of the subject, including comparison with experiment.
30. L. Pauling and C. D. Coryell, *Proc. Natl. Acad. Sci. U.S.A.* **22**, 159 and 210 (1936).
31. L. Pauling, *J. Am. Chem. Soc.* **53**, 1367 (1931).
32. H. A. Jahn and E. Teller, *Proc. R. Soc. London* **161**, 220 (1937).
33. J. H. Van Vleck, *J. Chem. Phys.* **7**, 61 and 72 (1939).
34. C. J. Gorter, *Physica (Utrecht)* **3**, 503 (1936) and other later papers in that journal; *Paramagnetic Relaxation* (Elsevier, Amsterdam, 1947).
35. I. Waller, *Z. Phys.* **79**, 370 (1932).
36. W. Heitler and E. Teller, *Proc. R. Soc. London* **155**, 629 (1936); M. Fierz, *Physica (Utrecht)* **5**, 433 (1938); R. de L. Kronig, *ibid.* **6**, 33 (1939).
37. J. H. Van Vleck, *Phys. Rev.* **57**, 426 (1940); *ibid.*, p. 1052.
38. ———, *ibid.* **59**, 724 (1940); *ibid.*, p. 730.
39. W. F. Giauque and D. P. Mac Dougall, *ibid.* **43**, 768 (1933); *ibid.* **47**, 885 (1935); F. Simon, *Nature (London)* (1935), p. 763.
40. E. M. Purcell and M. H. Hebb, *J. Chem. Phys.* **5**, 338 (1937).
41. W. H. Zachariasen, *ibid.* **3**, 197 (1935).
42. A. H. Cooke, R. Lazenby, M. J. Leask, *Proc. Phys. Soc. London* **85**, 767 (1965).
43. E. H. Erath, *J. Chem. Phys.* **34**, 1985 (1961).
44. T. H. Maiman, *Nature (London)* **187**, 493 (1960).
45. R. Finkelstein and J. H. Van Vleck, *J. Chem. Phys.* **8**, 790 (1940); J. H. Van Vleck, *ibid.*, p. 787. These papers relate to chrome alum whereas Maiman used chromium embedded in Al_2O_3, but the spectroscopic properties of the chromium ion are similar in the two cases [see (*29*), p. 238].
46. See especially P. L. Scott and C. D. Jeffries, *Phys. Rev.* **127**, 32 (1962); R. R. Ruby, H. Benoit, C. D. Jeffries, *ibid.*, p. 51.
47. E. M. Purcell, H. C. Torrey, R. V. Pound, *ibid.* **69**, 37 (1946); F. Bloch, W. W. Hansen, M. Packard, *ibid.*, p. 127.
48. J. E. Pake and E. M. Purcell, *ibid.* **74**, 1184 (1948); *ibid.* **75**, 534 (1948). See also (*50*).
49. J. H. Van Vleck, *ibid.* **74**, 1168 (1948).
50. N. Bloembergen, E. M. Purcell, R. V. Pound, *ibid.* **73**, 679 (1948).
51. C. J. Gorter and J. H. Van Vleck, *ibid.* **72**, 1128 (1947). See also (*49*).
52. For a very complete review of all the different models and their limitations see C. Herring, in *Magnetism*, J. E. Rado and H. Subl, Eds. (Academic Press, New York, 1966), vol. 4.
53. E. C. Stoner, *Philos. Mag.* **21**, 145 (1936); *Proc. R. Soc. London Ser. A* **165**, 372 (1938); *ibid.* **169**, 339 (1939).
54. J. H. Van Vleck, *Rev. Mod. Phys.* **25**, 220 (1953); in *Quantum Theory of Atoms, Molecules and the Solid State*, P. O. Lowdin, Ed. (Academic Press, New York, 1966), pp. 475–484.
55. For an excellent discussion of the theory of the magnetic ordering in rare earth metals see R. J. Elliott, in *Magnetism*, J. E. Rado and H. Suhl, Eds. (Academic Press, New York, 1966), vol. 2, part A.
56. R. M. Bozorth, A. E. Clark, J. H. Van Vleck, *Int. J. Magn.* **2**, 19 (1972).

Irreversible Gene Repression Model for Control of Development

Lessening of developmental potential may result from progressive repression of previously active genes.

Arnold I. Caplan and Charles P. Ordahl

Summary. As the pluripotent cells of early embryos differentiate, each progressively loses the potency to develop into several phenotypes. Ultimately, each cell becomes irreversibly restricted to the expression of a single phenotype. Although in many instances details regarding those restriction events are well known, there is little information concerning the nature of the gene transcription changes involved. A model that accounts for the diminution of developmental potential as resulting from progressive, irreversible repression of previously active genes is presented. A scheme of progressive gene repression, rather than selective gene activation, is most consistent with observations from experimental embryology as well as from more recent biochemical experimentation.

Despite our relatively detailed understanding of molecular biology, the processes which control the development of a multicellular organism from a single cell, the fertilized egg, are almost completely unknown. The conceptual framework and basic understanding for analyzing these processes are just now becoming available. In this article we endeavor to construct a testable model for analyzing these complex developmental events at the molecular level. The model takes into account the strict constraints of the biologically observable phenomena and attempts to relate these phenomena to transcriptional events. Although this model takes the major biological phenomena into account, it is a model of the extreme: as a basic condition of the arguments developed here, we must confess complete ignorance of the primary event (or events) controlling molecular differentiation. It is clear that a wide variety of posttranscriptional and posttranslational (*1*) events play a role in cell differentiation and development of a multicellular organism. Although many mechanisms must be operating, the viewpoint developed is that the modification of the genome (chromatin) must ultimately be involved in the elaboration and eventual stabilization of these developmental events. Thus, the simplified model developed below concentrates on events at the level of transcription and presents the argument that the changes in the intra- and intercellular "environment" constitute the primary trigger for the architectural changes that must be mediated at the transcription level in order to support and direct complicated developmental events.

Transcriptional Mechanisms

Two mechanisms in the extreme can be used to explain how differential gene transcription can be involved in the control of the progression of cellular differentiation and development: the "activation-repression" (*2*) and the "irreversible repression" mechanisms. In the activation-repression mechanisms, genes that have not previously been transcribed are activated (or derepressed), and their transcripts and subsequent translational products appear for the first time in a newly differentiated cell. The second mechanism is one in which at early stages of development all genes are active or transcriptionally accessible and progressive, selective repression is the primary mechanism governing the differentiation regime. According to this model, as development proceeds, there are events that cause irreversible repression (*3*) of some of these genes while oth-

Arnold I. Caplan is an associate professor of biology and codirector of the Developmental Biology Center at Case Western Reserve University, Cleveland, Ohio 44106. Charles P. Ordahl is an assistant professor of anatomy at Temple University Medical School, Philadelphia, Pennsylvania 19140.

0036-8075/78/0714-0120$02.00/0

ers become stabilized in the unrepressed state and represent genes specific to the newly established phenotype. We present the case for the irreversible gene repression model because it best fits the available information from both embryological studies and studies of transcriptional changes. We also discuss a molecular model that provides an experimental framework for attempting to discriminate between these mechanisms.

Biological Case for the Irreversible Repression Model of Gene Regulation

Starting with the fertilized egg, development proceeds by replication of the entire genome and is followed by cell division. Thus, two important segregation events occur: (i) the segregation of the two daughter genomes into separate compartments, and (ii) the segregation of different cytoplasms from the fertilized egg into two different compartments. It is well known that in most eukaryotes, the daughter cells of the first cleavage are identical in their developmental potential (*4*). These cells, when separated from one another, are capable of producing normal offspring, which are about half the normal size (*5*). The next cleavage, that from the two- to the four-cell stage, also does not generally restrict developmental potential (*6*); in some animals each blastomere produces identical offspring if grown separately. After the second or third cleavage, however, many of the cells are no longer totipotent and are no longer capable of producing complete, normal animals. Thus, cleavage segregates the cytoplasm giving rise to cells of different developmental potentials. This important and fundamental developmental phenomenon produces daughter cells that have a restricted developmental potential. This restriction is irreversible in the sense that, under normal circumstances, only a part of the organism's genome will be expressed as development proceeds. Therefore, an important irreversible restriction of genomic potential has occurred during or directly after the cleavage of the egg cytoplasm.

As examples of the developmental importance of segregating nuclei into specialized cytoplasms, two specific cases of developmental programs, which are dependent upon these events, are cited: (i) the fertilized egg of the snail *Ilyanassa* and (ii) the developing fruit fly *Drosophila* eggs. In the case of *Ilyanassa*, a small part of the cytoplasm (referred to as the polar lobe) of the fertilized egg annexes itself to one quadrant of the cell

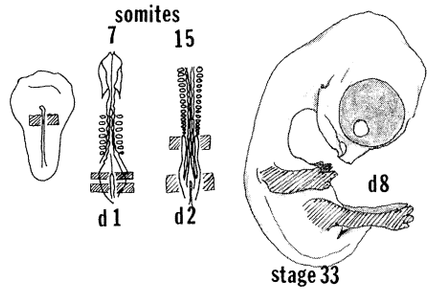

Fig. 1. Chick limb development. The relative locations of the "limb fields" in the developing chick embryo. Embryos at the primitive streak stage, 7-somite stage (*d1*), 15-somite stage (*d2*), and stage 33 (*d8*) are depicted with the shaded area indicating the approximate limb field.

prior to division. Subsequent to each division, this polar cytoplasm is reincorporated into one of the daughter cells. This process continues for a number of divisions. If the polar lobe is amputated, the embryo will continue to develop but will fail to form entomesoblast, and the resulting larvae completely lack mesoderm (*7*). Thus, the expression of those genes necessary to form an entire germ layer that will give rise to diverse cell types is wholly dependent on exposure of cell nuclei to a small portion of egg cytoplasm. In the absence of such exposure, the information for the formation of this germ layer remains quiescent.

