Chapter 2

COMPLEXITY AND SELF-ORGANIZATION IN BIOLOGICAL DEVELOPMENT AND EVOLUTION

Stuart A. Newman
New York Medical College, Valhalla, NY 10595
Gabor Forgacs
University of Missouri, Columbia, MO 65211

1. Introduction: Complex Chemical Systems in Biological Development and Evolution

The field of developmental biology has as its major concern *embryogenesis*: the generation of fully-formed organisms from a fertilized egg, the *zygote*. Other issues in this field, organ *regeneration* and tissue *repair* in organisms that have already passed through the embryonic stages, have in common with embryogenesis three interrelated phenomena: *cell differentiation*, the production of distinct cell types, *cell pattern formation*, the generation of specific spatial arrangements of cells of different types, and *morphogenesis*, the molding and shaping of tissues [1]. The cells involved in these developmental processes and outcomes generally have the same genetic information encoded in their DNA, the *genome* of the organism, so that the different cell behaviors are largely assocated with *differential gene expression*.

Because of the ubiquity and importance of differential gene expression during development, and the fact that each type of organism has its own unique genome, a highly gene-centered view of development prevailed for several decades after the discovery of DNA's capacity to encode information. In particular, development was held to be the unfolding of a "genetic program" specific to each kind of organism and organ, based on differential gene expression. This view became standard despite the fact that no convincing models had ever been presented for how genes or their products (proteins and RNA molecules) could alone build three

dimensional shapes and forms, or even generate populations of cells that utilized the common pool of genetic information in different ways. By "alone" is meant without the help of physics and chemical dynamics, the scientific disciplines traditionally invoked to explain changes in shape, form, and chemical composition in nonliving material systems.

It has long been recognized that biological systems are also physicochemical systems, and that phenomena first identified in the nonliving world can also provide models for biological processes. Indeed, at the level of structure and function of biomolecules and macromolecular assemblages such as membranes, biology has historically drawn on ideas from chemistry and physics. More recently, however, there has been intensified cooperation between biological and physical scientists at the level of complex biological systems, including developmental systems. Applications of results from these fields at the systems level, however, had to wait until it became clear to modern biologists (as it was to many before the "gene revolution") that organisms are more than programmed expressions of their genes, and that the behavior of systems of many interacting components is neither obvious nor predictable on the basis of the behavior of their parts. These new approaches also could not have occurred until physical scientists began fully engaging with experimentally-derived details of the phenomena in question (a departure from much of earlier "theoretical biology"), and developed theoretical tools and computational power sufficient to model systems of great complexity. This new systems biology has been emerging over the last decade.

Since biology, with the help of chemistry and physics, is increasingly studied at the systems level, there has also been new attention to the origination of such systems. In many instances, for example, it is reasonable to assume that complexity and integration in living organisms has evolved in the context of forms and functions that originally emerged (in evolutionary history) by straightforward physicochemical means. Thus the elaborate system of balanced, antagonistic signaling interactions that keep cell metabolism homeostatic, and embryogenesis on-track, can be seen as the result of accretion, by natural selection, of stabilizing mechanisms for simpler physicochemical generative processes that would otherwise be less reliable.

The richness of the phenomena of development prohibits any comprehensive review of physicochemical approaches to their analysis in the space of a single chapter. But a rough separation can be made between those processes that are largely the province of physics—the folding and stretching

of cell sheets and the molding and separation of cell masses—and those that are the province of chemical dynamics—switching between alternative stationary compositional states leading to cell differentiation, chemical oscillations resulting in segmental organization of tissues, and breaking of symmetry by reaction-diffusion coupling leading to cellular pattern formation. It is the latter group of processes and phenomena that we will consider below. But even in this narrowed set it will possible to describe only a limited selection of biological examples and proposed explanatory models. Finally, we will describe a hypothesized scenario by which a chemical-dynamical mechanism of development that plausibly originated a class of embryonic patterns early in the history of multicellular life was transformed over the course of evolution into a more reliable genetically-programmed mechanism for producing the same forms.

In many cases, specific molecules that participate in the chemical-dynamic mechanisms we discuss will be indicated, in recognition both of the enormous progress that has been made in recent years in identifying the genes and gene products involved in complex developmental processes, and of the credibility of proposed theoretical frameworks insofar as they deal with characterized, experimentally-accessible components. But readers mainly interested in the formal aspects of the processes under discussion may look past the molecular names with little disadvantage.

2. Dynamic Multistability and Cell Differentiation

The early embryos of multicellular organisms are referred to as *blastulae*. These are typically (but not invariably) hollow clusters of several dozen to several hundred cells. While the zygote, the fertilized egg that gives rise to the blastula, is ”totipotent,” that is, it has the potential to give rise to any of the more than 200 specialized cell types (e.g., the various types of muscle, blood, and nerve cells) of the mature human body, by the blastula stage the cells are no longer identical. In most species, the first few cell divisions in the embryo generate cells that are “pluripotent”—capable of giving rise to only a limited range of cell types. These cells, in turn, diversify into cells with progressively limited potency, ultimately generating all the (generally unipotent) specialized cells of the body [1].

The transition from wider to narrower developmental potency is referred to as *determination*. This stage of cell specialization generally occurs with

no overt change in the appearance of cells. Instead, subtle modifications, only discernable at the molecular level, set the altered cells on new and restricted developmental pathways. A later stage of cell specialization, referred to as *differentiation*, results in cells with vastly different appearances and functional modifications—electrically excitable neurons with extended processes up to a meter long, bone and cartilage cells surrounded by solid matrices, red blood cells capable of soaking up and disbursing oxygen, and so forth. Cells have generally become determined by the end of blastula formation, and successive determination increasingly narrows the fates of their progeny as development progresses. When the developing organism requires specific functions to be performed, cells will typically undergo differentiation.

Since each cell of the organism contains an identical set of genes (except for the egg and sperm and their immediate precursors, and some cells of the immune system), a fundamental question of development is how the same genetic instructions can produce different types of cells. This question pertains to both determination and differentiation. Since these two kinds of cell specialization are formally similar and probably employ overlapping set of molecular mechanisms, we will refer to both as "differentiation" in the following discussion, unless confusion would arise. Multicellular organisms solve the problem of specialization by activating only a type-specific subset of genes in each cell type.

The *biochemical state* of a cell can be defined as the list of all the different types of molecules contained within it, along with their concentrations. The dynamical state of a cell, like that of any dynamical system, resides in a multidimensional space, the "state space," with dimensionality equal to the number of system variables (e.g., chemical components) [2]. During the cell division cycle (i.e., the sequence of changes that produces two cells from one), also called simply the cell cycle, the biochemical state changes periodically with time. (This, of course, assumes that cells are not undergoing differentiation.) If two cells have the same complement of molecules at corresponding stages of the cell cycle, then, they can be considered to be of the same differentiated state. The cell's biochemical state also has a spatial aspect—the concentration of a given molecule might not be uniform throughout the cell. We will discuss this in Section 4, below. Certain properties of the biochemical state are highly relevant to understanding developmental mechanisms. The state of differentiation of the cell (its *type*) can be identified with the collection of proteins it is capable of making.

Of the estimated 20,000-25,000 human genes [3], a large proportion constitutes the "housekeeping genes," involved in functions common to all or most cells types [4]. In contrast, a relatively small number of genes—possibly fewer than a thousand—specify the type of determined or differentiated cells; these genes are *developmentally regulated* (i.e., turned on and off in an embryonic stage- and embryonic position-dependent fashion) during embryogenesis. The generation of cell type diversity during development is based to a great extent on sharp transitions in the dynamical state of embryonic cells, particularly with respect to their developmentally-regulated genes.

2.1. Cell states and dynamics

When a cell divides it inherits not just a set of genes and a particular mixture of molecular components, but it also inherits a dynamical system at a particular *dynamical state*. Dynamical states of many-component systems can be transient, stable, unstable, oscillatory, or chaotic [2]. The cell division cycle in the early stages of frog embryogenesis, for example, is thought to be controlled by a *limit cycle* oscillator [5]. A limit cycle is a continuum of dynamical states that define a stable orbit in the state space surrounding an unstable *node*, a node being a stationary (i.e., time-independent) point, or steady state, of the dynamical system. The fact that cells can inherit dynamical states was demonstrated experimentally by Elowitz and Leibler [6]. These investigators used genetic engineering techniques to provide the bacterium *Escherichia coli* with a set of feedback circuits involving transcriptional repressor proteins such that a biochemical oscillator not previously found in this organism was produced. Individual cells displayed a chemical oscillation with a period longer than the cell cycle. This implied that the dynamical state of the artificial oscillator was inherited across cell generations (if it had not, no periodicity distinct from that of the cell cycle would have been observed). Because the biochemical oscillation was not tied to the cell cycle oscillation, newly divided cells in successive generations found themselves at different phases of the engineered oscillation.

The ability of cells to pass on dynamical states (and not just "informational" macromolecules such as DNA) to their progeny has important implications for developmental regulation, since continuity and stability of a cell's biochemical identity is key to the performance of its role in

a fully developed organism. Inheritance of alternative cell states that does not depend on sequence differences in inherited genes is called "epigenetic inheritance" [7], and the biological or biochemical states inherited in this fashion are called "epigenetic states." Although epigenetic states can be determined by reversible chemical modifications of DNA [8], they can also represent alternative steady states of a cell's network of active or expressed genes [9]. The ability of cells to undergo transitions among a limited number of discrete, stable epigenetic states and to propagate such decisions from one cell generation to the next is essential to the capacity of the embryo to generate diverse cell types.

All dividing cells exhibit oscillatory dynamical behavior in the subspace of the full state space whose coordinates are defined by cell cycle–related molecules. In contrast, cells exhibit alternative stable steady states in the subspace defined by molecules related to states of cell determination and differentiation. During development, a dividing cell might transmit its particular system state to each of its daughter cells, but it is also possible that some internal or external event accompanying cell division could push one or both daughter cells out of the "basin of attraction" in which the precursor cell resided and into an alternative state. (The basin of attraction of a stable node is the region of state space surrounding the node, in which all system trajectories, present in this region, terminate at that node [2].)

It follows that not every molecular species needs to be considered simultaneously in modeling a cell's transitions between alternative biochemical states. Changes in the concentrations of the small molecules involved in the cell's housekeeping functions such as energy metabolism and amino acid, nucleotide, and lipid synthesis, occur much more rapidly than changes in the pools of macromolecules such as RNAs and proteins. The latter are indicative of the cell's gene expression profile and can therefore be considered against an average metabolic background. And even with regard to gene expression profile, most of a cell's active genes are kept in the "on" state during the cell's lifetime, since they are also mainly involved in housekeeping functions. The pools of these "constitutively active" gene products can often be considered constant, with their concentrations entering into the dynamic description of the developing embryo as fixed parameters rather than variables. (See Goodwin [10] for an early discussion of separation of timescales in cell activities.)

As mentioned above, it is primarily the regulated genes that are important to consider in analyzing determination and differentiation. And of these

regulated genes, the most important ones for understanding developmental transitions are those whose products control the activity of other genes.

2.2. Epigenetic multistability: the Keller autoregulatory transcription factor network model

Genes are regulated by a set of proteins called *transcription factors*. Like all proteins, these factors are themselves gene products, with the specific function of turning genes on and off. They do this by binding to specific sequences of DNA usually (but not always) "upstream" of the gene's transcription start site, called the "promoter." Developmental transitions are frequently controlled by the relative levels of transcription factors [11].Because the control of most developmentally-regulated genes is a consequence of the synthesis of the factors that regulate their transcription, transitions between cell types during development can be driven by changes in the relative levels of a fairly small number of transcription factors. We can thus gain insight into the dynamical basis of cell type switching (i.e., determination and differentiation) by focusing on molecular circuits, or networks, consisting solely of transcription factors and the genes that specify them. Networks in which the components mutually regulate one another's expression are termed "autoregulatory." During development, cells contain a variety of autoregulatory transcription factor circuits. It is obvious that such circuits will have different properties depending on their particular "wiring diagrams" that is the interaction patterns between components.

Transcription factors can be classified as either *activators*, which bind to a site on a gene's promoter, and enhance the rate of that gene's transcription over its basal rate, and *repressors*, which decrease the rate of a gene's transcription when bound to a site on its promoter. The basal rate of transcription depends on constitutive transcription factors, which are distinct from those in the autoregulatory circuits that we consider below. Repression can be competitive or noncompetitive. In the first case, the repressor will interfere with activator binding and can only depress the gene's transcription rate to the basal level. In the second case, the repressor acts independently of any activator and can therefore potentially depress the transcription rate below basal levels.

Keller [12] used simulation methods to investigate the behavior of several autoregulatory transcription factor networks with a range of wiring diagrams (Fig. 2.1). Each network was represented by a set of n coupled

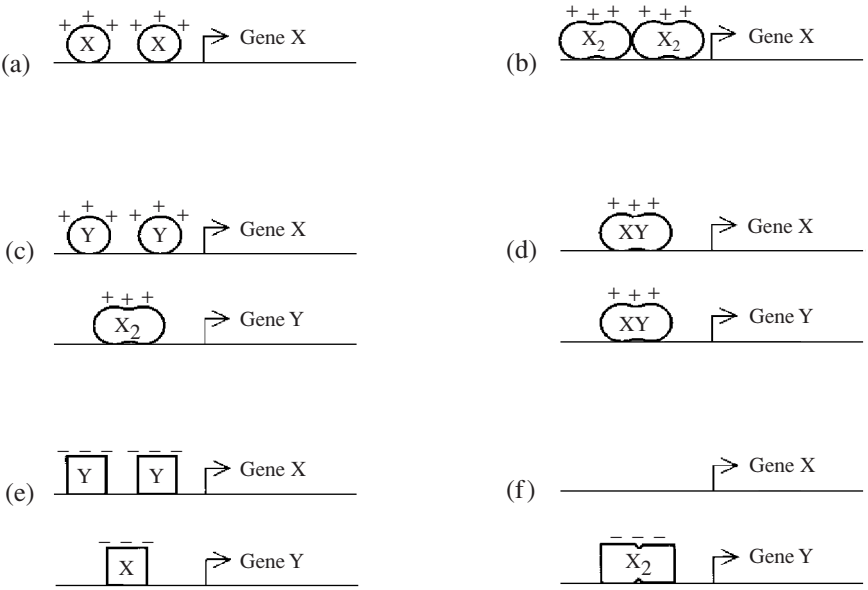


Figure 2.1. Six model genetic circuits discussed by Keller [12]. (a) Autoactivation by monomer X; (b) autoactivation by dimer X_2; (c) mutual activation by monomer Y and dimer X_2; (d) autoactivation by heterodimer XY; (e) mutual repression by monomers X and Y; (f) mutual repression by dimer X_2 and heterodimer XY. Activation and repression functions are represented respectively by $+$ and $-$. The various transcription factors bind to the promoters of the genes. Figure used by permission of Elsevier Publishing Co.

ordinary differential equations—one for the concentration of each factor in the network—and the steady-state behaviors of the systems were explored. The questions asked were: how many stationary states exist; are they stable or unstable?

