W. L. McGuire, P. A. Middleton, *Nature* **218**, 1249 (1968).
26. R. B. Church and B. J. McCarthy, *J. Mol. Biol.* **23**, 459 (196→ V. R. Glisin, M. V. Glisin, P. Doty, *Proc. Nat. Acad. Sci. U.S.* **56**, 285 (1966); A. H. Whitely, B. J. McCarthy, H. R. Whitely, *ibid.* **55**, 519 (1966).
27. H. Denis, *J. Mol. Biol.* **22**, 285 (196→ E. H. Davidson, M. Crippa, A. E. Mirsky, *Proc. Nat. Acad. Sci. U.S.* **60**, 152 (1968).
28. D. F. Poulson, *Amer. Natur.* **79**, 340 (1945).
29. Some of the most striking examples are the *t* mutant series in the mouse [L. C. Dunn and S. Gluecksohn-Waelsh, *Genetics* **38**, 261 (1953); P. Chesley and L. C. Dunn, *ibid.* **21**, 525 (1936)]; the creeper gene in the chicken [W. Landauer, *J. Genet.* **25**, 367 (1932)]; the control of a variety of morphogenetic patterns in insects such as bristle formation and wing color pattern [see review by D. H. Waddington, *New Patterns in Genetics and Development* (Columbia Univ. Press, New York, 1962), chap. 6]. A spectacular case of organizational control loci is provided by the "bithorax" series in drosophila [E. B. Lewis, in *The Role of Chromosomes in Development*, M. Locke, Ed. (Academic Press, New York, 1964), p. 231].
30. S. Hadorn, *Brookhaven Symp. Biol.* **18**, 148 (1965).
31. B. McClintock, *ibid.*, p. 162.
32. ———, *Develop. Biol.* (*Suppl.*) **1**, 84 (1967).
33. Since the relationship is commonly imperfect and nothing is yet known of the functional importance of sequence multiplicity, the phrase "redundant sequences" cannot be applied in its strict sense. Nevertheless redundant and repeated are used interchangeably since no properly applicable word exists.
34. R. J. Britten, in preparation.
35. ——— and D. E. Kohne, *Carnegie Inst. Wash. Year B.* **66**, 83 (1968).
36. C. Rouiller, Ed., *The Liver: Morphology, Biochemistry, Physiology* (Academic Press, New York, 1963), vol. 1, chaps. 9–11.
37. T. H. Hamilton, *Science* **161**, 649 (1968).
38. M. W. Noall and W. M. Allen, *J. Biol. Chem.* **236**, 2987 (1961).
39. R. J. Laquens, *Ultrastructures Res.* **10**, 578 (1964).
40. A. Csapo, *Ann. N.Y. Acad. Sci.* **75**, 740 (1959).
41. G. Pincus, Ed., *The Hormones: Physiology, Chemistry, and Applications* (Academic Press, New York, 1964), vol. 5, pp. 787–793.
42. Y. Aizawa and G. C. Mueller, *J. Biol. Chem.* **236**, 381 (1961).
43. G. C. Mueller, J. Gorski, Y. Aizawa, in *Mechanism of Action of Steroid Hormones*, C. A. Ville and L. L. Engel, Eds. (Pergamon, Oxford, 1961), pp. 181–184.
44. J. A. Nicolette and J. Gorski, *Arch. Biochem. Biophys.* **107**, 279 (1964).
45. H. Sinohara and H. H. Sky-peck, *ibid.* **106**, 138 (1964).
46. R. Soeiro, H. C. Birnboim, J. E. Darnell, *J. Mol. Biol.* **19**, 362 (1966).
→ R. B. Church and B. J. McCarthy, *Proc. Nat. Acad. Sci. U.S.* **58**, 1548 (1967).
48. G. Attardi, H. Parnas, M-I. H. Hwang, B. Attardi, *J. Mol. Biol.* **20**, 145 (1966).
49. R. A. Weinberg and S. Penman, *ibid.* **38**, 289 (1968).
→ J. Bonner and J. Widholm, *Proc. Nat. Acad. Sci. U.S.* **57**, 1379 (1967); I. Bekhor, J. Bonner, G. K. Dahmus, *ibid.*, in press.