In the case of *Drosophila*, another specific section of cytoplasm (also referred to as the "polar cytoplasm") influences the nuclei that migrate through it to differentiate into the reproductive cells. If this polar cytoplasm is removed from the egg so that migrating nuclei do not come in contact with such specialized cytoplasm, then the germ cells of *Drosophila* are absent, producing a sterile animal (*8*). Here again, polar cytoplasm is capable of influencing the genetic expression of nuclei that migrate through it, and therefore the segregation of this cytoplasm plays an important, if not determinant, role in the development of *Drosophila*.

As will be discussed below, the general state of the cell and the activity of the cell's cytoplasm provide important signals for the developmental programs. These signals, which are the sum of extrinsic and intrinsic processes, are, in all probability, the elusive embryonic inducers long sought by experimenters, but never found. It seems clear that the genome of a developing organism responds to positional or environmental cues (*9*); the responses to these cues are built into the genome and represent the genetic and evolutionary biases of that species. That these responses also bring about progressive restriction of the genetic potential, with these restrictions becoming irreversible, must also be a part of the genetic biases of the organism, with different species exhibiting different sequence and timing of such restrictive events, as is discussed below. The segregation of cytoplasm during cleavage establishes extranuclear environments that are deterministic to the developing organism and provide environmental cues that stimulate evolutionarily fixed genetic responses and programs.

Apart from the segregation of different cytoplasms with newly synthesized daughter nuclei, the further developmental restriction of groups of blastomeres can be observed. Just as individual blastomeres of the second or third cleavage are not capable of differentiating into complete organisms, groups of cells at later stages of development are not capable of differentiating into specific tissues. Detailed fate maps that depict this restriction of developmental potential have been drawn (*10*). As development proceeds, quadrants or groups of cells become "fated" to differentiate and develop into cells of a specific organ. For example, in the developing chick embryo, the cells predestined to form the limb can be recognized as a distinct group of cells very early in development (Fig. 1). These cells are incapable of forming tissue of the heart, for example, although their progenitor cells earlier in development were capable of forming either limb or heart tissue. Thus, as the development of an organism proceeds, developmental potential of each cell (that is, its ability to differentiate into a number of different phenotypes) is severely restricted. This restriction seems to be irreversible, and (as is discussed below) there is no evidence to support the general concept that normal cells, once committed to a specific pathway, are capable of "backtracking" to a more undifferentiated point in their developmental history where they are then capable of having a pathway choice again.

The work of Gurdon and his colleagues (*11*) and others (*12*) is often used to support the general contention that nuclei from highly differentiated tissues are still capable of supporting the development of a whole organism. However, these observations can also be used to support the general idea that irreversible gene repression is a common mechanism in the development and differentiation of tissues and organs. The Gurdon-type experiment (*11*) is to transplant a nucleus from a donor cell into an enucleated fer-

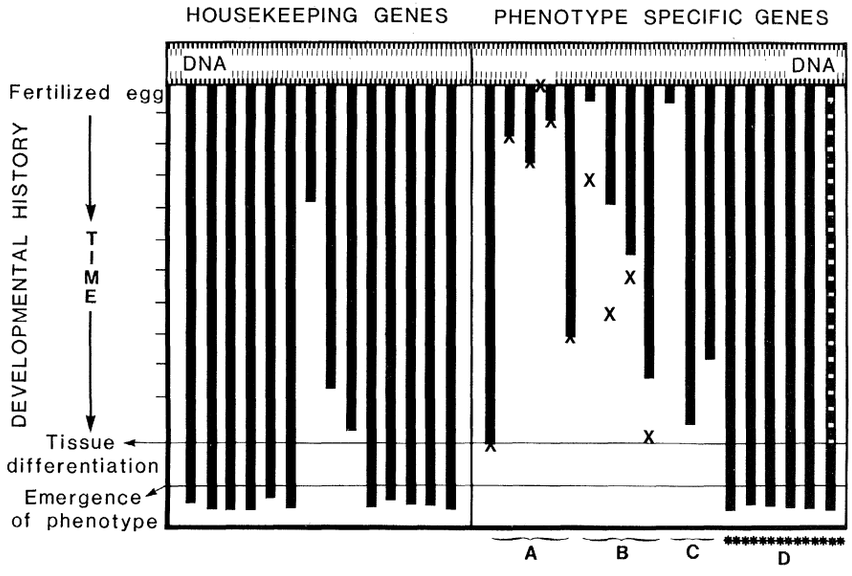

Fig. 2. The hypothetical transcription of "housekeeping" and "phenotype specific" genes is compared as a function of developmental time. As development proceeds, transcription is progressively restricted in both gene sets with X signifying irreversible repression of specific genes or gene sets. The term "tissue differentiation" refers to the time of irreversible commitment to a specific phenotype, with the subsequent recognition of those events noted as "emergence of phenotype."

tilized egg. When this experiment is done with nuclei obtained from early embryos, a very high percentage of the recipient eggs develop into complete, normal organisms. The cytoplasmic environment of a fertilized egg is therefore capable of interacting with a nucleus and allowing that nucleus to divide and to express information for a complete and normal organism. When developmentally older and more differentiated donor cells are used in this kind of experiment, the success rate of this nuclear transplantation experiment decreases to the point that, if intestinal cells from a young animal are used, only about one in 100 such transplantations result in a normal organism. Since the success rate is so high with nuclei from early embryos, we may assume that the experiment is not more difficult in terms of the nuclear transplantation manipulations with older tissues, but rather that a large fraction of the nuclei from cells of more developmentally advanced tissues have restricted potential and that the percentage of cells which retain a relatively unlimited potential progressively decreases as development proceeds. Implicit in this interpretation would be the view that, within any tissue of a developing organism, there persists a pool of cells with relatively unrestricted or uncommitted nuclei and that this pool decreases in size as the organism becomes more highly developed (*13*).

Therefore, one possible conclusion from the work of Gurdon (*11*) and others (*12*) seems to be that the developmental potential of any given cell or group of cells diminishes as the organism matures. We submit that this diminution of developmental potential is a result of irreversible gene repression and that, rather than gene activation, gene repression is generally responsible for the fixing of fate and the eventual expression of single phenotypes. The alternative explanation that repressed genes are activated in the transplanted nuclei, offers no explanation for why the percentage of successful transplants progressively decreases as donor nuclei from progressively older sources are used.

Relative to our naive state of knowledge about the molecular changes in chromatin which are correlated with a specific differentiation or developmental event, it is at present impossible for us to articulate or even envision, in a precise manner, the changes involved in the receiving and processing of developmental signals. However, it is unlikely that embryological events are stimulated or initiated by single factorial signals. The sum of many cytoplasmic and extracellular factors probably dictates or defines the developmental state of the cells (*14*). It is probable that a primary event in a cell's developmental history is a change in its external and subsequently internal environment, which is then communicated to the genetic material (*15*). The consequence of this transfer of environmental information to the genetic material is eventually an altered transcriptional pattern that changes the fate and future of that cell. In this context, environmental changes can be viewed as primary events, whereas transcription is a secondary or perhaps even tertiary event. The altered transcriptional pattern can change the cell and specifically cause it to become sensitive or insensitive to an environmental change and thus initiate a more complicated developmental processes all over again. Environmental changes may be as simple as the increased ratio of cell membrane to cytoplasm that follows a cell division. This increased ratio permits an increase in oxygen transport to cytoplasmic enzymes which, in turn, change the energy transducing capacities of the cell. This new developmental state, brought about by a cleavage division, may dictate an altered pattern of gene transcription. Another example would be a daughter cell that finds itself on the inner aspects of the embryo because of the position of its parent's cleavage plane and may be specifically sensitive or insensitive to this environmental change as determined by its present tissue position, history, and developmental potential.

Because of the lack of information concerning the ways in which primary environmental changes affect genomic function, we are unable to make models of control systems governing these events. What we can do, however, is to attempt to make models of the subsequent transcription events to explain the coarse controls at the level of the cell's genetic material. Even here the models are the "on-off" variety because (as is described below) our knowledge of how chromatin structure is related to gene expression is still rudimentary.

The Model

The experimental observations at the biological level place a few severe constraints or boundaries on the construction of a suitable model. First, the gradual restriction of developmental potential must be explained. Furthermore, a mechanism for making these restrictions irreversible must also be provided; such restrictions must correlate with cell "fating" or determination, which is known to begin as early as cleavage. Because developmental progression is an inherited process initiated during cleavage, we must assume that the control elements reside in the DNA and thus in the DNA-protein interaction related to chromatin structure and function. Second, the biological observations reveal

that the acquisition of phenotypic properties is a relatively late event and that this can be generally recognized by the synthesis and concentration of phenotype-specific macromolecules; in the case of lens development, the accumulation of alpha-crystalline proteins or in the case of pancreas development, the accumulation of insulin serve as examples. And third, all cells share common activities (for example, glycolysis, electron transport, and protein synthesis) and structures (such as mitochondria, endoplasmic reticulum, and plasma membrane) whose elements must be coded for by the same genes. These observations can be used to argue that at least two classes of gene products are present: (i) a group that codes for "housekeeping" proteins or molecules necessary for cell survival and common to all phenotypes (*16*) and (ii) a group coding for phenotype-specific proteins whose presence in sufficient concentrations dictates the phenotypic properties of a specific cell (*17, 18*).

