



## Gene Networks Capable of Pattern Formation: From Induction to Reaction–Diffusion

ISAAC SALAZAR-CIUDAD\*†‡, JORDI GARCIA-FERNÁNDEZ† AND RICARD V. SOLÉ\*

*\*Complex Systems Research Group, Department of Physics, FEN-UPC Campus Nord, Mòdul B5 08034 Barcelona, Spain and †Facultat de Biologia, Departament de Genètica, Universitat de Barcelona, Diagonal 645, 08028, Spain*

*(Received on 29 October 1999, Accepted in revised form on 2 May 2000)*

One of the main aims of developmental biology is to understand how a single and apparently homogeneous egg cell achieves the intricate complexity of the adult. Here we present two models to explain the generation of developmental patterns through interactions at the gene level. One model considers direct-contact induction between cells while the other takes into account diffusion of hormones. We show that sets of cells involving identical gene networks and communicating through hormones spontaneously exhibit ordered patterns. We have characterized these patterns and the specific networks responsible for them. The models allow to (i) compare diffusion and direct-contact induction processes as mechanisms of pattern generation; (ii) identify the possible range of behaviour of real gene networks and (iii) suggest causal mechanisms to generate known patterns. The evolutionary implications are discussed.

© 2000 Academic Press

### 1. Introduction

Pattern formation is one of the most interesting and fundamental problems in development. Particularly, the crucial question of developmental biology is unveiling how a single, apparently homogeneous zygote cell becomes a multicellular organism. Molecular genetics allows us to explore the nature of the minimal functional and heritable elements but a large number of questions concerning the dynamics of pattern formation remain open. Theoretical studies focus on the search for plausible mechanisms that may allow an increase in the organization of develop-

mental systems compatible with experimental data. Real epigenetic systems are only a subset of possible systems and we wonder whether the observed regulatory networks, and the patterns that they generate are the result of historical accidents, natural selection, dynamic constraints or a mixture of all these factors. Although theoretical embryology is still a young discipline, some ideas have already taken on mathematical form. As Slack said: “a few years ago such models were often dismissed as mere speculation, but now no serious student of the subject can afford to ignore them” (Slack, 1991).

The knowledge of what kind of patterns can be generated by systems involving a particular gene network architecture and cell-to-cell interactions can provide very useful information, especially if the number of possible patterns is small. Most previous models in development are described

‡Author to whom correspondence should be addressed: Complex Systems Research Group, Department of Physics, FEN-UPC Campus Nord, Mòdul B5 08034 Barcelona, Spain.

E-mail: isaac@complex.upc.es

by means of a specific reaction–diffusion or mechano-chemical system (Dumais & Harrison, 2000; Owen *et al.*, 1999; Collier *et al.*, 1996; Hunding & Engelhardt, 1990; Oster *et al.*, 1989; Meinhardt, 1982; Meinhardt & Klingler, 1987; Goodwin & Trainor, 1985; Glass & Kauffman, 1973). In this paper we present a statistical analysis of the pattern formation capacity of groups of cells involving genetic networks and to this goal we study a large ensemble of networks. What we intend is to find which patterns are achievable with intercellular gene networks modelled by means of a simple and biologically reasonable set of interactions between genes in the same or different cells. A clear criticism of this ensemble approach would be that the real genome is too complex to be properly described by this model. But by using this type of approach, we hope to be able to detect some fundamental commonalities at some coarse-grained level. This type of approach has been used by several authors in relation with the problem of cell differentiation (Kauffman, 1993; Glass & Kauffman, 1972; Lewis & Glass, 1991; Solé & Luque, 1995), homeostasis and redundancy in gene networks (Wagner, 1999) and the effect of gene duplications (Wagner, 1994). Previous theoretical studies have also explored the dynamics generated by sets of cells each one involving the same gene network. Two basic approaches have been used, from standard Boolean network models (Jackson *et al.*, 1986) to continuous models close to neural-like connectionist architectures (Mjolsness *et al.*, 1991) or biochemical networks of interacting chemicals (Kaneko & Yomo, 1994, 1997).

These models allow analysis of some aspects of early developmental processes. In the Jackson–Johnson–Nash (JJN) model, small numbers of cells arranged in a one-dimensional growing array are able to communicate with nearest cells through diffusion of some gene products. But in fact interactions are described not in terms of continuous concentrations but as Boolean interactions. The JJN model is completed by means of a schematic cell cycle which governs the conditions under which the cell will divide or cease division. Eventually, the system settles down into a final state; in other words, the tissue falls into an attractor (i.e. some recurrent pattern of gene expression). One of the interesting results of this

study was that cell–cell interactions allowed generation of new cell types and such a capacity was shown to be maximal when about 20% of the genes were involved in such interactions. Beyond the specific (inevitable) simplifications introduced by the JJN model, some general and probably robust consequences are obtained. Most fundamentally, that induction of novel cell types and the establishment of spatial heterogeneity are “virtually inherent in almost any genomic system” (Kauffman, 1993).

In the same context, but using continuous equations, the connectionist model developed by Mjolsness, Sharp and Reinitz (MSR) provides a systematic method for discovering correlations in experimental data on gene expression in developmental processes (Mjolsness *et al.*, 1991). This is in fact a generalized reaction–diffusion model with sigmoidal saturation. The basic model is defined by a set of differential equations

$$\frac{dg_{ij}}{dt} = -\lambda_j g_{ij} + \Phi \left[ \sum_{l=1}^{N_g} W^{jl} g_{il} + \theta_j \right] + D_j \nabla^2 g_{ij}, \quad (1)$$

where (a) the concentration of the  $j$ -th gene product in the  $i$ -th cell is indicated by  $g_{ij}$ ; (b) interactions among genes are described by means of a random matrix  $W^{jl}$  with both positive and negative values; (c) thresholds in the gene response  $\theta_j$  are considered and (d) diffusion between nearest cells takes place at a rate  $D_j$  and is described by means of the last term (the Laplacian) in the previous equation. The function defining the interactions among genes,  $\Phi(z)$ , is assumed to be a sigmoidal one. Specifically, the MSR model uses  $\Phi(z) = 1/(1 + \exp(-\mu z))$  but other choices lead to similar results. This model is a very general approach to the real behaviour of gene networks in development. Using this model and searching over parameter space by means of an appropriate computational optimization process (simulated annealing), these authors have been able to reproduce the gap gene expression patterns along the antero–posterior axis of *Drosophila* (Reinitz *et al.*, 1992). It is worth mentioning that the MSR model is very close to previous models of neural networks. The statistical analysis of such networks has clearly shown

that many relevant features of these systems may not be dependent on many specific details (Domany *et al.*, 1995).

The previous models thus show that some general, simple approximations to real gene networks can provide some insight into both generic properties of development (as the generation of novel cell types due to cell-cell interactions) and in specific pattern-forming networks of real systems (such as in *Drosophila*). An interesting problem, which we explore in the present paper, is the possible existence of a limited number of basic gene network topologies leading to the observed range of spatial patterns in living organisms. In this context, previous studies based on reaction-diffusion and related models have clearly identified a number of basic mechanisms leading to pattern formation. Concepts such as bistability, symmetry breaking, phase separation, lateral inhibition, Turing instabilities or morphogen gradients are very common to most of the available literature on theoretical embryology. When looking at the best-known theoretical models, we can see a limited range of basic dynamical processes. Additionally, the number of possible patterns attainable from a given mechanism is clearly constrained. The relevance of such constraints to the possible repertoire of logical types has important consequences for evolution (Conway Morris, 1998; Goodwin, 1994).