Here we discuss in detail one such network, designated as the "Mutual repression by dimer and heterodimer" (MRDH), shown in Fig. 2.1f. It comprises the gene encoding the transcriptional repressor, X, and the gene encoding the protein, Y, and thus represents a two component network. Below are listed the salient points of the MRDH model.

The rate of synthesis of a transcription factor is proportional to the rate of transcription of the gene encoding that factor. Transcription of a gene, in turn, depends on the specific molecules that are bound to sites in the gene's promoter. These can be monomeric protein molecules (X, Y), homodimers

(X_2, Y_2) or heterodimers (XY). Thus a promoter can be in various configurations, with respective relative frequencies, depending on its occupancy. Specifically, in the case of the MRDH network, as characterized by Keller (Fig. 2.1f) the promoter of gene X has no binding site for any activator or repressor molecule; its only configuration is the empty state, whose frequency therefore can be taken as 1. The basal rate of synthesis of X is denoted by S_{X_B}. Protein X is a non-competitive repressor of Y, whose promoter contains a single binding site for X_2. The monomeric forms of proteins X and Y and the heterodimer XY cannot bind DNA. Thus, while protein Y does not act directly as a transcription factor, it affects transcription since it antagonizes the repressor function of X by interfering with the formation of X_2. The promoter of gene Y can therefore be in two configurations: its binding site for X_2 is either occupied or not, with respective relative frequencies $K_X K_{X_2}[X]^2/(1 + K_X K_{X_2}[X]^2)$ and $1/(1 + K_X K_{X_2}[X]^2)$. Here K_X and K_{X_2} are, respectively, the binding affinity of X_2 to the promoter of gene Y and the dimerization rate constant for the formation of X_2 (we used the relationship $K_X[X_2] = K_X K_{X_2}[X]^2$). Synthesis of Y in both configurations is activator-independent with a rate denoted by S_{Y_B}. To incorporate the fact that X_2 reduces the rate of transcription of Y in a non-competitive manner, in the occupied Y promoter configuration $S_{Y_B}$is replaced with ρS_{Y_B}, with $\rho \leq 1$.

The overall transcription rate of a gene is calculated as the sum of products. Each term in the sum corresponds to a particular promoter occupancy configuration and is represented as a product of the frequency of that configuration and the rate of synthesis resulting from that configuration. In the MRDH network this rate for gene X is thus $1 \times S_{X_B}$, because it has only a single (empty) promoter configuration. The promoter of gene Y can be in two configurations (with rates of synthesis S_{Y_B} and ρS_{Y_B}, see above), therefore its overall transcription rate is $(1 + \rho K_X K_{X_2}[X]^2) S_{Y_B}/(1 + K_X K_{X_2}[X]^2)$.

The rate of decay of a transcription factor is a sum of terms, with each being proportional to the concentration of a particular complex in which the transcription factor participates. This is equivalent to assuming exponential decay. For the transcriptional repressor X in the MRDH network these complexes include the monomer X, the homodimer X_2 and the heterodimer XY. Denoting the corresponding decay constants as d_X, $d_{X_2}$and d_{XY}, the overall decay rate of X is given by $d_X[X] + 2d_{X_2} K_{X_2}[X]^2 + d_{XY} K_{XY}[X][Y]$, with K_{XY} being the rate constant for the formation of the heterodimer

XY (i.e., $[XY] = K_{XY}[X][Y]$). The analogous quantity for protein Y is $d_Y[Y] + d_{XY}K_{XY}[X][Y]$.

With the above ingredients, the steady-state concentrations of X and Y in the MRDH network, are determined by

$$\frac{d[X]}{dt} = S_{X_B} - \left\{d_X[X] + 2d_{X_2}K_{X_2}[X]^2 + d_{XY}K_{XY}[X][Y]\right\} = 0 \tag{2.1}$$

$$\frac{d[Y]}{dt} = \frac{1 + \rho K_X K_{X_2}[X]^2}{1 + K_X K_{X_2}[X]^2} S_{Y_B} - \{d_Y[Y] + d_{XY}K_{XY}[X][Y]\} = 0. \tag{2.2}$$

Keller found that if in the absence of the repressor X the rate of synthesis of protein Y is high, in its presence, the system described by Eqs. (2.1) and (2.2) exhibits three steady states, as shown in Fig. 2.2. Steady states 1 and 3 are stable, thus could be considered as defining two distinct cell types, while

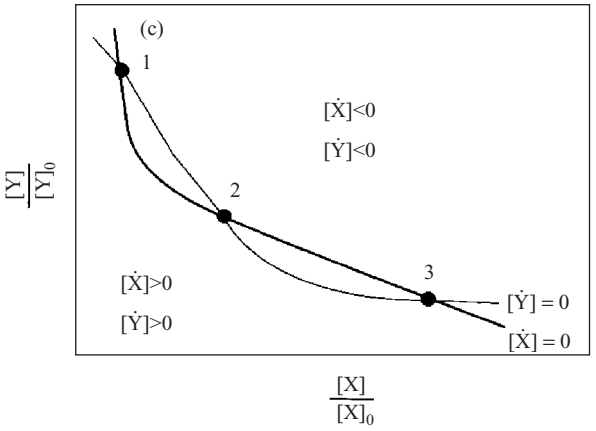


Figure 2.2. The solutions of the steady-state Eqs. 1 and 2, given in terms of the steady state solutions $[X_0]$ and $[Y_0]$. Here $[X_0]$ is defined as the steady state cellular level of monomer X produced in the presence of steady state cellular level $[Y_0]$ of monomer Y by the rate of transcription $[S_{X_0}]$. Thus by definition (see Eq. 1), $S_{X_0} = d_X[X_0] + 2\,d_{X_2}\,K_{X_2}\,[X_0]^2 + d_{XY}\,K_{XY}\,[X_0]\,[Y_0]$. Since along the thick and thin lines, respectively, $d[X]/dt \equiv [X^Y] = 0$ and $d[Y]/dt \equiv [\dot{Y}] = 0$, the intersections of these curves correspond to the steady state solutions of the system of equations, Eqs.1 and 2. Steady states 1 and 3 are stable, while steady state 2 is unstable. From Keller (1995) [12]. Used by permission of Elsevier Publishing Co.

steady state 2 is unstable. In an example using realistic kinetic constants, the steady-state values of [X] and [Y] at the two stable steady states differ substantially from one another, showing that the dynamical properties of these autoregulatory networks of transcription factors can provide the basis for generating stable alternative cell fates during early development.

The biological validity of Keller's model depends on whether switching between these alternative states is predicted to occur under realistic conditions. By this criterion the model works: the microenvironment of a cell, containing an autoregulatory network of transcription factors, could readily induce changes in the rate of synthesis of one or more of the factors via signal transduction pathways that originate outside the cell ("outside-in" signaling [13]). Moreover, the microenvironment can also affect the activity of transcription factors in an autoregulatory network by indirectly interfering with their localization in the cell's nucleus, where transcription takes place [14]. In addition, cell division may perturb the cellular levels of autoregulatory transcription factors, particularly if they or their mRNAs are unequally partitioned between the daughter cells. Any jump in the concentration of one or more factors in the autoregulatory system can bring it into a new basin of attraction and thereby lead to a new stable cell state.

2.3. Dependence of differentiation on cell-cell interaction: the Kaneko-Yomo "isologous diversification" model

The Keller model shows that the existence of multiple steady states in an embryonic cell's state space makes it possible, in principle, for more than one cell type to arise among its descendents. However this capability does not, by itself, provide the conditions under which such a potentially divergent cell population would actually be produced and persist as long as it is needed.

Experimental observations suggest that cell differentiation depends on properties of multicellular aggregates rather than simply those of individual cells. For example, during muscle differentiation in the early frog embryo, the muscle precursor cells must be in contact with one another throughout gastrulation (the set of rearrangements that establish the body's main tissue layers) in order to develop into terminally differentiated muscle [15, 16].

The need for cells to act in groups in order to acquire new identities during development has been termed the "community effect" [15]. This

phenomenon is a developmental manifestation of the general property of cells and other dynamical systems of assuming one or another of their possible internal states in a fashion that is dependent on inputs from their external environment. In the case noted above the external environment consists of other cells of the same genotype.

Kaneko, Yomo and co-workers [17-20] have described a previously unknown chemical-dynamic process, termed "isologous diversification," by which replicate copies of the same dynamical system (e.g., cells of the same initial type) can undergo stable differentiation simply by virtue of exchanging chemical substances with one another. This differs from the model described above in that the final state achieved exists only in the phase space of the collective "multicellular" system. Whereas the distinct local states of each cell within the collectivity are mutually reinforcing, these local states are not necessarily attractors of the dynamical system representing the individual cell, as they are in Keller's model. The Kaneko-Yomo system thus provides a model for the community effect.

The following is a simple version of the model, based on Kaneko and Yomo [17]. Improvements and generalizations of the model presented in subsequent publications [18-20] do not change its qualitative features.

Kaneko and Yomo consider a system of originally identical cells with intra- and inter-cell dynamics, which incorporate cell growth, cell division and cell death. The dynamical variables are the concentrations of molecular species ("chemicals") inside and outside the cells. The criterion by which differentiated cells are distinguished is the average of the intracellular concentrations of these chemicals (over the cell cycle). As a vast simplification only three chemicals, A, B and S, with respective time-dependent intracellular ($x_i^A(t)$, $x_i^B(t)$, $x_i^S(t)$ in the ith cell) and intercellular ($X^A(t)$, $X^B(t)$, $X^S(t)$) concentrations are considered in each cell and the surrounding medium. One of those (S) serves as source for the others. The model has the following features:

The source chemical S is catalyzed by a constitutive enzyme to produce chemical A, which in turn is catalyzed by a regulated enzyme to produce the chemical B. Chemical B on one hand is catalyzed by its own regulated enzyme to produce A, on the other hand controls the synthesis of DNA. This sequence of events is schematically shown in Fig. 2.3. The concentration of the constitutive enzyme is assumed to have the same constant value E^S in each cell, whereas those of the regulated enzymes in the ith cell, E_i^A and E_i^B are both taken to be proportional to the concentration x_i^B of the

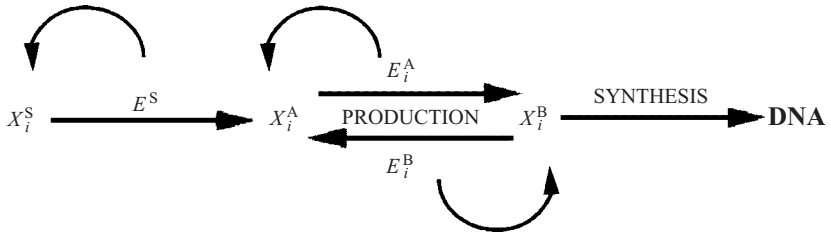


Figure 2.3. Schematic representation of the intracellular dynamics used in the model of Kaneko and Yomo [17]. Curved arrows symbolize catalysis. The variables $x_i^{\mathrm{A}}(t)$, $x_i^{\mathrm{B}}(t)$, $x_i^{\mathrm{S}}(t)$ and E_i^{A}, E_i^{B}, E^{S} denote respectively the concentrations of chemicals A, B and S and their enzymes in the i-th cell, as explained in the text.

chemical B (and therefore dependent on time) in that cell, $E_i^A = e_A x_i^B$ and $E_i^B = e_B x_i^B$ (e_A and e_B are constants). Thus, in terms of chemicals A and B the intracellular dynamics is described by

$$\begin{aligned} \frac{dx_i^A}{dt} &= (e_B x_i^B) x_i^B - (e_A x_i^B) x_i^A + E^S x_i^S \\ \frac{dx_i^B}{dt} &= (e_A x_i^B) x_i^A - (e_B x_i^B) x_i^B - k x_i^B \end{aligned} \tag{2.3}$$

Here the factor k accounts for the decrease of B due to its role in DNA synthesis (see Fig. 2.3). Note the non-linear character of these equations.

Cells are assumed to interact with each other through their effect on the intercellular concentrations of the chemicals A and B. Chemicals are transported in and out of cells. The rate of their transport into a cell is proportional to their concentration outside, but also depends on the internal state of the cell, as characterized by the intracellular concentrations of A and B. Kaneko and Yomo assume that the rate of import of chemical M (i.e., A, B or S) into the ith cell, denoted by $Transp_i^M$, has the form

$$Transp_i^M(t) = p\left(x_i^A + x_i^B\right)^3 X^M \tag{2.4}$$

Here p is a constant. As long as the dependence of *Transp* on the intracellular concentrations is nonlinear, the choice of the exponent (taken to be 3 above) leads to the same qualitative result.

Besides active transport described by Eq. (2.4) chemicals also enter the cells by diffusion through the membrane. The corresponding rate is

taken as

$$Diff_i^M(t) = D\left[X^M(t) - x_i^M(t)\right] \tag{2.5}$$

where D is a (diffusion) constant.

