

Underlying assumptions of developmental models

(embryonic development/signaling/nuclear factors/units/local information)

ROY J. BRITTEN*

Division of Biology, California Institute of Technology, 101 Dahlia Avenue, Corona del Mar, CA 92625

Contributed by Roy J. Britten, June 1, 1998

ABSTRACT These 10 obvious propositions make a model of the specification of form, intended to expose underlying assumptions of developmental biology for examination and future experimentation. (I) The control of development is by means of local interactions, rather than global control mechanisms. (II) A macromolecule near a specific site will bind by mass action. (III) Starting with a precursor cell, all cells are assembled automatically by specifically binding macromolecules. (IV) At the surface of cells are specific adhesion sites that determine how all cells bind to each other. (V) An organism will assemble automatically from parts (macromolecules, structures, and cells) specified by nuclear control factors. (VI) The nuclear control factors in each cell are from precursor cells and factors derived by signaling from other cells. (VII) The macromolecules that determine specific binding, cell adhesion, and signaling are controlled by nuclear control factors, and in a grand feedback the cell adhesion and signaling systems determine the nuclear factor patterns. (VIII) The embryonic precursor cells for organs, termed “precursor groups,” are linked by adhesion and signaling relationships. (IX) The precursor groups include precursors for regions of an organ and boundary cells between regions having few cell types, growing without additional specific cell-to-cell relationships. (X) Organs are held together by cell adhesion in functional relationships. Thus the form and function of the organism is specified entirely by local control mechanisms. Without global control systems, information for form is in the genes for structural proteins, adhesion molecules, control factors, signaling molecules, and their control regions.

There are a number of assumptions about processes such as self-assembly of cells from macromolecules and of organs from cells that are built into the current view of developmental biology but are not examined together in current literature. Rather than arguing each of these assumptions individually, I consider how they can be used to build an abstract model of the processes from germ cell to adult and how they logically fit together. This model is an attempt toward that goal in a very brief compass. It is a result of my cogitation over the location and storage mechanism of the information that specifies form in eukaryotes. I use a simplified vocabulary, and the descriptions differ from those of current developmental biology. The purpose is to compactly describe the ideas regarding development and not to review the evidence and original work. The model is not intended to be useful in describing the structure or pattern of any organism, but to bring out the logic of embryonic development.

About the Model

It is certain that development and growth occur by the assembly of macromolecules into cells and the assembly of cells

into organs, individually properly timed. But what are the controls and what features are logically required? This paper is an attempt to summarize current knowledge avoiding overarching regulatory concepts for which the detailed processes are not describable. Logically the way to do that is to follow the process of assembly molecule by molecule. Assembly based entirely on the binding properties of the added macromolecules implies that the information for form is in the genes that specify each macromolecule, including its specific binding properties. In the model, during oogenesis, the egg is automatically assembled from macromolecules (synthesized in the oocyte or nurse cells) depending on their ability to bind to parts of the pre-existing germ cell and to each other. Each binding step is specific, though we cannot yet describe most of the binding processes. Specific does not necessarily mean individual because, for example, many identical molecules may be destined for the membrane. Automatic is taken to mean that supplying the elements that go into a structure as well as possible cofactors causes assembly without any overall guidance.

Embryonic development surely depends on the control of transcription by means of nuclear control factors that bind to regulatory sites adjacent to genes and many succeeding regulatory processes. The resulting temporal and spatial pattern of gene expression is clearly central, but the role of this pattern in the control of embryonic form is presently a major subject of research. In the model, a cell-specific pattern of nuclear factors controls the expression of macromolecules. As macromolecules are synthesized within a cell they are modified and ready to join a structure. They bind specifically to the pre-existing partial structure in the correct locations and orientations, because the capability of specific binding is genetically determined. To reach their binding sites is a complex process probably involving the Golgi apparatus, vesicles, and possibly organized transport of vesicles. Diffusion and mass action are possibly restricted to the final binding process. Though intricate, the process is automatic in the sense just mentioned. When bound, each unit and adjacent units then expose one or more binding sites where the next macromolecule(s) will specifically bind. The succeeding extended series of binding events establishes each cell, with its specific adhesion and signaling receptors. The many types of cells produced in this way then assemble specifically into organs and structures. In other words, the internal control processes in the cells determine their adhesion elements and thus the form of the organs. The concept of automatic self-assembly is old and obvious, but is absent from current literature except in reference to specific molecules such as tubulins and the ciliate *Stentor* that can undergo self-assembly after surgical dissection (1), which is rare among eukaryotes. In this model, self-assembly is the logical replacement for potential overarching regulatory concepts. It is the adhesion and signaling among the cells that determine their interactions with each other and establish the

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. §1734 solely to indicate this fact.

