George Marnellos and Eric D. Mjolsness. In Modeling Neural Development, edited by Arjen van Ooyen. The MIT Press, Cambridge, Massachusetts, 2003.

Gene Network Models and Neural Development

January 11, 2002

Abstract

Rapid advances in molecular biology and genomics in recent years have highlighted the need for theoretical tools to analyze and integrate a flood of data. Models, both at detailed biochemical and at more abstract levels, could help make sense of experimental observations, fragmentary by nature, and formulate new hypotheses for testing, on a broad range of questions surrounding the development and physiology of organisms. In this review we present mathematical and computational models of molecular processes underlying biological development, and concentrate on the role of gene and their interactions.

We describe in greater detail a gene network modeling formalism, based on neural networks, that was introduced by Mjolsness *et al.* to study gene regulatory interactions during development; and the application of this framework in a computational model of *Drosophila* early neurogenesis. Although such models are only small steps in the elucidation of how genes orchestrate the complex patterns of neural development, they could provide directions to subsequent research, and could be refined to deal with more powerful data as they become available.

1 Genomic Advances Document Unity of Life

From bacteria and their limited repertoire of multicellular forms to primates and their elaborate nervous systems, the diversity of form and function and the interdependence between the two have occupied biologists for a long time:

how are morphological and behavioral patterns of individual organisms coded for and what forces bring them about? how is the tremendous richness of these patterns generated in nature, what rules or constraints determine its generation and what purpose might it serve? Evolutionary biology provides of course a framework for addressing questions like these, but now, with the advent of molecular techniques and their widespread application, we have the unique opportunity to explore at a mechanistic level how life's diversity is generated and why.

In the last 30-40 years, molecular methods for manipulating primarily genetic material but also other cellular components have allowed researchers to investigate hitherto inaccessible processes both at the cell level and at higher levels of integration. Workers in the biological sciences, from basic research to agriculture and medicine, now have a common language to communicate their experiences in distant fields. Genes and entire biochemical pathways isolated in one experimental system (yeast, for instance, or *C. elegans*) can now almost routinely be identified in many other organisms, not only providing clearer answers to questions originally investigated in individual species but also revealing previously unimagined levels of homology.

Examples from many biological systems abound. A particularly striking one is that of *homeotic* genes which are necessary for specifying animal body patterns; they are transcription factors, containing a conserved region that binds to DNA, and determine regional identity. Their characteristic mutant phenotypes, which can lead to body part duplication or replacement of one body part by another (e.g. legs in place of antennae), had long puzzled geneticists who had studied them in *Drosophila*; in fact in 1978 one of those researchers, E.B.Lewis, formulated a detailed hypothesis as to how homeobox genes might have contributed to the evolution of the fly body plan, thus anticipating many related discoveries in the following decade [51]. These genes were first cloned in *Drosophila* [63, 86], and because of the conserved region, the *homeobox*, their homologues were subsequently isolated in many other different phyla, from cnidarians to vertebrates [3, 47, 72, 85, 50] and were shown to be organized in similar gene groupings and expressed in the same order along the body axis as in *Drosophila*. Such findings have prompted the re-evaluation of homologies between body parts across phyla and evolutionary relationships in general, and have provided insights into the evolution of developmental mechanisms, pointing to the importance of regulatory gene evolution in morphological diversity [14].

Other examples of genes and pathways conserved across species include,

to name but a few: the Ras family of G proteins which integrate inputs from a wide range of cellular pathways and control many aspects of cell proliferation and differentiation in yeast, flies, nematodes and mammals, by transmitting signals from tyrosine kinases at the plasma membrane through a kinase cascade to the nucleus [7]; the mitogen-activated protein kinase (MAPK) pathway, which consists of a cascade of three classes of kinases that deliver signals from membrane to nucleus, and which was first identified in the mating response pathway of yeast (*Saccharomyces cerevisiae*) and then in many metazoans (in metazoans, the MAPK cascade conveys inputs of Ras proteins, among others) [74]; the homologous pathways which regulate fly and vertebrate limb development, as in the case of the *hedgehog* family of proteins involved in the anteroposterior patterning of the limb [49, 20, 83, 21]; the family of nuclear receptors, such as steroid, retinoid and thyroid hormone receptors, which mediate the potent effect of hormones on gene expression and which are dominant regulators of organ physiology as well as of insect morphogenesis [6, 41, 57, 97].

2 Need for Theoretical Tools

The list of such discoveries grows daily, supplying strong evidence about the unity of life forms and their biological building blocks, which could not have been anticipated, strengthening the conviction that description at the gene level will provide a unifying principle to explain diverse biological phenomena, and exposing the need of analytical methods to study these phenomena.

The situation is perhaps comparable to what happened in physics during the first decades of this century, when the structure of the atom was probed and this opened the door to the discovery of subatomic particles and a unified way of viewing matter in all its forms. However there is a crucial difference between today's biology and the state of physics at the beginning of the century: whereas physicists at the time had access to a formidable chest of mathematical tools, which had been the language of physicists for centuries, the theoretical means at the disposal of biologists seem limited in comparison. There have been of course major contributions, like the work of Lotka and Volterra on population biology and predator-prey models [56, 100], Fisher on gene flow dynamics in populations [22], Turing on the role of reaction and diffusion of substances in pattern formation during development [98], Thom on the topological aspects of morphogenesis [96] and Hodgkin and Huxley on

generation and propagation of action potentials in nerve axons [35]. These pioneering studies have helped establish whole fields of research in ecology, population genetics, developmental biology and electrophysiology, but the fact remains that, on the whole, theory in biology has not occupied the same place in the interpretation of data and the designing of experiments as it has in physics.

The need of theoretical methods in biology has become more apparent in the last decade as data from molecular experiments are pouring in at an ever increasing rate, especially from whole genome sequencing projects; people have started responding to this realization, most notably in the area of computational molecular biology, at the border between biology and biochemistry, where mathematical tools (for instance, statistical analysis and combinatorial optimization) have been applied to the study of structure and function in genomes, genes and gene products in order to answer specific biological questions, like, for example, discovering genes in sequences, detecting gene homologies between species, building phylogenetic trees, and predicting protein and nucleic acid secondary and tertiary structure from sequence information.

But the intensity of theoretical effort in genomic research will have to be extended to the study of development and the phenotypic variation that it generates and selection acts upon; in particular, it will have to be directed to the study of genes and their involvement in these processes, since molecular data will contain a wealth of information on developmental and evolutionary questions.

3 Models of Molecular Processes and Development

REACTION-DIFFUSION. There has already been a considerable amount of work on mathematical models of development. A large part of this work has been in the tradition of Turing's reaction-diffusion approach, by researchers like H. Meinhardt and J.D. Murray, who have modeled the stable patterns that can emerge when chemical substances, *morphogens* (usually two in number, an *activator* and an *inhibitor*), diffuse and react with each other over a morphogenetic field, in systems like insect segmentation, sea shells, animal coats and butterfly wings ([26, 65, 66, 68, 69], for overview of reaction-

diffusion and related models in mathematical biology see [70]). These efforts have generally dealt with abstract quantities and have not attempted to make explicit connections between these and interactions of specific genes, although in some cases subsequent experimental work has provided candidate molecules, as in the case of Meinhardt's model of how insect leg proximo-distal coordinates are set up [64, 5, 18].

MECHANICAL MODELS. Other researchers have considered cell movements and mechanical properties of cells and tissues and modeled processes like gastrulation and neurulation [77], cartilage condensation in limb morphogenesis and patterning of feather primordia [71, 25], aggregation of *Dictyostelium* amoebae [87], cell intercalation and sorting [95, 102, 2, 29, 30] and skin generation [94].

LINDENMAYER SYSTEMS. Drawing inspiration from formal languages, Lindenmayer has modeled development using sets of rules, *grammars*; rules describe cellular processes like growth, division and differentiation, and are applied to modify strings that represent organisms [53]. These, so called, Lindenmayer systems have been used to model growth and branching patterns of plants [79]; they have also been extended to include cell-cell interactions, through the use of context-sensitive grammars, as well as 2-D and 3-D cells which can change shape, as in some of the mechanical models above [55, 54].