That different models with apparently different rules but the same basic qualitative formal structure can generate a similar set of patterns is well known in physics (Nicolis & Prigogine, 1977; Mikhailov, 1990; Mikhailov & Loskutov, 1991). To show just an example, let us consider the MSR model with only two gene products (morphogens) and the problem of pattern formation due to phase separation. As discussed by Newman, phase separation is an important underlying mechanism involved in the establishment of distinct spatial domains in some morphoetic processes (Newman & Comper, 1990; Newman, 1994). Phase separation in a spatially extended system is easy to model and it typically involves the competition between two types of components. Not surprisingly, such a process has been used in the study of spatial segregation of two competing species (Solé *et al.*, 1992). Let us

consider this process, as described by the previous set of equations, assuming  $N_g = 2$ :

$$\begin{aligned} \frac{dg_{i1}}{dt} = & -\lambda_1 g_{i1} + \Phi[W^{11}g_{i1} + W^{12}g_{i2} + \theta_1] \\ & + D_1 \nabla^2 g_{i1}, \end{aligned} \tag{2}$$

$$\begin{aligned} \frac{dg_{i2}}{dt} = & -\lambda_2 g_{i2} + \Phi[W^{21}g_{i1} + W^{22}g_{i2} + \theta_2] \\ & + D_2 \nabla^2 g_{i2}, \end{aligned} \tag{3}$$

where we restrict ourselves to the one-dimensional case. In the MRS model, we have a sigmoidal function  $\Phi(x)$  which saturates at high and low values of the argument  $x$  (taking one and zero values at these two limits, respectively). But there is no reason to use this specific form for the reaction term. In this paper we will consider a different function, following a Michaelis-Menten (MM) behaviour. Explicitly, we have used  $\Phi_\delta(x) = \delta x/(1+x)$  if  $x \geq 0$  and zero otherwise. In other words, if only inhibition is at work, no changes (nor negative values) in  $x$  take place. We need to define a matrix of interaction values  $\Omega = (W_{ik})$  and here we use

$$\Omega = \begin{pmatrix} -1 & -1 \\ -1 & 1 \end{pmatrix} \tag{4}$$

for the MRS model and we use

$$\Omega = \begin{pmatrix} 1 & -1/2 \\ -1/2 & 1 \end{pmatrix} \tag{5}$$

for the MM model. These choices are inspired in ecological competition processes (Solé *et al.*, 1992) and basically introduce a symmetric positive feedback and a negative cross-reaction that leads to mutual inhibition. In the MM case, we avoid some problems derived from the sigmoidal function used by MRS. A single cell with no inputs into the  $j$ -th gene product will show a dynamics for this gene given by  $dg_j/dt = -\lambda_j g_j + 1/2$ , which means that a non-zero level of activity is present:  $g_j^* = 1/2\lambda_j$ . In the MM

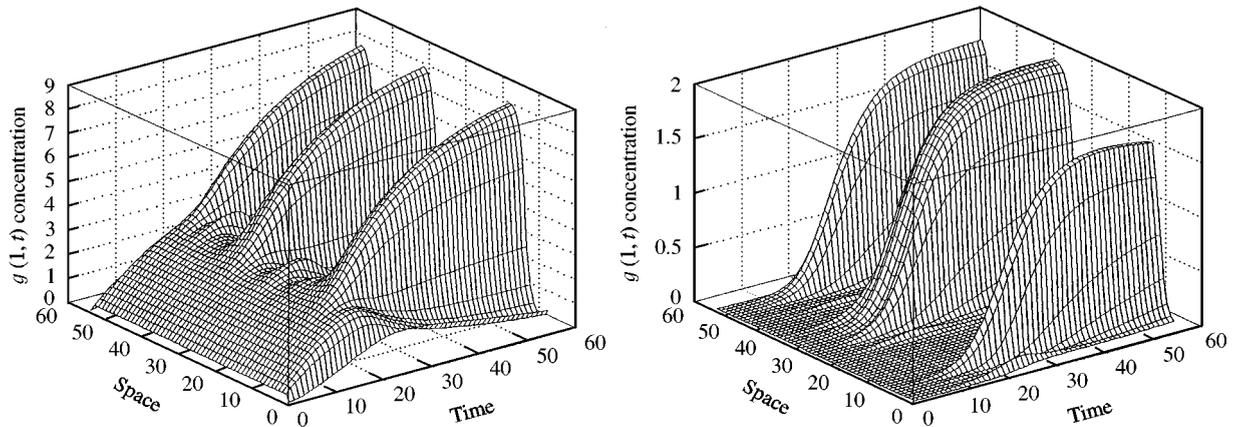


FIG. 1. Spatial structures emerging from phase-separation processes in two different models involving two morphogens. Here the (a) Mjølness–Reinitz–Sharp and (b) the Michaelis–Menten models (see text) have been used on a one-dimensional spatial domain defined by  $N_s = 55$  cells with nearest-neighbor communication and zero-flux boundary conditions. The spatiotemporal dynamics of the concentration for  $g_1$  is shown (see text for parameters). Stable structures are formed with an average number of peaks, strongly dependent on initial conditions.

approximation this problem does not arise (although the results obtained using the MRS model are basically the same).

In Fig. 1 we show the dynamics of these models for the  $10^{-3}$  previous matrices and  $\lambda_j = 0.1$ ,  $D_j = 0.2$ ,  $\mu_j = 2.5$  and  $\theta_j = 0$  for the MRS model and  $\lambda_j = 0.1$ ,  $\mu_j = 3$ ,  $D_j = 0.2$  and  $\theta_j = 0$  for the MM model, respectively. The system has zero-flux boundary conditions and the initial condition is  $g_{ij}(0) = 10^{-2} + \eta(i)$ . So each cell  $i$  receives a small random value  $\eta(i) \in \{-10^{-3}, +10^{-3}\}$ . The patterns formed are slightly different since the specific form of each reaction term is different. In the MRS model there is a transient increase of the concentration of both morphos before the competition process fully operates. Such a phenomenon is not present in the MM counterpart. In both cases, the final pattern is strongly dependent on initial conditions, displays an average number of final peaks and each morpho is depleted in those regions where the other morpho is more abundant.

Similar results can be obtained by looking at other pattern-forming mechanisms (such as simple gradient-forming networks). Both models are flexible enough to obtain the range of patterns found in other reaction–diffusion models. But of course there are many potential candidates of gene network interactions leading to stable spatial heterogeneities (although unstable

dynamics and chaos are also possible as shown in our study). And it is possible that some new, different pattern formation processes would be found by playing with the parameters and network topologies. In particular, we might formulate a number of questions concerning the pattern-forming capacity of these models:

1. Among all the possible networks, are pattern-forming networks common?
2. Are there a very large number of possible types of spatial patterns?
3. Are there differences between the patterns obtained from models with diffusion (i.e. long-range interactions) and those generated from cell-to-cell inductive interactions?
4. Are the networks responsible for such patterns organized in some special sets sharing similar characteristics?
5. Are the pattern-forming networks related with real, known gene networks involved in developmental processes?
6. Can we infer some aspects about the organization of the whole developmental process and about their evolution from the answers to the previous questions?

In this paper these questions are explicitly explored by means of an extensive study of randomly generated networks of gene interactions

with both diffusion and direct inductive, cell-to-cell interactions. Although we only show the results obtained from the MM approach, the same basic conclusions are obtained by using the MRS model. This agreement suggests that the results presented here are robust and not dependent on the specific features of the model.

### 2. Theoretical Model

We constructed a large ensemble of networks, in which genes were connected randomly: in this study we have analysed the behaviour of  $10^6$  different networks. In order to address some specific questions we have studied more accurately a smaller number of networks. The model considers  $N_c$  cells arranged in a line or  $N_c \times N_c$  cells on a square lattice. In each cell, the relationship between the  $N_g$  genes are the same (because they are supposed to be coded in the genome). Genes in these networks interact according to a small set of simple rules that simulate the more relevant mechanisms by which real genes regulate each other (see Fig. 2).