© 1998 by The National Academy of Sciences 0027-8424/98/959372-6\$2.00/0
PNAS is available online at www.pnas.org.

Abbreviation: pgroup, precursor group.

*To whom reprint requests should be addressed. e-mail: rbritten@etna.bio.uci.edu.

set of factors and character of each of the many cells that are parts of many lineages that sequentially divide and form each part of every organ of the organism. It is an extraordinary process of control that, depending on local interactions, establishes all of these parallel chains of events.

Simplified Terminology. There are good sources of appropriate references and customary terminology (2). Organisms are formed primarily of macromolecules termed in this model units that make up the structure of the cells and extracellular parts. Here they are named units if they can be or are bound into a structure, regardless of additional biochemical or other capabilities. The unconventional usage "assembly of units" means "growth from egg to adult." Local cell-to-cell relationships are primarily adhesion to and signaling between adjacent cells. The word signaling is used comprehensively to cover all possible processes whereby adjacent and more distant cells and tissues can affect the pattern of gene expression in a nucleus. The word adhesion has its usual meaning for specific binding between cells, assuming that all cells have specific adhesion molecules on their surfaces that locate them by binding to matching sites on other cells. These are not always fixed and cells can migrate in controlled fashion. A precursor group (pgroup), defined under proposition VIII, is the effective precursor of an organ in the embryo. The word organ covers all of the familiar organs and any structures requiring pgroups. Region refers to a part of an organ that is uniform with few kinds of cells. The terms organ and pgroup may be applied to, say, the whole brain or any describable part that contains within it regions with simple cell content. That is not vagueness in the model definitions but results from the arbitrariness of anatomical terms and the complexity of the real structure of organisms. The pgroups when determined will express the actual developmental pathways. The overlapping and sharing of the pgroups at various stages of development reflects the relationships of the organs and ultimately their distinct or shared evolutionary origins.

The 10 Propositions. I: *The control of development is by means of local interactions. Local means binding between macromolecules and interactions between adjacent or nearby cells and does not exclude morphogen or hormone interactions. It does exclude potential global control mechanisms containing much information specifying form.* That is a requirement because only mechanistically understood kinds of control are acceptable and overarching regulatory principles are not. A few long-range diffusing molecules have important roles, for example in sex determination. There exist important gradients such as bicoid in *Drosophila* eggs that act over distances that are many times cell dimensions and contain information important to development. Although not precisely local such gradients do not contain global information specifying the form of the embryo. As an example of other large-scale processes, the environment of the embryo may supply minerals, nutrition, and hormones but again does not specify the form of the embryo, although abnormal environment can damage the embryo or upset control processes. The extent of plant growth and leaf size and other aspects respond to conditions, pretty much as a choice between alternatives, each of which is controlled by local cell and tissue interactions. There is no evidence for global control processes specifying form nor have any complete theoretical models been devised, so this model is restricted to local control except for certain well-known, long-range processes such as hormone action and control factor gradients, each not including large amounts of information that specify form.

II: *A macromolecule near a specific site will bind by mass action.* This finding is obvious, mentioned because of its central role.

III: *Starting with a precursor cell, all cells are assembled automatically by specifically binding new macromolecules.* That is logically required. There is little direct evidence but no

alternatives are known. During cleavage the units present in the precursor cell become available for the daughter cells and are bound in place. Often the daughter cells differ in detail from each other. All of these controlled events are, in the model, caused by the specific genes expressed in each of the cells in all of the lineages.

IV: *At the surface of cells are specific adhesion sites that determine how all cells bind to each other.* This finding derives from a wide range of studies (3–7).

V: *Both the molecular and cellular binding processes are specific, and an organism will assemble automatically when the parts (macromolecules, structures, and cells) appear as specified by nuclear control factors.* Automatic assembly with cells as building blocks is how the model works. The adhesion sites present on the cells are determined by the state of the factors so no other information is required for assembly. The process is not mass action because many cell locations are the result of their duplication in place and other cells move in controlled fashion. For this reason and others a dissociated organ will not usually self-assemble again. Plants may lack specific adhesion and cell motion, but the assembly during growth is automatic in any case.

VI: *The set of nuclear control factors in each cell is a combination of inherited factors from precursor cells and factors derived by signaling from other cells.* This proposition is obvious, but it does introduce a powerful concept. Adjacent cell signaling will set most patterns but diffusible ligands or factors are also important. In early cleavage the control factors are mostly maternal. Later, signaling control becomes more important as cells become differentiated.