BIOCHEMICAL KINETICS. The models above do not attempt to make any connections to specific genes or biochemical pathways involved in the processes modeled, and in fact most of them do not make any reference to such factors at all. There have been models incorporating such molecular elements, although in this case modeling has been restricted to processes within single cells. Savageau, applying methods from chemical kinetics, has modeled biosynthetic pathways and the regulation of gene expression in prokaryotes; he analyzed the fixed points and periodic behaviors of these systems and also considered questions of optimality in the design of the pathways [84]. Other workers have also examined dynamical features of metabolic pathways and gene expression [28, 37, 31, 99], Ca^{2+} signalling [19] and the complex interactions of progression through the cell-cycle [75].

In a similar framework, Bray has pointed out the similarities of biochemical signal cascades to neural networks that might be performing some kind of pattern recognition within cells. It has in fact been shown that chains of chemical reactions can be viewed as neural networks that can be reduced to Hopfield nets [34, 33, 9]. Bray has modeled networks of cell-signalling reactions [8, 12] and has optimized the reaction parameters to achieve a desired

mode of functioning or output of a pathway. He has used this method to simulate the signalling cascades involved in bacterial chemotaxis and find reaction parameter values such that the simulated system exhibits various chemotactic behaviors of known mutant phenotypes [11, 10].

In a biochemically more concrete look at morphogen gradients during development (cf. section on Reaction-Diffusion models above), Kerszberg and Changeux [48] examine how different assumptions about transcription factor dimers, autocatalytic feedback and competition for regulatory binding sites by these dimers lead to different patterns of transcription and protein concentration; the authors use results of the model to interpret morphogen gradients of Bicoid and Hunchback in the *Drosophila* blastoderm (see also section on blastoderm below).

von Dassow *et al.* [101] have more recently also looked at developing *Drosophila* embryos, using a pretty realistic biochemical interaction model: they examined interactions between segment polarity genes and found that the whole system is dynamically robust and can resist changes to its kinetic parameters.

PHAGE λ. In work concentrating on gene interactions, Shea and Ackers [89, 1] have developed a detailed quantitative model of regulation of certain genes of bacteriophage λ that are involved in maintenance of the lysogenic state (when the prophage gets integrated into the DNA of the host) and induction of lysis (when the virus actively replicates). In constructing their model, the authors stayed very close to biochemical facts concerning the structure of the genes involved and their promoters, binding constants, dimerization, cooperative interactions and so on. In a hybrid modeling approach, McAdams and Shapiro have also looked at the lysogeny-lysis switch of phage λ [62]. by integrating chemical kinetics with an electrical circuit simulation of the genes and regulatory interactions that control this switch.

BOOLEAN NETS. In a different vein, abstracting away from biochemical detail, Kauffman has introduced networks of elements with binary states to model gene regulatory interactions [42]. These *boolean networks* are intended to be idealizations of continuous dynamical systems with elements that behave in a sigmoidal fashion (as is the case with many cellular and biochemical processes); they are believed to capture a skeleton of the dynamical structure of such continuous systems (see Chapter 5 in [45] and references therein). Since boolean networks have finitely many different states, they are guaranteed to have fixed points and state cycles, which could be viewed as corresponding to stable differentiation states and periodic behaviors of

cells — and such parallels are being explored in specific cases where known gene expression patterns appear to be consistent with this description [92]. Kauffman has explored the stability of the dynamics of these boolean nets, which depends on the number of inputs of each element, and on the number of elements [42, 43, 44]. He has also ascribed fitness values to different configurations of boolean nets and investigated features of the fitness landscapes that result in such configuration spaces, like number, similarity and accessibility of fitness peaks [46].

4 Gene Net Framework

Genes being a natural module for the description of living systems, they also appear to be a natural level of abstraction for integrated biological models. Starting from this premise, Mjolsness, Sharp and Reinitz [67] have introduced a modeling framework for the study of development, centered around genes and their interactions. This framework shares features with the models described above but in a combination that is not found in any of the others. It incorporates features that allow modeling of processes at a tissue level, like the reaction-diffusion and mechanical models, for instance, but unlike the biochemical kinetics and phage λ models; and unlike the former but similar to the latter, it also includes a description of molecular processes, like gene expression. The modeling framework consists of two major components:

- **1.** A neural network representation of molecular level interactions; gene interactions, as well as other molecular signalling and regulatory events, are modeled as a particular kind of neural nets, *recurrent* nets with connections allowed in both directions between any pair of nodes ([36], see also [32] for a survey of recurrent nets and neural nets in general); in this formulation, gene product concentrations correspond to node activation levels and connection weights to gene interaction strengths.

- **2.** A Lindenmayer-system-like grammar of rules [53, 79], *L-grammar*, which describe cell-cell interactions and changes in number, type and state of cells.

4.1 Dynamics

In more detail, genes in such networks interact as nodes in a recurrent neural net, summing inputs from other genes at any given time t

$$u_a(t) = \sum_b T_{ab} v_b(t) \tag{1}$$

where T is the matrix of gene interactions and $v_b(t)$ gene product concentrations within the cell; if we include interactions with neighboring cells, this becomes

$$u_a(t) = \sum_b T_{ab} v_b(t) + \sum_i \sum_b \hat{T}_{ab} \hat{v}_b^i(t) \tag{2}$$

where $\hat{T}$ is the matrix of gene interactions with neighboring cells and the $\hat{v}_b^i(t)$ gene product levels in neighboring cell i. Concentration $v_a(t)$ of the product of gene a then changes according to

$$\frac{dv_a}{dt} = R_a g(u_a(t) + h_a) - \lambda_a v_a(t) \tag{3}$$

where $u_a(t)$ is the linear sum of Eq. 1, R_a the rate of production of gene a's product, h_a the threshold of activation of gene a and λ_a the rate of decay of gene a product; function g is a monotonic, non-linear function, usually a sigmoid, like the following one which we have used in gene net models and which is centered at 0 and takes values between 0 and 1:

$$g(x) = 0.5(1 + \frac{x}{\sqrt{(}1 + x^2)}). \tag{4}$$

Levels of gene products should be thought to correspond to gene product activities in the biological system rather than to actual concentrations, and gene interactions should be thought to correspond closer to genetic rather than specific biochemical (transcriptional etc.) interactions. The form of Eq. 3 can be justified as follows: if we consider gene a as a producer molecule that can be either in an activated/producing state or in inactivated/non-producing one, depending on the concentrations of other gene products that can bind at its regulatory regions, then the amount of species a produced is proportional to the fraction of time that gene a spends in the activated state (or equivalently to the fraction of producer molecules in that state); species

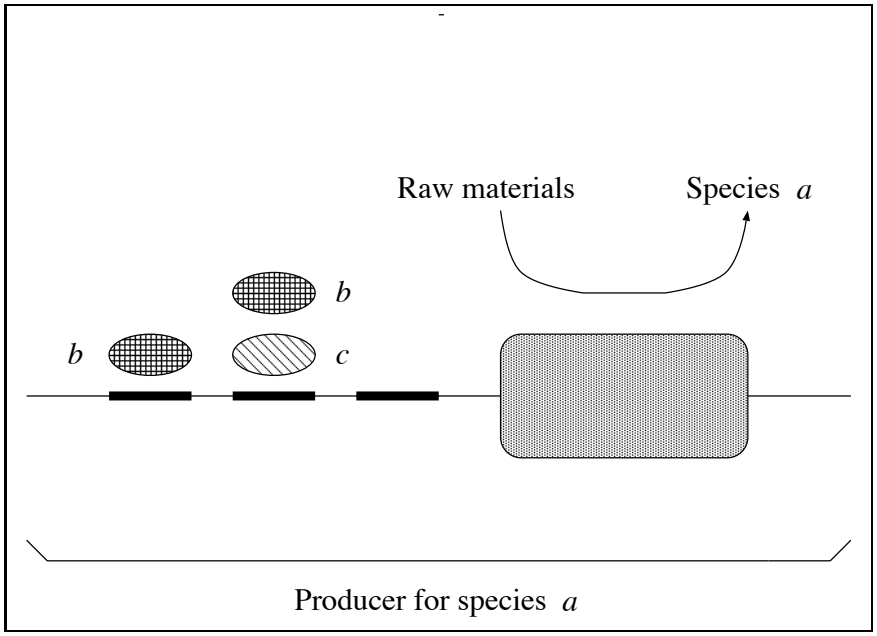

Figure 1: Illustration of the biochemical model incorporated in Eqs. 3 and 5: production of species a depends on the activation of gene a, which is determined by the binding of gene products b and c at regulatory regions of gene a and by interactions of these with the transcription apparatus at the gene. Adapted from Fig. 2 in [67].

a also decays at a rate λ_a independent of gene product concentrations. This is expressed in the following equation (which has the same form as Eq. 3):

$$\frac{dv_a}{dt} = R_a[FractionActivated] - \lambda_a v_a \tag{5}$$

where $[FractionActivated]$ is the fraction of time that gene a is activated. It is this fraction, which depends on the concentrations of other gene products, that is approximated by the recurrent net formulation of Eq. 3. See also Fig. 1. For a more detailed biochemical rationale of why gene expression kinetics can be approximated by Eq. 3 see Section 4 in [67].