The numeric values for gene activity levels in the model are indicated as  $g_{ij}$  where  $i = 1, \dots, N_c$  indicates cell number (ordered in space) and  $j = 1, \dots, N_g$  stands for gene number. These values represent gene product concentrations inside cells or the concentration of active protein forms.  $N_h$  of the genes code for hormones (meaning molecules that affect gene expression in other cells),  $N_r$  genes code for cellular or nuclear

receptors of the hormones and the rest are transcription factors or proteins involved in signal transduction inside cells. For most simulations, we set  $N_h = N_g/3$ ,  $N_r = N_g/3$ . The genes that are not hormones nor receptors can also affect the global dynamic of the systems because they can produce various intracellular behaviours (for example periodic oscillations). *A priori*, thus, the model cannot be reduced to include only hormones and receptors. It does not preclude that some networks may be formed only by hormones and receptors. Gene products interact by binding (or by other potentially activating changes such as phosphorylation, proteolytic cleavage, etc.) to other gene products or by binding to *cis*-regulatory sequences on promoters. The model assumes that the change in the activity or chance of transcription induced by interaction follows a saturating function of Hill shape (a class of functions widely used for binding processes at the molecular level (Cornish-Bowden, 1979). However, our results, as will be discussed, also hold for most growing and saturating functions. Enhancers and protein-binding sites are characterized by a value  $W_{jk}$  that weights both the affinity of the transcriptional factor for the enhancer and the intensity of the response produced by the binding. The model considers two possibilities: the hormone diffuses (diffusive model; DM) or the hormone is membrane-bound (direct contact induction model; DCIM) and then can only affect the immediate neighbors. Hormone values in DM represent the concentration of the hormone in the extracellular matrix surrounding the cell (the DM assumes that the hormone is constantly secreted while being transcribed). It needs to be noted that DCIM is not equivalent to a DM even if hormones diffuse so poorly that they only allow communication between immediate neighbours. One of the main differences is that in DM the diffusing hormone always affects the cell producing it. This difference causes, as we will see, that many patterns that can be attained by DCIM cannot be attained by DM. In the direct contact induction model,  $W_{hi}^*$  is the same as  $W_{jk}$  but for the binding of hormones and receptors.  $W_{jk}$  and  $W_{hi}^*$  are random,  $W_{jk} \in [-1, +1]$  and  $W_{hi}^* \in [0, 1]$ . Between this range any value is possible, this allows a more fine-grained exploration of the ensemble. Inside such a range any value is

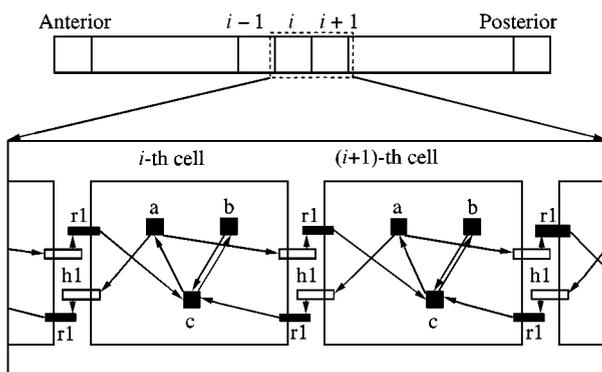


FIG. 2. Diagram showing the interactions involved in the DCIM model. Large squares are cells. Small squares inside cells are gene products. Arrows indicate gene product interactions.

possible. We performed the study using uniform and normal distributions. These random-generated networks are modelled by a dynamic system obeying the following set of equations:

(a) Reaction–diffusion model:

*Non-hormone gene products*

$$\frac{\partial g_{ij}}{\partial t} = \frac{\Phi[\sum_{k=1}^{N_g} W_{jk} g_{ik}]}{K_m + \Phi[\sum_{k=1}^{N_g} W_{jk} g_{ik}]} - \mu g_{ij}, \quad (6)$$

where  $i = 1, \dots, N_c$  and  $j = N_h + 1, \dots, N_g$ .

*Hormones*

$$\frac{\partial g_{il}}{\partial t} = \frac{\Phi[\sum_{k=1}^{N_g} W_{lk} g_{ik}]}{K_m + \Phi[\sum_{k=1}^{N_g} W_{lk} g_{ik}]} - \mu g_{il} + D_l \nabla^2 g_{il}, \quad (7)$$

where  $i = 1, \dots, N_c$  and  $l = 1, \dots, N_h$ .

(b) Direct contact induction model:

*Non-receptor gene products*

$$\frac{\partial g_{ij}}{\partial t} = \frac{\Phi[\sum_{k=1}^{N_g} W_{jk} g_{ik}]}{K_m + \Phi[\sum_{k=1}^{N_g} W_{jk} g_{ik}]} - \mu g_{ij}, \quad (8)$$

where  $i = 1, \dots, N_c$  and  $j = N_h + 1, \dots, N_g - N_r$ .

*Receptors*

$$\frac{\partial g_{il}}{\partial t} = \frac{\Phi[\sum_{k=1}^{N_g} W_{jk} g_{ik} + \sum_{h=1}^{N_h} W_{hl}^*(g_{i-1,l} + g_{i+1,l})]}{K_m + \Phi[\sum_{k=1}^{N_g} W_{jk} g_{ik} + \sum_{h=1}^{N_h} W_{hl}^*(g_{i-1,l} + g_{i+1,l})]} - \mu g_{il}, \quad (9)$$

where  $i = 1, \dots, N_c$  and  $l = N_g - N_r + 1, \dots, N_g$ .

$\Phi$  is the Heaviside function, i.e.  $\Phi(x) = x$  for  $\forall x > 0$  and  $\Phi(x) = 0$  otherwise. It is introduced to ensure that inhibiting interactions does not lead to active gene product degradation.  $\mu$  is the intrinsic rate of degradation affecting all gene products.  $K_m$  is the binding rate constant and  $D_l \in [0, 1]$  is the diffusion rate constant of the  $i$ -th hormone and is randomly generated. The operator  $\nabla^2$  is implemented in our discrete space

model as  $(g_{i+1,j} + g_{i-1,j} - 2g_{i+1,j})$  for the one-dimensional case and as  $(g_{i+1,m,j} + g_{i-1,m,j} + g_{i,m+1,j} + g_{i,m-1,j} - 4g_{i+1,j})$  for the two-dimensional system (where  $m = 1, \dots, N_c$ ), the distance between neighbour cells is thus a spatial unit. Zero-flux boundary conditions and periodic boundary conditions are used in all cases. The equations can be transformed by using  $\delta = 1/(K_m \mu)$ . The equations then become

(a) Reaction–diffusion model:

*Non-hormone gene products*

$$\frac{\partial g_{ij}}{\partial t} = \delta \frac{\Phi[\sum_{k=1}^{N_g} W_{jk} g_{ik}]}{1 + \Phi[\sum_{k=1}^{N_g} W_{jk} g_{ik}]} - g_{ij}. \quad (10)$$

*Hormones*

$$\frac{\partial g_{il}}{\partial t} = \frac{\Phi[\sum_{k=1}^{N_g} W_{lk} g_{ik}]}{K_m + \Phi[\sum_{k=1}^{N_g} W_{lk} g_{ik}]} - \mu g_{il} + D_l \nabla^2 g_{il}. \quad (11)$$

(b) Direct contact induction model:

*Non-receptor gene products*

$$\frac{\partial g_{ij}}{\partial t} = \delta \frac{\Phi[\sum_{k=1}^{N_g} W_{jk} g_{ik}]}{1 + \Phi[\sum_{k=1}^{N_g} W_{jk} g_{ik}]} - g_{ij}. \quad (12)$$

*Receptors*

$$\frac{\partial g_{il}}{\partial t} = \delta \frac{\Phi[\sum_{k=1}^{N_g} W_{jk} g_{ik} + \sum_{h=1}^{N_h} W_{hl}^*(g_{i-1,l} + g_{i+1,l})]}{1 + \Phi[\sum_{k=1}^{N_g} W_{jk} g_{ik} + \sum_{h=1}^{N_h} W_{hl}^*(g_{i-1,l} + g_{i+1,l})]} - g_{il}. \quad (13)$$

Although the model does not include all the subtleties of real systems, it keeps in a reasonable way a set of rules of interaction that are widespread among many of the gene products involved in intercellular communication. Some interesting properties of our equations need to be mentioned. First, most gene products have different binding rate constants and degradation rates ( $K_m$  and  $\mu$ ). However, we have found that using different  $\delta$  values for different gene products does

not change our basic conclusions. Although the same  $\delta$  value is used in all equations, the matrices  $(W_{jk})$  and  $(W_{hl}^*)$  effectively modify the maximum value that can be attained by each gene.