VII: *The macromolecules that determine specific binding, cell adhesion, and signaling are produced as specified by the nuclear control factors, and in a grand feedback the cell adhesion and signaling systems determine the nuclear factor patterns that control the expression of these macromolecules.* The pattern of transcription is established by the transcription factors that are determined by the cell–cell interactions, signaling, and diffusible factors. The factor pattern is changed from outside the cells and in turn affects the intercellular binding sites. The expression of the units, their binding, and the set of transcription control factors operate as a linked set, different, of course, in each part of the developing embryo. Any of the elements in a feedback loop are subject to experimental manipulation and a molecule might seem to be in charge when, in fact, no single part of the feedback loop is in control (8, 9).

Propositions I–VI are general descriptions of the processes that underlie development. The most important control processes are part of VII and include the hottest areas of research at the heart of development. The next propositions deal with the formation of the precursors of organs and ultimately the organs themselves. These propositions are based on a very large body of specific examples where the factors and signaling systems have been observed and the stages of morphogenesis followed. There are great differences among the model animal systems. Here is a summary from a review (2): "In some embryos, specification depends on intercellular interaction during cleavage, while in others this cannot be so since specification occurs while nuclei are syncytial; some rely on invariant cell lineages, while others develop from populations of migratory cells of no fixed lineage; some generate autonomously specified founder cells, while others have none; and so forth." This model applies to all; autonomous specification and syncytial cases need comment and some rephrasing is required. In insects the syncytial nuclei are formed in place and the specific cell-to-cell adhesion is delayed until cellularization. The nuclei do influence each other's factor pattern and before cellularization many specific interactions are formed, in *Drosophila* for example, that lead to the "stripes" that establish the precursors of segments. The anterior-posterior and dorsal-ventral axes also are established. It is likely that the nuclei

directly influence their neighbors by passing factors and signaling elements, possibly during nuclear membrane breakdown. The details of the signaling need examination, and the well-organized cytoskeletal structures may be important. For plants the next three sections should be rewritten because some processes and the uses of the terms embryo and organ and their nature are distinct from those for animals.

VIII: The embryonic precursor cells of organs termed pgroups are linked by specific adhesion and signaling relationships. In the model during early development the initiation of a multicellular structure or organ occurs as follows. At an appropriate stage a pgroup of cells for an organ is assembled depending on their intercellular adhesion and signaling. These cells are influenced in differentiation and structure by the signals from each other, leading to what I term the "embryonic precursor group," called for short the pgroup. The pgroup grows by cell duplication. Pgroups remain highly organized and have specific sets of adhesion and control signal patterns through the whole pathway as the precursor of an organ. In the *Drosophila* syncytium during early development the nuclei that are being committed are not in cells and thus the term precursor group is used instead of precursor cell group. Often, for simplicity, I will use the word cell where I should say cell or nucleus to properly include the syncytial cases. There is clear evidence that cells that will not ultimately contribute to the organ nevertheless have strong influence on the cell lineages that do contribute. The pgroup is shorthand for a complex process in which at early times particular groups of cells are precursor to more than one organ or region of an organ. As duplication occurs and lineages branch the pgroups branch and separate. They increase in number and effectively narrow down to their ultimate roles as precursors of organs or regions. These cells are members of cell lineages, have been specified for this task, and are specifically connected by adhesion and signaling to members of other lineages. A description as cell lineages is not significantly different, although the pgroup description may supply additional insights because it focuses on intercellular interactions. There is no equivalent term in common use, but the requirements of generalization in this summary have forced the invention of the term pgroup.

In the quotation above is the line "...some generate autonomously specified founder cells, while others have none..." that serves as an introduction to the large subject of how much of early development is autonomous and how much is conditional. Autonomous describes that various experiments do not affect the fate of certain cells or pgroups whereas conditional indicates that the fate is influenced easily and presumably in normal development other cells have influence on the development of the pgroups. The difference can be interpreted as the degree of effect of signaling from other cells on the factor set present and thus on the characteristics of succeeding cells in the lineage. The distinction is not always clear, and an originally autonomous lineage is likely to show conditional response later. The distinction between autonomous and conditional does not affect the generalization that groups of functionally linked cells (pgroups) are the precursors of organs, whatever the early history of the pgroups. The ability of natural selection to establish all of the series of specific relationships of signaling and adhesion that carry a cell lineage through the many duplications and steps to a functional location in an adult organ has a great fascination. This selection is simultaneously performed for many lineages, leading to perhaps more than 10^{10} ultimate cells, and the relationships among these cells maintain their function in the adult. That is the essence of the model.