4.2 L-grammar

The gene net framework allows for cell transformations in the models; for instance, cells may change their state (i.e. the levels of gene products or

other state variables), change type, give birth to other cells, die. These transformations are represented by a set of grammar rules, the *L-grammar*, as in Lindenmayer Systems [53, 79]. Rules are triggered depending on the internal state of each cell (or perhaps also of other cells) and are of two kinds, discrete and continuous time rules. Transformations that happen gradually (smoothly) over time are described by continuous time rules, while processes that occur as abrupt, discontinuous changes are given by discrete time rules, which are instantaneous. Rules may involve one or more cells, representing intracellular processes and cell-cell interactions, respectively (see Fig. 2). How rules are triggered depends on the internal state of each cell (or perhaps

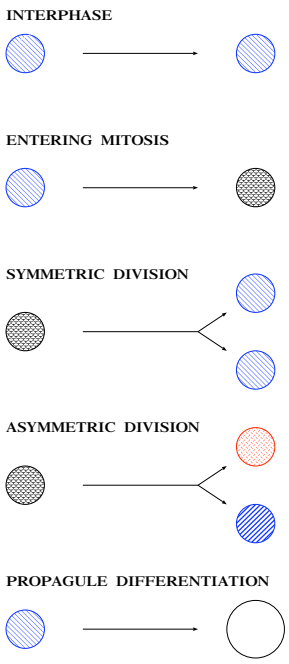

Figure 2: Graphical representation of the set of rules that are applied to cells in the model. Blue and red disks denote cells in interphase, black disks cells in mitosis and the larger circle a propagule with its stored reserves. Note that the asymmetric cell division rule produces two cells which differ in their gene product concentrations and are depicted by different colors, whereas the symmetric division rule produces identical cells.

also of other cells). There are some constraints as to what rules may be active in a cell at any given time: only a single continuous one-cell rule is allowed at a time but several continuous two-cell rules may operate simultaneously;

only one birth or death rule may occur at a time, but other discrete rules that bring about changes in cell type may happen simultaneously, as long as they all transform a cell to the same cell type. A set of binary variables C keep track of what rules are active in any particular cell at any given time. Vector $\mathbf{u}$ of Eq. 2 is therefore more accurately given, for a cell i, by

$$\mathbf{u}_i = \sum_r C_i^r \mathbf{T}_1^r \cdot \mathbf{v}_i + \sum_r C_i^r \sum_j \Lambda_{ij} \mathbf{T}_2^r \cdot \mathbf{v}_j \tag{6}$$

where $\mathbf{T}_1^r$ is the interaction strength matrix for one-cell rule r; $\mathbf{v}_i$ the state variable (gene product concentration) vector for cell i; $\mathbf{T}_2^r$ the interaction strength matrix for two-cell rule r (r, of course, in both sums of Eq. 6 is just a a dummy variable that stands in for the actual names of the rules, which could be, for instance, *mitosis*, *cell-death*, *interphase* and so on); $\mathbf{v}_j$ the state variable vector for cell j, located in the neighborhood of cell i, the type of neighborhood being specified in particular models; Λ_{ij} a factor that modifies the influence of cell j on cell i and depends on the geometry of cells and their positions in a model. Variables C_i^r determine which rules in Eq. 6 operate at any given time; if $C_i^r = 1$ then the corresponding rule is active, while if $C_i^r = 0$ the rule is inactive; these C variables encode the constraints on rule activity that were described above.

When two rules cannot be both active at the same time, the rule of highest *strength* at that time wins: strength, S_i^r, of rule r in cell i signifies the likelihood that rule r will be triggered in that cell at that time, and depends on the internal state of the cell (and perhaps also on interactions with neighboring cells). Considering only dependence on the internal state of the cell itself (i.e. on the state variable vector $\mathbf{v}_i$), rule strength is given by

$$S_i^r = \mathbf{v}_i \cdot \mathbf{s}_r + \theta_r \tag{7}$$

where $\mathbf{s}_r$ is a vector that describes how each state variable of the cell contributes to the strength of rule r and θ_r is the default likelihood that rule r will be triggered.

Rule strengths together with the constraints determine which rules are active at a given time, and along with the parameters of state variable dynamics, $\mathbf{T}, \mathbf{R}, \mathbf{h}$, and λ of Eqs. 3 and 2, and geometry factors $\mathbf{\Lambda}$, completely specify how the modeled system develops.

4.3 Optimization

Models using the gene net framework can be formulated as optimization tasks that seek values for the model parameters such that the model optimally fits biological data or behaves in a certain desired manner. Such requirements can be captured in a so called *cost* function (or *energy* function) $E(\mathbf{v})$, which depends on the state variable values $\mathbf{v}$ during development of the system; E, of course, ultimately depends on the values of model parameters. A common example of a cost function is a least-squares cost function

$$E = \sum_{cells,genes,times} (v^i_{aMODEL}(t) - v^i_{aDATA}(t))^2, \tag{8}$$

which is the squared difference between gene product concentrations in the model and those in the data, summed over all cells and over all gene products and times for which data is available.

A quadratic penalty term on these parameters, of the form

$$Penalty = \sum_{parameters} w_i p_i^2 \tag{9}$$

where p_i are parameter values and w_i weights (usually the same for all parameters), is added to E to produce the final, *objective*, function which is optimized:

$$Objective = w_E E(\mathbf{v}(\mathbf{p})) + w_p Penalty(\mathbf{p}) \tag{10}$$

$\mathbf{p}$ being the vector of model parameters and w_E, w_p weights for the energy and penalty terms of the objective function, respectively. The penalty term prevents optimized parameters from growing excessively large and hence saturating the sigmoid functions of the model or causing overflow errors in computer simulations. It effectively restricts the search space and thus may facilitate the optimization; however, if the restricted search space does not contain the sought optima or if, depending on the optimization algorithm used, parts of the parameter space are not equally accessible from all other parts of the space, this may adversely affect the optimization search — see [24] for a discussion, in the context of genetic algorithms, of this and the more general problem of optimizing objective functions that have many components.

The objective functions in gene net models typically have a large number

of variables, are highly nonlinear and cannot be solved analytically or readily optimized with deterministic methods. We have therefore used numerical, stochastic techniques to optimize them, namely *simulated annealing* (SA) and *genetic algorithms* (GA). Both of these optimization methods have a number of parameters that can affect their performance and need to be tuned for each individual problem.

For more details on this connectionist framework and its application to lateral inhibition models, see Mjolsness et al. [67] and Marnellos (1997) ([58] section 2.1).

4.4 Overview

This combination of differential equations and grammatical structure is intended to make models computationally feasible and yet maintain a wide repertoire of behaviors at the molecular and tissue levels. Grammars can be thought to summarize aspects of the intracellular and intercellular dynamics of the system being modeled, which would otherwise require a large number of extra state variables and model parameters to describe. So grammars offer a concise and thus computationally tractable representation. The neural net idealization representing molecular interactions is at a similar level of abstraction as the biochemical kinetics models; the phage λ models, in contrast, incorporate much greater biochemical detail and it would be very expensive computationally to have that much detail in models of multicellular development.