We use random initial conditions: a randomly chosen gene in a randomly chosen cell takes a small random value between 0 and 1, all other genes in this chosen cell and in the rest of the cells of the systems are set to zero. Uniform initial conditions in which a gene takes identical values in all the cells of the system are also used. These initial conditions could be interpreted either as a random fluctuation or a developmental cue induced from another part of the embryo. More complex initial conditions have been explored qualitatively.

An exhaustive search for patterns in the ensemble was performed. We consider that there is a pattern if different cells express the same gene at different levels, then we call this gene the pattern gene (of course there can be more than one pattern gene). To define two patterns as different we use the following rule. For each cell  $i$ , we compare the value of the pattern gene in such a cell with that in the rest of the cells in the array. We define a variable  $\phi_j(i)$  for each cell  $i$  that is equal to the number of cells in the system  $j$  that have a value of the pattern gene equal to or lower than that in cell  $i$ . Two patterns are considered to be equal if cells in analogous positions have the same values of  $\phi$  (that is  $\phi_1(i) = \phi_2(i) \forall i \in N_c$ ). For example, in a three-cell system with the pattern gene taking the values  $g_1 = 0.1$ ,  $g_2 = 0.9$ ,  $g_3 = 0.1$  the pattern will be characterized by  $\phi(1) = 2$ ,  $\phi(2) = 3$ ,  $\phi(3) = 2$ . For each generated network the behaviour over time was simulated until a steady stationary behaviour was found or until  $10^5$  time steps were attained. Equations were numerically integrated by means of the Euler method, with  $\delta t = 0.001$ .

For  $10^6$  networks we just explored the spectrum of possible patterns. For 8000 such gene networks we performed a structural study. Once a pattern was found we analysed the network that generates it to identify the minimal gene network responsible for the pattern. We used an algorithm that eliminates interactions in the original gene networks that do not change the pattern. With this method a large number of patterns were analysed and the gene networks

generating them were compared. Only gene networks with less than 20 genes were used in this analysis.

Some biologically relevant properties of each pattern have been studied (and are summarised in Table 3). For 4000 gene networks we studied what happens to the generated pattern: (i) when different initial conditions are brought to the same set of networks; (ii) when the same networks are modelled in systems of different number of cells; (iii) when perturbations in gene values (noise) are introduced once the pattern is found; and (iv) when zero-flux or periodic boundary conditions are used.

### 3. Results

Although networks were constructed randomly many of them lead to stable heterogeneous, ordered patterns. In this sense, we can give an affirmative answer to the first question formulated in the introduction. Thus, pattern formation seems to be frequent in our gene networks for both the diffusion and the induction mechanisms (see Tables 1 and 2, for the one-dimensional case). A rich, but limited number of basic patterns are found, from single stripes to gradients and spots, but also many combinations of these well-known basic motifs (see Fig. 3 for examples of simple patterns and Fig. 4 for examples of combined patterns). Some of the basic types found in our study are shown in Fig. 3 (see below). There are a large number of possible patterns. Not all are shown. A first relevant observation from the extensive statistical analysis of 8000 networks of one-dimensional systems is that inductive interactions lead to a larger number of temporally stable patterns than diffusive models. This is shown in Tables 1 and 2: as the number of genes  $N_g$  increases (and thus the dimensionality of our dynamical system and parameter space) both the DM and the DCIM models show a decrease in the number of spatially homogeneous patterns. However, although both display an increase in the number of unstable (i.e. time-dependent) structures such an increase is rapid in the DM case (more than 60% of the patterns are unstable) but remains small in the DCIM model (around 11% of cases). These observations respond to the third question: although both

TABLE 1

Probabilities of finding the different patterns in relation with the number of genes  $N_g$  (here  $N_c = 4, 10, 20, 30$ ) in DCIM. In each case  $N_h = N_g/2$  and  $N_r = N_g/2$ . For each  $N_g$  value 1000 networks were constructed and the frequencies of the different patterns were computed. We used a uniform distribution for the matrices  $(W_{jk})$  and  $(W_{hi}^*)$ . Here: SS = single stripe; P = plateau

$N_g$	Homogeneous	SS + P	Chessboard	Periodic	Stripes	Chaotic
4	0.9101	0.0589	0.0260	0.0050	0.0000	0.0000
10	0.9809	0.0364	0.1040	0.0087	0.0243	0.0104
20	0.8151	0.0239	0.0895	0.0199	0.0110	0.0378
30	0.5973	0.0133	0.1770	0.0398	0.0531	0.1195

TABLE 2  
As in Table 1 but for DM

$N_g$	Homogeneous	One stripe	Stripes	Chaotic
4	0.9460	0.0060	0.0000	0.0480
10	0.7276	0.0636	0.0199	0.1889
20	0.4931	0.0963	0.0393	0.3713
30	0.2683	0.0488	0.0610	0.6220

pattern-forming mechanisms can generate similar types of structures (see below) the inductive mechanism is much more robust and less parameter-dependent, thus allowing to generate stable patterns over a wide domain of parameter space.

The next step in our study is to identify the underlying gene network responsible for the observed patterns generated by means of both mechanisms. We call *module* a subset of the gene network (a sub-network) characterized by a particular topology and set of signs of gene-gene interactions (then we refer to a module as an abstraction of the network structure or topology; a directed graph) (see Table 1). A large number of different modules can generate the patterns found. A wide exploration of the model allows us to find the simplest network module associated with each spatial pattern (Table 1). Although the number of patterns is large most networks are due to the combination of networks that include few simple modules. Thus, most patterns are in fact combinations of a small number of basic patterns. We also found that all the modules that produce the same pattern share interesting

structural properties with the minimal module that generates this concrete pattern. Some modules are equal to minimal modules in which the intracellular interaction between two genes is mediated indirectly through an additional gene (that do not interact with other genes) instead of directly as in the minimal module. We refer to such modules as *topologically equivalent*. All the networks causing a given pattern include a sub-network that is topologically equivalent to the minimal module of such a pattern. Thus, we have identified a close relationship between all the patterns found and the family of gene networks producing each of them. When the pattern is generated by the combination of simpler patterns the network includes the minimal modules of each of the forming patterns. These conditions are necessary but not sufficient: the possession of a minimal module does not imply that the gene network may produce the pattern. In some cases the same pattern can be generated by a network including a minimal module (or a topologically equivalent module) or by the combination of networks derived from various minimal models.

All the modules found can be classified into four basic types depending on some structural properties of their connections. One result of our model is that the four biological properties of the patterns that we have studied turn out to be the same within the networks of a given type.