IX: The pgroups include precursors for regions of an organ and boundary cells between regions. The regions and boundaries ultimately have few types of cells and grow to very many cells without additional kinds of specific cell-to-cell relationships. This proposition is a logically required one to build a model that

describes how large organs are formed entirely on the basis of the limited number of possible specific cell-to-cell relationships. An organ can easily have more than 10^9 cells and if all of these cells had different specific relationships just including their adjacent cells 10^{10} such connections might be required. It is very unlikely that 10^{10} different and effective adhesion and signaling relationships exist. The solution is to divide organs into regions with few different kinds of cells. Within these regions a limited set of cell-to-cell relationships (adhesion and signaling) determine the functional state of the cells and in combination with boundary cells determine their shape. Logically if the organ is to be determined by a modest number of local and specific cell-to-cell contacts the pgroup must be started fairly early in development. The pgroup cells can map the regions of an organ by cell-to-cell contact and signaling before the number of cells is too large. Then a transition must occur as the embryo grows. Regions made up of a limited set of cells form, and these are bounded by other regions containing cells of a different but region-specific type.

Within a region the limited set of cells have a limited set of adhesion molecules and signaling molecules, and they all bind to and control each other to maintain their functions, but only a small number of specific controlling relationships is present in the cells of the region. There is no size limit to such a region in principle. There are also boundary cells between one region and another and these prevent the control pattern of one region extending into another region. This pattern allows large structures to be specified by a modest number of specific cell-to-cell relationships. Early the pgroup is organized so that precursors of boundary cells are located between the regional precursor cells and when they divide they continue to remain between cells of different types. They remain in interstitial positions and divide at the required rate, much slower than the regional cells. In this way the pgroup can control many different regions of an organ, all separated by boundary cells. A region often contains cells of more than one type such as a variety of glial cells and neurons in brain regions. Growth occurs by expanding the number of cells in each region and the boundary cells. The regions and boundaries are required for assembling large organs. Diffusible substances control cell division and growth and other aspects of their function.

X: Organs are held together by cell adhesion in functional relationships. Thus the form and function of the organism is specified entirely by local control mechanisms that establish the organs. The form of the organs is decided by the complex cell-to-cell relationships and the adhesion molecules that bind the cells. The organs are formed in association with each other and the adhesion characteristics of the appropriate parts bind them to each other, establishing the form of the organism. Some will contribute more to the form than others, because they may be larger or external. In a sense the division of an organism into organs is arbitrary and the pgroups are considered to apply to the actual requirements for assembly, and there is much to be learned. For indirect development the organs may become parts of the larva or pupa, whereas a subset of lineages form the pgroups of the imaginal disks or rudiment and become parts of the adult.

Not Only Form But Function As Well. The organism carries out all of the biochemistry and the entire set of inter-organ hormone and homeostatic relationships. The neural systems and the brain and underlying behavioral patterns also are formed, including those that change during growth and maturation to keep the individual fit and adapted to each stage. Healing and resistance to microbial invasion should not be neglected along with other things there is not enough space to list. The outcome is that natural selection operates on the health and behavior including the ability to mate and produce offspring that in turn are fertile. Any genetic variation in all of the steps required for organ formation and establishment of all of the features mentioned (or not mentioned) is selected by

this outcome. The model does not depend for its survival on our knowledge or ignorance of all of the details. It depends on the required logic and our knowledge of clues regarding the major processes.

Topics Regarding the Model

Self-Assembly. Self-assembly is commonly observed *in vitro* with partial systems. For example, histones assemble with naked DNA to form nucleosomes and chromatin (10). Tubulins form microtubules by self-assembly (11). Bacterial cell wall subunits self-assemble to form wall-like structure. Viral capsids self-assemble *in vitro* (12). Factors bind to specific sites in regulatory regions of genes and complex assemblies of regulatory protein molecules form, all by mass action. At every cell division the nuclear membrane disassembles and reassembles, except in yeast. Self-assembly is an explanation for some normal biological processes in fungi (13). Ribosomes are assembled and all of the steps in mRNA translation and protein synthesis operate by mass action. Self-assembly is a common process. It is no stretch of the imagination that as a cell grows and divides the macromolecules of a cell membrane or any other part of its structure diffuse into place where they are bound. Some processes do help place the units in the correct domains for assembly. Automatic self-assembly is an explanation of the formation of differentiated cells, depending on specific binding of units. Included are the tubulin, actin, and intermediate filament proteins that have an important role in cell structure. In the model the cells self-assemble, leading to the formation of pgroups and ultimately the organs. The specific adhesion characteristics of the cells underlie the assembly but, as mentioned, the process is not mass action because many cell locations are the result of duplication in place and some cells move in controlled fashion. Probably the adhesive binding as cells approach each other closely is analogous to self-assembly by mass action. The following comments indicate that there are different opinions about self-assembly (14): "There is no master coordinator, so large-scale organization has to arise from individual contingency relationships. Morphogenesis of cells into complicated shapes and the construction of complex organs like the brain on the basis simply of the principles of self-assembly is not feasible." The authors emphasize variation, taking highly variable angiogenesis as an example, and propose that "natural selection" occurs throughout development. However, I believe that it is feasible to build complex structures by self-assembly. In most circumstances the variation is limited, but not forbidden, by the binding and signaling relationships among cells. A computer program could symbolically make a structure of any sort, even variable, by using the principles of this model. Proof that this is the way nature works is, as always, required.