Combining intracellular (Eq. 2.3) and intercellular (Eqs. 2.4 and 2.5) dynamics, the rate equations for the intracellular chemicals become

$$\begin{aligned}
\frac{dx_i^S}{dt} &= -Ex_i^S + Transp_i^S + Diff_i^S \\
\frac{dx_i^A}{dt} &= (e_B x_i^B)x_i^B - (e_A x_i^B)x_i^A + Ex_i^S + Transp_i^A + Diff_i^A \\
\frac{dx_i^B}{dt} &= (e_A x_i^B)x_i^A - (e_B x_i^B)x_i^B - kx_i^B + Transp_i^B + Diff_i^B
\end{aligned} \tag{2.6}$$

It is further assumed that only the source chemical is supplied by a flow from an external tank to the chamber containing the cells. Since it must be transported across the cell membrane to produce chemical A (Eqs. 2.4 and 2.5), the intercellular dynamics of the source chemical is described by

$$\frac{dX^S}{dt} = f\left(\overline{X^S} - X^S\right) - \sum_{i=1}^{N}\left(Transp_i^S + Diff_i^S\right) \tag{2.7}$$

Here $\overline{X^S}$ is the concentration of the source chemical in the external tank, f is its flow rate into the chamber and N is the total number of cells in the system.

Kaneko and Yomo [17] consider cell division to be equivalent to the accumulation of a threshold quantity of DNA. DNA is synthesized from chemical B and therefore the ith cell, born at time t_i^0, will divide at $t_i^0 + T$ (T defines the cell cycle time) when the amount of B in its interior (proportional to x_i^B) reaches a threshold value. Mathematically this condition is expressed as $\int_{t_i^0}^{t_i^0+T} dt x_i^B(t) \geq R$ in the model, with R being the threshold value.)

To avoid infinite growth in cell number, a condition for cell death has to also be imposed. It is assumed that a cell will die if the amount of chemicals A and B in its interior is below the "starvation" threshold S (which is expressed as $\lfloor x_i^A(t) + x_i^B(t) \rfloor < S$).

Simulations based on the above model and its generalizations, using a larger number of chemicals [18-20], led to the following general features, which are likely to pertain to real, interacting cells as well:

1. As the model cells replicate (by division) and interact with one another, eventually multiple biochemical states corresponding to distinct cell types appear. The different types are related to each other by a hierarchical structure in which one cell type stands at the apex, cell types derived from it stand at subnodes, and so on. Such pathways of generation of cell type, which are seen in real embryonic systems, are referred to as developmental lineages.
2. The hierarchical structure appears gradually. Up to a certain number of cells (which depends on the model parameters), all cells have the same biochemical state (i.e., $x_i^A(t)$, $x_i^B(t)$ and $x_i^S(t)$ are independent of i). When the total number of cells rises above a certain threshold value, the state with identical cells is no longer stable. Small differences between cells first introduced by random fluctuations in chemical concentrations start to be amplified. For example, synchrony of biochemical oscillations in different cells of the cluster may break down (by the phases of $x_i^A(t)$, $x_i^B(t)$, $x_i^S(t)$ becoming dependent on i) (Fig. 2.4a). Ultimately, the population splits into a few groups ("dynamical clusters"), with the phase of the oscillator in each group being offset from that in other groups, like groups of identical clocks in different time zones.
3. When the ratio of the number of cells in the distinct clusters falls within some range (depending on model parameters), the differences in intracellular biochemical dynamics are mutually stabilized by cell-cell interactions.
4. With further increase of cell number, the average concentrations of the chemicals over the cell cycle become different (Fig. 2.4b). That is to say, groups of cells come to differ not only in the phases of the same biochemical oscillations, but also in their average chemical composition integrated over the entire lifetimes of the cells. After the formation of cell types, the chemical compositions of each group are inherited by their daughter cells (Fig. 2.4c).

In contrast to the Keller model described above, in which different cell types represent a choice among basins of attraction for a multi-attractor system, with external influences having the potential to bias such preset alternatives, in the Kaneko-Yomo model interactions between cells can give rise to stable intracellular states which would not exist without such interactions. Isologous diversification thus provides a plausible model for

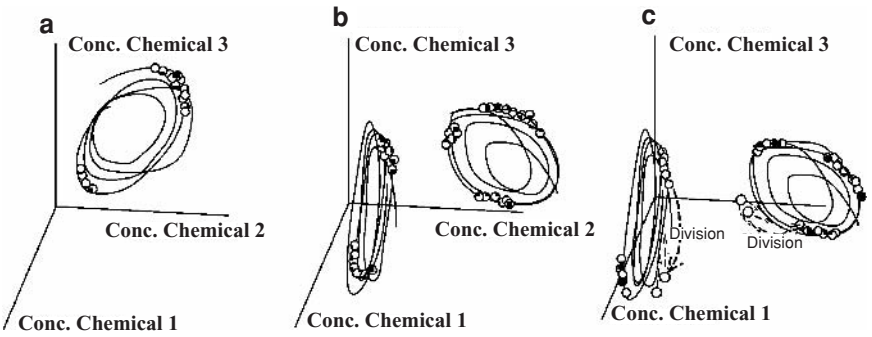


Figure 2.4. Schematic representation of some stages of the differentiation scenario in the model of Kaneko and Yomo [17]. Stage 1 (not shown): Synchronous divisions with synchronous oscillations of the chemicals. Up to a certain number of cells dividing cells arising from a single precursor cell have the same characteristics. Although each cell division is not exactly identical due to the accompanying fluctuation in the biochemical composition, the phase of oscillation in the concentrations, and as a result, the timing of cell division, remains synchronous for all cells; (a) Stage 2: Clustering in the phases of oscillations. When the number of cells rises above a certain (threshold) value, the state with identical cells is no longer stable. Small differences introduced by fluctuations start to be amplified, until the synchrony of the oscillations is broken. The cells then split into a few groups, each having a different oscillation phase. The cells within each group are identical in phase. This diversification in the phases, however, is not equivalent to cell differentiation, because the time average of the biochemical concentrations reveals that the cells are almost identical; (b) Differentiation in chemical composition. With the further increase of the cell number, the average concentrations of the biochemicals over the cell cycle become different. The orbits of chemical dynamics plotted in the phase space of biochemical concentrations lie in distinct regions within the phase space, while the phases of oscillations for cells within each group remain different; (c) "Breeding true" of the differentiated cells. This is indicated by new system points, generated by cell division, residing on trajectories (dashed lines) that track the parental trajectories. After the formation of cell types, the chemical compositions of each group are inherited by their daughter cells. In other words, chemical compositions of cells are recursive over subsequent divisions. Adapted, with changes, from Kaneko (2003) [20], with permission.

the community effect [15], described above. It is reasonable to expect that both intrinsic multistability of a dynamical system of the sort analyzed by Keller, and interaction-dependent multistability, as described by Kaneko, Yomo, and coworkers, based as they are on generic properties of complex dynamical systems, are utilized in initiating developmental decisions in various contexts in different organisms.

3. Biochemical Oscillations and Segmentation

A wide variety of animal types, ranging across groups as diverse as insects, annelids (e.g., earthworms), and vertebrates, undergo *segmentation* early in development, whereby the embryo, or a major portion of it, becomes subdivided into a series of tissue modules [1].These modules typically appear similar to each other when initially formed; later they may follow distinct developmental fates and the original segmental organization may be all but obscured in the adult form. Somite formation (or somitogenesis) is a segmentation process in vertebrate embryos in which the tissue to either side of the central axis of the embryo (where the backbone will eventually form) becomes organized into parallel blocks of tissue.

Somitogenesis takes place in a sequential fashion. The first somite begins forming as a distinct cluster of cells in the anterior region (towards the head) of the embryo's body. Each new somite forms just posterior (towards the tail) to the previous one, budding off from the anterior portion of the unsegmented presomitic mesoderm (PSM) (Fig. 2.5). Eventually, 50 (chick), 65 (mouse), or as many as 500 (certain snakes) of these segments will form.

3.1. Oscillatory dynamics and somitogenesis

In the late 19th century the biologist William Bateson speculated that the formation of repetitive blocks of tissue, such as the somites of vertebrates or the segments of earthworms might be produced by an oscillatory process inherent to developing tissues [21]. More recently, Pourquié and coworkers made the significant observation that the gene *c-hairy1*, which specifies a transcription factor (see previous section), is expressed in the PSM of avian embryos in cyclic waves whose temporal periodicity corresponds to the formation time of one somite [22, 23] (Fig. 2.5). The *c-hairy1* mRNA and protein product are expressed in a temporally-periodic fashion in individual cells, but since the phase of the oscillator is different at different points along the embryo's axis, the areas of maximal expression sweep along the axis in a periodic fashion.

Experimental evidence suggests that somite boundaries form when cells which have left a posterior growth zone move sufficiently far away from a source of a diffusible protein known as fibroblast growth factor 8 (FGF8) in the tailbud at the posterior end of the embryo [24]. The FGF gradient thus

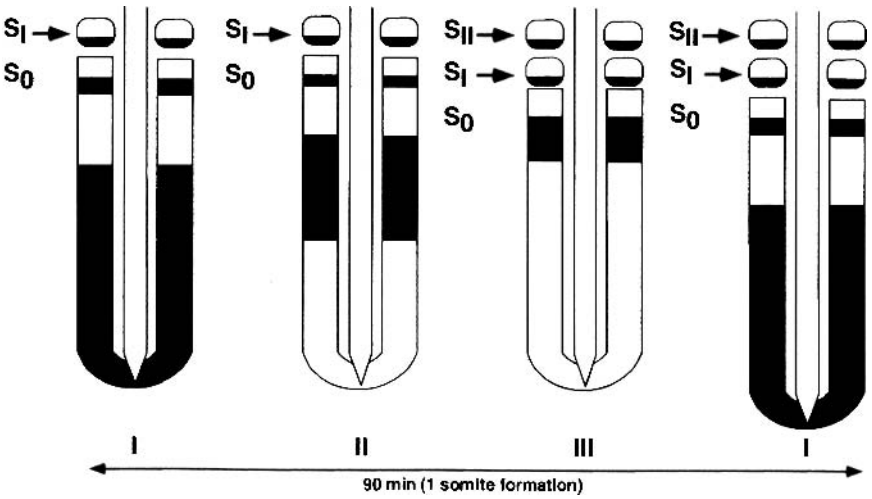


Figure 2.5. Formation of somites, segmented blocks of tissue along the main body axis, in chicken embryos, associated with traveling waves of expression of a regulatory protein (*c-hairy1*). The protein distribution is represented as black regions in this schematic drawing. *c-hairy1* is expressed in a temporally-periodic fashion in individual cells, but since the phase of the oscillator is different at different points along the embryo's axis the areas of maximal expression sweep along the axis in a periodic fashion. Expression is confined to the caudal (toward the tail) half of each somite, where it plays a functional role in causing separation from adjacent, presegmented tissue. S_0, S_I and S_{II} represent successively-forming somites. Reprinted, with modifications, from *Cell*, Vol. 91, I. Palmeirim, D. Henrique, D. Ish-Horowicz & O. Pourquié, Avian hairy gene expression identifies a molecular clock linked to vertebrate segmentation and somitogenesis [22], p. 642, Copyright 1997, with permission from Elsevier.

acts as a "gate" that, when its low end coincides with a particular phase of the segmentation clock, results in formation of a boundary [24, 25]. The general features of this mechanism (called the "clock and wavefront" model) were predicted on the basis of dynamical principles two decades before there was any direct evidence for a somitic oscillator [26].

3.2. The Lewis model of the somitogenesis oscillator

During somitogenesis in the zebrafish a pair of transcription factors known as *her1* and *her7*, which are related to chicken *hairy1* (see above),

oscillate in the PSM in a similar fashion, as does the cell surface signaling ligand deltaC. Lewis [27] and Monk [28] have separately suggested that *her1* and *her7* constitute an autoregulatory transcription factor gene circuit of the sort treated by Keller [12] (see section 2, above), and are the core components of the somitic oscillator in zebrafish. Lewis also hypothesized that deltaC, whose signaling function is realized by activating the Notch receptors on adjacent cells, is a downstream effector of this oscillation. The two *her* genes negatively regulate their own expression [29, 30] and are positively regulated by signaling via Notch [20, 31]. Certain additional experimental results [32] led Lewis to the conclusion that signaling by the Notch pathway, usually considered to act in the determination of cell fate [33], in this case acts to keep cells in the segment-generating growth zone in synchrony [27]. Such synchrony has been experimentally confirmed in chicken embryos [34, 35].