The basic modules corresponding to each pattern, together with their dependence with respect to perturbations and size effects are summarized

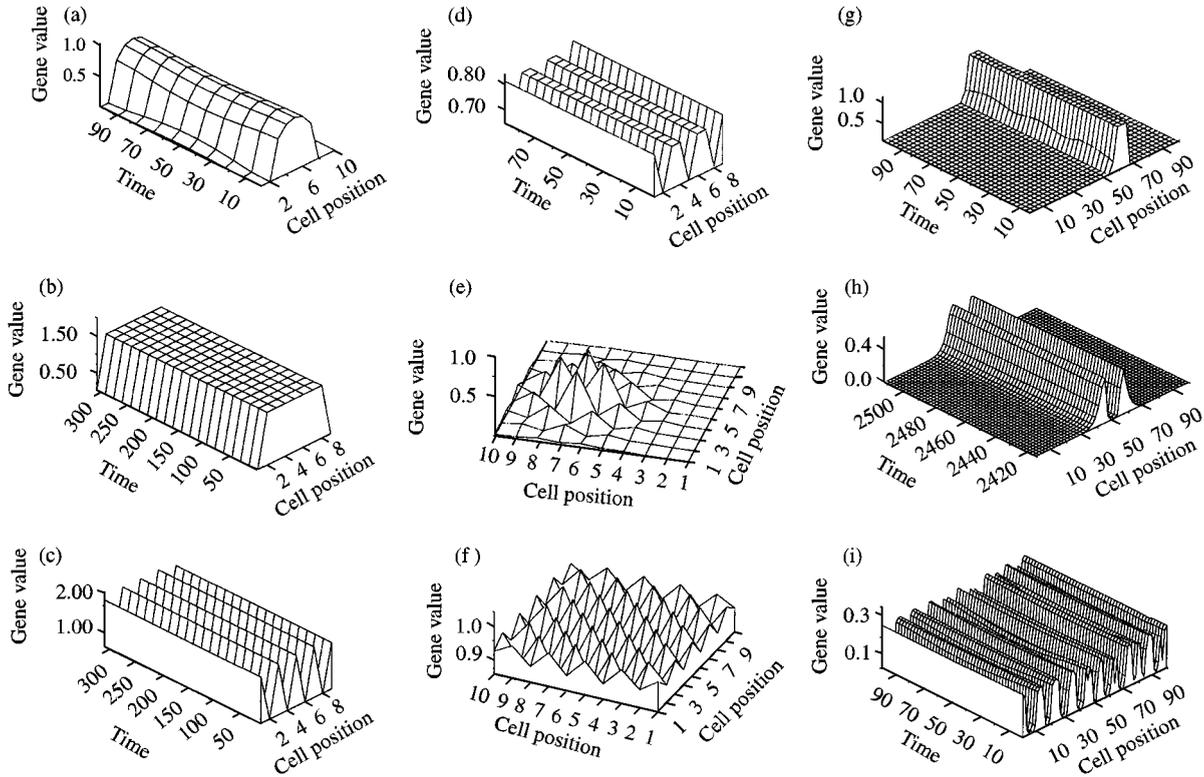


FIG. 3. Some examples of one-dimensional [(a)–(d) and (g)–(i)] and two-dimensional patterns found in the model [eqns (4a–d)] for periodic boundary conditions [except in (f)]. Examples (a)–(f) belong to the direct contact induction model (DCIM) while g, h and i belong to the diffusive model (DM). The patterns shown are (a) single stripe:  $\delta = 10$ ,  $W_{14} = 0.60$ ,  $W_{15} = -0.75$ ,  $W_{23} = 0.29$ ,  $W_{33} = 0.94$ ,  $W_{55} = 0.78$ ,  $W_{31}^* = 0.91$ ,  $W_{42}^* = 0.19$ ,  $W_{52}^* = 0.99$ ; (b) plateau:  $\delta = 10$ ,  $W_{14} = 0.60$ ,  $W_{15} = -0.75$ ,  $W_{23} = 0.29$ ,  $W_{55} = 0.78$ ;  $W_{31}^* = 0.91$ ,  $W_{42}^* = 0.19$ ,  $W_{52}^* = 0.99$ ; (c) chessboard:  $\delta = 10$ ,  $W_{13} = 0.38$ ,  $W_{23} = 0.28$ ,  $W_{24} = -0.56$ ,  $W_{31}^* = 0.65$ ,  $W_{42}^* = 0.98$ ; (d) stripes:  $\delta = 20$ ,  $W_{14} = 0.10$ ,  $W_{24} = 0.12$ ,  $W_{23} = 0.13$ ,  $W_{25} = -0.77$ ,  $W_{34} = 0.20$ ,  $W_{36} = -0.92$ ,  $W_{41}^* = 0.83$ ,  $W_{52}^* = 0.21$ ,  $W_{62}^* = 0.57$ ,  $W_{73}^* = 0.93$ ; (e) hierarchic pattern:  $\delta = 20$ ,  $W_{11} = 0.96$ ,  $W_{24} = -0.56$ ,  $W_{31}^* = 0.65$ ;  $W_{42}^* = 0.98$ ; (f) chessboard:  $\delta = 10$ ,  $W_{13} = 0.38$ ,  $W_{23} = 0.38$ ,  $W_{24} = -0.86$ ,  $W_{31}^* = 0.65$ ,  $W_{42}^* = 0.98$ ; (g) one stripe:  $\delta_1 = 10$ ,  $W_{11} = 0.83$ ,  $W_{21} = 0.57$ ,  $\mu_1 = 0.41$ ,  $\mu_2 = 0.037$ ,  $D_2 = 0.81$ ,  $K_{M2} = 0.1$ ; (h) multiple (two) stripes:  $W_{13} = 0.20$ ,  $W_{23} = 0.62$ ,  $W_{33} = 0.50$ ,  $W_{41} = 0.89$ ,  $W_{32} = 0.48$ ,  $W_{54} = -0.64$ ,  $\mu_1 = 0.99$ ,  $\mu_2 = 0.19$ ,  $\delta_3 = 10.3$ ,  $\delta_4 = 50$ ,  $\delta_4 = 32.2$ ,  $D_1 = 0.18$ ,  $D_2 = 0.38$ ,  $K_{M1} = 0.1$ ,  $K_{M2} = 0.1$ ,  $K_{M3} = 0.1$ ; (i) stripes:  $W_{13} = 0.44$ ,  $W_{23} = 0.39$ ,  $W_{31} = 0.32$ ,  $W_{34} = -0.81$ ,  $W_{42} = 0.62$ ,  $W_{43} = -0.56$ ,  $\mu_1 = 0.37$ ,  $\mu_2 = 0.87$ ,  $\delta_3 = 13$ ,  $\delta_4 = 34$ ,  $D_1 = 0.54$ ,  $D_2 = 0.78$ ,  $K_{M1} = 0.1$ ,  $K_{M2} = 0.1$ .

in Tables 3 (for DCIM) and 4 (DM model). The basic types found are

1. *Hierarchical*: Those networks in which a hormone is never affected, directly or indirectly, by the gene under their influence [see Fig. 5(a)]. These are composed of hierarchic chains of interactions between hormones. As more hormones are included in the network more cells express the pattern gene. These modules require that some gene in the module is self-activating. For the DM the patterns are simple stripes and multiple stripes (the number of stripes depending on the number of hormones) that have as a focus the cells that were active at the beginning. For

DCIM the patterns are characterized by stripes (usually one cell wide) or spots (in two dimensions and also one cell wide) of different heights (see Fig. 3(e) for an example) spaced by non-active cells. *Initial conditions*: The pattern generated is dependent on initial conditions. The patterns can generally be described as a frozen damped wave (see Fig. 3) around the cell that received the initial condition, so different initial conditions can just change the localization of the wave. The dependence on initial conditions is stronger when the frozen damped wave reaches more cells. *Effect of noise*: The pattern is homeostatic to gene value perturbations (noise) affecting genes in a cell that are active (it is