The neural net idealization has the following additional advantage: neural nets can be "trained" to produce desired outputs. This property of the neural net formalism has been extensively studied [32] and there are algorithms to perform the training — for instance, in the case of sufficiently simple recurrent neural nets, there are even deterministic methods (like those described in [78, 103]) to do the training. In the gene net framework, training corresponds to fitting experimental observations, or having the simulated system behave in a desired fashion, by optimizing the adjustable parameters of the gene nets, i.e. gene interaction strengths, activation thresholds etc.; this is similar to what Bray has done in his bacterial chemotaxis models mentioned above [11, 10], which of course are models of single cells only.

The scope of biological questions that can be addressed with the gene net framework is comparable to that of Kauffman's boolean nets which are computationally less expensive than gene nets; however, because of the binary

way in which they represent molecular events and because they do not represent cells, tissues or such entities, boolean nets cannot be readily used to interpret a large body of molecular and cell-level experimental observations.

A framework very similar in scope, structure, expressiveness and level of biochemical detail to the gene net method has also been proposed by Fleischer [23]. Its main differences with the gene net framework are that it does not have grammar rules to represent state changes in cells, but instead uses conditional terms in the ordinary differential equations that describe how state variables change; state variable dynamics are not neural-net-like but of a more arbitrary form that can be specified by the user; and it has mainly been used to simulate artificial configurations of cells and not interpret biological observations.

5 Applications of the Gene Net Framework

DROSOPHILA BLASTODERM. Reinitz, Mjolsness and Sharp [81, 82] have applied the gene net framework to an early stage of development of the *Drosophila* embryo (the *blastoderm* stage). They have looked at the well-characterized hierarchy of regulatory genes that control the early events of *Drosophila* embryogenesis by setting up their expression patterns along the embryo's length and dividing it into segments. This includes maternal gene products, *bicoid* (*bcd*) and *hunchback* (*hb*), expressed in broad gradients along the anteroposterior axis of the embryo, and so called *gap, pair-rule* and *segment polarity* genes, which end up segmenting the whole length of the embryo into stripes, each a single cell wide [76, 38, 93]. As the expression of these genes does not vary along the dorsoventral axis of the blastoderm and since there are no separate cells at the blastoderm stage but the embryo is a syncytium of nuclei arranged at its surface like a shell, the authors have modeled the system as a single row of nuclei which are the sites of gene expression and which interact with each other through the diffusion of gene products.

They have investigated questions of positional specification in the blastoderm and their model has yielded predictions and interpretations of experimental observations: it predicted that Bicoid and Hunchback proteins cooperatively determine position in the anteroposterior axis [81], which has subsequently been confirmed by experiment [90]; and offered insights into the spatiotemporal expression pattern of pair-rule gene *even-skipped (eve)*, on questions like which domains of gap gene expression set the boundaries of

eve stripes, and on the timing and order of appearance of these stripes [82]. Moreover, the model provided an explanation for a cell-biological observation, namely that pair-rule mRNA's are apically localized: it showed that *eve* stripes do not form unless Eve protein has very low diffusivity, which could presumably result from its apical localization.

Drosophila Neurogenesis. Marnellos and Mjolsness [60, 61] have worked on early neurogenesis in Drosophila and constructed models to simulate how neuroblasts and sensory organ precursor (SOP) cells differentiate from proneural clusters of equivalent cells (see more below). These neurogenesis models have made predictions about the dynamics of cluster resolution and how the interplay of factors like proneural cluster shape and size, gene expression levels, and strength of cell-cell signalling determine the timing and position of appearance of neuroblasts and SOP cells; and about the robustness of this process and the effects of perturbations in gene-product levels on cell differentiation.

Xenopus Ciliated Cells. In a more recent model Marnellos and Mjolsness [59] probed lateral inhibitory signalling through the Delta-Notch pathway and its role in the emergence of Xenopus ciliated cells in a salt-and-pepper pattern on the, initially uniform, epidermis. The model reproduced the phenotypes observed experimentally under the assays tested. Statistical analysis of "genotypes" in the model suggested that the model could account for the variability of embryonic responses to the experimental assays, and highlighted a component of lateral inhibition that may be the chief source of this variability.

6 A Gene Net Model of Early Neurogenesis in Drosophila

6.1 Background

In Drosophila, neuroblasts and sensory organ precursor (SOP) cells differentiate from epithelia to give rise to the central nervous system in the fly embryo and to epidermal sensory organs in the peripheral nervous system of the adult fly, respectively. Neuroblasts are neural precursor cells that divide to form neurons and glia; they segregate from the ventral neuroectoderm of the embryo in a regular segmental pattern. SOPs appear at stereotypical positions on imaginal discs of fly third instar larvae and divide to produce a

neuron and three other cells that form Drosophila's sensory organs, like the bristles on its thorax.

The activities of two main sets of genes working in opposite directions are thought to underlie this differentiation process: one promoting neural development and the other preventing it and favoring epidermal development.

Neuroblasts and SOPs differentiate from apparently equivalent clusters of cells expressing genes of the *achaete-scute* complex, so called proneural genes. Eventually only one cell from each proneural cluster retains proneural gene expression and becomes a neuroblast or SOP, in a process referred to as cluster resolution (see Fig. 3). Proneural genes thus promote the neuronal fate. The other set includes a number of genes also encoding nuclear proteins,

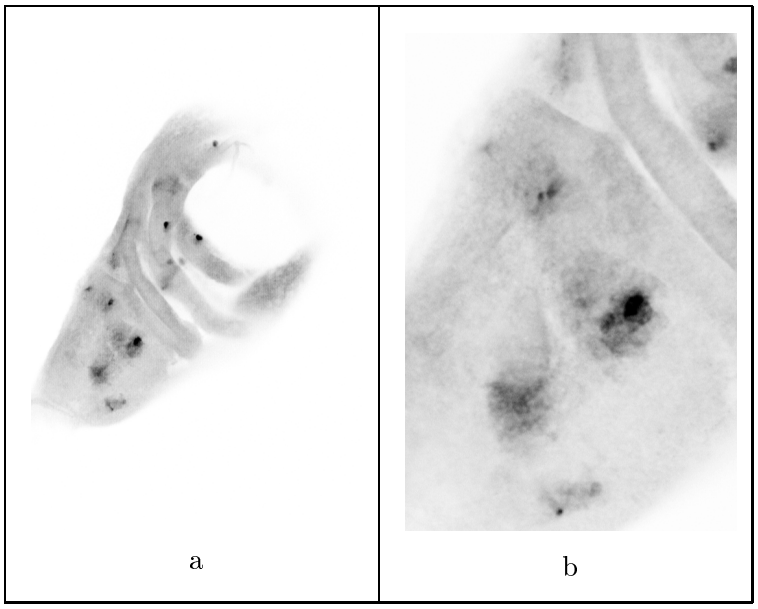

Figure 3: (a) Proneural gene expression in clusters in a *Drosophila* wing disc (the appendage of the fly larva that gives rise to the wing and the back of the adult). The *lacZ* reporter gene indicates *achaete* expression (*achaete* is one of the proneural genes). (b) Detail of (a), note cluster on lower left that has not yet resolved; other clusters appear to be at a more advanced stage of resolution.

like genes of the *Enhancer-of-split* (*E(spl)*) complex and *hairy* as well as other genes for membrane and cytoplasmic proteins; all these tend to suppress neurogenesis and promote epidermal development. In this paper we refer

to this set of genes as *epithelial* genes — in the literature they are called "neurogenic" genes, because loss-of-function mutations of these genes lead to overproduction of neurons.

Cluster resolution and the singling out of neural precursors from within proneural clusters is brought about by inhibitory lateral signalling between adjacent cells, through the signalling pathway of receptor Notch and its ligand Delta; the neural fate is promoted in the future neuroblasts and SOPs and suppressed in other cells.

Several other genes that are involved in this specification of cell-fate are also expressed in characteristic spatial and temporal patterns during the process (for reviews see [13, 73, 4]).

Despite the number of empirical observations that have been gathered, several features of this system remain unexplained: a precise characterization of lateral signalling is still lacking; we do not understand dynamical aspects of the system, for example, whether and how the shape and size of proneural clusters determine how cluster resolution proceeds; it is not clear what the role, if any, of cell delamination, which accompanies neuroblast differentiation in the fly embryo, is.

6.2 Model

In our model, cells are represented as overlapping circles in a 2-dimensional hexagonal lattice; the extent of overlap determines the strength of interaction between neighboring cells. Cells in the model express a small number of genes corresponding to genes that are involved in neuroblast and SOP differentiation. In the work presented here we have used networks with four genes (one corresponding to the proneural group, another for the epithelial group and two for the ligand and receptor, respectively, mediating cell-cell signalling).