The broad outlines of the model are diagramed in Fig. 2. The two classes of genes, those coding for housekeeping molecules and those coding for phenotypic specific molecules, are indicated as separate regions of the genome. Although these regions are diagrammatically depicted as separate blocks, it is unlikely, from available information, that these gene regions are present as contiguous blocks in eukaryotic DNA (*19*). Developmental time follows the vertical axis starting with a fertilized egg and progressing through the emergence of a single specific phenotype; thus a similar but individual pattern would exist for each different phenotype. In the case of most of the housekeeping genes, transcripts are continually being produced throughout the developmental history of all cells.

There are other genes that, because of the environmental circumstances in which the cell finds itself, are not necessary for cellular function and transcription of these sequences is very infrequent at different stages during development. However, the lack of frequent transcription of these genes is not irreversible and, under appropriate circumstances, the genes again can be transcribed at high frequency. These are indicated by the vertical bars which do not pass through the tissue differentiation or phenotypic emergence stage. An example of this might be a lactate dehydrogenase (LDH) isozyme that may be virtually absent during a relative aerobic phase of tissue growth but may be needed in quantity as the aeration state of the tissue is modified (*20*).

The phenotype-specific genes are classed into three general groups. The first group, indicated as group A, are genes whose transcription is repressed at various stages during development and the irreversible repression of these genes (indicated by X) is coincident with the cessation of their transcription. A second set of genes, represented by group B, are genes whose transcription is terminated during various stages of development, but the irreversible restriction of this transcription takes place at some time after the genes have ceased being actively transcribed. The third group, group C, of phenotype-specific genes represents genes whose transcription has been restricted but the restriction never becomes irreversible. These group C genes, for example, might be reactivated at some later stage of cell function either in response to hormonal stimulation or as a regeneration response to injury. The last group of genes, group D, are genes whose transcription is never repressed and whose eventual presence in a cell determines its phenotype. The last bar in this gene set (represented by the dashed line in Fig. 2) indicates the possibility that a gene has been available for transcription but has been transcribed at low frequency through earlier stages in development; the increase in frequency of transcription is coincidental with the tissue differentiation phase as the cells pass through the commitment phase (*21*).

Thus, as a cell emerges into a new specific phenotypic compartment, there are two groups of transcripts which can be found: (i) one that represents the housekeeping gene transcripts and (ii) those that represent transcripts from genes that are specifically responsible for the phenotypic characteristics of that cell. Irreversible repressional events have excluded the transcripts of some genes in the cell and since these restrictions are irreversible, the developmental potential of the cell is correspondingly restricted.

Transcriptional Considerations

This model predicts that differential gene expression during development arises by restricting or narrowing the transcriptional potential of cells by selectively and irreversibly repressing genes that were previously active. Thus, a cell's developmental potential is directly related to the diversity of transcriptional accessibility in its genome. The epitome of developmental potentiality is the fertilized egg and, therefore, according to this model, the genome of the fertilized egg should have all genes accessible to the transcriptional machinery. It is essential to state that, while the model predicts that all genes are accessible to transcription in the fertilized egg, it cannot predict that a transcript from every gene be present in a single fertilized egg. This distinction is important because, as is discussed below, it is unrealistic to expect any single cell to possess a single transcript from every gene in addition to the multiple (up to hundreds of thousands of) transcripts from some genes, such as ribosomal RNA genes. It may seem trivial to make a distinction between a "repressed gene" and an "untranscribed unrepressed" gene. However, the model attempts to equate *developmental potential* with *transcriptional potential*, and this latter term pertains to the developmental state of the chromatin rather than to its functional products at any specific time. To be testable, of course, this model must predict that the unrepressed genes (categories C and D in Fig. 2) are transcribed to some extent and that, given a sufficiently large population of fertilized eggs, there should be at least one transcript from each gene.

The fertilized egg, in this sense would be comparable to the bacterium *Escherichia coli*, which transcribes at least 96 percent of its genetic information but transcribes approximately a third of this information only once per 300 to 1000 cells (*22*). These very rare transcripts in *E. coli* are probably from "uninduced" genes, for which it is advantageous to maintain transcriptional accessibility in case environmental conditions change. Thus, in *E. coli* at least, genes not essential at a given moment are accessible for transcription at low levels.

The question arises as to whether it is valid to draw an analogy between the manner in which *E. coli* meets environmental uncertainties and the way embryonic cells meet developmental uncertainties. One straightforward approach to test both this analogy and our model would be to measure the RNA diversity at various stages of embryonic development to determine whether transcriptional diversity decreases as cells become developmentally restricted. At present, a limited number of measurements of this type which address the prediction of the model have been made (*16, 23–27*). The most comprehensive analyses have been made with sea urchin embryos by Britten, Davidson, and coworkers showing that polysome associ-

Table 1. Computation of cellular RNA populations.

RNA class	Species* (No.)	Percentage of total cell RNA mass†	Mass per cell in nucleotides‡	Diversity in nucleotides	Copies of each species per cell	Percent of genome sense strand§
Ribosomal RNA	3	80	8×10^9	7×10^3	1.14×10^6	7×10^{-4}
Transfer RNA	50	15	1.5×10^9	5×10^3‖	3×10^5	5×10^{-4}
Diverse RNA	~10^6¶	5	5×10^8	1×10^9	0.5	100
Prevalent copy dRNA	400¶	4	4×10^8	4×10^5	1×10^3	0.04
Single copy dRNA	10,000¶	1	1×10^8	1×10^7	10	1
Rare copy dRNA	1×10^6¶	0.1	1×10^7	1×10^9	0.01	100

*Characterization of the number of species is approximate and does not reflect complex families of tRNA or rRNA. †Based on estimates from (*37*). ‡Based on chick cell with approximately 1×10^{10} nucleotides of RNA mass per cell and a nonrepetitive sense strand genome size of 1×10^9 nucleotides [1.15×10^9 nucleotide pairs (haploid) (*36*)]; 85 percent of which is nonrepetitive DNA (*25*). §Nonrepetitive. ‖Here 100 nucleotides are used as an average-sized tRNA. ¶Using 1000 nucleotides as an average transcript size.

ated messenger RNA (mRNA) diversity is highest at the gastrula stage and progressively decreases in later stages of development. Other RNA-DNA hybridization analyses with specific embryonic tissues show that, with the possible exception of the mammalian brain (*27*), all tissues lose transcript diversity as they proceed to differentiate (*16, 23–27*). While most of these data are consistent with the prediction of the model they cannot, in themselves, support the model because all studies so far could not have detected transcripts from genes transcribed at low frequency which are of principle interest in the context of this model. Three problems must be addressed.

1) How many transcribed genes are there in eukaryotic genomes?

2) Assuming all diverse genes are transcribed, how many cells would be required to provide at least one transcript of each gene?

3) Given the expected relative abundance of the low frequency transcripts, would they be detectable by current technology?

The problem of the number of transcribed genes in a eukaryote has been the subject of debate (*28–30*). The problem arises in part from the observation that some amphibians have more DNA per nucleus than man (*31*), which is contrary to the usual concept of the relationship between "complexity" and DNA content, since it is clear that man is more "complex" than the frog. This has led some to postulate that most of the DNA in eukaryotes serves as a mutation sink to provide genetic flux outside the sphere of selection pressure (*30*). Similar arguments were put forth for prokaryotes a decade ago (*32*) when it was thought that only 10 to 20 percent of the *E. coli* genome was expressed. The suggestion that only a fraction of the eukaryotic genome is expressed is also supported, however, by the finding that the sequence divergence of genes coding for presumptive messenger RNA (mRNA) is evolutionarily slower than that of the bulk of the nonrepetitive DNA (*33*). These arguments, in the context of a "gene" defined only in terms of coding of an amino acid placement in polypeptides, indicate that, as an upper limit, eukaryote genomes contain 100,000 genes (*28–30*), which is equivalent to less than 10 percent of mammalian sense strand DNA. However, mouse brain polysomes contain enough sequence diversity to code for 100,000 different average size proteins (*34*). Thus, measurement of gene diversity transcribed in a single, albeit complex, tissue exceeds the upper limit, based on these theoretical considerations.

This problem is more complex if the definition of a gene is relaxed to include any diverse DNA sequence that is transcribed in vivo. Eukaryotic cells transcribe enormously diverse RNA populations (*24, 25, 34, 35*) which, in some cases, is equivalent to at least 40 percent of the total genomic content. The fact that we know virtually nothing about the function of most of those diverse transcripts does not lessen their probable importance to the cell, which expends considerable energy in maintaining their turnover. If mouse brain transcribes 40 percent of its genome (*34*), then it must be considered an open question as to whether all the diverse sense strand DNA sequence codes for RNA at some time during ontogeny.

Assuming then that all diverse (nonrepetitive) sequences in eukaryotic DNA are transcribed at some time and that in very early embryonic cells all the DNA is accessible to transcription, then how frequently are these diverse genes transcribed? The constraints of RNA mass become important here because at least 90 to 95 percent of the cellular RNA is rRNA and transfer RNA (tRNA) (*36*). The diverse transcript population must fit within approximately 5 percent of the cellular RNA mass.