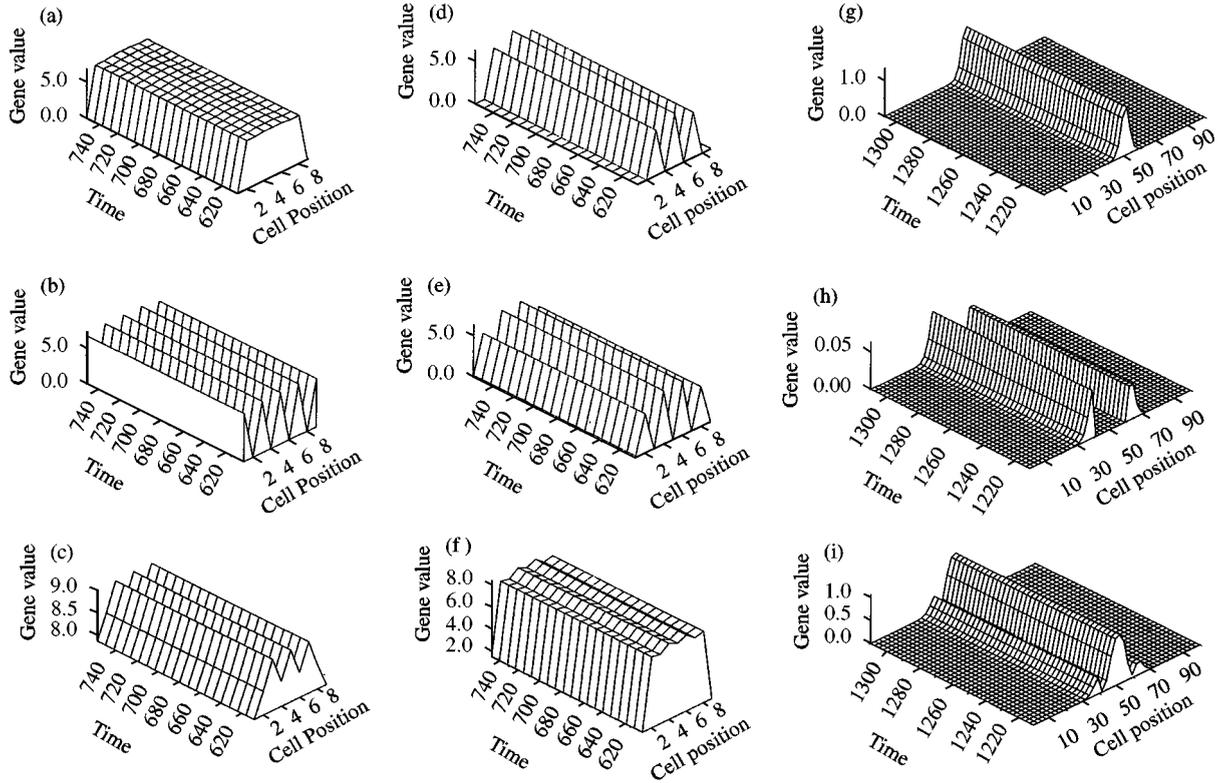


FIG. 4. Some examples of patterns that are generated by the combination of basic networks. All are one-dimensional patterns. (a)–(f) belong to the direct contact induction model (DCIM) while (g), (h) and (i) belong to the diffusive model (DM). In all cases flux-zero boundary conditions are used. The pattern in (c) is generated by combining the networks that generate the patterns shown in (a) and (b), in all cases the initial condition is given to the gene 9. The pattern in (f) is generated by combining the networks that generate the patterns shown in (d) and (e), in all cases the initial condition is given to the gene 9. The pattern in (g) is generated by combining the networks that generate the patterns shown in (g) and (i), in all cases the initial condition is given to the gene 7. The patterns shown are (a) plateau:  $\delta = 10$ ,  $W_{34} = 0.07$ ,  $W_{38} = -0.44$ ,  $W_{39} = 0.70$ ,  $W_{47} = 0.49$ ,  $W_{77} = 0.75$ ,  $W_{88} = 0.53$ ,  $W_{73}^* = 0.34$ ,  $W_{84}^* = 0.81$ ; (b) chessboard:  $\delta = 10$ ,  $W_{15} = 0.84$ ,  $W_{19} = 0.41$ ,  $W_{21} = 0.34$ ,  $W_{26} = -0.75$ ,  $W_{51}^* = 0.87$ ,  $W_{62}^* = 0.07$ ; (c) combined pattern, the same parameters as in A and B and:  $W_{10,6} = 0.68$ ,  $W_{10,83} = 0.83$ ; (d) hierarchic pattern:  $\delta = 10$ ,  $W_{11} = 0.96$ ,  $W_{25} = 0.72$ ,  $W_{19} = 0.70$ ,  $W_{19} = 0.38$ ,  $W_{51}^* = 0.92$ ,  $W_{62}^* = 0.14$ ; (e) chessboard:  $\delta = 10$ ,  $W_{37} = 0.83$ ,  $W_{39} = 0.38$ ,  $W_{43} = 0.81$ ,  $W_{48} = -0.26$ ,  $W_{73}^* = 0.15$ ,  $W_{84}^* = 0.81$ ; (f) combined pattern, the same parameters as in D and E:  $W_{10,6} = 0.38$ ,  $W_{10,83} = 0.78$ ; (g) single stripe:  $W_{77} = 0.37$ ,  $W_{17} = 0.55$ ,  $W_{41} = 0.49$ ,  $\mu_1 = 0.15$ ,  $\delta_7 = 12.5$ ,  $\delta_4 = 17$ ,  $D_1 = 0.20$ ,  $K_{M4} = 0.1$ ,  $K_{M7} = 0.1$ ; (h) multiple (two) stripes:  $W_{77} = 0.37$ ,  $W_{27} = 0.86$ ,  $W_{32} = 0.70$ ,  $W_{52} = 0.25$ ,  $W_{63} = 0.64$ ,  $W_{65} = -0.94$ ,  $\mu_2 = 0.37$ ,  $\mu_3 = 0.54$ ,  $\delta_5 = 76$ ,  $\delta_6 = 80$ ,  $\delta_7 = 12.5$ ,  $D_2 = 0.87$ ,  $D_2 = 0.98$ ,  $K_{M5} = 0.1$ ,  $K_{M6} = 0.1$ ,  $K_{M7} = 0.1$ ; (i) combined pattern, the same parameters as in (g) and (h):  $W_{84} = 0.32$ ,  $W_{86} = 13.1$ ,  $K_{M8} = 0.1$ .

different from zero). However, perturbations affecting non-expressed genes can induce a new different pattern. *Effect of size and boundary conditions*: If the number of cells expressing the pattern gene is large, different patterns may appear for different values of  $N_c$  and different boundary conditions.

2. *Non-expansive emergent (NEEM)*: In these networks, a given hormone [h1 in Fig. 5(c)] in a cell affects one receptor in its neighbours that indirectly activates this hormone in the neighbour cells. Additionally, it is required that the

same hormone activate a second genetic cascade that activates a second hormone [this time in the neighbour cells; h2 in Fig. 5(c)] that inhibits the h1 hormone in the first input cell. In fact, the h1 hormone activates two extracellular loops that share some genes. NEEM form stripes, single stripe and plateau patterns in DCIM (see Table 3 and Fig. 3). In fact, we can define a different minimal module depending on the width of the stripe. These are all, however, very similar and only differ in the possession of a different number of hormones in the activatory loop.