Genes interact as nodes in recurrent neural nets, as described above in Equations 2 and 3. The matrix T of gene interactions has the structure depicted in the table below; columns in this table are for input genes and rows for genes affected (empty boxes signify zero interaction strength. i.e. no interaction):

Intracellular Interactions				
	Proneural	Epithelial	Receptor	Ligand
Proneural	♦	♦		
Epithelial	♦	♦		
Receptor	♦	♦		
Ligand	♦	♦		

This table shows that we have allowed only proneural and epithelial gene products to directly regulate the expression of other genes (themselves included), since these two genes correspond to transcription factors in the real biological system.

We have modeled lateral interactions between cells by the binding of ligand to the receptor in the neighboring cell and subsequent regulation of the epithelial gene by the active ligand-receptor complex — this corresponds to the signal relayed from activated Notch receptor to epithelial gene *E(spl)*. In more detail, the ligand-receptor reaction is assumed to be governed by mass-action type kinetics:

$$L + R \leftrightarrows L \circ R \tag{11}$$

where L is ligand (on one cell), R receptor (on a neighboring cell) and $L \circ R$ the active receptor-ligand complex; the rate of the reaction to the right is, say, k_1 and to the left k_2. If v_L is ligand concentration, v_R receptor concentration and $[L \circ R]$ concentration of the receptor-ligand active complex, then we have that

$$\frac{d[L \circ R]}{dt} = k_1 v_L v_R - k_2 [L \circ R] \tag{12}$$

$$\frac{dv_L}{dt} = \frac{dv_R}{dt} = -k_1 v_L v_R + k_2 [L \circ R]. \tag{13}$$

This reaction is assumed to take place at a much faster timescale than gene expression and to have reached a steady state before influencing gene expression. Thus the epithelial gene in a cell receives input from receptor-ligand complexes activated by ligand in the six surrounding cells (the lattice is hexagonal).

We optimize on gene interaction strengths in order to fit gene expression patterns described in the literature; the cost function optimized is a least-squares one, as in Eq. 8. We have used a stochastic algorithm, simulated annealing, for this optimization. For more details on the model and the optimization method used see [58, 61].

The gene expression datasets we optimize on, the *training* datasets, are adapted from schematic results described in the experimental literature [17, 91, 39]; they specify the initial pattern of concentrations of the gene products, i.e. the proneural clusters, the desired intermediate pattern, and the desired final pattern when the proneural clusters have resolved to single cells expressing the proneural gene at high levels (see Fig. 4); it is left to the optimization to find the right model parameters so that the system develops from the initial state through the intermediate one to the desired final one. The initial concentrations of receptor and ligand are uniform for all cells and their subsequent concentrations are not constrained by the dataset (in this respect, they are comparable to hidden units in neural nets). With parameter values derived by optimization on the dataset of Fig. 4, the model makes predictions about how the interplay of factors like proneural cluster shape and size, gene expression levels, and strength of cell-cell signalling determines the timing and position of appearance of neuroblasts and SOP cells; and about the robustness of this process and the effects of gene product level perturbations on cell differentiation.

The figures below (taken from [61]) illustrate the robustness of model results to small perturbations (Fig. 5) and the kind of predictions that the model makes (Fig. 6); these can also be viewed as *in silico* tests of possible biological manipulations and experiments.

7 Discussion

In this paper we have presented an overview of computational models of gene regulation that have been applied to questions in biological development, including fairly detailed models of biochemical reactions and their dynamics as well as more abstract reaction-diffusion and boolean network models. In particular we have concentrated on a gene network model based on Mjolsness *et al.* framework, as it has been applied to early neurogenesis in Drosophila.

Although different in detail, all such models offer insights into the pro-

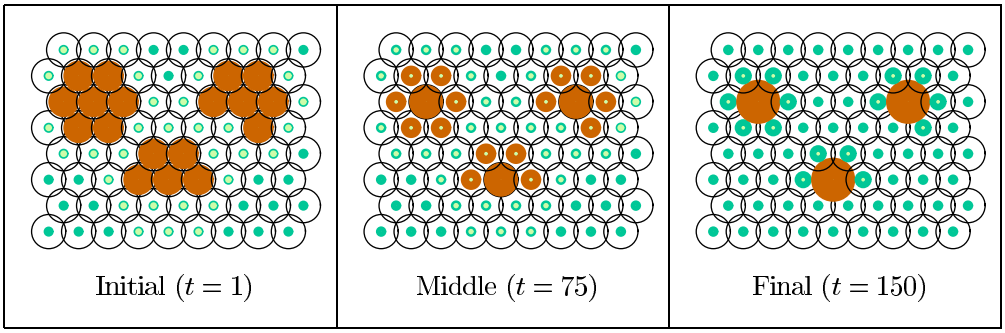

Figure 4: Cells are modeled as circles on a hexagonal lattice. Gene expression is represented by disks, proneural expression in brown, epithelial in green, and where the two overlap in yellow-green; disk radius is proportional to level of expression. This figure shows the training dataset: on the left, the initial concentrations of the gene products — there is only proneural gene expression in three symmetrical clusters; in the middle, the desired intermediate pattern of expression; on the right, the desired final pattern of gene expression — proneural expression is retained only in the central cell of each cluster, the future neuroblast or SOP, whereas all other cells express the epithelial gene. Times t indicate the points in the run when the desired expression pattern is compared with the actual one (see Eq. 8); at $t = 1$ there is of course only initialization and no comparison. Initial concentrations of ligand and receptor are not shown.

cesses under study, by posing biological questions in more concrete terms and testing the logical consistency and inferences of underlying assumptions. The gene network approach has the advantages that, through its grammar strucrture, it can represent a large spectrum of molecular and tissue-level processes, while at the same time being computationally tractable, and, because of its neural net dynamics, it is suitable for having its adjustable parameters trained on experimental data or optimized to produce other desired behaviors. However, as with other models with parameters fitted on experimental data, available regression methods (like optimization through simulated annealing, used in the neurogenesis model presented above) are often inadequate for fitting the parameters of gene network models; this could limit the scalability of such models.

Neurogenesis, which we have examined in this review using our gene network model, encompasses general questions of cell differentiation in epithelia and tissue level dynamic interactions, that are relatively simple to formulate and yet complex enough to be of theoretical and experimental interest,

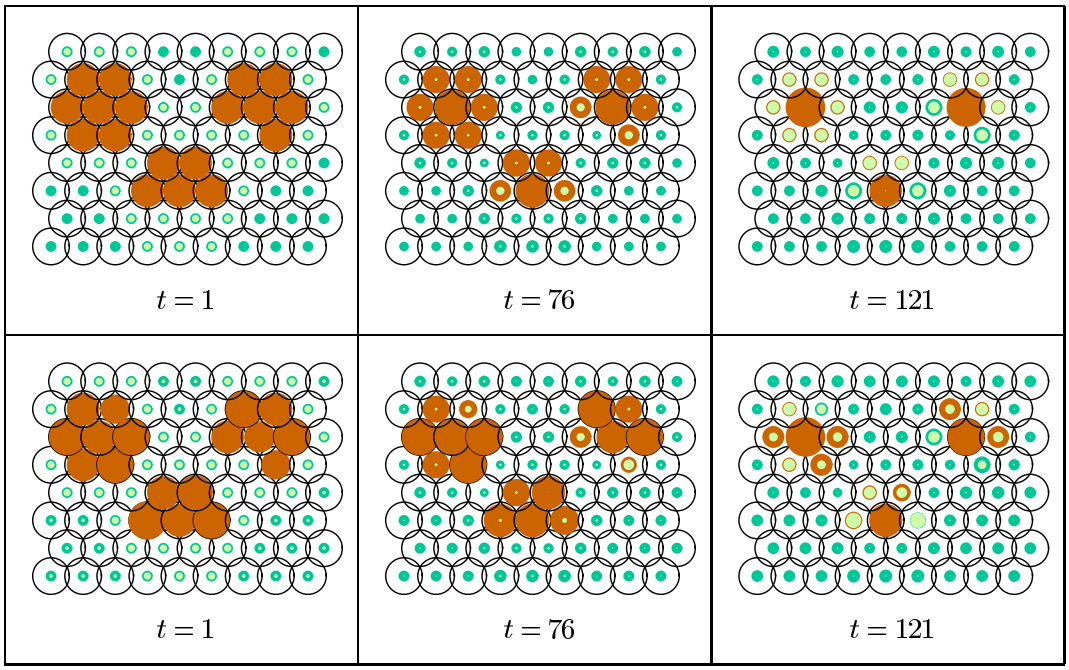

Figure 5: Computer simulation of neural precursor differentiation with parameter values in the model derived by optimization on the dataset of Fig. 4. From left to right, different time frames of the evolution of gene product concentrations. Top row: run with identical initial gene product concentrations for all cells in each proneural cluster. Bottom row: initial proneural concentrations vary by about 10-15% between cells in each cluster. In both runs the clusters resolve in the same way, as the comparison of the two panels at $t = 121$ shows; in the bottom run, the clusters take slightly longer to resolve. This illustrates the robustness of cluster resolution to small changes in initial gene expression levels in proneural clusters. Conventions as in Fig. 4.