In order to make estimates of the probable abundance of diverse RNA in the cell, we modeled an avian cellular RNA population (Table 1). Basically, there are three distinct populations; rRNA, tRNA, and diverse RNA (dRNA). Both the rRNA and tRNA are predominately in the cell cytoplasm but the dRNA is distributed in both the nucleus and cytoplasm. The relative mass of RNA in each population corresponds to the values of Soeiro *et al.* (*37*) and, if anything, is an underestimate of the amount of dRNA in an embryonic cell. If all genes were transcribed to an equal extent, the relative amount of dRNA could accommodate a single transcript from the entire nonrepetitive DNA fraction per two cells. However, not all genes are transcribed to the same extent in a given cell and transcripts from some genes are present in thousands of copies per cell. To obtain a more reasonable picture of the cellular dRNA population, we compartmentalized it into "abundant copy," "single copy," and "rare copy" classes (Table 1). The relative abundance and diversity of the abundant- and single-copy classes correspond to what might be expected for mRNA in adult or cultured eukaryotic cells (*38–40*). As a class, the total cell dRNA is probably even larger in mass and far more diverse (*34, 41*). What is of particular interest is that, even with a tissue culture cell, transcripts from the entire nonrepetitive fraction of the genome could be present. This fraction, if present at an average transcript copy frequency of one per 100 cells, would constitute only 0.1 percent of the total RNA, and, if present as one copy per ten cells, would constitute only 1 percent of the total cell RNA mass.

These calculations only establish the possibility that RNA transcripts from the entire nonrepetitive sense strand DNA could exist in eukaryotic cell populations. The relative mass of this population (<1 percent of the total RNA) would be too small to affect measurements of cellular RNA content or base composition. It is also unlikely that sequences present as one copy per ten or per 100 cells would be detected by typical saturation hybridization experiments (*26, 42*), although the existence of such a diverse class of RNA is suggested by some experiments. These arguments, while they do not argue in favor of the model, are important because they show

that such considerations do not preclude transcription of the entire genome in eukaryotic cell populations.

Although available estimates of transcript diversity of RNA from adult and embryonic cells that have been made by saturation hybridization to nonrepetitive DNA (*16, 23–27*) generally support the broad outline of the proposed model, these measurements probably only detect RNA present at concentrations greater than one copy per cell. Using the cell RNA populations modeled in Table 1, we would predict that the rare-copy dRNA would go virtually undetected in most, if not all, of these experiments. Table 2 shows the expected hybridization $C_{R0}t$ values (*28, 43*) necessary to hybridize maximally dRNA. These calculations were based on published values for hybridization kinetics as a function of sequence complexity (*44–46*). Those RNA species present in total cell RNA at a frequency of once per 100 cells (as in Table 1) would saturate 90 percent of complementary DNA sequences at $C_{R0}t > 10^6$. Very few saturation hybridization experiments achieve $C_{R0}t$ values greater than 5×10^4, and hence, at best, probably quantitate only single-copy dRNA.

However, diverse species may be up to 20- to 50-fold more abundant in RNA prepared from nuclei (*34, 37*) because of the elimination of much rRNA. This increase in relative concentration would commensurately lower the $C_{R0}t$ values necessary to detect these transcripts by hybridization (Table 2). Experiments with nuclear RNA (*26, 34, 41*) reveal levels of diversity up to ten times higher than when cytoplasmic RNA is analyzed. Analysis of nuclear RNA from mouse brain (*34*) indicates that up to 40 percent of sense strand nonrepetitive DNA is transcribed in this organ. Taking into account the high $C_{R0}t$ values obtained ($\sim 5 \times 10^4$) and the enrichment of dRNA in the nucleus, it is possible that these measurements approach full saturation for sequences present as one copy per 100 nuclei (Table 2). Whether these measurements reflect the full diversity of RNA in mouse brain remains to be seen. Since, however, the adult brain is a highly diverse organ, then 40 percent sense strand transcription is not an unrealistic maximum transcription level in the context of the model we postulate.

In sea urchin gastrulas, measurement of nuclear RNA complexity reveals a diverse class of RNA ten times more complex than the cytoplasmic mRNA class (*26*). Analysis of the hybridization kinetics of this RNA class indicates that a class of dRNA appearing as three copies per 100 cells (*26*) could exist with transcripts complementary to the entire nonrepetitive fraction of the genome. At present, the available high-resolution analysis of nuclear RNA diversity of this type is too limited to provide critical assessment of the validity of the model we postulate. However, these available data are consistent with the predictions of the model, and further experiments may provide sufficient data to permit conclusive evaluation.

Table 2. Predicted saturation of sense strand nonrepetitive DNA with maximally diverse RNA.

RNA source	Fraction of total cell RNA mass*	Relative mass of maximally diverse RNA†	Expected $C_{R0}t$ value‡	
			Chick	Mouse
Total cell	1.00	0.001	1.9×10^6	3.5×10^6
Cytoplasm	0.95	0.001	1.9×10^6	3.5×10^6
Polysomes	0.95	0.001	1.9×10^6	3.5×10^6
Nucleus	0.02	0.05	3.7×10^4	6.9×10^4

*Relative mass in RNA populations based on values from (*37*). †Presumptive relative mass of maximally dRNA (see Table 1) is based on the assumption that all of the maximally dRNA is contained in the RNA class in question. ‡Expected $C_{R0}t$ for 90 percent hybridization is based on (i) the pseudo first-order (RNA driven) hybridization constant of 231 for the ϕX174 genome (*44*) and (ii) haploid sense strand nonrepetitive genome sizes of ϕX174, chick, and mouse of, respectively, 5374 nucleotides (*45*), 1.0×10^9 nucleotides as in Table 1, and 1.87×10^9 nucleotides. These estimates are for assay of hybrid duplexes by hydroxyapatite chromatography (*44*); hybridization assays with the use of resistance to single-strand nucleases would necessitate somewhat higher $C_{R0}t$ values to achieve 90 percent hybridization (*46*). We chose 90 percent hybridization as a reasonable value to indicate the presence of full complementarity. Hybridization to the 99 percent level would necessitate $C_{R0}t$ values slightly more than two times higher.

New Technology

The advent of the use of complementary DNA (cDNA) technology (*38, 47, 48*) may provide a somewhat more rapid approach for a partial test of this model. We indicate that the test would be partial because synthetic cDNA is complementary to only the polyadenylated RNA while nonpolyadenylated RNA goes undetected. However, the ability to radioactively label cDNA's to high specific activity in vitro, as well as the lack of cDNA self-annealing, considerably reduces the problem of high "background" levels of DNA-DNA renaturation during RNA-DNA hybridization experiments (*16, 26, 38, 48*). Thus, lower ratios of RNA to DNA are needed to detect very rare species of RNA. The hybridization $C_{R0}t$ values are fixed physical parameters determined by the relative abundance of RNA and will not change regardless of whether cDNA or nonrepetitive DNA is used as a hybridization probe. Thus, the specific radioactivity of the probe and its relative diversity are the two limiting factors in these types of measurements.

Besides synthetic cDNA probes to specific mRNA's or populations of polyadenylated RNA's, it is now theoretically possible to use the new plasmid technology (*49*) to obtain phenotype specific probes. In this case, either synthetic cDNA's to specific mRNA's or populations of polyadenylated RNA's or (more importantly) selected pieces of genomic DNA can be cloned and selected. Although several experimental restrictions exist, the possibility of using this approach is being tested for sea urchin, drosophila, and frog development (*50*), and the results of these measurements can be used to test this model.

Two published experiments, with cDNA probes, can be interpreted in terms of the model presented here. The cDNA to ovalbumin mRNA has been used to measure the concentration of ovalbumin in oviduct tissue at different developmental ages and under the influence of estrogen stimulation (*51*). These measurements show that estrogen modulates the transcription of the gene for ovalbumin mRNA. However, in unstimulated immature oviduct, and in estrogen withdrawn oviduct, there is evidence for the presence of ovalbumin mRNA at very low relative concentrations (*51*).

It would be more convincing in terms of the model, however, if a specific gene transcript was found in cells that would never be expected to express that gene. Such a finding is reported by Humphries *et al.* (*52*) who show that globin mRNA is detected in the nuclei of nonerythroid cells and in cultured cell lines. In both of these cases, the presence of the "phenotypically inappropriate" transcript indicates that these regions of the genome are transcriptively accessible. In addition, these transcripts appear to be present at two to 100 copies per nucleus. These results are encouraging because they suggest not only that the basic tenet of the model of transcriptional accessibility may be correct, but also that the estimate for relative abundance of diverse RNA outlined in Table I may be too severely restricted. If so, the problems of detection may be considerably easier to overcome.

To summarize, the transcriptional considerations of this model dictate that

developmental potential is equitable to transcriptional accessibility and that, in a cell with maximum development potential, all genes should be transcribed at a low level. The available data from prokaryotes suggest that this means of maintaining diversity of genomic response does occur in nature. Available measurements with eukaryotes regarding transcriptional changes during embryogenesis give limited support to the model, but definitive evaluation requires further experimental information. The essential point of these considerations is that no available evidence precludes the basic assumptions of the model (*53*).

Specific Developmental System for Possible Test of the Model

The strongest information concerning molecular events of development comes from studies of terminal differentiation systems. These systems represent branch points where decisional events involve the synthesis and morphological structuring of complicated end organs and tissues. For a molecular model to be relevant, it must function at both ends of development: the beginning events involving restriction of developmental potential and the final events when morphogenesis is taking place. The developing limb system is a suitable example for discussion because it represents a dramatic branch point and because, in the discussion of limb regeneration, we can make the strong point that dedifferentiation and subsequent redifferentiation do not exist.

Also, the normal and regenerating limb provide a biological and molecular development system that may be of eventual use in testing the irreversible gene repression model. (i) The limbs are external appendages and readily accessible to experimentation. (ii) The sequence and temporal events involved in the differentiation and development of the limb appear to be relatively synchronous (*54*). (iii) The two prominent phenotypes, muscle and cartilage, develop from the limb mesenchymal cell and are phenotypically distinct and minimally overlapping. For example, one phenotype precipitates contractile proteins intracellularly while the other deposits complex mucopolysaccharides extracellularly; one is aerobic while the other is anaerobic; one has cell division irreversibly inhibited while the other allows cell division to proceed albeit at a low rate. (iv) Rudimentary evidence suggests that the progenitor cells to muscle and cartilage may be expressing low levels of molecules for both phenotypes with the eventual accentuation of one phenotype (*55–57*). (v) The developmental changes represent the final developmental decision these cells must make and, for the model to be valid, the end of the developmental spectrum should also fit the restraints dictated by the model.