TABLE 3

For each pattern the minimal module is shown as well as the more relevant characteristics. It is assumed in this case that hormones only can affect receptors. DCIM = Direct contact induction model

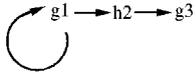
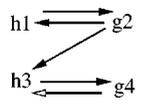
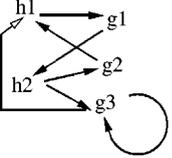
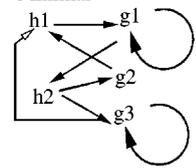
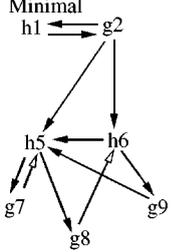
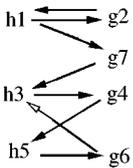
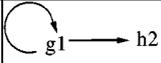
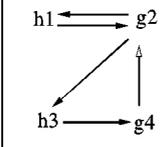
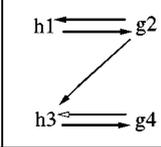
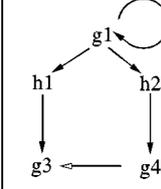
Pattern name	Module	Dependence on initial conditions	Effect of size	Homeostasis to value perturbations	Gene with spatial pattern
Hierarchic	Example 	Yes	None	Yes	g3, h2 and g1
Chessboard EEM	Minimal 	No	None: all the systems are like checkboard	Yes	h3 and g4
Plateau NEEM	Minimal 	Yes	None	Yes	g3
Single stripe NEEM	Minimal 	Yes	None	Yes	g3
Stripes NEEM	Minimal 	No	Yes: more stripes are produced in larger systems. Stripes use to form in the same place	Yes	h5
Chaotic	Minimal 	Yes	Yes	No	h3

TABLE 4

As in Table 3 but for DM relevant characteristics. It is assumed in this case that hormones only can affect receptors. DCIM = Direct contact induction model

Name	Module	Dependence on initial conditions	Effect of size	Homeostasis to value perturbations	Gene with spatial pattern
Hierarchic		Yes	None	Yes	
Stripes		No	As in DCIM stripes	Yes	g2
Chaotic		Yes	Yes	No	h3
Twin peaks		Yes	None	Yes	g3

*Initial conditions:* The exact pattern is dependent on initial conditions. *Effect of noise:* The pattern shows the same homeostatic properties to perturbations exhibited by hierarchical modules. *Effect of size and boundary conditions:* If the number of cells expressing the pattern gene is large, the pattern may depend on  $N_c$  and on the type of boundary conditions.

3. *Expansive emergent:* As in the previous case but the loops do not share genes. For the DCIM model, chessboard patterns are due to such modules. As far as we know, this is the first time that stripes and spots are found in models considering direct contact induction. In DM stripe patterns are found, just as in reaction-diffusion models. *Initial conditions:* The existence of a self-activating loop causes that for most initial conditions the same pattern (or a very similar pattern) appears. In DM the pattern can exhibit a more strong dependence on initial conditions. *Effect of noise:* The pattern is not sensitive to any kind of external perturbations in gene values. *Effect*

*of size and boundary conditions:* The existence of a loop leads to the propagation of the initial conditions to the whole system. The patterns appearing for different values of  $N_c$  are very similar. In DM some patterns exhibit a dependence on the size of the system (those that also exhibit a dependence on initial conditions). The self-activating loop also causes that homogeneous non-zero initial conditions for zero-flux boundary conditions (it needs to be noted that zero-flux boundary conditions in lines and squares of cells imply that the cells in the extremes have few neighbours) typically lead to the same result as random initial conditions. In both cases, the properties of the patterns and the structure of the modules are similar to standard reaction-diffusion models. The patterns produced under different boundary conditions are similar, except for some specific differences arising from the zero-flux case.

4. *Gradient generating:* Those modules that have a hormone affected by the gene product it

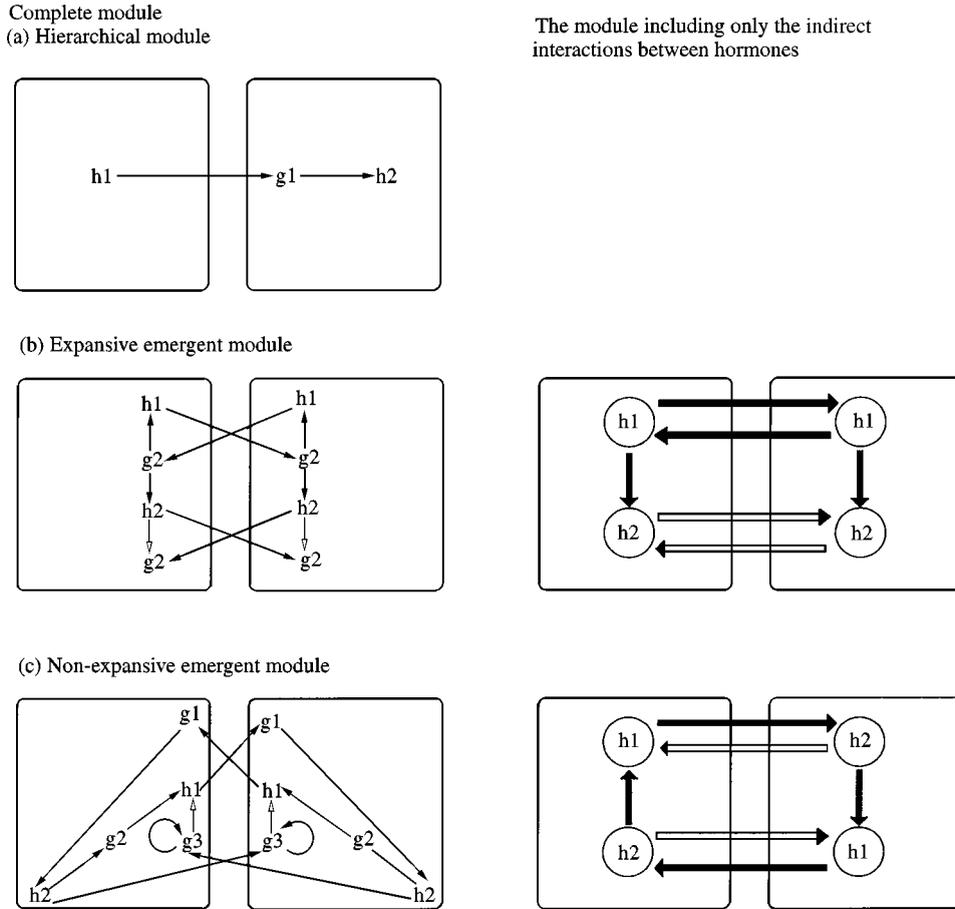


FIG. 5. The figure shows one example of each type of module. For simplicity only two cells are shown. In the first row the entire module is shown. The non-hormone gene products are indicated by  $g$  and a number while hormones are indicated by  $h$  and a number. Black arrows indicate activating interactions while white arrows indicate inhibitory interactions. In the second row only the interactions (direct or indirect) between hormones are shown: (a) hierarchic module: only the entire module is shown, (b) expansive emergent module and (c) non-expansive emergent module.

influences but that do not include any of the structural properties of the other types of modules. These patterns appear only in DCIM with zero-flux boundary conditions. The pattern is always a gradient focused in the cell in the middle of the lattice or array of cells.

All the modules found form patterns in one and two dimensions. In the two cases the described relationships between the gene network and the pattern and its properties hold. In DM and DCIM chaotic behaviour can also be found for some modules (see Table 3).

The use of normal distributions or uniform distributions does not affect most of our conclusions since we are making predictions about

gene network topologies. The probabilities of the different patterns are dependent on the type of distribution used. We have measured such probabilities only for the uniform distribution. We expect that, as in random neural networks, the use of a normal distribution increases the probability of patterns and decreases the probability of chaos (Sompolinsky *et al.*, 1988). This effect is due to the higher variance of uniform distribution (for the same range of values).

We have also explored the effect of using complex initial conditions. As previously stated, the expansive emergent networks give the same results irrespective of the initial condition. In DM square lattices of cells the initial condition can bias the polarity of stripes (especially when the

initial condition is a gradient), an effect similar to the one reported in another study (Varea *et al.*, 1997). Hierarchic and NEEM give many different patterns when various types of initial conditions are used. In most cases, initial conditions that differ in the cells that have genes expressed (it is with values over zero) give different patterns. All such patterns exhibit the properties previously outlined. In any case, the use of a different initial condition provokes a network to exhibit patterns with different properties. A more detailed study of such properties in a more evolutionary framework will be presented in the future.