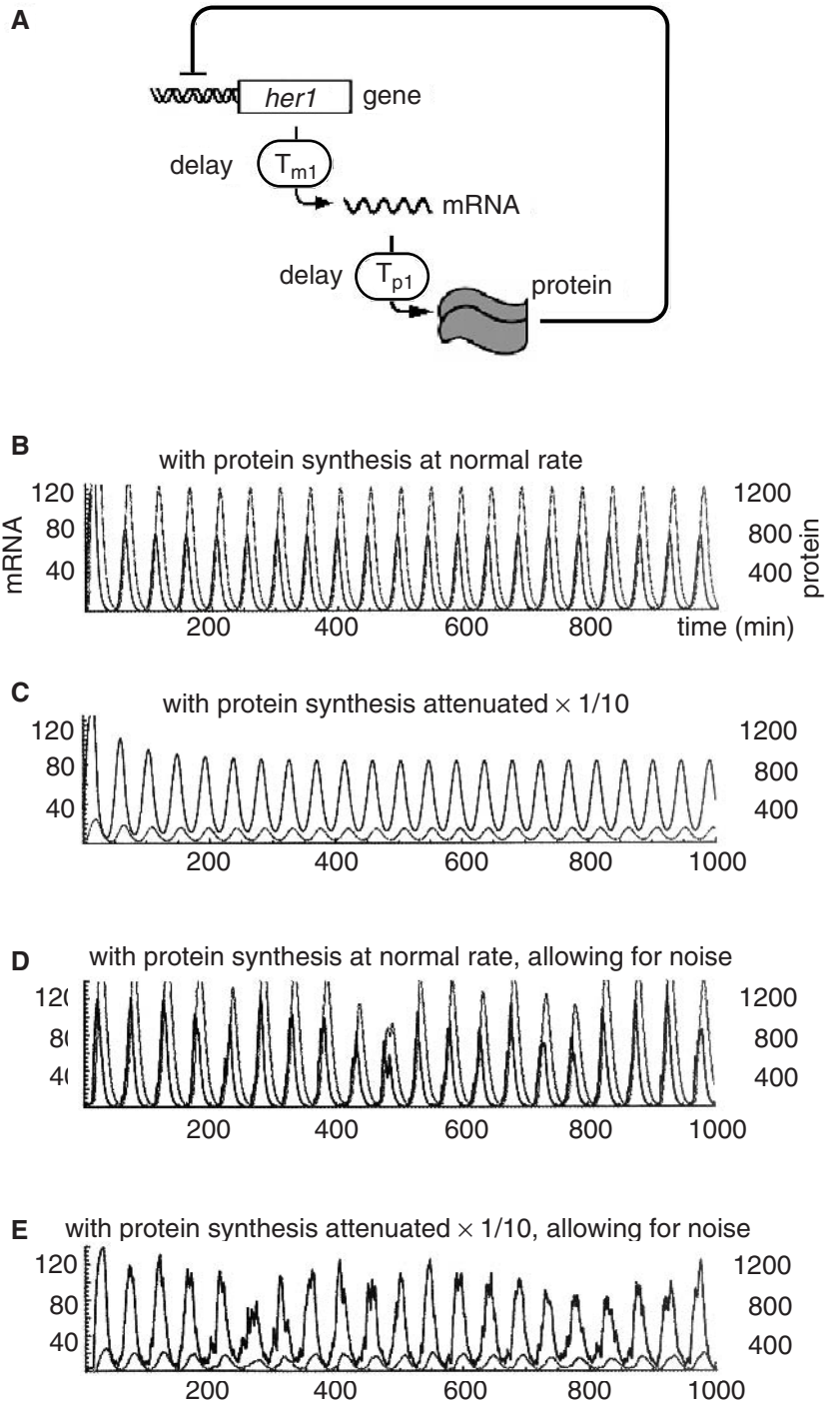
Lewis provides a simple mechanism for the oscillatory expression of the *her1* and *her7* genes, which we briefly summarize here. The model is based on the reasonable assumption that there exists a feedback loop in which the Her1 and Her7 proteins directly bind to the regulatory DNA of their own genes to inhibit transcription. Also incorporated into the model is the recognition that there is always a delay between the initiation of transcription and the initiation of translation, T_m (since it takes time for the mRNA molecule to translocate into the cytoplasm), as well as between the initiation of translation and the emergence of a complete functional protein molecule, T_p.

For a given autoregulatory gene, let $m(t)$ be the number of mRNA molecules in a cell at time t and let $p(t)$ be the number of the corresponding protein molecule. The rate of change of m and p, are then assumed to obey the following equations

$$\frac{dp(t)}{dt} = am(t - T_p) - bp(t) \tag{3.1}$$

$$\frac{dm(t)}{dt} = f\left[p(t - T_p)\right] - cm(t). \tag{3.2}$$

Here the constants b and c are the decay rates of the protein and its mRNA, respectively, a is the rate of production of new protein molecules and $f(p)$ is the rate of production of new mRNA molecules. The function $f(p)$ is assumed to be a decreasing function of the amount of protein. (The form Lewis and Monk used is $f(p) = k/(1 + p^2/p_0^2)$, with constants


A
her1 gene
delay T_{m1}
mRNA
delay T_{p1}
protein
B
with protein synthesis at normal rate
mRNA
protein
time (min)
C
with protein synthesis attenuated × 1/10
D
with protein synthesis at normal rate, allowing for noise
E
with protein synthesis attenuated × 1/10, allowing for noise

k and p_0, to represent the action of an inhibitory protein, assumed to be a dimer. However, results were surprisingly insensitive to the specific form of $f(p)$.)

The above delay differential equations were numerically solved for *her1* and *her7* (for which Lewis was able to estimate the values of all the model parameters in Eqs. 3.1 and 3.2 using experimental results). The solutions indeed exhibit sustained oscillations in the concentration of Her1 and Her7, with the predicted periods close to the observed ones (Fig. 2.6). The important conclusions from the analysis of Lewis are that no oscillations are possible if delay is not incorporated (i.e., $T_m = T_p = 0$) and that the oscillators are quite insensitive to blockade of protein synthesis (i.e., to the value of a in Eq. 3.1). Furthermore, Lewis showed that incorporation into the model of the inherently noisy nature of gene expression (by adding stochastic effects to the deterministic equations, Eqs. 3.1 and 3.2) reinforces continued oscillations. (Without noise oscillations are eventually damped, which would upset normal somite formation beyond the first few.)

Figure 2.6. Cell autonomous gene expression oscillator. A. Molecular control circuitry for a single gene, *her1*, whose protein product acts as a homodimer to inhibit *her1* expression. In case of a pair of genes (i.e. *her1* and *her7*) the analogous circuit would contain an additional branch with coupling between the two branches. B. Computed behavior for the system in A (defined by Eqs. 8 and 9) in terms of the number mRNA molecules per cell in black and protein molecules in gray. Parameter values were chosen appropriate for the *her1* homodimer oscillator (see the form of the function $f(p)$) based on experimental results: $a =$ 4.5 protein molecules per mRNA molecules per minute; $b = c = 0.23$ molecules per minute, corresponding to protein and mRNA half-lives of 3 minutes; $k = 33$ mRNA per diploid cell per minute, corresponding to 1000 transcripts per hour per gene copy in the absence of inhibition; $\mu_0 = 40$ molecules, corresponding to a critical concentration of around 10^{-9} M in a 5 μm diameter cell nucleus; $T_m \approx 20.8$ min; $T_p \approx 2.8$ min. C. Decreasing the rate of protein synthesis ($a = 0.45$) causes little or no effect in the period of oscillation. All the other parameters are the same as in A. D. Computed behavior for system in A when the noisy nature of the gene expression is taken into account. To model stochastic effects Lewis introduced one more independent parameter, the rate constant k_{off} for the dissociation of the repressor protein (i.e., Her1) from its binding site on the regulatory DNA of its own gene (*her1*). Results are shown for $k_{off} = 1$ min^{-1}, corresponding to a mean lifetime of 1 min of the repressor bound state. E. The same as in D except for the rate of protein synthesis, which is as in C. The parameter values not mentioned explicitly in C-E are the same as in B. Protein concentration is represented by the upper curve B and D and the lower curve in C and E. Adapted, with changes, from Lewis (2003) [27], with permission.

4. Reaction-Diffusion Mechanisms and Embryonic Pattern Formation

The nonuniform distribution of any chemical substance, whatever the mechanism of its formation, can clearly provide spatial information to cells. For at least a century embryologists have considered models for pattern formation and its regulation that employ diffusion gradients [1]. Only in the last decade, however, has convincing evidence been produced that this mechanism is utilized in early development. The prime evidence comes from studies of mesoderm *induction*, a key event preceding gastrulation in the frog *Xenopus*. Nieuwkoop [36] originally showed that mesoderm (the middle of the three "germ layers" in the three-layered gastrula—the one that gives rise to muscle, skeletal tissue, and blood) only appeared when tissue from the upper half of an early embryo ("animal cap") was juxtaposed with tissue from the lower half of the embryo ("vegetal pole"). By themselves, animal cap and vegetal pole cells, respectively, only produce ectoderm, which gives rise to skin and nervous tissue, and endoderm, which gives rise to the intestinal lining. Later it was found that several released, soluble factors of the TGF-β protein superfamily and the FGF protein family could substitute for the inducing vegetal pole cells (reviewed by Green [37]). Both TGF-β [38] and FGFs [39] can diffuse over several cell diameters.

None of this proves beyond question that simple diffusion of such released signal molecules (called "morphogens") between and among cells, rather than some other, cell-dependent mechanism, actually establishes the gradients in question. Kerszberg and Wolpert [40] for example, assert that capture of morphogens by receptors impedes diffusion to an extent that stable gradients can never arise by this mechanism. They propose that morphogens are instead transported across tissues by a "bucket brigade" mechanism in which a receptor-bound morphogen on one cell moves by being handed off to receptors on an adjacent cell.

Lander and co-workers [41] use quantitative estimates of the spreading of morphogens in an insect developmental system [42, 43] and conclude, using a mathematical model of morphogen spread and reversible binding to receptors, that free diffusion is indeed a plausible physical mechanism for establishing embryonic gradients. The model of Lander et al. [41], as well as more complex ones describing formation of the embryo's primary (anteroposterior) axis and generation of left-right asymmetry can be

considered examples of a generalized reaction-diffusion system, which we will now describe.

4.1. Reaction-diffusion systems

The rate of change in the concentrations of n interacting molecular species (c_i, $i = 1, 2, \ldots n$) is determined by their reaction kinetics and expressed in terms of ordinary differential equations

$$\frac{dc_i}{dt} = F_i(c_1, c_2 \ldots c_n). \tag{4.1}$$

The explicit form of the functions F_i in Eq. (4.1) depends on the details of the reactions. Spatial inhomogeneities also cause time variations in the concentrations even in the absence of chemical reactions. If these inhomogeneities are governed by diffusion, then in one spatial dimension,

$$\frac{\partial c_i}{\partial t} = D_i \frac{\partial^2 c_i}{\partial x^2}. \tag{4.2}$$

Here D_i is the diffusion coefficient of the ith species. In general, both diffusion and reactions contribute to the change in concentration and the time dependence of the c_is is governed by reaction-diffusion equations

$$\frac{\partial c_i}{\partial t} = D_i \frac{\partial^2 c_i}{\partial x^2} + F_i(c_1, c_2 \ldots c_n). \tag{4.3}$$

Reaction-diffusion systems exhibit characteristic parameter-dependent bifurcations (the "Turing instability"), which are thought to serve as the basis for pattern formation in several embryonic systems, including butterfly wing spots [44], stripes on fish skin [45], distribution of feathers on the skin of birds [46], the skeleton of the vertebrate limb [47, 48], and the primary axis of the developing vertebrate embryo [49], which we discuss in detail, below.

4.2. Axis formation and left-right asymmetry

Gastrulation in the frog embryo is initiated by the formation of an indentation, the "blastopore," through which the surrounding cells invaginate, or tuck into the hollow blastula. Spemann and Mangold [50] discovered that the anterior blastopore lip constitutes an *organizer*: a population of cells that directs the movement of other cells. The action of the

Spemann-Mangold organizer ultimately leads to the formation of the notochord, the rod of connective tissue that first defines the anteroposterior body axis, and to either side of which the somites later form (see Section 3, above). These investigators also found that an embryo with an organizer from another embryo at the same stage transplanted at some distance from its own organizer would form two axes, and conjoined twins would result. Other classes of vertebrates have similarly acting organizers.

A property of this tissue is that if it is removed, adjacent cells differentiate into organizer cells and take up its role. This indicates that one of the functions of the organizer is to suppress nearby cells with similar potential from exercising it. This makes the body axis a partly self-organizing system. The formation of the body axis in vertebrates also exhibits another unusual feature: while it takes place in an apparently symmetrical fashion, with the left and right sides of the embryo seemingly equivalent to one another, at some point the symmetry is broken. Genes such as *nodal* and *lefty* start being expressed differently on the two sides of the embryo [51], and the whole body eventually assumes a partly asymmetric morphology, particularly with respect to internal organs, such as the heart.

Turing [52] first demonstrated that reaction-diffusion systems like that represented in Eq. (4.3) will, with appropriate choice of parameters and boundary conditions, generate self-organizing patterns, with a particular propensity to exhibit symmetry breaking across more than one axis. Using this class of models, Meinhardt [49] has presented an analysis of axis formation in vertebrates and the breaking of symmetry around these axes.

4.3. Meinhardt's models for axis formation and symmetry breaking

The first goal a model of axis formation has to accomplish is to generate an organizer de novo. For this high local concentrations and graded distributions of signaling molecules are needed. This can be accomplished by the coupling of a self-enhancing feedback loop acting over a short range with a competing inhibitory reaction acting over a longer range. The simplest system that can produce such a molecular pattern in the x-y plane consists of a positively autoregulatory activator (with concentration $A(x, y; t)$) and an inhibitor (with concentration $I(x, y; t)$). The activator controls the production of the inhibitor, which in turn limits the production of the activator.

This process can be described by the following reaction-diffusion system [49]

$$\frac{\partial A}{\partial t} = D_A \left(\frac{\partial^2 A}{\partial x^2} + \frac{\partial^2 A}{\partial y^2} \right) + s \frac{A^2 + I_A}{I \left(1 + s_A A^2\right)} - k_A A \tag{4.4a}$$

$$\frac{\partial I}{\partial t} = D_I \left(\frac{\partial^2 I}{\partial x^2} + \frac{\partial^2 I}{\partial y^2} \right) + s A^2 - k_I I + I_I \tag{4.4b}$$

The A^2 terms specify that the feedback of the activator on its own production and that of the inhibitor in both cases is non-linear. The factor $s > 0$ describes positive autoregulation, the capability of a factor to induce positive feedback on its own synthesis. This may occur by purely chemical means ("autocatalysis"), which is the mechanism assumed by Turing [52] when he first considered systems of this type. More generally, in living tissues, positive autoregulation occurs if a cell's exposure to a factor it has secreted causes it to make more of the same factor [53]. The inhibitor slows down the production of the activator (i.e., the $1/I$ factor in the second term in Eq. 4.4a). Both activator and inhibitor diffuse (i.e., spread) and decay with respective diffusion (D_A, D_I) and rate constants (k_A, k_I). The small baseline inhibitor concentrations, I_A and I_I can initiate activator self-enhancement or suppress its onset, respectively, at low values of A. The factor s_A, when present, leads to saturation of positive autoregulation. Once the positive autoregulatory reaction is under way, it leads to a stable, self-regulating pattern in which the activator is in dynamic equilibrium with the surrounding cloud of the inhibitor.

The various organizers and subsequent inductions leading to symmetry breaking, axis formation and the appearance of the three germ layers in amphibians during gastrulation, can all, in principle, be modeled by the reaction-diffusion system in Eqs. (4.4a) and (4.4b), or by the coupling of several such systems. The biological relevance of such reaction-diffusion models depends on whether there exist molecules that can be identified as activator-inhibitor pairs. Meinhardt's model starts with a default state, which consists of ectoderm. Patch-like activation generates the first "hot spot", the *vegetal pole organizer*, which induces endoderm formation (simulation in panel A in Fig. 2.7). A candidate for the diffusible activator in the corresponding self-enhancing loop for endoderm specification is the TGF-β-like factor Derriere, which activates the VegT transcription factor [54].