About Nuclear Control Factors. This brief note indicates the evolutionary complexity and enormous power of the relationships implied when factors are used in the model. (A) Factors work in teams. A single gene may have 50 or more sites in the 5' control region to which a dozen or more different factors bind, and the 5' region can be dissected into modules with specific control function, subject to quantitative modeling (15). Enhancers and other control devices may bind in other parts of the gene region as well. (B) There are networks of expression control among factors operating positively and negatively. In *Drosophila* dozens of such relationships are known, which is just the beginning because most of the control pathways for downstream nonregulatory genes are as yet unknown (<http://www.mssm.edu/molbio/genet/>). There is a hierarchy of factors based on "upstream and downstream" relationships but many downstream factors partially control upstream factor expression. (C) Many members of this network are nuclear factors but others like the Hh and Wnt families operate externally and cytoplasmically as part of the signaling system

(16, 17). (D) Some factors are highly conserved in evolution and carry out comparable roles in distant organisms (18). Others are conserved but carry out different roles, as if they were devices useful for many purposes (19, 20). (E) There are dozens of homeobox genes found in all species, a family that probably became widespread by duplication and evolution of new functions (21). (F) The clusters of Hox (Antp) occur in the genome in about the order they are expressed from end to end of the embryo in mouse or *Drosophila* (22, 23), although the reason for this conservation is unknown. (G) The number of factors is potentially very large but they can be placed in 30 classes based on their mechanism of function, and presumably their evolutionary branching (<http://transfac.gbf.de/>). This list (as of mid-1997) contains 2,285 factors of which 686 are in one species (human). There are 4,602 sites recognized in 929 genes or five sites per gene. Signaling elements and receptors are not included. (H) Some molecules, such as beta-catenin, carry out signaling, are involved in adhesion, and can act as nuclear regulatory factors in combination with others.

Note on Formation of the Structure of Cells. During assembly each of the units has to be bonded into the pre-existing structure with binding valences specific to adjacent units. The whole of the structural information leading to the form is expressed as assembly occurs. Each unit has the capability of specifically binding to a site and then forms binding sites for one or more other units that bind in succeeding events. It is the sequence of assembly events that forms the cells. Transcription regulation assures that all of the units are expressed along with other required molecules, that they are present in correct concentrations and properly modified for assembly. The phrase mass action needs qualification because many units are produced en masse and reach the membrane in vesicles. The details of what happens upon their release are not known but presumably there is short-distance diffusion to their target sites. There are a number of specific vesicles (24), but the total range of different vesicles is not known. The vesicles are not free and the cytoplasmic structure and tubulin links affect their motion (25). In the model there are more kinds of differentiated cells than the usual meaning. That is, there are a very large number of different types of cells as a result of specific recognition and adhesion sites compared with the few hundred types of differentiated cells in the usual sense.

Note on Oogenesis. All living systems except a few parasites (viruses and DNA elements) have an unbroken lineage from cell to cell going back to some cell precursor, perhaps part of the origin of life, many billions of years ago. So the starting point can only be a complete cell. Thus in oogenesis the first units are bound to pre-existing structure and the pre-existing parental cell informs the early stages. In the model each molecule that will be a part of the egg structure binds into place onto existing specific binding valences and, in turn, exposes its own binding valences for succeeding molecules to bind onto. In the sea urchin *Strongylocentrotus purpuratus* oogenesis the germ-line cell grows from a few microns to 80 microns and perhaps 10^{12} macromolecules are laid down in their correct relationships. Some units enter the egg in many copies, as for example membrane structural proteins. The membrane proteins have affinities for each other and are specifically modified by glycosylation, and after modification they bind into the lipid layers. They cannot each have individual specific binding sites. Presumably vesicles from the Golgi apparatus deliver large numbers that all can enter the membrane in an organized process. The mass assembly of membrane is interrupted to permit sperm receptors, etc., to be assembled. Among many steps the vitelline layer and the cortical granules must be assembled and the growth terminated. In the model this process is automatic and incorporation results from expression of units under control of the set of nuclear factors in the egg and nurse cells that changes during the course of oogenesis, for example when vitellogenesis starts.