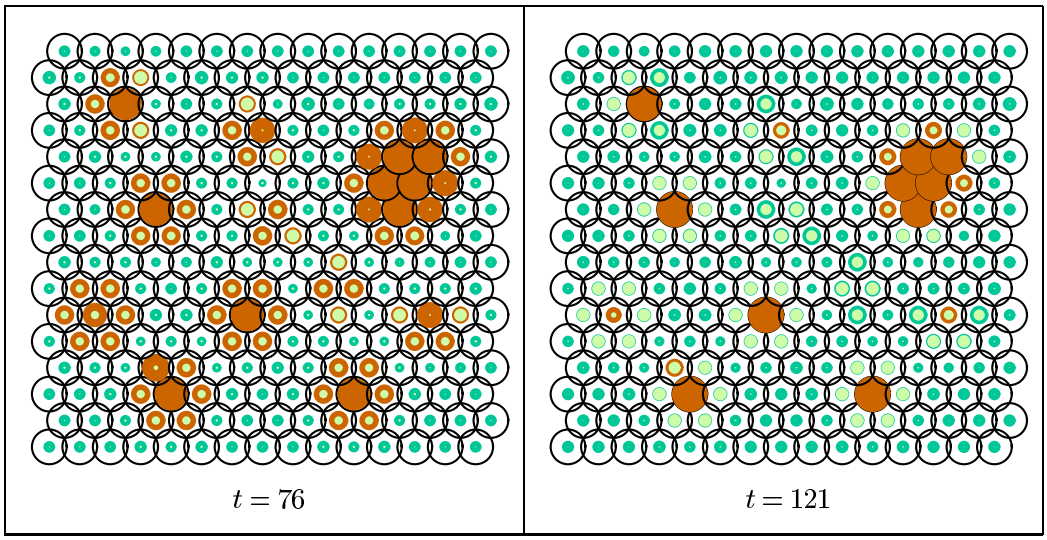

Figure 6: Simulation with perturbations of gene expression in individual cells of two symmetrical, 7-cell clusters. The clusters are the two ones in the lower left corner of the dataset. At $t = 60$ the level of epithelial expression in the central cell of the upper one of these two clusters is instantaneously increased, while in the lower cluster proneural expression is increased in a peripheral cell. Both perturbations can be detected in the left panel ($t = 76$). The first perturbation abolishes cluster resolution, while the second has no effect on resolution, as can be seen in the right panel.

since such processes occur repeatedly in metazoan development. We have been able to extract predictions about the dynamics of proneural cluster resolution and draw conclusions about mechanisms that may be sufficient or necessary for neuroblast and sensory organ precursor differentiation. We have also tested how robustly various genotypes lead to cluster resolution, under various perturbations.

The kinds of questions that can be posed with the model described above are not only of relevance to neurogenesis in *Drosophila* but are common to many developing organisms, especially in view of the fact that homologues to genes involved in *Drosophila* neurogenesis have been isolated in many species from worms to mammals, participating in a variety of developmental processes. In vertebrate neurogenesis such homologues act in ways similar to those of the *Drosophila* genes to regulate the number of neurons generated (see [16], [52]); one would therefore expect that a theoretical and empirical understanding of *Drosophila* neurogenesis would provide insights into neurogenesis in higher vertebrates, for instance, into questions surrounding neuronal proliferation in the developing mammalian cortex (see [15], [80]). It would be interesting to see an extension of the neurogenesis model presented here to mammalian neurogenesis, now that molecular data are becoming increasingly available in this area too.

Considering ways to move ahead with developmental modeling approaches more generally, one has to observe that models described in this review have been primarily concerned with the dynamics of development, i.e. how cells develop in space and time. Such models can, of course, be constrained by the large amounts of genomic data becoming available, but, still, they cannot incorporate all the available data without becoming intractable. A good portion of these biological data is of a qualitative nature, so cannot be readily mapped to precise interaction strengths between the components of a model. Other methods are needed to achieve a more comprehensive biological knowledge representation. Researchers have recently experimented with graphical models and pathway/model databases. Graphical models [27] attempt to systematically describe relationships (edges) between elements (nodes) in biological systems, bringing together data on mRNA expression, protein interactions, environmental conditions, etc.; these descriptions are probabilistic with probabilities conditioned on existing data, allow inferences from these data, and point to areas where more data are needed. Pathway/model databases [40] describe metabolic and gene-regulatory networks, enzymes and other proteins, and try to present a more global picture of many

interacting processes within an organism; they are based on an ontology – which is a database structure or schema that captures important features of the underlying system and precisely defines their relationships – and include theories about how the organism works that can be derived from the ontology. For such databases to extract knowledge from large datasets, methods for efficiently generating mathematical model , storing them in the databases, comparing them to each other, and validating them against existing data will also be needed. In connection with this last point, see, for instance, Cellerator [88], a program that allows users to specify a set of biochemical reactions, group them in a hierarchical graph structure that corresponds to the process modeled, translate them to ordinary differential equations and solve them.

Both graphical models and pathway/model databases appear to be attractive vehicles for storing the more detailed kind of dynamical models presented in this review, for comparing them to each other and to available data and for assessing how well they fit in the more global ontology of an organism.

References

[1] G.K. Ackers, A.D. Johnson, and M.A. Shea. Quantitative model for gene regulation by λ phage repressor. *Proceedings of the National Academy of Sciences USA*, 79:1129–1133, 1982.

[2] P. Agarwal. Cellular segregation and engulfment simulations using the cell programming language. *Journal of Theoretical Biology*, 176:79–89, 1995.

[3] M. Akam. Hox and HOM: Homologous gene clusters in insects and vertebrates. *Cell*, 57:347–349, 1989.

[4] S. Artavanis-Tsakonas, K. Matsuno, and M.E. Fortini. Notch signaling. *Science*, 268:225–232, 1995.

[5] K. Basler and G. Struhl. Compartment boundaries and the control of *Drosophila* limb pattern by Hedgehog protein. *Nature*, 368:208–214, 1994.

[6] M. Beato, P. Herrlich, and G. Schütz. Steroid hormone receptors: Many actors in search of a plot. *Cell*, 83:851–857, 1995.

[7] M.S. Boguski and F. McCormick. Proteins regulating ras and its relatives. *Nature*, 366:643–654, 1994.

[8] D. Bray. Intracellular signalling as a parallel distributed process. *Journal of Theoretical Biology*, 143:215–231, 1990.

[9] D. Bray. Protein molecules as computational elements in living cells. *Nature*, 376:307–312, 1995.

[10] D. Bray and R.B. Bourret. Computer analysis of the binding reactions leading to a transmembrane receptor-linked multiprotein complex involved in bacterial chemotaxis. *Molecular Biology of the Cell*, 6:1367–1380, 1995.

[11] D. Bray, R.B. Bourret, and M.I. Simon. Computer simulation of the phosphorylation cascade controlling bacterial chemotaxis. *Molecular Biology of the Cell*, 4:469–482, 1993.

[12] D. Bray and S. Lay. Computer simulated evolution of a network of cell-signaling molecules. *Biophysical Journal*, 66:972–977, 1994.