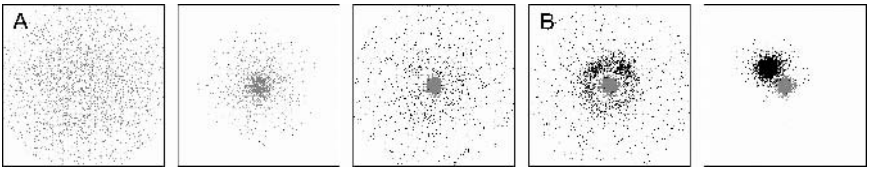


Figure 2.7. Axial pattern formation and induction of two "hot spots" in Meinhardt's model of axis formation. (A) The interaction of a short ranging positive feedback loop (activator, gray) and a long ranging inhibitory substance (not shown) constitutes an unstable system. In the simulation shown a small initial elevation of the activator leads to a focal activation. In a system without pre-localized determinants, such a reaction could be responsible for the formation of the vegetal pole. (B) A second such system (black) forms a second hot spot next to the first if it is activated over a long range and locally repressed by the first activator. This process can lead to symmetry breaking. According to the model, this corresponds to the Nieuwkoop center at a position displaced from the pole. Adapted, with changes from Meinhardt (2001) [49]. Used with permission.

VegT expression remains localized to the vegetal pole, but not because of lack of competence of the surrounding cells to produce VegT [55]. These findings provide circumstantial evidence for the existence of the inhibitor required by the reaction-diffusion model. Subsequently, a second feedback loop forms a second hot spot in the vicinity of the first, in the endoderm. This is identified with the "Nieuwkoop center," a second organizing region, which appears in a specific quadrant of the blastula (see Fig. 2.7). A candidate for the second self-enhancing loop is FGF together with Brachyury [56]. Interestingly, the inhibitor for this loop is hypothesized to be the first loop itself (i.e., the vegetal pole organizer), which acts as local repressor for the second. As a result of this local inhibitory effect, the Nieuwkoop center is displaced from the pole (simulation in panel B in Fig. 2.7). With the formation of the Nieuwkoop center the spherical symmetry of the embryo is broken. In Meinhardt's model this symmetry breaking "propagates" and thus forms the basis of further symmetry breakings, in particular the left-right asymmetry (see below).

By secreting several diffusible factors, the Nieuwkoop center induces the formation of the Spemann-Mangold organizer [57]. (If the second feedback loop, responsible for the Nieuwkoop center is not included in the model, two Spemann-Mangold organizers appear symmetrically with respect to the animal-vegetal axis and no symmetry breaking occurs.) With the formation of the Spemann-Mangold organizer the developmental process is in

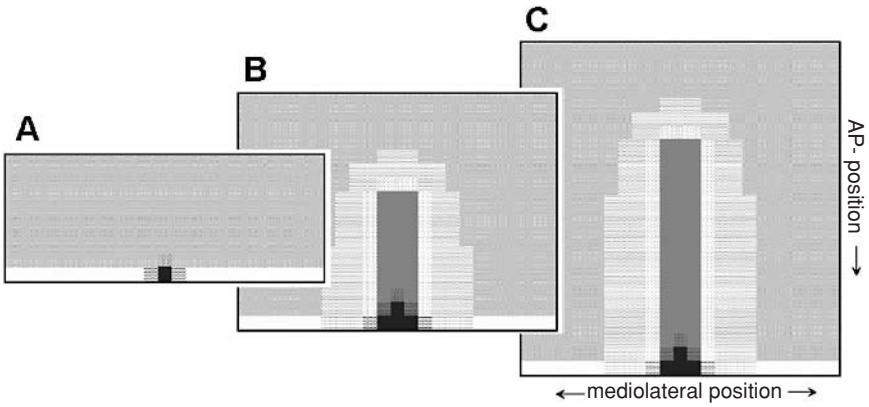


Figure 2.8. Formation of the midline and enfolding of the anteroposterior axis in Meinhardt's model of axis formation. In this simplified simulation a system that is tuned to make stripes (dark gray) is triggered by the organizer, i.e., a system that is activated in a spot-like manner (black). Because the stripe system (corresponding to the notochord) also repels the spot system (corresponding to the Spemann organizer), the spot system is shifted in front of the tip of the stripe, causing its straight elongation. Cells therefore temporarily acquire organizer quality before participating in midline formation. Due to saturation in the self-enhancement, the stripe system does not disintegrate into individual patches, but instead establishes the midline. This, in turn, generates positional information for the dorsoventral or mediolateral axis by acting as a sink for a ubiquitously produced substance (medium and light gray, e.g., BMP-4). The local concentration of the latter is a measure of the distance from the midline. See Meinhardt (2001) [49] for additional details. (Figure adapted with changes from Meinhardt (2001) [49]. Used with permission.)

full swing. The organizer triggers gastrulation, in the course of which the germ layers acquire their relative positions and the notochord forms. This long thin structure marks the midline of the embryo, which itself inherits organizer function and eventually establishes the primary (AP) embryonic axis. A simulation of midline formation, based on Meinhardt's model is shown in Fig. 2.8.

Finally, the breaking of left-right symmetry can be understood again as a competition between already existing and newly developing self-enhancing loops, similarly to the formation of the Nieuwkoop center. The molecule that best fulfils the role of the "left" activator in Meinhardt's model is the product of the *nodal* gene, which is a diffusible, positively autoregulatory member of the TGF-β superfamily. Nodal induces expression from the embryonic midline of another TGF-β related molecule, Lefty, and Lefty

antagonizes Nodal production [58-60]. Because Nodal and Lefty are antagonistic diffusible signals that differ in the range of their activities [59, 61, 62], the ingredients for a symmetry breaking event along the primary embryonic axis are present [51].

Although reaction-diffusion mechanisms like those described are indifferent to which side is left and which side is right, this decision evidently makes a difference biologically: While total inversion of the symmetry of internal organs (*situs inversus totalis*) usually has little adverse impact on health, partial inversions, in which the heart, for example, is predominantly on the right, are highly deleterious [63]. Perhaps for this reason vertebrate embryos have means to ensure that 99.99% of humans, for example, have standard left-right asymmetry.

In nonliving systems in which patterns self-organize by reaction-diffusion mechanisms (see, for example, Castets et al. [64] and Ouyang and Swinney [65]) the local initiators of activation typically arise by random fluctuations; patterns of spots and stripes of chemical concentration with generically similar appearance will form in such cases, but the detailed patterns will differ. In embryonic systems there is clearly a premium on having pattern forming mechanisms that generate the same results in each successive generation. In the axis-forming system of amphibians, birds and mammals just discussed, *monocilia* (motile extensions of the cell surface), at a localized site on the embryo midline (the Spemann organizer or primitive node), beat in an anticlockwise direction, creating fluid currents that bias the distribution of nodal and thus determine the direction of the broken symmetry [66]. The left-right asymmetry of the body follows from this early embryonic event that mobilizes a chemical-dynamic instability to produce a reliable morphological outcome.

5. Evolution of Developmental Mechanisms

Many key gene products that participate in and regulate multicellular development emerged over several billions of years of evolution in a world containing only single-celled organisms. Less than a billion years ago multicellular organisms appeared, and the gene products that had evolved in the earlier period were now raw material to be acted upon by physical and chemical-dynamic mechanisms on a more macroscopic scale [67]. The mechanisms described in the previous sections—chemical multistability,

chemical oscillation, and reaction-diffusion-based symmetry breaking, in conjunction with physical mechanisms such as changes in aggregate viscoelasticity and surface free energy, and adhesive differentials (both between cell types and across the surface of individual cells), would have caused these ancient multicellular aggregates to take on a wide range of biological forms. Any set of physicochemical activities that generated a new form in a reliable fashion within a genetically uniform population of cell aggregates would have constituted a primitive developmental mechanism.

But most modern-day embryos develop in a more rigidly programmed fashion: the operation of cell type- and pattern-generating physical and chemical-dynamic mechanisms is constrained and focused by hierarchical systems of coordinated gene activities—so-called "developmental programs." These programs are the result of eons of molecular evolution that occurred mainly *after* the origination of multicellularity. The manner in which the morphological outcomes of physical and chemical-dynamic mechanisms may have been "captured" during early evolution to produce animal body plans, has been considered in earlier writings [7, 67, 68]. In the remainder of this chapter we will consider instead a model for how two different modes of segmentation in insects arose by variations in a common underlying chemical-dynamic mechanism.

5.1. Segmentation in insects

A major puzzle in the field of evolutionary developmental biology ("EvoDevo") is the fact that evolutionarily-related organisms such as beetles ("short germ-band" insects) and fruit flies ("long germ-band" insects) have apparently different modes of segment formation. Similarly to somitogenesis in vertebrates (see Section 3), in short germ-band insects [69] (as well as in other arthropods, such as the horseshoe crab [70]), segmental primordia are added in sequence from a zone of cell proliferation ("growth zone") (Fig. 2.9). In contrast, in long germ-band insects, such as the fruit fly *Drosophila*, a series of chemical stripes (i.e., parallel bands of high concentration of a molecule) forms in the embryo, which at this stage is a *syncytium*, a large cell with single cytoplasmic compartment containing about 6000 nuclei arranged in a single layer on the inner surface of the plasma membrane [71]. These stripes are actually alternating, evenly-spaced bands of transcription factors of the "pair-rule" class. The pair-rule genes include *even-skipped*, *fushi tarazu*, and *hairy*, which is the insect

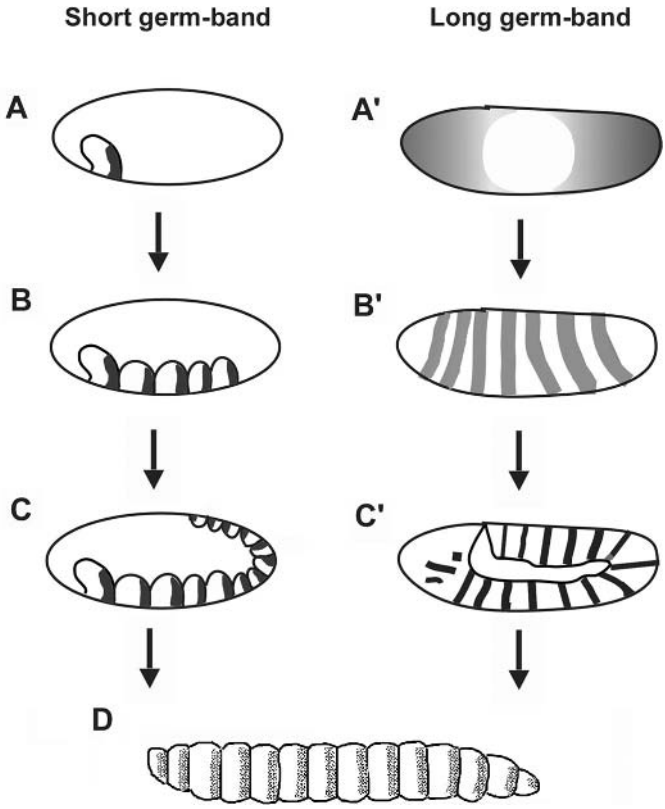


Figure 2.9. Schematic summary of segmentation modes in short germ-band and long germ-band insects. (Left, A) In short germ-band insects, one or groups of a few segments appear in succession. Black patches indicate expression of a segment polarity gene such as *engrailed*. (B) More segments appear posteriorly from a zone of proliferation. (C) The remainder of the segments form sequentially, as in B. (D) Idealized insect larva showing full array of segments. (Right, A') Long germ-band embryo with gradients of expression of maternal genes (e.g., *bicoid* and *nanos*) shown schematically. For simplicity, the patterns of gap gene expression (e.g., *hunchback*, *Krüppel*), intervening between steps A' and B' in *Drosophila*, are not shown. (B') Expression of pair-rule genes (e.g., *eve*, *ftz*, *hairy*) shown schematically in gray. (C') Expression of segment polarity genes (e.g., *engrailed*) shown in black. Adapted, with changes, from Salazar-Ciudad et al. (2001) [89].

homolog of the *c-hairy1* gene expressed in a periodic fashion during vertebrate somitogenesis (see Section 2). When cellularization (the enclosure of each nucleus and nearby cytoplasm in their own complete plasma membrane) takes place shortly thereafter, the cells of the resulting blastoderm will have periodically-distributed identities, determined by the particular mix of transcription factors they have incorporated. The different cell states are later transformed into states of differential adhesivity [72], and morphological segments form as a consequence.

No individual cell-based ("cell autonomous") oscillations have thus far been identified during segmentation of invertebrates, unlike the case in vertebrates such as the mouse, chicken, and zebrafish. However, the sequential appearance of gene expression stripes from the posterior proliferative zone of short germ-band insects and other arthropods such as spiders, has led to the suggestion that these patterns in fact arise from a segmentation clock like that found to control vertebrate somitogenesis [73] (see Section 2.3).

On theoretical [74, 75] and experimental [76] grounds it has long been recognized that the kinetic properties that give rise to a chemical oscillation (such systems exhibit the "Hopf instability"; Section 3), can, when one or more of the components is diffusible, also give rise to standing or traveling spatial periodicities of chemical concentration (the "Turing instability"; Section 4). Considering embryonic tissues as excitable chemical-dynamic media can potentially unify the different segmentation mechanisms found in short and long germ-band insects. This would be quite straightforward if the *Drosophila* embryo were patterned by a reaction-diffusion system, which can readily give rise to a series of chemical standing waves ("stripes").

In reality, however, *Drosophila* segmentation is controlled by a hierarchical system of genetic interactions that has little resemblance to the self-organizing pattern forming systems associated with reaction-diffusion coupling. The formation of overt segments in *Drosophila* (see St Johnston and Nusslein-Volhard [77] and Lawrence [71] for reviews) requires the prior expression of a stripe of the transcription factor product of the *engrailed* (*en*) gene in the cells of the posterior border of each of 14 presumptive segments [78]. The positions of the engrailed stripes are largely determined by the activity of the pair-rule genes *even-skipped* (*eve*) and *fuhsi-tarazu* (*ftz*), which exhibit alternating, complementary seven stripe patterns prior to the formation of the blastoderm [79, 80].