Sea Urchin Micromeres. An example of autonomous development is worth reviewing because it shows important higher-level relationships. At the 16-cell stage of sea urchin cleavage four smaller cells termed micromeres are present that lead to a population of cells called the primary mesenchyme (PMC). The descendants of these cells are fated to form the skeletal structure of the sea urchin larva. If cultured as individual cells they go through much of what would have been their fate in the larva including expressing specific proteins of the skeletal structures and forming bits of skeleton. The micromeres can be considered a pgroup for skeletal structure in the embryo. In early cleavage in sea urchins the asymmetric cell divisions can be considered the starting point of the pgroup, and the member cells have considerable autonomy. Careful studies of this process (26) show that sequential changes in the binding affinities occur as the PMC go through stages and finally fuse to form a syncytium and begin skeleton production. In a remarkable example of feedback control, if the micromeres are removed from the gastrula stage embryo other mesenchyme cells take over their role and ultimately form skeleton. During normal development the PMC release active substances to prevent this replacement process (27). This finding is a remarkable insight into the upper-level relationships where a pgroup and its progeny have to “defend their right” to carry out normal function. Though apparently autonomous early in development they apparently require contact with ectoderm to complete the whole species-specific skeletal structure (28). They are also powerful inducers and if surgically implanted in the animal cap of a normal sea urchin embryo they cause the formation of a second archenteron. Then development continues and two skeletal structures are formed (29).

More Autonomous Examples and Ectopic Structures. The *ey* (eyeless) gene is presumably involved in the control of the eye pgroup of *Drosophila* along with five other regulatory genes that cause the eyeless defect and cause ectopic eyes to form (9). Transformation with pairs of these genes causes a synergistic interaction that induces great numbers of eyes in many locations, for example, *So* (30) or *dachshund* (31) combined with the *eyes-absent* gene. In the model these regulatory genes activate genes that produce one or more units that normally occur in cells of the initial eye pgroup as part of feedback loops. When these precursor or key units for the eye are produced in high concentrations in ectopic locations in transformation experiments they influence a cell to start the pgroup for eye formation. The synergism of multiple factors is accounted for by assuming that each factor is a part of the feedback relationships in the pgroup. According to the model, signaling among the pgroup cells establishes the differentiated cells of the initial parts of the eye structure. This activity is followed by further cell formation and cell–cell interactions, leading to the eye structure in an ectopic location. The evolutionary conservation of the role of factors in some pgroups is remarkable. The chicken *ey* analog will replace *ey* of *Drosophila* and produce *Drosophila*-type eyes rather than chicken-type eyes in the fly. Because of this conservation the *ey* gene was named a master regulator (8, 32), but it is better described as one of several regulators of the pgroup that can act as an initiator. Another germane case is the antennapedia mutation of *Drosophila* where a leg replaces an antenna. A mutation occurs in a factor or signal control of the normal antenna pgroup in the imaginal disk and the related leg pgroup replaces it. This evidence indicates that pgroups for different structures may be evolutionarily related and have conserved many elements. Transdetermination observations and many gain-of-function mutants in *Drosophila* lead to a similar conclusion. In the sea urchin embryo it is remarkably easy to initiate a pgroup unless there are specific suppressors present and the same probably holds for *Drosophila*.

Conditional (Regulative) Ability. After a number of divisions the sea urchin embryo hatches and proceeds at about 27 hr to gastrulate. The regulative abilities at this stage are worth sketching, and the following part of an abstract is germane (27). “Gastrulation in the sea urchin involves an extensive rearrangement of cells of the archenteron giving rise to secondary mesenchyme at the archenteron tip followed by the foregut, midgut, and hindgut. To examine the regulative capacity of this structure, pieces of the archenteron were removed or transplanted at different stages of gastrulation. After removal of any or all parts of the archenteron, the remaining veg 1 and/or veg 2 tissue regulated to replace the missing parts. Endoderm transplanted to ectopic positions also regulated to that new position in the archenteron. . . We propose that this regulative ability requires extensive and continuous short-range communication between cells of the archenteron to reorganize the tissues and position the boundaries of this structure even after experimental alterations.” Of course the foregut, midgut, and hindgut listed include or are the pgroups for these parts of the larval organs. Apparently the pgroup cells, as in the case of micromeres, are signaling to prevent other cells from taking over their roles and when removed other cells do just that. The formation of pgroups when cells are surgically removed indicates that many or all cells have enormous capacities that normally are suppressed. This summary version of the model can only point to the subtlety of early decisions of cell membership in pgroups. That is particularly evident in mammalian embryos where given cells may form lineages with different fates in different individuals.