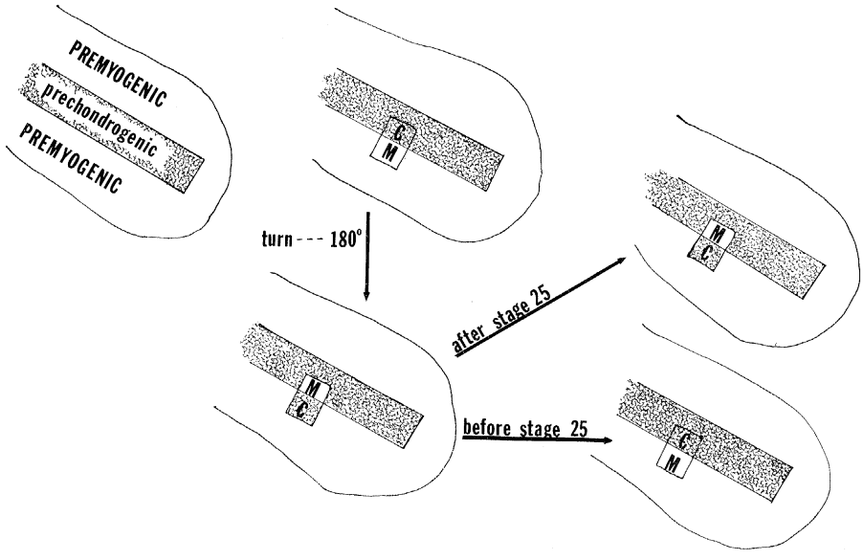

Fig. 3. Experiments by Searls and his colleagues (*56, 62–64*) in which a block of tissue that includes cells from the prechondrogenic and the premyogenic zone are rotated 180° and transplanted to recipient hosts. Before stage 25, the tissue would exhibit chondrogenic or myogenic properties as dictated by its new position. If this experiment is done after stage 25, the tissue would exhibit myogenic or chondrogenic properties as indicated by its position of origin. These data indicate that stage 25 is a critical stage in the differentiation process, resulting in myogenic or chondrogenic phenotypes.

Normal and Regenerate Limb Development

The development of cartilage and muscle in the avian limb has a number of common features with the appearance of specific phenotypes in other developing tissue. Cells predestined to become incorporated into limb tissue arise in a discrete area in the flank of the developing embryo (*58*). Eventually, one can identify two areas in the flank region which become the anterior and posterior limbs (Fig. 1). As development proceeds, differential rates of cell division between flank mesenchyme and presumptive limb mesenchyme produce bulges referred to as limb buds. As limb buds become recognizable, the limb ectoderm changes and becomes specific for limb and may no longer be exchanged for other ectoderm in other areas, such as the flank (*59*). At the same time, the fate of the limb mesenchyme becomes even more restricted. For example, wing mesenchyme transplanted into leg ectoderm "jacket" will always differentiate into wing structures; the form component and thus the three-dimensional juxtaposition of muscle and cartilage phenotypes is predetermined into these tissues (*60, 61*).

Limb mesenchymal cells are a common pool of progenitor cells giving rise to muscle and cartilage (eventually bone). At early limb stages a single mesenchymal cell probably has the capacity to differentiate into either muscle or cartilage cells, but commits itself to either muscle or cartilage differentiation at a specific time. In the chick wing, this phenotypic commitment or stabilization takes place at stage 24 to 25 (about $4^1/_2$ days of incubation) (*62*). Mesenchymal cells microsurgically transferred (Fig. 3) from the premyogenic area to the prechondrogenic area prior to this special time of commitment or stabilization will differentiate as determined by their new local environment (that is, cartilage). However, if such a transfer from the prechondrogenic area to premyogenic area is made after the stabilization event, the phenotype that develops will be determined by the site of origin of that transplanted tissue; in this case, chondrogenic elements will develop in the myogenic area (*63*). Such observations support the view that limb mesenchymal cells before the stabilization or commitment event seem to have the ability to elaborate both phenotypes.

In context of the irreversible gene repression model, the uncommitted limb mesenchymal cells probably have the

macromolecular machinery for both phenotypes, muscle and cartilage. Enzymes involved in the synthesis of cartilage extracellular matrix materials are present in very early limb buds a full 24 to 36 hours prior to stabilization and commitment events (*54, 57*); also muscle myosin may be present prior to the stabilization and commitment events (*57*). Whether each mesenchymal cell synthesizes macromolecules specific both to muscle and cartilage is not yet known. However, such a possibility has precedence in the case of muscle development in which specific repression of one phenotypic-specific macromolecule seems to be well documented. The breast muscle and leg muscle of the chick embryo have a very similar spectrum of macromolecules with the important exception that myoglobin is present in leg but absent in breast muscle. Myoglobin is synthesized in the early breast muscle development, but its synthesis is shut off during the later phases of this process. Data presented by Heywood (*64*) further indicate that the mRNA from myoglobin is present during the early synthesis phase of this process, but absent later, suggesting the possibility that its transcription had been developmentally repressed. These observations indicate that macromolecules and machinery specific for the two divergent phenotypes are present long before these phenotypes are fully expressive or developmentally stabilized.

Phenotype-specific macromolecules in progenitor cells prior to divergent expressional and commitment events is not an exclusive property of chick limb mesenchymal cells. For example, in the development of neural crest cells into elements of the sensory and sympathetic nervous systems, the progenitor cells contain both cholinergic- and adrenergic-specific molecules. As these cells further develop, the synthesis of one or the other of these compounds is terminated coincident with the emergence of specific sensory or sympathetic phenotypes.

Taken together, these observations support the interpretation that in limb mesoderm cells, the genes coding for muscle- and cartilage-specific macromolecules are open and transcribed, albeit at relatively low rates. If this extreme interpretation is correct, then differentiation in this case could be viewed as the selective stabilization of transcription of one set of genes with the irreversible repression of others. The assertion that a gene coding for one phenotype is irreversibly repressed is based on the fact that in no case has a cell been observed to dedifferentiate (revert to a more embryonic state) and then proceed into a different developmental compartment.

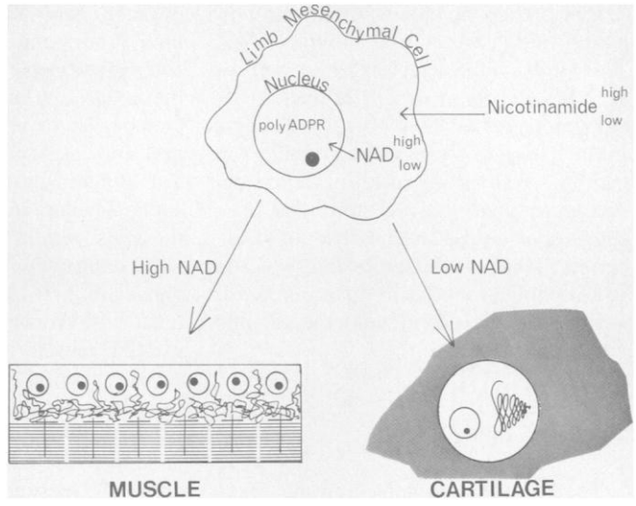

Fig. 4. The influence of NAD on chick limb mesenchymal cell differentiation into muscle and cartilage cells. The external concentration of nicotinamide influences the intercellular pool size of NAD, which seems to influence the choice between muscle and cartilage development (*65, 66*).

Experimentation in our laboratory has centered around the question of what controls the choice of phenotypic expression of limb mesenchymal cells to muscle and cartilage phenotypes (*65, 66*). As is discussed above, the environment in which the cells find themselves is phenotypically deterministic. Our experiments indicate that the intracellular nicotinamide adenine dinucleotide (NAD) pool plays a controlling role in this choice mechanism governing the differentiation of cartilage or muscle phenotypes with high internal pool sizes of NAD dictating muscle differentiation and low internal pool sizes of NAD dictating chondrogenic expression (Fig. 4). The internal pool sizes of NAD within mesenchymal cells are directly determined by the extracellular concentration of its precursor, nicotinamide, which is transported from the yolk to distal areas like the limb by the developing circulatory system.

In the limb, the developing vascular system establishes two different nutrient areas. One is highly vascularized in the premyogenic area, which results in high nutrient and nicotinamide levels, and a second area has low nicotinamide and nutrient levels. The vascular system differentiates and establishes these two separate nicotinamide-rich and nicotinamide-poor environments long before stage 24 to 25 (*66*). The mesenchymal cell genome seems to be geared to responding to such environmental cues by differentiating into the phenotype as dictated by their intracellular NAD pool size, which are a result of the differential delivery of nicotinamide by the vascular system.

Mesenchymal cell chromatin responds to (or senses) cytoplasmic NAD pool sizes by synthesizing poly(adenosine diphosphate–ribose), which is covalently bound to both histone and nonhistone proteins (*67*). The synthesis of poly(ADP-ribose) is directly correlated with the differentiation and stabilization event, and the newly synthesized poly(ADP-ribose) is bound to a discrete subfraction of chromatin (*68*). It is clear from the emphasis of this review that we would favor the role of newly synthesized poly(ADP-ribose) in specifically and irreversibly repressing certain gene sequences.

Although hybridization data are not available for the transcriptional events taking place during the decisional process of limb mesenchymal cells, two hybridization studies are available for events involved in the muscle development pathway.