However, despite the appearance of being produced by a reaction-diffusion system (Fig. 2.9), the stripe patterns of the pair-rule genes in *Drosophila* are generated rather by a complex set of interactions among transcription factors in the syncytium that encompasses the entire embryo. The formation of *eve* stripe number 2, for example, requires the existence of sequences in the *eve* promotor that switch on the *eve* gene in response to a set of spatially-distributed morphogens that under normal circumstances have the requisite values only at the stripe 2 position (Fig. 2.10) [81-83]. In particular, these promoter sequences respond to specific combinations of products of the "gap" genes (e.g., *giant*, *knirps*, the embryonically-produced version of *hunchback*). These proteins are transcription factors that are expressed in a spatially-nonuniform fashion and act as activators and competitive repressors of the pair-rule gene promoters [84] (also see discussion of the Keller model in Section 2). The patterned expression of the gap genes, in turn, is controlled by the responses of their own promoters to particular combinations of products of "maternal" genes (e.g., *bicoid*, *staufen*), which are distributed as gradients along the embryo at even earlier stages (Fig. 2.9). As the category name suggests, the maternal gene products are deposited in the egg during oogeneis.

While the expression of *engrailed* along the posterior margin of each developing segment seems to be a constant theme during development of arthropods, including those other than *Drosophila* (e.g., grasshoppers [69, 85], beetles [86]), the expression patterns of pair-rule genes is less well-conserved over evolution [87, 88]. The accepted view is that the short germ-band "sequential" mode is the more ancient way of making segments, and that the long germ-band "simultaneous" mode seen in *Drosophila,* which employs pair-rule stripes, is more recently evolved. The existence of "intermediate germ-band" insects, in which segmentation is sequential in one region of the embryo and simultaneous in another, suggests that *Drosophila* was derived from a short germ-band ancestor via such intermediate forms. Why and how cellularization of the blastula in some ancestral insects was impeded so as to produce a syncytium, is unknown.

5.2. Chemical dynamics and the evolution of insect segmentation

As noted above, the kinetic properties that give rise to a limit cycle chemical oscillation, can, when one or more of the components is diffusible, also

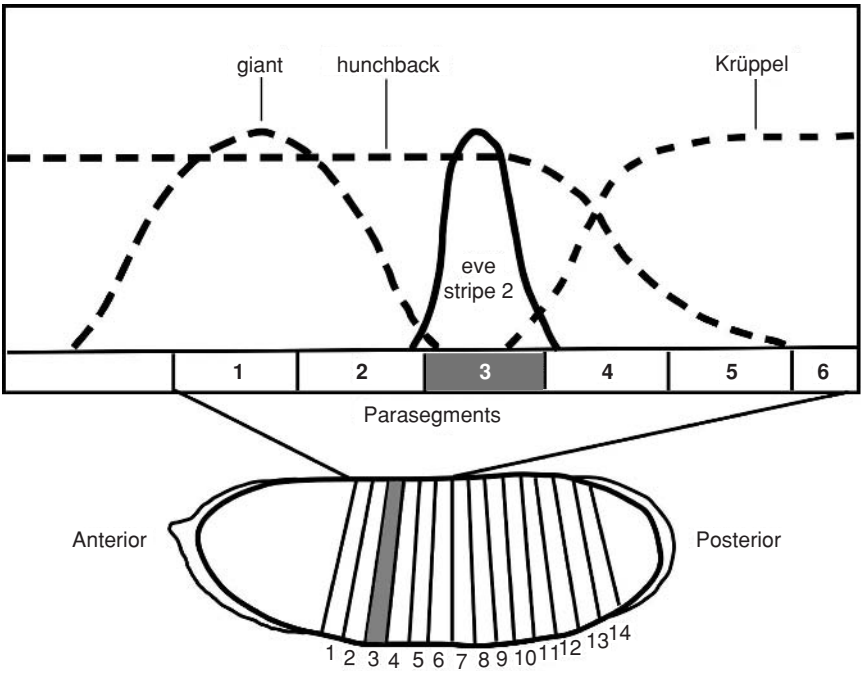


Figure 2.10. Specification of the second stripe of transcription of the *Drosophila even-skipped gene*. The *even-skipped (eve)* gene contains stripe-specific promoters responsive to different concentrations of the protein products of the gap-class genes. Stripe 2 of *eve* expression coincides with the third parasegment of the *Drosophila* embryo, the parasegments comprising the posterior region of one future morphological segment along with the anterior region of the future segment just posterior to it. Products of the gap genes, *giant, hunchback*, and *Krüppel* form gradients within the syncytium that activate *eve* expression at the stripe 2 location, while suppressing it to either side [81, 82]. Other *eve* stripes, expressed from the same gene, are specified by promoters responsive to different levels of gap gene products.

give rise to standing or travelling spatial periodicities of chemical concentration. Whether a system of this sort exhibits purely temporal, spatial, or spatiotemporal periodicity depends on particular ratios of reaction and diffusion coefficients. An important requirement of both these kinetic schemes is the presence of a direct or indirect positive autoregulatory circuit. A simple dynamical system that exhibits temporal oscillation or standing waves, depending on whether or not diffusion is permitted, is described by Salazar-Ciudad et al. [89] (Fig. 2.11).

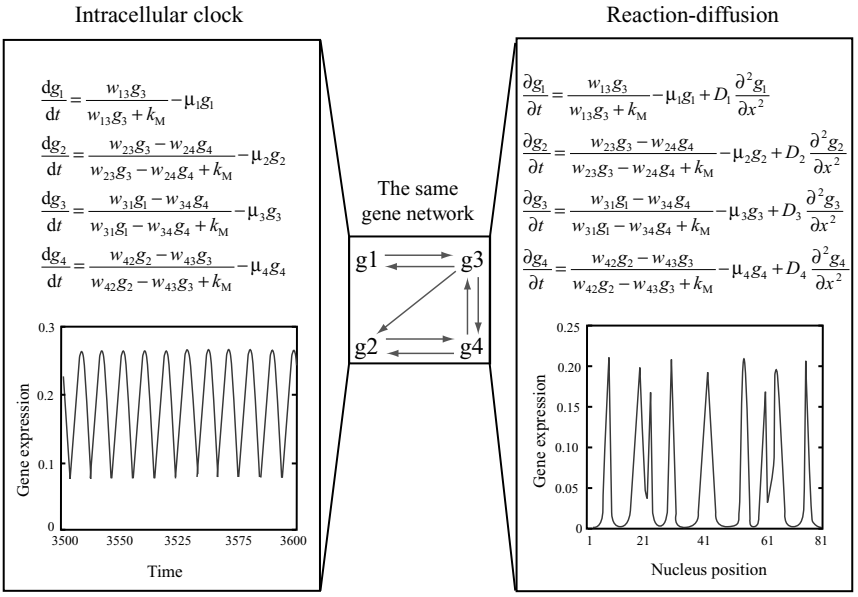


Figure 2.11. An example of a network that can produce (for the same parameter values) sequential stripes when acting as an intracellular biochemical clock in a cellularized blastoderm with a posterior proliferative zone, and simultaneously-forming stripes when acting in a diffusion-permissive syncytium. The network, based on the model of Salazar-Ciudad et al. (2001) [90] is shown in the central box. Black arrows indicate positive regulation and white arrows negative regulation. In the upper boxes, the equations governing each of the two behaviors are shown. The four genes involved in the central network diagram, as well as their levels of expression, are denoted by g1, g2, g3, and g4. In the reaction-diffusion case, g1 and g2 can diffuse between nuclei (note that the two set of equations differ only in the presence of a diffusion term for the products of genes 1 and 2). The lower boxes indicate the levels of expression of gene 2 for the two systems. For the intracellular clock the x-axis represents time, whereas in the reaction-diffusion system this axis represents space. In the 1 = D pattern shown, the initial condition consisted of all gene values set to zero except gene 1 in the central of 81 nuclei, which was assigned a small value (the exact value did not affect the pattern). The patterns shown were found when the following parameter values were used: $k_M = 0.01; w_{13} = 0.179; w_{23} = 0.716; w_{24} = -0.704; w_{31} = 0.551; w_{34} = -0.466; w_{42} = 0.831; w_{43} = -0.281; \mu_1 = 1.339; \mu_2 = 2.258; \mu_3 = 2.941; \mu_4 = 2.248$. For the reaction-diffusion case, the same parameter values are used but in addition: $D_1 = 0.656$ and $D_2 = 0.718$. From Salazar-Ciudad et al. [89], © Blackwell Publishing Co. Used with permission.

Based on the experimental findings, theoretical considerations and evolutionary inferences described above, it was hypothesized that the ancestor of *Drosophila* generated its segments by a reaction-diffusion system [89, 91], built upon the presumed chemical oscillator underlying short germ-band segmentation. Modern-day *Drosophila* contains (retains, according to the hypothesis) the ingredients for this type of mechanism. Specifically, in the syncytial embryo several of the pair-rule proteins (e.g., eve, ftz) diffuse over short distances among the cell nuclei that synthesize their mRNAs, and positively regulate their own synthesis [92-94].

The hypothesized evolutionary scenario can be summarized as follows: the appearance of the syncytial mode of embryogenesis converted the chemical oscillation-dependent temporal mechanism found in the more ancient short germ-band insects into the spatial standing wave mechanism seen in the more recently evolved long germ-band forms. Of course, the pair-rule stripes formed by the proposed reaction-diffusion mechanism would have been equivalent to one another. That is, they would have been generated by a single mechanism acting in a spatially-periodic fashion, not by stripe-specific molecular machinery (see above). Thus despite suggesting an underlying physical connection between modern short germ-band segmentation and segmentation in the presumed anscestor of long germ-band forms, this hypothesis introduces a new puzzle of its own: Why does modern-day *Drosophila* not use a reaction-diffusion mechanism to produce its segments?

5.3. Evolution of developmental robustness

Genes are always undergoing random mutation, but morphological change does not always track genetic change. Particularly interesting are those cases in which the outward form of a body plan or organ does not change, but its genetic "underpinning" does. This can result from particular kind of natural selection, termed "canalizing selection" by Waddington [95] (see also Schmalhausen [96]). Canalizing selection will preserve those randomly acquired genetic alterations that happen to enhance the reliability of a developmental process. Development would thereby become more complex at the molecular level, but correspondingly more resistant ("robust") to external perturbations or internal noise that could disrupt non-reinforced physical mechanisms of determination. Indeed, the patterns formed by reaction-diffusion systems are notoriously sensitive to

temperature and domain size [97, 98], and any developmental process that was solely dependent on this mechanism would be unreliable.

If the striped expression of pair-rule genes in the ancestor of modern *Drosophila* was generated by a reaction-diffusion mechanism, this inherently variable developmental system would have been a prime candidate for canalizing evolution. The elaborate systems of multiple promoter elements responsive to pre-existing, nonuniformly distributed molecular cues (e.g., maternal and gap gene products), seen in *Drosophila* is therefore not inconsistent with this pattern having originated as a reaction-diffusion process.

The question was raised above as to why modern-day *Drosophila* does not use a reaction-diffusion mechanism to produce its segments. In light of the discussion in the previous paragraphs, we can consider the following tentative answer: pattern-forming systems based on reaction-diffusion may be inherently unstable evolutionarily (for the same reasons of sensitivity to parameters and domain size that make them dynamically unstable), and would therefore have been replaced, or at least reinforced, by more hierarchically-organized genetic control systems under pressure of natural selection. A corollary of this hypothesis is that the patterns produced by modern, highly evolved, pattern-forming systems would be robust in the face of further genetic change.

Similar canalizing evolution may also have occurred in short-germ band insects—not as much is known about the underlying mechanism of segmentation in these groups. However, it is also possible that oscillatory mechanisms (which continue to be used in vertebrates [23], and may be used in short germ-band insects) are less easy to replace by more reliable alternatives (i.e., hierarchies of genetic regulation) than are standing wave mechanisms. (This issue is discussed in greater detail in Salazar-Ciudad et al. [89].)

The evolutionary hypothesis of self-organization followed by canalization has been examined computationally in a simple physical model by Salazar-Ciudad and coworkers [90]. The model consists of N_c nuclei arranged in a row within a syncytium. Each nucleus has the same genome (i.e., the same set of N_g genes), and the same *epigenetic system* (i.e., the same activating and inhibitory relationships among these genes). Genes in these networks interact according to a set of simple rules that embody unidirectional interactions in which an upstream gene activates a downstream one, as well as reciprocal interactions, in which genes feed back

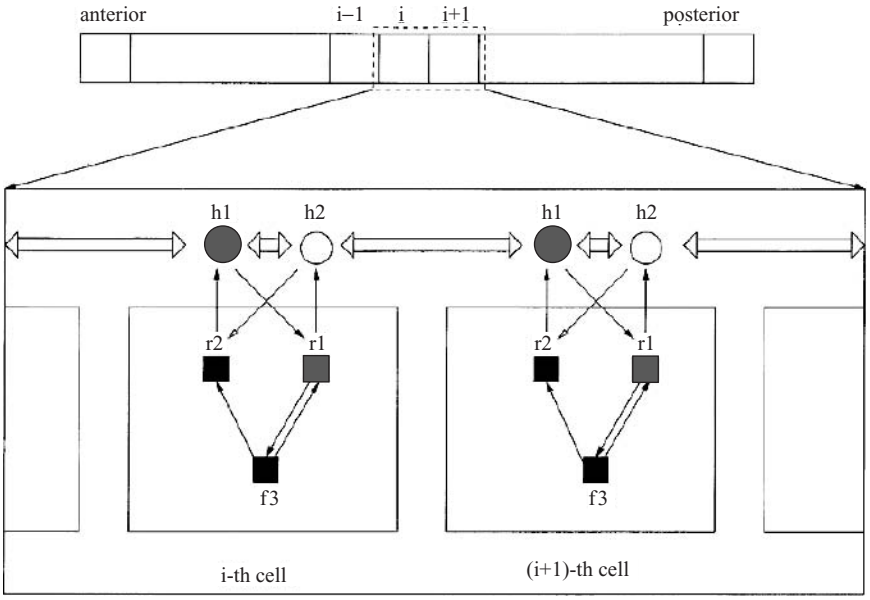


Figure 2.12. Schematic representation of the kinds of interactions included in the gene network evolution model of Salazar-Ciudad et al. The boxes denote genes acting inside cells whereas circles denote diffusible paracrine factors. Abbreviations: h designates a hormone or paracrine factor, r a receptor, and f a transcription factor. Arrows with black arrowheads indicate positive interactions and arrows with white arrowheads indicate inhibitory interactions. Thick horizontal arrows indicate diffusion. From Salazar-Ciudad et al. (2001) [90]. Used with permission.