Discussion

Focus on the Molecules. The focus of the model is on the units including control factors and signaling systems and the local specific relationships between cells and their neighbors. That is like saying the foot soldiers of an army are responsible for its victories. In a sense that is always so, but what about the generals and staff officers? The army wouldn't run without them. As far as I can tell there are none in biological development. That leads to a conclusion expressed by an example. Follow in detail a given pgroup and the cell lineage within it that leads through all of its branches to an adult neuron with a crucial role in the brain. The pgroup shaped the region of the brain and all of the steps and interactions of the lineage leading to this one neuron are just what are required, and all have been honed by selection and that is all there is to it. Certainly there is much feedback and contributions by other cells and tissues at each stage so that the development of this lineage was only a part of a system with many strong requirements. That adds up to a very abstract answer to the question: How do all of the marvelous biological structures we see take their form?

Confirmation: Features of the Model that Need Testing. (i) The process of cell self-assembly in the model requires many structural units that specifically bind, and a search needs to be made for a large class of such molecules. The model predicts that genetic defects that affect development would result from mutations in the genes of many such units. (ii) At present developmental processes are identified by signaling molecules or factors neglecting the feedback loop between cell surface signaling and factors and gene expression, and it appears that complete feedback descriptions would lead to new insights. (iii) The proposal that development results from self-assembly without global control mechanisms needs direct demonstration. An approach might be to examine by transformation all of the successive steps in the formation of an organ, considering all possible parallel or redundant stages. If there were global overarching controls a different sort of persistence of the organ would be expected in the face of damage to the individual stages. (iv) The major focus of present research is on

the steps in development and their control mechanisms. The time is ripe for carrying an analysis step by step all the way to an adult organ. (v) A direct examination of regions and boundaries is called for describing the simplicity of the intercellular relationships in the regions and how the boundaries function as well as their specific growth rates.

Major Conserved Features. A question is whether the large-scale characteristics of an animal can be formed without large-scale guidance systems. Is anything needed in addition to all of the binding and regulatory information of the model? How do the four limbs of tetrapods or 5-fold symmetry of sea urchins form? They are formed from the details and fate of individual cells. It would be satisfactory for theorists of development to be able to point to a biological feature of regulation, which corresponds to some large-scale structure such as tetrapodism, and point out why it was conserved in evolution, but that has not been done. The model suggests nothing of the kind because it uses only local interactions. The explanation for conservation of such a major feature is that it is hard to escape from. It is difficult to make so many consistent changes. The evolutionary evidence suggests that minor modifications can occur as in the evolution of forelegs to wings in birds and to arms in primates, and the slow elimination of legs as in whales and dolphins while leaving traces in the skeleton. It is logical that pgroups could be slowly modified in this way. Transdetermination and the conditional (regulative) events mentioned also suggest that major events of switching are possible that change the evolutionary fates of pgroups. The organization of pgroups or potential pgroups is powerful. In the absence of suppression they run out their course, completing intricate steps leading to eyes or legs or whatever. This process happens also ectopically, and the path also can be easily switched in transdetermination. Therefore changes of pgroups have striking possibilities that supply opportunities in evolution. Slight changes in fundamental control linkages could lead to new forms of the organ specified. This hypothesis opens up the possibility of macroevolution, restricted by the number of internal links in the pgroups but supported by the organizational power of the pgroups and the flexibility of conditional development.

Where Is the Information Stored? This was my original question and a summary of the answer under the model follows. The structure of the germ cell contains required information. It cannot be bypassed in animals but we don't know just what is critical (in addition to its genome). Plants do bypass the need for a germ line but a cell always is required. The maternal contribution includes at least structure, mRNA, factors, and localization information as well as the material for the cleavage stages of early development. That arises from gene expression during oogenesis including nurse cells. The germ cell or growing plant cell ultimately is derived from gene expression. Thus it all goes back to the DNA though the route may be more or less indirect. What part of the DNA matters? All we know about are the coding regions and the associated regulatory sequences in the DNA. The important genes include all of the units, the factors, the signaling molecules, receptors, and the adhesion molecules. Obviously the binding sites and binding regions of the units required for cell assembly are important. The information for form is coded in all of these

genes, including regulatory regions, and their relationships. As yet decoding requires an embryo.