1) Using saturation hybridization to unique sequence DNA, we have shown that the transcript diversity stays relatively high and constant during the major cytological phase of limb muscle development. A 50 percent decrease in the RNA diversity is observed after day 16 of embryonic development (*25*). Thus, a high level of transcript diversity, most of which is probably nuclear-restricted, is observed during the prominent cytological events of muscle development. The final phases of muscle development are marked by a massive repression event seen as a decrease of transcript diversity.

2) On a mass basis, 80 to 85 percent of the transcripts that contain polyadenylate [poly(A)] are present in approximately the same frequency throughout all muscle developmental stages and also in other embryonic tissues (*69*). Briefly, a

cDNA probe was synthesized from 9- or 14-day leg muscle total poly(A)-RNA. This probe hybridizes 80 to 85 percent with RNA from all stages of muscle development tested as well as embryonic brain. These analyses indicate that transcript diversity differences as well as developmental events are controlled by a relatively limited population of transcripts. Thus, rather subtle differences in RNA populations seem to account for striking changes in developmental and cytological events.

The Regenerating Limb

The regenerating newt limb has been used to suggest that dedifferentiation was possible. Dedifferentiation implies that a committed and expressive cell can reenter its more embryonic progenitor compartment and then be capable of a phenotypic choice once again. In the case of the newt regenerate, for example, this concept would predict that a cartilage cell could dedifferentiate and divide, and its daughter cells could enter the muscle compartment. Our model maintains that this does not happen and the newer experimental evidence seems to support this view (*70, 71*).

When the distal quadrant of the newt limb is amputated, a healing response takes place in the skin, and subsequently the underlying tissue (of mesoderm origin) begins to lose structural organization and large numbers of cells can be seen dividing at the distal tip, referred to as the developing blastema. The origin of the cells within this blastema is somewhat controversial. The two prominent views are, first, that the cells of the blastema arise from dedifferentiation of preexisting muscle and cartilage cells and, second, that, although some cells within the blastema have their origins in preexisting phenotypes, the majority of cells arise from the selective proliferation of a group of "mesenchymal" cells that have not yet committed themselves to a specific phenotype. Experiments done to date do not support the view that cells can dedifferentiate into a more embryonic type which is then capable of differentiating into one or more other phenotypes. For example, a chondrocyte can have its surrounding extracellular matrix removed and it can undergo multiple divisions. However, the daughter cells of these divisions clearly differentiate into chondrocytes again and not into myogenic elements (*70*). Likewise, it is not known whether nuclei from multinucleated muscle cells give rise to chondrogenic elements. However, a discrete subpopulation of mononucleated cells associated with multinucleated muscle cells are capable of dividing and then differentiating into either muscle or cartilage phenotype. These muscle "satellite cells" are probably undifferentiated mesenchymal cells which are still capable of choosing the cartilage or muscle developmental pathway and have not yet committed themselves to a specific route.

Thus, the regenerating limb must form an "embryonic-like" population of cells in response to initial surgical insult in concert with other environmental signals. The muscle and cartilage cells which differentiate from this population arise only from cells committed to one phenotype or the other prior to regeneration, and thus no reversibility of phenotypic commitment or fate occurs. Each cell type, however, retains the information for restructuring the limb component, both in terms of molecular components and three-dimensional form components. This undoubtedly calls upon reexpression of genes not expressed in the adult limb, but this reexpression occurs within a relatively narrow spectrum of genes specific to each phenotype. If genes were generally either simply repressed or derepressed, this would provide no compelling explanation of why phenotypic commitment is retained by regenerating cells. The fact that phenotypic commitment is retained supports the concept that some genes become irreversibly repressed during normal development. Thus, the fact that regeneration occurs at all suggests that some genes are available for expression during such a response, but these genes are not repressed to the same degree as genes specific to a completely different phenotype.

The developing and regenerating limb system represents one possible system for testing the model and has advantages over the fertilized egg and early embryo systems because of the limited number of possible phenotypes and greater mass of tissue being analyzed. Together, these two systems (fertilized egg and limb) represent the very beginning and very end of a specific developmental spectrum and yet possess many features in common such as the sensitivity to external environment and the possibility that phenotype-specific sequences are transcribed prior to further development. If our proposed model has general validity, it should apply equally well to either extreme of a developmental spectrum.

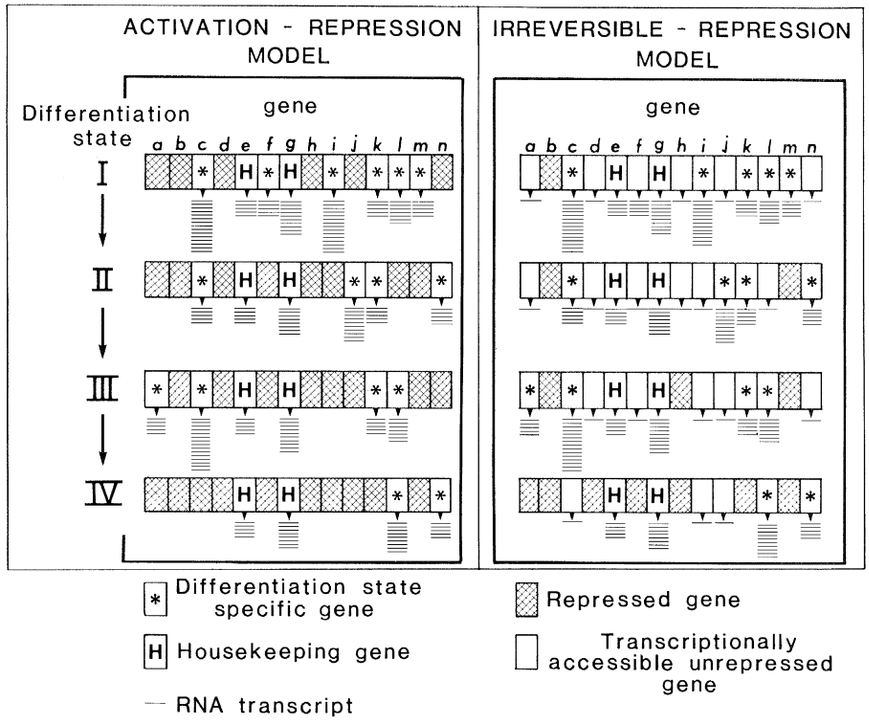

Fig. 5. Comparison of the activation-repression model and the irreversible-repression model. In each case, the complexity and frequency of the abundant transcripts are the same, with the exception that, in the irreversible-repression model, some genes are transcriptionally accessible and are transcribed at very low frequency. It is assumed that the majority of the genome is not repressed in the fertilized egg in the irreversible-repression model while the activation of previously repressed genes accounts for the change in transcript complexity in the other model.

Conclusions

With the above considerations in mind, the irreversible gene repression model is compared to the gene activation-repression model in Fig. 5, which shows the progressive changes in the transcriptional status of the genome for both models as early embryonic cells (line I) progress through successive developmental states (lines II and III) to a fully differentiated adult cell (line IV). The genomes are similar in that housekeeping genes (designated H) and phenotype-stage-specific genes (designated *) are transcribed frequently to give an abundant class of RNA that is identical in complexity in both cases. However, many genes that are repressed in the activation-repression model are accessible to infrequent transcription in the irreversible repression model, and this gives rise to a small but highly diverse population of RNA in the latter.

Given that the final differentiation state (IV) is identical in the two models, the chief difference lies in the path by which the genome achieves the final stage of differentiation and how the genome can modulate its differentiation state to alter its differentiation path or to assume a more "embryonic status" during a healing or regeneration response. For the activation-repression model, the mechanisms for these changes and responses are simple, at least in concept. When transcription of a gene is necessary, it is "turned on"; when unnecessary, it is "turned off." This straightforward mechanism needs no further conceptual explanation (*72*).

The conceptual explanation of the irreversible repression model is more complicated because the status of the genome should explain the cascade of developmental fate described earlier, as well as the production of prevalent gene products. In the irreversible repression model, at early states of development (state I) there are also genes that are "turned on" and those which are "turned off"; but the majority of genes are accessible for transcription, although many are transcribed only to a limited extent. Gene *b* represents a gene that is always repressed and never available for transcription in this cell regardless of the state of differentiation. Gene *a*, on the other hand, is one whose products are necessary to define developmental state III. In the activation-repression model, this gene is turned on at state III but is repressed at earlier and later stages; in the irreversible repression model, it is transcribed frequently at state III but is available and transcribed at low frequency at states I and II. As development proceeds from state III to state IV, this gene is irreversibly repressed.

Gene *c* is one whose expression is necessary throughout the earlier stages of development (states I, II, and III) but not in the final state IV. In this case, the activation-repression model would dictate that this gene be repressed in the final differentiated state while irreversible repression model does not make such a stringent demand upon the transcription apparatus and this gene can be open but transcribed to only a minor extent. Genes *e* and *g* represent housekeeping genes whose transcription is constant and continuous during all stages of development. The products of gene *h* are not essential at any states of development and is irreversibly repressed in the final state of differentiation and development. Gene *h* could, for example, represent a gene which permits a developmental choice at some earlier stage of development. If this choice is not exercised by a certain time, the developmental fate is fixed and this gene is repressed irreversibly.

In a line-by-line analysis of these models, the key difference between the two models lies in the fact that the activation-repression model requires that each specific gene be either "on" or "off," while the irreversible gene repression model dictates that the majority of genes are accessible to transcription early in development with the eventual irreversible repression of these genes as development proceeds. In each case as modeled in this figure, the complexity of the abundant transcripts is approximately the same with the only difference being the rare transcripts of transcriptively accessible genes in the irreversible model.