(via their products) on each other's activities (Fig. 2.12). This formalism was based on a similar one devised by Reinitz and Sharp [99], who considered the specific problem of segmentation in the *Drosophila* embryo. Gene products interact by binding to or modifying one another, or by binding to cis-acting (i.e., located on the same DNA strand as the gene) promoter DNA sequences, as described above in the Keller model [12]. A subset of the genes (N_{p}) specify diffusible (paracrine) factors. The model assumes that the change in gene expression induced by an interaction follows a saturating Hill function (a class of function widely used for molecular binding processes [100]). Promoters are characterized by the values of coupling constants W_{jk} that weight both the affinity of the transcriptional factor for the promoter and the intensity of the response produced by the binding.

The gene regulatory networks considered by Salazar-Ciudad and coworkers are modeled by a dynamic system obeying the following set of equations, which describe the change in the concentration in each of the N_c nuclei (numbered by the first index) of each of the N_g gene product (numbered by the second index), due to the effect of all other genes:

Non-diffusible gene products:

$$\frac{\partial g_{ij}}{\partial t} = \frac{\Phi\left(h_{ij}\right)}{k_M + \Phi\left(h_{ij}\right)} - \mu_j g_{ij}, \quad h_{ij} = \sum\nolimits_{k=1}^{N_g} W_{jk}\ g_{jk} \qquad (i = 1, 2 \ldots, N_c,\ j = N_p + 1, \ldots, N_g) \tag{5.1}$$

Diffusible gene products:

$$\frac{\partial g_{il}}{\partial t} = \frac{\Phi\left(h_{il}\right)}{k_M + \Phi\left(h_{il}\right)} - \mu g_{il} + D_l \nabla^2 g_{il}, \quad h_{il} = \sum\nolimits_{k=1}^{N_g} W_{lk}\ g_{lk} \qquad (i = 1, 2, \ldots, N_c, l\ = 1, \ldots, N_p) \tag{5.2}$$

The Heaviside function, Φ ($\Phi(x) = x$ for $x > 0$ and $\Phi(x) = 0$ otherwise), ensures that inhibiting interactions do not lead to negative concentration values. D_l and μ_l are, respectively, the diffusion coefficient and the intrinsic rate of degradation for gene product g_l, and k_M is a Michaelis-type constant defining the rate of response to an activator or inhibitor.

One may ask whether a system of this sort, with particular values of the gene-gene coupling constants W_{jk}, can form a pattern. Salazar-Ciudad and coworkers performed simulations on systems containing 25 nuclei [90]. Zero-flux boundary conditions were used (i.e., the boundaries of the domain were impermeable to diffusible morphogens), and initial conditions were set such that at $t = 0$ the levels of all gene products had zero value except for that of an arbitrarily chosen gene, which had a non-zero value in the nucleus at the middle position. A pattern was considered to arise if after some time different nuclei stably expressed one or more of the genes at different levels. Isolated single nuclei, or isolated patches of contiguous nuclei expressing a given gene, are the 1-dimensional analogue of isolated stripes of gene expression in a 2-dimensional sheet of nuclei, such as that in the *Drosophila* embryo prior to cellularization [90].

It had earlier been determined that the core mechanisms responsible for stable patterns fell into two non-overlapping topological categories [101]. Mechanisms that form patterns by virtue of the unidirectional influence

of one gene on the next, in an ordered succession, as in the "maternal gene induces gap gene induces pair-rule gene" scheme described above for aspects of *Drosophila* segmentation, were termed "hierarchical." In contrast, mechanisms in which reciprocal positive and negative feedback interactions give rise to the pattern, were termed "emergent." Emergent systems are equivalent to dynamical systems, such as the transcription factor networks discussed in Section 2 and the reaction-diffusion systems discussed in Section 2.4.

This is not to imply that an entire embryonic patterning process—the formation of a pair-rule stripe in *Drosophila*, for example—is wholly hierarchical or emergent. Rather, it indicates that the process can be decomposed into modules that are unambiguously of one or the other topology [101].

In their computational studies of the evolution of developmental mechanisms Salazar-Ciudad and co-workers identified emergent and hierarchic networks that produced a particular pattern—e.g., three "stripes" [90]. (Note that patterns themselves are neither emergent or hierarchical—these terms apply to the mechanisms that generate them.) They asked if given networks would "breed true" phenotypically, despite changes to their underlying circuitry. That is, would their genetically altered "progeny" exhibit the same pattern as the unaltered version? Genetic alterations in these model systems consisted of point mutations (i.e., changes in a W_{jk} value), duplications, recombinations (i.e., interchange of various W_{jk} values between genes) and the acquisition of new interactions (i.e., a W_{jk} that was initially equal to zero was randomly assigned a small positive or negative value).

To evaluate the consequence of such alterations it was necessary to define a metric of "distance" between different patterns. Roughly, this was done by specifying the state of each nucleus in a model syncytium in terms of the value of the gene product forming a pattern. Two patterns were considered to be equivalent if all nuclei in corresponding positions were in the same state. The degree of divergence from such equivalence could then be quantified [90].

It was found that emergent networks were much more likely to diverge from the original pattern than hierarchical networks after undergoing such simulated evolution (Fig. 2.13). In other simulations the pattern itself was held constant (i.e., only those genetic variants that had the same number of "stripes" as the original were retained after each iteration) and it was found that networks that started out as emergent could be converted into hierarchical networks.

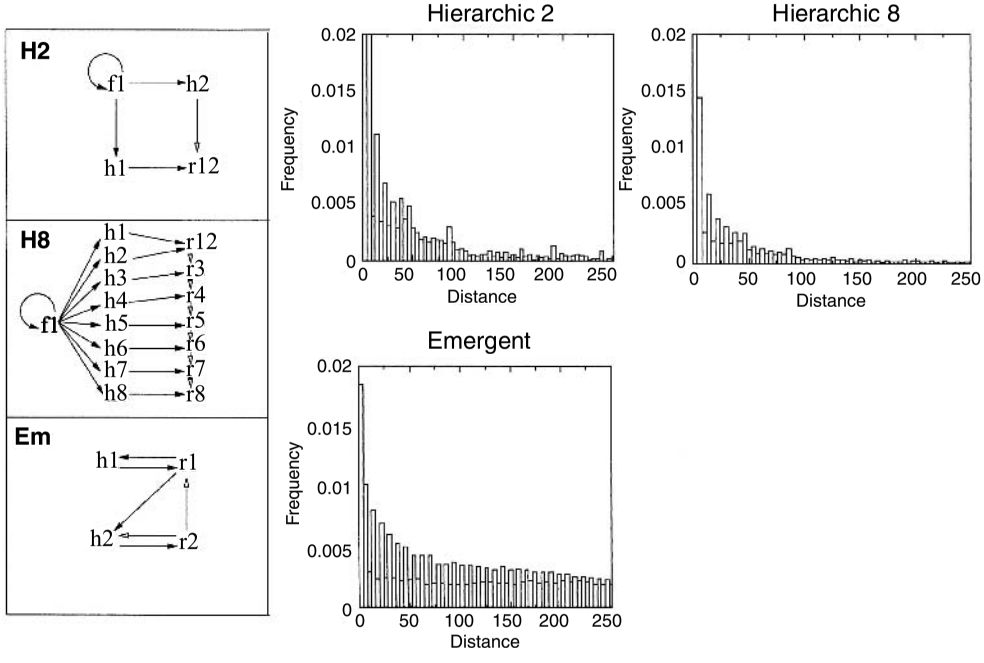


Figure 2.13. (*left panels*) Three network topologies studied in simulated evolution experiments by Salazar-Ciudad et al. [90]. The terms *h*, *r* and *f*, and arrows are as defined in the legend to Figure 2.12. "Hierarchic 2" (H2) and "Hierarchic 8" (H8) refer to hierarchical network topologies with two paracrine factors and one receptor, and eight paracrine factors and seven receptors, respectively. "Emergent" (Em) is a network topology with two paracrine factors and two receptors. (*right panels*) Estimation of the relationship between genotype and phenotype in the three networks topologies shown on the left. The *x* axis represents the distance between a network and a mutated version. (See text and original article [90] for definition of distance measure.) The *y* axis represents the relative frequency of such distances between the 100,000 networks of each topology analyzed and 30 randomly mutated versions of each. Adapted, with changes, from Salazar-Ciudad et al. (2001) [90]. Used with permission.

Subject to the caveats to what is obviously a highly schematic analysis, the implications of these computational experiments for the evolution of segmentation in long germ-band insects are the following: (i) if the ancestral embryo indeed generated its seven-stripe pair-rule protein patterns by a reaction-diffusion mechanism, and (ii) if this pattern was sufficiently well-adapted to its environment so as to provide a premium on breeding true, then (iii) genetic changes that preserved the pattern but converted the underlying network from an emergent one to a hierarchical one (as seen in present-day *Drosophila*) would have been favored.

6. Conclusions

Biological development depends on genes and their products, and biological evolution depends on genetic change. But it is a mistake to consider development as being the result of programmed gene expression changes alone, or to consider evolution to be caused by genetic variation and selection exclusively. The cell types and tissue patterns and forms that arise during development are manifestations of complex systems, and system behaviors are not simply a matter of inventories of components and their interconnections. Moreover, the changes in form and function induced by genetic mutation are also determined by the system properties of the organism, particularly those in play when it is taking form, i.e., during development. Therefore, even though genetic mutation may be random and natural selection opportunistic, evolutionary change has preferred directions.

None of this will be surprising to physical scientists who study complex systems. Despite earlier precedents, (e.g., Rashevsky [102], Bertalanffy [103, 104]), that were before their time in that the molecular components of the cell physiological and developmental systems discussed were unknown, the systems way of thinking is only now becoming part of the biological mainstream, though there have been notable landmarks along the way (e.g., Goodwin [10], Winfree [105], Meinhardt [97]).

The models we have presented in the preceding sections are all unified in that they deal with chemical-dynamical aspects of multicellular development and evolution. Other physical aspects of the same questions are reviewed elsewhere [106-109]. The examples considered in this chapter represent several different aspects of the new systems biology. The most

abstract of the models are those dealing with cell differentiation. While the autoregulatory transcription factor network model of Keller [12] deals with general categories of gene-gene interactions rather than specifically-named molecules, all the categories considered, including the one described in detail in Section 2.2, are based on molecular biological findings. Multistability is essentially a direct consequence of these interactions. The representation of cells and molecules in the Kaneko-Yomo isologous diversification model is more distant from any experimental data, but does not contradict anything known about the formal properties of such systems. The unique thing about this model is that as an exercise in dynamical systems theory motivated by biological questions, it has actually uncovered a dynamical phenomenon—interaction-dependent attractors—not previously appreciated. If this phenomenon indeed proves relevant to developmental properties such as the community effect [15], it would be a rare case (Turing's 1952 description of the reaction-diffusion instability [52] is another) where biologically-motivated mathematics led to a fundamental finding as well as a solution to a biological problem.

As hypothesized by Bateson [21] and Cooke and Zeeman [26], cell autonomous chemical oscillations are clearly relevant to vertebrate segmentation [22]. The biochemical oscillator analysis of Lewis [27] and Monk [28] is based on highly detailed knowledge of the molecular and cellular interactions occurring during zebrafish somitogenesis. The introduction of a time-lag, a key component in causing the oscillator to behave realistically, is a mathematical maneuver that is virtually dictated by the molecular biology of gene expression [27]. Along with Meinhardt's molecularly-cognizant analysis of the amphibian axis-forming system as excitations of an excitable medium [49], the analyses of Lewis, as well as the important work, not discussed here, of Odell and coworkers [110, 111] on the chemical dynamics of developmental modules, represent the leading edge of an integrated experimental-mathematical approach to systems phenomena during development.

The models of Lewis and Meinhardt, while having the formal structure of chemical-dynamic schemes, make use of chemically nonexplicit ("black box") cell response functions to describe the synthesis and breakdown of their gene product components. For this reason, Hentschel et al. [49] who, like Meinhardt [49], use a Turing mechanism-inspired approach to model an aspect of embryogenesis (in their case the skeletal pattern of the vertebrate limb), propose that the term "reactor-diffusion" be

used instead of reaction-diffusion to describe these hybrid cell-chemical systems.

In analyzing the complex question of the evolution of a developmental mechanism, Salazar-Ciudad et al. [89, 90] resort to schematic simplifications analogous to those used by Keller [12] and Kaneko and coworkers [17-20]. Like those models, that of Salazar-Ciudad and coworkers employs gene-gene product interactions consistent with experimental knowledge of the relevant systems, in this case a simplified version of a molecularly-based model for the formation of pair-rule stripes in *Drosophila* [99]. The evolutionary simulations performed with this model suggest a plausible scenario for the increasingly-recognized fact that developmental mechanisms may evolve, even when the morphologies they generate remain static over vast periods [7]. In fact, canalizing selection can reasonably have transformed an embryonic pattern that may have originated as a set of reactor-diffusion processes analogous to the chemical-dynamic mechanisms discussed earlier in this chapter (i.e., the pair-rule stripes), into the hierarchical genetic program seen in the modern fruit-fly.

Sections 2.2–2.4 of this chapter indicated the operation of chemical-dynamic processes in differentiation, segmentation and axis formation of modern-day organisms. Nonetheless, the enormous degree of apparently "hard-wired" molecular integration seen in some developmental systems has been used in the past to support the premise that these physicochemical determinants are of only marginal relevance. The discussion in Section 2.5 suggests that by taking evolution into account, we can reconcile developmental systems with the dynamics of the less constrained systems from which they originated.

Acknowledgment

We acknowledge Nation Science Foundation grant No. IBN - 0083653 under the Biocomplexity Initiative, for support.