[13] S. Campuzano and J. Modollel. Patterning of the *Drosophila* nervous system - the *achaete-scute* gene complex. *Trends in Genetics*, 8:202–208, 1992.

[14] S.B. Carroll. Homeotic genes and the evolution of arthropods and chordates. *Nature*, 376:479–485, 1995.

[15] V.S. Caviness, T. Takahashi, and R.S. Nowakowski. Numbers, time and neocortical neuronogenesis: a general developmental and evolutionary model. *Trends in Neurosciences*, 18:379–383, 1995.

[16] A. Chitnis, D. Henrique, J. Lewis, D. Ish-Horowicz, and C. Kintner. Primary neurogenesis in *Xenopus* embryos regulated by a homologue of the *Drosophila* neurogenic gene *Delta*. *Nature*, 375:761–766, 1995.

[17] P. Cubas, J.-F. de Celis, S. Campuzano, and J. Modolell. Proneural clusters of *achaete-scute* expression and the generation of sensory organs in the *Drosophila* imaginal wing disc. *Genes and Development*, 5:996–1008, 1991.

[18] F.J. Diaz-Benjumea, B. Cohen, and S.M. Cohen. Cell interaction between compartments establishes the proximal-distal axis of *Drosophila* legs. *Nature*, 372:175–179, 1994.

[19] G. Dupont and A. Goldbeter. Oscillations and waves of cytosolic calcium: Insights from theoretical models. *Bioessays*, 14:485–493, 1992.

[20] Y. Echelard, D.J. Epstein, B. St-Jacques, L. Shen, J. Mohler, J.A. McMahon, and A.P. McMahon. Sonic hedgehog, a member of a family of putative signaling molecules, is implicated in the regulation of CNS and limb polarity. *Cell*, 75:1417–1430, 1993.

[21] M.J. Fietz, J.-P. Concordet, R. Barbosa, R. Johnson, S. Krauss, A.P. McMahon, C. Tabin, and P.W. Ingham. The *hedgehog* gene family in *Drosophila* and vertebrate development. *Development*, n.SUPPL:43–51, 1994.

[22] R.A. Fisher. *The Genetical Theory of Natural Selection*. The Clarendon Press, Oxford, 1930.

[23] K. Fleischer and A.H. Barr. A simulation testbed for the study of multicellular development: The multiple mechanisms of morphogenesis. In C.G. Langton, editor, *Artificial Life III : Proceedings of the Workshop on Artificial Life, held June 1992 in Santa Fe, New Mexico*. Addison-Wesley, Reading, MA, 1994.

[24] C.M. Fonseca and P.J. Fleming. An overview of evolutionary algorithms in multiobjective optimization. *Evolutionary Computation*, 3:1–16, 1995.

[25] Oster. G.F., J.D. Murray, and A.K. Harris. Mechanical aspects of mesenchymal morphogenesis. *Journal of Embryology and Experimental Morphology*, 78:83–125, 1978.

[26] A. Gierer and H. Meinhardt. A theory of biological pattern formation. *Kybernetik*, 12:30–39, 1972.

[27] D.K. Gifford. Blazing pathways through genetic mountains. *Science*, 293:2049–2051, 2001.

[28] B.C. Goodwin. Oscillatory behaviour in enzymatic control systems. *Advances in Enzyme Regulation*, 5:425–438, 1965.

[29] F. Graner. Can surface adhesion drive cell rearrangement? Part I: Biological cell-sorting. *Journal of Theoretical Biology*, 164:455–476, 1993.

[30] F. Graner and Y. Sawada. Can surface adhesion drive cell rearrangement? Part II: A geometrical model. *Journal of Theoretical Biology*, 164:477–506, 1993.

[31] S.P. Hastings, J.J. Tyson, and D. Webster. Existence of periodic solutions for negative feedback control systems. *Journal of Differential Equations*, 25:39–64, 1977.

[32] J.A. Hertz, R.G. Palmer, and A.S. Krogh. *Introduction to the Theory of Neural Computation*. Addison-Wesley, 1991.

[33] A. Hjelmfelt and J. Ross. Chemical implementation and thermodynamics of collective neural networks. *Proceedings of the National Academy of Sciences USA*, 89:388–391, 1992.

[34] A. Hjelmfelt, E.D. Weinberger, and J. Ross. Chemical implementation of neural networks and turing machines. *Proceedings of the National Academy of Sciences USA*, 88:10983–10987, 1991.

[35] A.L. Hodgkin and A.F. Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *Journal of Physiology (London)*, 117:500–544, 1952.

[36] J.J. Hopfield. Neurons with graded response have collective computational properties like those of two-state neurons. *Proceedings of the National Academy of Sciences USA*, 81:3088–3092, 1984.

[37] A. Hunding. Limit-cycles in enzyme systems with nonlinear negative feedback. *Biophysics of Structure and Mechanism*, 1:47–54, 1974.

[38] P.W. Ingham. The molecular genetics of embryonic pattern formation in *Drosophila*. *Nature*, 335:25–34, 1988.

[39] B. Jennings, A. Preiss, C. Delidakis, and S. Bray. The Notch signalling pathway is required for *Enhancer of split* bHLH protein expression during neurogenesis in the *Drosophila* embryo. *Development*, 120:3537–3548, 1994.

[40] P.D. Karp. Pathway databases: A case study in computational symbolic theories. *Science*, 293:2040–2044, 2001.

[41] P. Kastner, M. Mark, and P. Chambon. Nonsteroid nuclear receptors: What are genetic studies telling us about their role in real life? *Cell*, 83:859–869, 1995.

[42] S.A. Kauffman. Metabolic stability and epigenesis in randomly connected nets. *Journal of Theoretical Biology*, 22:437–467, 1969.

[43] S.A. Kauffman. Differentiation of malignant to benign cells. *Journal of Theoretical Biology*, 31:429–451, 1971.

[44] S.A. Kauffman. The large-scale structure and dynamics of gene control circuits: An ensemble approach. *Journal of Theoretical Biology*, 44:167–190, 1974.

[45] S.A. Kauffman. *The Origins of Order*. Oxford University Press, 1993.

[46] S.A. Kauffman and S. Levin. Towards a general theory of adaptive walks on rugged landscapes. *Journal of Theoretical Biology*, 128:11–45, 1987.

[47] C. Kenyon and B. Wang. A cluster of Antennapedia-class homeobox genes in a nonsegmented animal. *Science*, 253:516–517, 1991.

[48] M. Kerszberg and J. Changeux. A model for reading morphogenetic gradients: Autocatalysis and competition at the gene level. *Proceedings of the National Academy of Sciences USA*, 91:5823–5827, 1994.

[49] S. Krauss, J.-P. Concordet, and P.W. Ingham. A functionally conserved homolog of the *Drosophila* segment polarity gene *hedgehog* is expressed in tissues with polarizing activity in Zebrafish embryos. *Cell*, 75:1431–1444, 1993.

[50] R. Krumlauf. *Hox* genes in vertebrate development. *Cell*, 78:191–201, 1994.

[51] E.B. Lewis. A gene complex controling segmentation in *Drosophila*. *Nature*, 276:565–570, 1978.

[52] J. Lewis. Neurogenic genes and vertebrate neurogenesis. *Current Opinion in Neurobiology*, 6:3–10, 1996.

[53] A. Lindenmayer. Mathematical models for cellular interaction in development, parts I and II. *Journal of Theoretical Biology*, 18:280–315, 1968.

[54] A. Lindenmayer. Models for plant tissue development with cell division orientation regulated by preprophase bands of microtubules. *Differentiation*, 26:1–10, 1984.

[55] A. Lindenmayer and G. Rozenberg. Parallel generation of maps: Developmental systems for cell layers. In V. Claus, H. Ehrig, and G. Rozenberg, editors, *Graph Grammars and Their Application to Computer Science; First International Workshop*, volume 73 of *Lecture Notes in Computer Science*, pages 301–316. Springer, Berlin, 1979.

[56] A.J. Lotka. *Elements of Physical Biology*. Williams and Watkins, Baltimore, MD, 1925.

[57] D.J. Mangelsdorf and R.M. Evans. The RXR heterodimers and orphan receptors. *Cell*, 83:841–850, 1995.

[58] G. Marnellos. *Gene Network Models Applied to Questions in Development and Evolution*. PhD thesis, Yale University, 1997.