The definitive test of this model would be an assay which could assess the transcriptional status of accessibility of each gene in the cell's genome. Several attempts of assaying for gene accessibility have been made with specific cDNA's to determine whether specific genes are transcribed either in vivo or in vitro when purified chromatin preparations are used (*73, 74*). In one case, when highly purified cDNA to ovalbumin mRNA was used, it appears that, in progenitor oviduct tissue and hormone-unstimulated tissues, there was infrequent transcription of the ovalbumin gene at a level of several orders of magnitude lower than found in fully adult oviduct that had been estrogen-stimulated. In the separate case of isolated chromatin and in vitro transcription assay (*73*), it seems clear that the ovalbumin gene is accessible to RNA polymerase, regardless of the state of activation of oviduct tissue.

Others have shown that fibroblasts transformed by infection with Rous sarcoma virus begin to transcribe RNA which hybridizes with the DNA complementary to hemoglobin mRNA (*75*). Two explanations are possible for such a finding: (i) transformation derepresses certain genes, or (ii) transformation increases the frequency of transcription of certain genes that were already accessible to transcription. To choose between these two possibilities, it would be necessary to obtain an estimate of the minimal number of transcripts of the hemoglobin gene in the fibroblast population prior to transformation. In this regard, Humphries *et al.* (*52*) demonstrate that the hemoglobin gene is transcribed at low frequency in fibroblast cells and other nonhematogenic tissue. This evidence would tend to support the latter hypothesis regarding the mechanism of transformation-mediated transcription of the hemoglobin gene.

This last-mentioned example illustrates a second important difference between the gene activation and irreversible repression model because, in the irreversible repression model, some genes may be accessible to the transcription machinery, although in the normal differentiated states these sequences are transcribed to a minor extent. Increased transcription of such genes as a response to environmental stimuli (either normal or abnormal) would thus be augmentation of an already established genomic state rather than specific activation of a repressed gene. Available data indicate that the abnormal function or growth (or both) of cells in response to the introduction of teratogen may be the result of potentiation of transcription of genes that are already accessible rather than being the result of specific gene activation. Teratogenic effects take place quickly and are usually the result of narrow changes in cell expression. It is possible that teratogenic effects are directly related to transcriptional accessibility and the abnormal accentuation of transcription of genes from the low-frequency group. Again, direct testing of this possibility can occur coordinately with a test of the irreversible repression model.

In summary, a model has been put forth that attempts to integrate the known biological and molecular features of embryological systems. This model is testable at a variety of levels, and its articulation may serve to stimulate such activity.