Thanks go to Eric Davidson for saying that what I wrote in a draft as being logically necessary was obvious and to Elliot Meyerowitz, Bill Klein, Geoffrey Graham, and Costas Flytzanis for comments. This work was supported by National Institutes of Health grants.

- de Terra, N. (1982) in *Developmental Order: Its Origin and Regulation*, eds. Subtelny, S. & Green, P. B. (Liss, New York), pp. 165–182.
- Davidson, E. H. (1994) *BioEssays* **16**, 603–615.
- Edelman, G. M., Cunningham, B. A. & Thiery, J. P., eds. (1990) *Morphoregulatory Molecules* (Wiley, New York).
- Brümmendorf, T. & Rathjen, F. G. (1994) *Protein Profile* (Academic, London), Vol. 1, No. 9.
- Jones, F. S., Kioussi, C., Copertino, D. W., Kallunki, P., Holst, B. D. & Edelman, G. M. (1997) *Proc. Natl. Acad. Sci. USA* **94**, 2632–2637.
- Miller, J. R. & McClay, D. R. (1997) *Dev. Biol.* **192**, 310–322.
- Miller, J. R. & McClay, D. R. (1997) *Dev. Biol.* **192**, 323–339.
- Callaerts, P., Halder, G. & Gehring, W. J. (1997) *Annu. Rev. Neurosci.* **20**, 483–532.
- Desplan, C. (1997) *Cell* **91**, 861–864.
- Hansen, J. C., van Holde, K. E. & Lohr, D. (1991) *J. Biol. Chem.* **266**, 4276–4282.
- Odde, D. J. (1997) *Biophys. J.* **73**, 88–96.
- McCarthy, M. P., White, W. I., Palmer-Hill, F., Koenig, S. & Suzich, J. A. (1998) *J. Virol.* **72**, 32–41.
- Wosten, H. A., Schuren, F. H. & Wessels, J. G. (1994) *EMBO J.* **13**, 5848–5854.
- Gerhart, J. & Kirschner, M. (1997) *Cells, Embryos, and Evolution* (Blackwell Science, Malden, MA), p. 146.
- Yuh, C.-H., Bolouri, H. & Davidson, E. H. (1998) *Science* **279**, 1896–1902.
- Ingham, P. W. (1995) *Curr. Opin. Genet. Dev.* **5**, 492–498.
- Nusse, R. & Varmus, H. E. (1992) *Cell* **69**, 1073–1087.
- Bürglin, T. R. (1994) in *Guidebook to the Homeobox Genes*, ed. Duboule, D. (Oxford Univ. Press, New York), pp. 25–71.
- Lowe, C. J. & Wray, G. A. (1997) *Nature (London)* **389**, 718–721.
- Davidson, E. H. (1997) *Nature (London)* **389**, 679–680.
- Duboule, D., ed. (1994) *Guidebook to the Homeobox Genes* (Oxford Univ. Press, New York).
- De Robertis, E. M. (1994) in *Guidebook to the Homeobox Genes*, ed. Duboule, D. (Oxford Univ. Press, New York), pp. 11–23.
- Carroll, S. B. (1995) *Nature (London)* **376**, 479–485.
- Matese, J. C., Black, S. & McClay, D. R. (1997) *Dev. Biol.* **186**, 16–26.
- Welte, M. A., Gross, S. P., Postner, M., Block, S. M. & Wieschaus, E. F. (1998) *Cell* **92**, 547–557.
- McClay, D. R., Wray, G. A., Alliegro, M. C. & Ettensohn, C. A. (1990) in *Morphoregulatory Molecules*, eds. Edelman, G. M., Cunningham, B. A. & Thiery, J. P. (Wiley, New York), pp. 533–546.
- McClay, D. R. & Logan, C. Y. (1996) *Development (Cambridge, U.K.)* **122**, 607–616.
- Armstrong, N. & McClay, D. R. (1994) *Dev. Biol.* **162**, 329–338.
- Ransick, A. & Davidson, E. H. (1993) *Science* **259**, 1134–1138.
- Pignoni, F., Hu, B., Zavitz, K. H., Xiao, J., Garrity, P. A. & Zipursky, S. L. (1997) *Cell* **91**, 881–891.
- Chen, R., Amoui, M., Zhang, Z. & Mardon, G. (1997) *Cell* **91**, 893–903.
- Halder, G., Callaerts, P. & Gehring, W. J. (1995) *Science* **267**, 1788–1792.