[59] G. Marnellos, G. A. Deblandre, E. Mjolsness, and C. Kintner. Delta-Notch lateral inhibitory patterning in the emergence of ciliated cells in Xenopus: Experimental observations and a gene-network model. In *Pacific Symposium on Biocomputing*, volume 5, pages 329–340, 2000.

[60] G. Marnellos and E. Mjolsness. A gene network approach to modeling early neurogenesis in Drosophila. In *Pacific Symposium on Biocomputing*, volume 3, pages 30–41, 1998.

[61] G. Marnellos and E. Mjolsness. Probing the dynamics of cell differentiation in a model of Drosophila neurogenesis. In *Artificial Life VI, Proceedings of the Sixth International Conference on Artificial Life*, volume 6, pages 161–170. MIT Press, 1998.

[62] H.H. McAdams and L. Shapiro. Circuit simulation of genetic networks. *Science*, 269:650–656, 1995.

[63] W. McGinnis, R.L. Garber, J. Wirz, A. Kuroiwa, and W. Gehring. A homologous protein-coding sequence in *Drosophila* homeotic genes and its conservation in other metazoans. *Cell*, 37:403–408, 1984.

[64] H. Meinhardt. Cell determination boundaries as organizing regions for secondary embryonic fields. *Developmental Biology*, 96:375–385, 1983.

[65] H. Meinhardt. Hierarchical inductions of cell states: A model for segmentation in *Drosophila*. *Journal of Cell Science (Supplement)*, 4:357–381, 1986.

[66] H. Meinhardt. A model for pattern generation on the shells of molluscs. *Journal of Theoretical Biology*, 126:63–89, 1987.

[67] E. Mjolsness, D.H. Sharp, and J. Reinitz. A connectionist model of development. *Journal of Theoretical Biology*, 152:429–453, 1991.

[68] J.D. Murray. On pattern formation mechanisms for lepidopteran wing patterns and mammalian coat markings. *Philosophical Transactions of the Royal Society, London (series B)*, 295:473–496, 1981.

[69] J.D. Murray. A pre-pattern formation mechanism for animal coat markings. *Journal of Theoretical Biology*, 88:161–199, 1981.

[70] J.D. Murray. *Mathematical Biology*. Springer-Verlag, Berlin, 2nd edition, 1993.

[71] J.D. Murray, G.F. Oster, and A.K. Harris. A mechanical model for mesenchymal morphogenesis. *Journal of Mathematical Biology*, 17:125–129, 1983.

[72] M.T. Murtha, J.F. Leckman, and F.H. Ruddle. Detection of homeobox genes in development and evolution. *Proceedings of the National Academy of Sciences USA*, 88:10711–10715, 1991.

[73] M.A.T. Muskavitch. Delta-Notch signalling and *Drosophila* cell fate choice. *Developmental Biology*, 166:415–430, 1994.

[74] A. Neiman. Conservation and reiteration of a kinase cascade. *Trends in Genetics*, 9:390–394, 1993.

[75] B. Novak and J.J. Tyson. Modeling the cell division cycle: M-phase trigger, oscillations, and size control. *Journal of Theoretical Biology*, 165:101–134, 1993.

[76] C. Nüßlein-Volhardt and E. Wieschaus. Mutations affecting segment number and polarity in *Drosophila*. *Nature*, 287:795–801, 1980.

[77] G.M. Odell, G. Oster, P. Alberch, and B. Burnside. The mechanical basis of morphogenesis. i. epithelial folding and invagination. *Developmental Biology*, 85:446–462, 1981.

[78] B.A. Pearlmutter. Learning state space trajectories in recurrent neural networks. *Neural Computation*, 1:263–269, 1989.

[79] P. Prusinkiewicz and A. Lindenmayer. *The Algorithmic Beauty of Plants*. Springer-Verlag, New Yrok, 1990.

[80] P. Rakic. A small step for the cell, a giant leap for mankind: a hypothesis of neocortical expansion during evolution. *Trends in Neurosciences*, 18:383–388, 1995.

[81] J. Reinitz, E. Mjolsness, and D.H. Sharp. Model for cooperative control of positional information in *Drosophila* by Bicoid and maternal Hunchback. *Journal of Experimental Zoology*, 271:47–56, 1995.

[82] J. Reinitz and D.H. Sharp. Mechanism of *eve* stripe formation. *Mechanisms of Development*, 49:133–158, 1995.

[83] R. Riddle, R.L. Johnson, E. Laufer, and C. Tabin. Sonic hedgehog mediates the polarizing activity of the ZPA. *Cell*, 75:1401–1416, 1993.

[84] M.A. Savageau. *Biochemical Systems Analysis*. Addison-Wesley, 1976.

[85] M. Schummer, I. Scheurlen, C. Schaller, and B. Galliot. HOM/HOX homeobox genes are present in Hydra (Chlorohydra viridissima) and are differently expressed during regeneration. *EMBO Journal*, 11:1815–1823, 1992.

[86] M.P. Scott and A.J. Weiner. Structural relationships among genes that control development: Sequence homology between the *Antennapedia*, *Ultrabithorax* and $fushi - tarazu$ loci in *Drosophila*. *Proceedings of the National Academy of Sciences USA*, 81:4115–4119, 1984.

[87] L.A. Segel. *Modeling Dynamic Phenomena in Molecular and Cellular Biology*. Cambridge University Press, 1984.

[88] B. Shapiro and E. Mjolsness. Developmental simulations with Cellerator. In *Second International Conference on Systems Biology (ICSB)*, 2001.

[89] M.A. Shea and G.K. Ackers. The O_R control system of bacteriophage lambda: A physical-chemical model for gene regulation. *Journal of Molecular Biology*, 181:211–230, 1985.

[90] M. Simpson-Brose, J. Treisman, and C. Desplan. Synergy between the Hunchback and Bicoid morphogens is required for anterior patterning in *Drosophila*. *Cell*, 78:855–865, 1994.

[91] J.B. Skeath and S.B. Carroll. Regulation of proneural gene expression and cell fate during neuroblast segregation in the *Drosophila* embryo. *Development*, 114:939–946, 1992.

[92] R. Somogyi and C.A. Sniegoski. Modeling the complexity of genetic networks: Understanding multigenic and pleiotropic regulation. *Complexity*, 1:45–63, 1996.

[93] D. St Johnston and C. Nüßlein-Volhardt. The origin of pattern and polarity in the *Drosophila* embryo. *Cell*, 68:201–219, 1992.

[94] D. Stekel, J. Rashbass, and E.D. Williams. A computer graphic simulation of squamous epithelium. *Journal of Theoretical Biology*, 175:283–293, 1995.

[95] D. Sulsky, S. Childress, and J. Percus. A model of cell sorting. *Journal of Theoretical Biology*, 106:275–301, 1984.

[96] R. Thom. *Stabilité Structurelle et Morphogenèse: Essai d'une Théorie Générale des Modèles*. W.A. Benjamin, Reading, Mass., 1972.

[97] C. Thummel. From embryogenesis to metamorphosis: The regulation and function of *Drosophila* nuclear receptor superfamily members. *Cell*, 83:871–877, 1995.

[98] A.M. Turing. The chemical basis of morphogenesis. *Philosophical Transactions of the Royal Society, London (series B)*, 237:37–72, 1952.

[99] J.J. Tyson and H.G. Othmer. The dynamics of feedback control circuits in biochemical pathways. *Progress in Theoretical Biology*, 5:1–62, 1978.

[100] V. Volterra. Variazioni e fluttuazioni del numero d'individui in specie animali conviventi. *Memorie (Accademia Nazionale dei Lincei; Classe di Scienze Fisiche, Matematiche e Naturali)*, ser. 6, vol. 2:31–113, 1926.

[101] G. von Dassow, E. Meir, E.M. Munro, and G.M. Odell. The segment polarity network is a robust developmental module. *Nature*, 406:188–192, 2000.

[102] M. Weliky and G. Oster. The mechanical basis of cell rearrangement: I. Epitelial morphogenesis during *Fundulus* epiboly. *Development*, 109:373–386, 1990.

[103] R.J. Williams and D. Zipser. A learning algorithm for continually running fully recurrent neural networks. *Neural Computation*, 1:270–280, 1989.