References and Notes

1. F. C. Kafatos and R. Gelinas, in *Biochemistry Series One*, J. Paul, Ed. (University Park Press, Baltimore, 1974), vol. 9, pp. 223–264; S. M. Heywood and D. S. Kennedy, *Biochemistry* **15**, 3314 (1976); A. J. Bester, D. Kennedy, S. M. Heywood, *Proc. Natl. Acad. Sci. U.S.A.* **72**, 1523 (1975); J. Ilan and J. Ilan, *Dev. Biol.* **25**, 280 (1971); J. E. Darnell, W. R. Jelinik, G. R. Molloy, *Science* **181**, 1215 (1973); J. E. Darnell, W. Jelinik, L. Puckett, E. Derman, S. Bachenheimer, in *The 34th Symposium of the Society for Developmental Biology*, J. Papaconstantinou, Ed. (Academic Press, New York, 1976), pp. 53–74; H. Lodish, T. Alton, R. P. Dottin, A. M. Weiner, J. P. Margolskee, in *ibid.*, pp. 75–103; E. K. F. Bautz and P. Karlson, *Regulation of Transcription and Translation in Eucaryotes* (Springer-Verlag, New York, 1973); D. Hendrick, *Wilhelm Roux' Arch. Entwicklungsmech. Org.* **179**, 33 (1976).
2. M. Ptashne, K. Backman, M. Z. Humayun, A. Jeffrey, R. Maurer, B. Meyer, R. T. Sauer, *Science* **194**, 156 (1976).
3. The term "irreversible repression" means that once a phenotype-specific gene or gene set is repressed during normal developmental events, it cannot be transcribed at some later time. This does not mean that crafty molecular biologists cannot find mechanisms to derepress these genes. On the contrary, such techniques are now emerging. What may be irreversible repression for animal cells may not be occurring in plant cells, which seem to be able to form a complete, new organism from a single cell. Irreversible repression is a model to explain the normal developmental flow, and, if this flow is experimentally uninterrupted, repressed genes will not be transcribed subsequent to their repression.
4. H. Driesch, *Z. Wiss. Zool.* **53**, 160 (1891); H. Spemann, *Wilhelm Roux' Arch. Entwicklungsmech. Org.* **12**, 224 (1901); *ibid.* **15**, 448 (1903); *ibid.* **16**, 551 (1903).
5. This is not the case for mammalian embryos.
6. G. A. Schmidt, *Wilhelm Roux' Arch. Entwicklungsmech. Org.* **129**, 1 (1933).
7. A. C. Clement, *J. Exp. Zool.* **149**, 193 (1962).
8. R. W. Hegner, *J. Morphol.* **25**, 375 (1914).
9. L. Wolpert, *J. Theor. Biol.* **25**, 1 (1969); B. C. Goodwin and M. H. Cohen, *ibid.*, p. 49.
10. W. Vogt, *Wilhelm Roux' Arch. Entwicklungsmech. Org.* **106**, 542 (1925); J. Pastells, *J. Exp. Zool.* **89**, 255 (1942); H. Stalsburg and R. L. DeHoan, *Dev. Biol.* **19**, 128 (1969).
11. J. B. Gurdon, *The Control of Gene Expression in Animal Development* (Harvard Univ. Press, Cambridge, Mass., 1974).
12. R. E. Briggs and T. J. King, *Proc. Natl. Acad. Sci. U.S.A.* **38**, 455 (1952); *J. Morphol.* **100**, 269 (1957); M. A. Bernardino and T. J. King. *Dev. Biol.* **15**, 102 (1966).
13. This rationale could serve as a framework to explain the general phenomena of aging, in which this pool of cells with relatively "undifferentiated" nuclei decreases below some minimally tolerable point relative to species and tissue viability.
14. Some cells can tolerate wide ranges of fluctuations in these factors, even the removal from the in vivo environment to in vitro cultures while continuing to express their developmental potential.
15. E. C. Zeeman, *Lect. Math. Life Sci.* 7, 69 (1974).
16. G. A. Galau, W. H. Klein, M. M. Dovis, B. J. Wold, R. J. Britten, E. H. Davidson, *Cell* 7, 487 (1976).
17. W. J. Rutter, R. L. Pictet, P. W. Morris, *Annu. Rev. Biochem.* **42**, 601 (1973).
18. Untranslated transcripts in all likelihood may also be grouped into these two categories, but further discussion of this class of transcripts must await identification of their function.
19. E. H. Davidson, B. R. Hough, C. S. Amenson, R. J. Britten, *J. Mol. Biol.* 77, 1 (1973); E. H. Davidson, B. R. Hough, W. H. Klein, R. J. Britten, *Cell* **4**, 217 (1975).
20. T. P. Fondy and N. D. Kaplan, *Ann. N.Y. Acad. Sci.* **119**, 888 (1965); D. M. Dawson, H. M. Eppenberger, N. O. Kaplan, *J. Biol. Chem.* **242**, 210 (1967); C. L. Market, *Isozymes III. Developmental Biology* (Academic Press, New York, 1975).
21. This example is distinctly different from specific gene activation in that the possibility exists that this specific gene has been accessible to the transcriptive machinery, but has been transcribed at very low frequency or essentially not transcribed at all. This is not a case of derepression but one of increased or preferential transcriptional accessibility.
22. W. E. Hahn, D. E. Pettijohn, J. Van Ness, *Science* **197**, 582 (1977).
23. I. R. Brown and R. B. Church, *Dev. Biol.* **29**, 73 (1972); B. W. Levy and G. H. Dixon, *Eur. J. Biochem.* **74**, 61 (1977).
24. D. M. Anderson, G. A. Galau, R. J. Britten, E. H. Davidson, *Dev. Biol.* **51**, 138 (1976).
25. C. P. Ordahl and A. I. Caplan, *ibid.* **54**, 61 (1976).
26. B. R. Hough, M. J. Smith, R. J. Britten, E. H. Davidson, *Cell* **5**, 291 (1975).
27. L. Grouse, M. D. Chitten, B. J. McCarthy, *Biochemistry* **11**, 798 (1972).
28. R. J. Britten and D. E. Kohne, *Science* **161**, 529 (1968).
29. C. D. Laird, *Chromosoma* **32**, 378 (1971); B. Lewin, *Cell* **1**, 107 (1974); *ibid.* **4**, 77 (1975); J. O. Bishop, *ibid.* **2**, 81 (1974).
30. T. Ohta and J. Kimura, *Nature (London)* **223**, 118 (1971); S. J. O'Brien, *ibid.* **242**, 52 (1973); S. Ohno, in *Evolution of Genetic Systems*, H. H. Smith, Ed. (Gordon & Breach, New York, 1973).
31. S. Ohno and N. B. Atkin, *Chromosoma* **18**, 455 (1966).
32. D. Kennel, *J. Mol. Biol.* **34**, 85 (1968).
33. M. Rosbash, M. S. Campo, K. S. Gummerson, *Nature (London)* **258**, 682 (1975).
34. J. A. Bantle and W. E. Hahn, *Cell* **8**, 139 (1976).
35. E. H. Davidson and R. J. Britten, *Q. Rev. Biol.* **48**, 565 (1973).
36. A. H. Sparrow, H. J. Price, A. G. Underbring, *Brookhaven Symp. Biol.* **23**, 451 (1971).
37. R. Soeiro, M. H. Vaughan, J. R. Warner, J. E. Darnell, *J. Cell Biol.* **39**, 112 (1968).
38. J. O. Bishop, J. G. Murton, M. Rosbash, M. Richardson, *Nature (London)* **250**, 199 (1974).
39. N. D. Hastie and J. O. Bishop, *Cell* **9**, 761 (1976).
40. G. U. Ryffel and B. J. McCarthy, *Biochemistry* **14**, 1379 (1975); M. J. Getz, G. D. Birnie, B. D. Young, E. MacPhail, J. Paul, *Cell* **4**, 121 (1975).
41. L. Kleiman, G. D. Birnie, B. D. Young, J. Paul, *Biochemistry* **16**, 1218 (1977).
42. F. A. Galau, R. J. Britten, E. H. Davidson, *Cell* **2**, 9 (1974).
43. $C_{R0}t$ = moles of nucleotide RNA per liter times seconds. All $C_{R0}t$ values are expressed in equivalence to reactions conducted in 0.18*M* Na^+ [R. J. Britten, D. E. Graham, B. R. Neufeld, *Methods Enzymol.* **29**, 363 (1974)].
44. G. A. Galau, R. J. Britten, E. H. Davidson, *Proc. Natl. Acad. Sci. U.S.A.* **74**, 1020 (1977).
45. B. G. Barrell, G. M. Ari, C. A. Hutchinson, *Nature (London)* **264**, 34 (1976).
46. M. J. Smith, R. J. Britten, E. H. Davidson, *Proc. Natl. Acad. Sci. U.S.A.* **72**, 4805 (1975).
47. Complementary DNA is DNA complementary to a specific RNA or group of RNA's. In this article, cDNA refers to sequences synthesized with reverse transcriptase.
48. I. M. Verma, G. F. Temple, H. Fan, D. Baltimore, *Nature (London) New Biol.* **235**, 161 (1972); J. Ross, H. Aviv, E. Scolnick, P. Leder, *Proc. Natl. Acad. Sci. U.S.A.* **69**, 264 (19 → T. Honjo, S. Packman, D. Swan, M. Nau, P. Leder, *ibid.* **71**, 3659 (1974); I. M. Verma, R. A. Firtel, H. F. Lodish, D. Baltimore, *Biochemistry* **13**, 3917 (1974); H. Diggelman, C. H. Faust, B. Mach, *Proc. Natl. Acad. Sci. U.S.A.* **70**, 693 (19 → S. E. Harris, A. R. Means, W. M. Mitchell, B. W. O'Malley, *ibid.*, p. 3776.
49. P. E. Lobban and A. O. Kaiser, *J. Mol. Biol.* **78**, 453 (1973); J. F. Morrow, S. N. Cohen, A. C. Chang, H. W. Boyer, H. M. Goodman, R. B. Helling, *Proc. Natl. Acad. Sci. U.S.A.* **71**, 1743 (19 → M. Grunstein and D. S. Hogress, *ibid.* **72**, 3961 (19 → R. Higuchi, G. V. Paddock, R. Wall, W. Salser, *ibid.* **73**, 3146 (1976).
50. J. D. Watson, F. P. Crick, P. Chambon, Eds. *Cold Spring Harbor Symp. Quant. Biol.*, in press.
51. G. S. McKnight, P. Pennequin, R. T. Schimke, *J. Biol. Chem.* **250**, 8105 (1975); S. E. Harris, J. M. Rosen, A. R. Means, B. W. O'Malley, *Biochemistry* **14**, 2071 (1975).
52. S. Humphries, J. Windass, R. Williams, *Cell* 7, 267 (1976).
53. Perlman *et al.* [S. M. Perlman, P. J. Ford, M. M. Rosbash, *Proc. Natl. Acad. Sci. U.S.A.* **74**, 3835 (1977)] reported detecting adult-form globin mRNA in *Xenopus* oocytes at concentrations comparable to the majority of polyadenylate [poly(A)] RNA species. Since adult hemoglobin is not present until after metamorphosis, they too suggest the possibility that "most or all of the structural gene sequences utilized during the lifetime of the organism are transcribed" in the oocyte. This view is consistent with our model.
54. R. L. Searls, in *Developmental Regulation: Aspects of Cell Differentiation*, S. J. Coward, Ed. (Academic Press, New York, 1973).
55. G. M. Crosby, thesis, Brandeis University (1967); J. Medoff, *Dev. Biol.* **16**, 118 (1967).
56. R. L. Searls, *Proc. Soc. Exp. Biol. Med.* **118**, 11720 (1965).
57. J. Medoff and E. Zwilling, *Dev. Biol.* **28**, 138 (1972).
58. B. I. Balinski, *An Introduction to Embryology* (Saunders, Philadelphia, 1970); M. Kieny, *C.R. Acad. Sci.* **270**, 2009 (1970); M. Pinot, *J. Embryol. Exp. Morphol.* **23**, 109 (1970); S. Chaube, *J. Exp. Zool.* **144**, 29 (1959).
59. M. Kieny, *J. Embryol. Exp. Morphol.* **8**, 457 (1960).
60. E. Zwilling, *J. Exp. Zool.* **128**, 423 (1955); *ibid.* **132**, 173 (1956); *ibid.*, p. 241.
61. The polarity of limb outgrowth as well as the general relationships between limb parts is affected by this mesenchymal-ectodermal interaction. These interactions are controlled by a variety of factors that originate in the limb in various quandrants and interact in a complicated manner which dictates limb topology but does not dictate phenotypic expression of underlying mesenchyme.
62. M. Y. Janners and R. L. Searles, *Dev. Biol.* **23**, 136 (1970); E. Zwilling, *J. Exp. Zool.* **176**, 397 (1971).
63. R. L. Searls and M. Y. Janners, *J. Exp. Zool.* **170**, 365 (1969).
64. W. C. Thompson, E. A. Buzash, S. M. Heywood, *Biochemistry* **12**, 4559 (1973); S. M. Heywood, A. Havaranis, H. Herrmann, *J. Cell. Physiol.* **82**, 319 (1973).
65. A. I. Caplan, *Exp. Cell Res.* **62**, 341 (1970); M. J. Rosenberg and A. I. Caplan, *Dev. Biol.* **38**, 157 (1974); *J. Embryol. Exp. Morphol.* **33**, 747 (1975).
66. A. I. Caplan and S. Koutroupas, *J. Embryol. Exp. Morphol.* **29**, 571 (1970).
67. M. Harris, Ed. *Poly(ADP-Ribose); An International Symposium* (NIH 74-477, Department of Health, Education, and Welfare, Washington, D.C., 1974).
68. A. I. Caplan and M. J. Rosenberg, *Proc. Natl. Acad. Sci. U.S.A.* **72**, 1852 (1975).
69. C. P. Ordahl and A. I. Caplan, in preparation.
70. T. P. Steen, *J. Exp. Zool.* **167**, 49 (1968).
71. The clear exception to the dogmatic assertion that dedifferentiation does not occur in higher eukaryotic cell development is exhibited by plant cells where a root cell or leaf cell can give rise to a whole plant. More information is required with regard to the transcriptional complexity of different plant tissues, and whether differential gene transcription plays any role in phenotypic expression of different tissue phenotypes. It is possible that gene repression also plays a role in the control of development of plants but that irreversibility of such repressional events is not required for expressional events.
72. Rutter and colleagues [W. J. Rutter, W. D. Ball, W. S. Bradshaw, W. R. Clark, T. G. Sanders, *Exp. Biol. Med.* **1**, 110 (1967); W. J. Rutter, J. B. Kemp, W. S. Bradshaw, W. R. Clark, R. A. Ronzio, T. S. Sanders, *J. Cell. Physiol.* **72**, 1 (1968)] have proposed a gene activation model for development of the pancreas in which the primary gene activation is followed by a "protodifferentiation" state allows for the synthesis of low levels of phenotype-specific macromolecules. The observation that phenotype-specific macromolecules can be detected in the pancreas rudiment even at the earliest times can be interpreted to support the basic arguments of the irreversible repression model proposed in this article. Unanswered is the question as to when in the history of the pancreas precursor cells are mRNA's for phenotype-specific molecules first present. From the earliest times observed, the pancreas precursor cells are synthesizing detectable levels of translational products. Whether the transcripts coding for these products were suddenly "turned on" or whether the DNA sequences coding for these RNA's had been transcriptionally accessible from the beginning of development is the point where the activation-repression model and the irreversible repression model differ.
73. H. Towle, M. Tsai, S. Y. Tsai, B. O'Malley, *J. Biol. Chem.* **252**, 2396 (1977).
74. M. M. Smith and R. C. C. Huang, *Proc. Natl. Acad. Sci. U.S.A.* **73**, 775 (1976); D. Sullivan *et al.*, *J. Biol. Chem.* **248**, 7530 (1973); P. R. Harrison, R. S. Gilmour, N. A. Affara, D. Conkie, J. Paul, *Cell Differ.* **3**, 23 (1974); A. W. Steggles, G. N. Wilson, J. A. Kantor, D. J. Picciano, A. K. Falvey, W. F. Anderson, *Proc. Natl. Acad. Sci. U.S.A.* **71**, 1218 (1974).
75. M. Groudine and H. Weintraub, *Proc. Natl. Acad. Sci. U.S.A.* **72**, 4464 (1975).

76. We thank Dr. Raymond Lasek and Dr. Gerald Karp for their interesting and helpful discussions and many other colleagues for their criticism. This work was partially supported by grants from the National Foundation–March of Dimes, the Muscular Dystrophy Association, NIH, NSF, and the Josiah Macy, Jr. Foundation.