References

1. S. F. Gilbert, M. S. Tyler, and R. N. Kozlowski, *Developmental Biology*, 7th ed., Sinauer Associates, Sunderland, MA (2003).
2. S. H. Strogatz, *Nonlinear Dynamics and Chaos With Applications to Physics, Biology, Chemistry, and Engineering*. Perseus, Cambridge, MA (1994).
3. International Human Genome Consortium. Finishing the euchromatic sequence of the human genome. *Nature,* **431,** 931–945.

4. M. J. Lercher, A. O. Urrutia, and L. D. Hurst, *Nat Genet* **31**, 180-183 (2002).
5. M. T. Borisuk, J. J. Tyson, *J Theor Biol* **195**, 69-85 (1998).
6. M. B. Elowitz and S. Leibler, *Nature* **403**, 335-338 (2000).
7. S. A. Newman and G. B. Müller, *J Exp Zool* **288**, 304-317 (2000).
8. E. Jablonka and M. J. Lamb, *Epigenetic Inheritance and Evolution*. Oxford University Press, Oxford, U.K. (1995).
9. S. A. Kauffman, *J Theor Biol* **22**, 437-467 (1969).
10. BC Goodwin, *Temporal Organization in Cells; A Dynamic Theory of Cellular Control Processes*. Academic Press, London, New York (1963).
11. M. Mannervik, Y. Nibu, H. Zhang, and M. Levine *Science* **284**, 606-609 (1999).
12. A. D. Keller, *J Theor Biol* **172**, 169-185 (1995).
13. F. G. Giancotti and E. Ruoslahti, *Science* **285**, 1028-1032 (1999).
14. C. Morisco, K. Seta, S. E. Hardt, Y. Lee, S. F. Vatner, and J. Sadoshima, *J Biol Chem* **276**, 28586-28597 (2001).
15. J. B. Gurdon, *Nature* **336**, 772-774 (1988).
16. H. J. Standley, A. M. Zorn, and J. B. Gurdon, *Development* **128**, 1347-1357 (2001).
17. K. Kaneko and T. Yomo, *Physica D* **75**, 89-102 (1994).
18. K. Kaneko and T. Yomo, *Bull Math Biol* **59**, 139-196 (1997).
19. C. Furusawa and K. Kaneko, *Bull Math Biol* **60**, 659-687 (1998).
20. K. Kaneko, Organization Through Intra-Inter Dynamics, in *Origination of Organismal Form: Beyond the Gene in Developmental and Evolutionary Biology,* G. B. Müller and S. A. Newman (eds.), MIT Press, Cambridge, MA (2003).
21. W. Bateson, *Materials for the Study of Variation*. London, Macmillan (1894).
22. I. Palmeirim, D. Henrique, D. Ish-Horowicz , and O. Pourquie, *Cell* **91**, 639-48 (1997).
23. O. Pourquié, A biochemical oscillator linked to vertebrate segmentation, in *Origination of Organismal Form: Beyond the Gene in Developmental and Evolutionary Biology.* Müller G. B. and Newman S. A. (eds.), Cambridge, MA, MIT Press (2003).
24. J. Dubrulle, M. J. McGrew, and O. Pourquié, *Cell* **106**, 219-232 (2001).
25. O. Pourquié, *Science* **301**, 328-330 (2003).
26. J. Cooke and E. C. Zeeman, *J Theor Biol* **58**, 455-476 (1976).
27. J. Lewis, *Curr Biol* **13**, 1398-1408 (2003).
28. N. A. Monk, *Curr Biol* **13**, 1409-1413 (2003).
29. S. A. Holley, R. Geisler, and C. Nusslein-Volhard, *Genes Dev* **14**, 1678-1690 (2000).
30. A. C. Oates and R. K. Ho, *Development* **129**, 2929-2946 (2002).
31. C. Takke and J. A. Campos-Ortega, *Development* **126**, 3005-3014 (1999).
32. S. A. Holley, D. Julich, G. J. Rauch, R. Geisler, and C. Nusslein-Volhard, *Development* **129**, 1175-1183 (2002).
33. S. Artavanis-Tsakonas, M. D. Rand, and R. J. Lake, *Science* **284**, 770-776 (1999).
34. C. D. Stern and R. Bellairs, *Anat Embryol (Berl)* **169**, 97-102 (1984).

35. D. R. Primmett, W. E. Norris, G. J. Carlson, R. J. Keynes RJ, and C. D. Stern *Development* **105**, 119-130 (1989).
36. P. D. Nieuwkoop, *Wilhelm Roux' Arch. Entw. Mech. Org.* **162**, 341-373 (1969).
37. J. Green, *Dev Dyn* **225**, 392-408 (2002).
38. N. McDowell, J. B. Gurdon, and D. J. Grainger, *Int J Dev Biol* **45**, 199-207 (2001).
39. B. Christen and J. Slack, *Development* **126**, 119-125 (1999).
40. M. Kerszberg and L. Wolpert, *J Theor Biol* **191**, 103-114 (1998).
41. A. D. Lander, Q. Nie, and F. Y. *Dev Cell* **2**, 785-796 (2002).
42. E. V. Entchev, A. Schwabedissen, and M. Gonzalez-Gaitan, *Cell* **103**, 981-991, (2000).
43. A. A. Teleman and S. M. Cohen, *Cell* **103**, 971-980 (2000).
44. H. F. Nijhout, *The Development and Evolution of Butterfly Wing Patterns*. Smithsonian Institute Press, Washington (1991).
45. S. Kondo and R. Asai, *Nature* **376**, 765-768 (1995).
46. T. Jiang, H. Jung, R. B. Widelitz, and C. Chuong, *Development* **126**, 4997-5009 (1999).
47. S. A. Newman and H. L. Frisch, *Science* **205**, 662-668 (1979).
48. H. G. E. Hentschel, T. Glimm, J. A. Glazier, and S. A. Newman, *Proc Roy Soc London B Biol Sci* **271**, 1713-1722 (2004).
49. H. Meinhardt, *Int J Dev Biol* **45**, 177-188 (2001).
50. H. Spemann and H Mangold, *Wilhelm Roux' Arch. Entw. Mech. Org.* **100**, 599-638 (1924).
51. L. Solnica-Krezel, *Curr Biol* **13**, R7-9 (2003).
52. A. Turing, *Phil Tran. Roy Soc Lonon B* **237**, 37-72 (1952).
53. E. Van Obberghen-Schilling, N. S. Roche, K. C. Flanders, M. B. Sporn, and A. Roberts, *J. Biol. Chem.* **263**, 7741-7746 (1988).
54. B. Sun, S. Bush, L. Collins-Racie, E. LaVallie, E. DiBlasio-Smith, N. Wolfman, J. McCoy, and H. Sive, *Development* **126**, 1467-1482 (1999).
55. D. Clements, R. V. Friday, and H. R. Woodland, *Development* 126, 4903-4911 (1999).
56. S. Schulte-Merker and J. C. Smith, *Curr Biol* 5, 62-67 (1995).
57. R. Harland and J. Gerhart, *Annu. Rev. Cell. Dev. Biol.* **13**, 611-667 (1997).
58. H. Juan and H. Hamada, *Genes Cells* **6**, 923-930 (2001).
59. W. W. Branford and HJ Yost, Lefty, *Curr. Biol.* 12, 2136-2141 (2002).
60. M. Yamamoto, Y. Saijoh, A .Perea-Gomez, W. Shawlot, R. R. Behringer, S. L. Ang, H. Hamada, and C. Meno, *Nature* **428**, 387-392 (2004).
61. Y. Chen and A. F. Schier, *Curr. Biol.* **12**, 2124-2128 (2002).
62. R. Sakuma, Y. Y. Ohnishi, C Meno, H. Fujii, H. Juan, J. Takeuchi, T. Ogura, E. Li, K. Miyazono, and H. Hamada, *Genes Cells* **7**, 401-412 (2002).
63. A. S. Aylsworth, *Am. J. Med. Genet.* **101**, 345-355 (2001).
64. V. Castets, E. Dulos, J. Boissonade, and P. DeKepper *Phys. Rev. Lett.* **64**, 2953-2956 (1990).
65. Q. Ouyang and H. Swinney, *Nature*, **352**, 610-612 (1991).

66. S. Nonaka, H. Shiratori, Y. Saijoh, and H. Hamada, *Nature* **418**, 96-99 (2002).
67. S. A. Newman, From Physics To Development: The Evolution of Morphogenetic Mechanisms, in *Origination of Organismal Form: Beyond the Gene in Developmental and Evolutionary Biology.* GB Müller and SA Newman (eds.), Cambridge, MA, MIT Press (2003).
68. S. A. Newman, *J. Ev. Biol.* **7**, 467-488 (1994).
69. N. H. Patel, T. B. Kornberg, and C. S. Goodman, *Development* **107**, 201-212 (1989).
70. T. Itow, *Roux's Arch. Dev. Biol* **195**, 323-333 (1986).
71. P. A. Lawrence, *The Making of a fly: The Genetics of Animal Design*, Blackwell Scientific, Oxford, Boston (1992).
72. K. D. Irvine and E. Wieschaus, *Development* **120**, 827-841 (1994).
73. A. Stollewerk, M. Schoppmeier, and W. G. Damen, *Nature* **423**, 863-365 (2003).
74. J. Boissonade, E. Dulos, and P. DeKepper, Turing Patterns: From Myth to Reality, in *Chemical Waves and Patterns*. R Kapral and K Showalter (eds.), Kluwer, Boston (1994).
75. C. B. Muratov, *Phys. Rev.* **E55**, 1463–1477 (1997).
76. I. Lengyel and I. R. Epstein, *Proc. Natl. Acad. Sci. USA* **89**, 3977-3979 (1992).
77. D. St Johnston and C. Nusslein-Volhard, *Cell* **68**, 201-219 (1992).
78. T. L. Karr, M. P. Weir, Z. Ali, and T. Kornberg, *Development* **105**, 605-612 (1989).
79. M. Frasch and M. Levine, *Genes. Dev.* **1**, 981-995 (1987).
80. K. Howard and P. Ingham, *Cell* **44**, 949-957 (1986).
81. S. Small, A. Blair, and M. Levine, *EMBO J.* **11**, 4047-4057 (1992).
82. S. Small, R. Kraut, T. Hoey, R. Warrior, and M. Levine, *Genes. Dev.* **5**, 827-839 (1991).
83. S. Small, A. Blair, and M. Levine, *Dev. Biol.* **175**, 314-324 (1996).
84. D. E. Clyde, M. S. Corado, X. Wu, A. Pare, D. Papatsenko, and S. Small, *Nature* **426**, 849-853, (2003).
85. N. H. Patel, EE Ball, and CS Goodman, *Nature* **357**, 339-342 (1992).
86. S. J. Brown, N. H. Patel, and R. E. Denell, *Dev. Genet.* **15**, 7-18 (1994).
87. R. Dawes, I. Dawson, F. Falciani, G. Tear, and M. Akam, *Development* **120**, 1561-1572, (1994).
88. S. J. Brown, R. B. Hilgenfeld, and R. E. Denell, *Proc. Natl. Acad. Sci. USA* **91**, 12922-12926, (1994).
89. I. Salazar-Ciudad, R. Solé, and S. A. Newman, *Evolution & Development* **3**, 95-103 (2001).
90. I. Salazar-Ciudad, S. A. Newman, and R. Solé, *Evolution & Development* **3**, 84-94 (2001).
91. S. A. Newman, *BioEssays* **15**, 277-283 (1993).
92. K. Harding, T. Hoey, R. Warrior, and M. Levine, *Embo. J.* **8**, 1205-1212 (1989).
93. D. Ish-Horowicz, S. M. Pinchin, P. W. Ingham, and H. G. Gyurkovics, *Cell* **57**, 223-232 (1989).

94. A. F. Schier and W. J. Gehring, *Embo. J.* **12**, 1111-1119 (1993).
95. C. H. Waddington, *The Strategy of the Genes*. Allen and Unwin, London (1957).
96. I. I. Schmalhausen, *Factors of Evolution*. Blackstone, Philadelphia (1949).
97. H. Meinhardt, *Models of Biological Pattern Formation*, Academic, New York, (1982).
98. L. J. Harrison, *Kinetic Theory of Living Form*. Cambridge, Cambridge University Press (1993).
99. J. Reinitz, E. Mjolsness, and D. H. Sharp, *J. Exp. Zool.* **271**, 47-56 (1995).
100. A. Cornish-Bowden, *Fundamentals of Enzyme Kinetics*, revised edition, Ashgate, London, Brookfield, VT (1995).
101. I. Salazar-Ciudad, J. Garcia-Fernandez, and R. V. Sole, *J. Theor. Biol.* **205**, 587-603 (2000).
102. N. Rashevsky, *Mathematical Biophysics*, revised editon, Univ. of Chicago Press, Chicago (1948).
103. L. V. Bertalanffy, *Vom Molekül zur Organismenwelt; Grundfragen der Modernen Biologie*, 2. verb. Aufl. ed. Akademische Verlagsgesellschaft Athenaion, Potsdam (1949).
104. L. V. Bertalanffy: *General System Theory; Foundations, Development, Applications*, G. Braziller, New York, (1969).
105. A. T. Winfree, *The Geometry of Biological Time*. Springer-Verlag, New York, (1980).
106. S. A. Newman and WD Comper, *Development* **110**, 1-18, (1990).
107. G. B. Müller and SA Newman, *Origination of Organismal Form: Beyond the Gene in Developmental and Evolutionary Biology*. MIT Press, Cambridge, MA, (2003).
108. H. Levine and E. Ben-Jacob, *Phys. Biol.* **1**, 2004, in press,.
109. G. Forgacs and S. A. Newman, *Biological Physics of the Developing Embryo*. Cambridge Univ. Press, Cambridge, MA, in press.
110. G. von Dassow, E. Meir, E. M. Munro, and G. M. Odell, *Nature* **406**, 188-192 (2000).
111. E. Meir, G. von Dassow, E. Munro, and G. M. Odell, *Curr. Biol.* **12**, 778-786 (2002).

http://www.springer.com/978-0-387-23264-5

Complexity in Chemistry, Biology, and Ecology
Bonchev, D.D.; Rouvray, D. (Eds.)
2005, XX, 348 p., Hardcover
ISBN: 978-0-387-23264-5