There is a clear example in real systems of expansive emergent modules, the Notch/Delta pathway (Skeath & Carroll, 1992; Kimble & Simpson, 1997; Robey, 1997), which is associated with a pattern similar to that generated in the model: that is, lateral inhibition. In this case no perturbation needs to expand, because the “hormone” (in this case Delta, a transmembrane protein with EGF domains) is homogeneously expressed from the beginning due to an intracellular loop in the transcriptional factors activating it (González-Gaitán & Jäckle, 1995). The genes involved in this network are widespread in Metazoa phyla (Kimble & Simpson, 1997) and in known cases, the relationship between these genes and the pattern they generate are the same. The chessboard pattern is dependent on the number of neighbours used but always an activated cell is surrounded by inactive cells.

In many cases real patterns are generated primarily in one axis and then other dimensions may be neglected. Moreover, two-dimensional models yield similar results to their one-dimensional counterparts when the initial perturbation is a stripe instead of a single point (most developmental systems have some kind of initial cue which is often more similar to a stripe than to a point, e.g. the distribution of maternal genes in *Drosophila* can be described as a simple decaying band in each of the extremes of the egg). Interestingly, unstable-in-time patterns are very similar to those reported by other authors to explain the shell ornamentation of some molluscs (Meinhardt & Klingler, 1987). Both models, Meinhardt's and ours include saturating reactions of diffusible morphos. However, the exact functions used by Meinhardt are very different from ours.

It supports our statement that the network topologies responsible for a pattern are not very dependent on the exact function used.

#### 4. Discussion

The main result of this study is that many gene networks in our model show ordered patterns (spontaneously or following an initial cue) and that the gene networks generating a pattern are topologically equivalent. Another important result is that both the patterns and the networks responsible for them can be categorized in a small number of types depending on some structural properties of the networks and some biologically interesting properties of the patterns they generate. Patterns emerge from some simple networks using the known rules of interaction between gene products. This is an important result because it states that gene networks have an intrinsic pattern formation capacity. The large number of patterns found is explained by the combination of basic simple networks. It is an interesting result because it may imply that real gene networks could be understood by studying such basic networks. The ensemble approach is in this sense very useful to estimate the likelihood of a combinatorial nature of the networks generating complex patterns, at least from a design point of view. The ensemble approach can provide an estimation of all the possible pattern formation networks. This has not been provided by other theoretical approaches. The ensemble approach may provide additional interesting results if other cellular properties are included (such as cellular motion and mitosis). In fact, however, the combination of existing networks is a plausible evolutionary way to increase complexity, and as we will discuss later, molecular cues suggest that this may be the case. In this respect, our study suggest that complex patterns may not require complex mechanisms.

It would also be expected that real modules resemble the minimal modules found in this study because the number of variations required at the DNA level to generate them are small (and thus may appear more frequently. It is thus a likely way of generating innovation). In addition, they have selective advantages because the chance that a mutation breaks the module is

small in the minimal modules (because they have less genes than non-minimal modules). The number of patterns found in models of development is small (Newman, 1994). In our model, although the number of networks explored is large the number of basic patterns is remarkably small. Although these observations may be due to the simplistic nature of models it may also suggest that the number of basic patterns amenable in development is small. Our study suggest that although the number of basic patterns is small they may appear frequently. Thus, in evolution, variations at the DNA level may easily generate networks capable of pattern formation. Thus, from an evolutionary perspective, logical variation may be produced by combining the simple modules found in the present models of development. Recent advances have provided new data that can be easily interpreted from this perspective. Currently, it is clear that the genes involved in development are widely conserved in metazoan. In many cases, the relationships between these genes are also conserved. This conservation also suggests that evolution proceeds very often by changing the relationships between modules of genes involved in similar processes (Huang, 1998). The final logy of an organism may correlate to the choice of the adequate pattern-formation module in each stage and location. As an example, the Notch/Delta system is fairly well reproduced by our model and properly illustrates this hypothesis. The set of genes involved in this system is expressed in different stages of development, and in different organisms (Kimble & Simpson, 1997) and mostly generates a chess-board pattern. Moreover, although these genes mainly play a role in neural differentiation, they work together in other cells and tissues to generate the same kind of pattern. Some researchers (Huang, 1998) have suggested that development may be structured in syntagmata performing conserved general functions and that evolution may proceed by combining such syntagmata. Our results suggest that an important set of such syntagmata would be pattern formation modules.

We expect that the four types found may have a more general validity and many other developmental systems may exhibit some of these four types. The topological characteristics of each

type are very clear and the properties of the patterns that each type generates are very different. Such types appear general enough to expect that systems interacting by other rules may have analogous categories. Anyway we believe that thinking about the general structural and phenomenological properties of gene networks may be useful in problems related to the evolution of development and to the general organization of development.

Some interesting inferences about the evolution and structure of development can be made by considering some of the studied properties. Non-expansive emergent and hierarchic modules have some evolutionary advantages due to the dependence on initial conditions. In hierarchic modules the lack of reciprocity may lead to a simpler relationship between genotype and phenotype that may facilitate evolution, at least for short evolutionary time-scales. Small changes in the location of the initial perturbation allow, in non-expansive modules, to an increase in the number of possible patterns (although all the patterns generated in this way used to be similar). Moreover, these changes can be more asymmetric since the networks do not allow expansion of the modification through all the systems, thus previous compartmentalization of the system is not erased. In contrast, hierarchic models have some obvious disadvantages. Among them, the capacity to project the changes forward can also make the system more sensitive to errors and mutations, and then large hierarchies produce large constraints and a high mutational cost. The mechanisms producing expansive emergent patterns also generate constraints but these are related to the generated pattern (Alberch, 1982).

The different possibilities (hierarchic, emergent) are compatible and the most interesting task is to find which mechanisms are more adequate in each case and why. Different networks can generate similar patterns. Expansive emergent modules, being simpler than the hierarchic modules (but producing more elaborate patterns) required to generate the same pattern, may have been generated or recruited easily and earlier in evolution. However, hierarchic networks may generate patterns similar to those made by emergent modules if they undergo combination between them. Thus, it is conceivable that the

mechanisms responsible for emergent patterns may easily appear in evolution but subsequently, they may have been replaced by more hierarchic ones (Newman, 1994). In the latter case, the driving force for this change would be a direct genotype–phenotype correlation. It is worth mentioning that, in expansive emergent patterns (at least in those considered in reaction–diffusion models and in our model) the number of different patterns that can be generated is small. Patterns with higher complexity than stripes can only be generated through other types of networks such as, for example, combined hierarchical and emergent modules. Thus, although emergent modules generate the most complex patterns if we consider a single module, hierarchic modules may be more capable to combining between them to generate more complex patterns. A good comparison would require a deeper and accurate study on an evolutionary framework.

Finally, we would like to summarize our conclusions by answering the questions we addressed in the introduction.

1. Our model suggests that pattern formation capacity is a generic property of gene networks.
2. Although there are many different patterns all of them are due to the combination of few simple patterns. When a complex pattern is found its network includes the minimal networks able to generate such simpler basic patterns.
3. DM and DCIM exhibit different patterns but they show the same four basic types of modules in spite of the non-trivial qualitative differences involved in their rules. Although both mechanisms lead to stable spatial patterns, these are much more common and diverse in the inductive case. This robustness is likely to have been relevant during the evolution of developmental pathways.
4. All the networks responsible for such patterns are organized in some special sets sharing similar structural characteristics. Moreover, the patterns generated by each class of networks also share many common characteristics.

5. We have some cases in which the model shows networks found also in real systems generating the same pattern. The model networks may be useful as an hypothesis for real patterns in which the network is not known.
6. Finally, our results seem to favour a modular organization of development based on such basic modules. Additionally, the existence of these basic types of modules suggests that we can take valuable insights about how the evolution of development proceeds by studying such types and the effects of their combination. This is a question that we would further address.

We are grateful to Gemma Marfany, Stuart Newman, Gunter P. Wagner, Stuart Kauffman, Brian Goodwin, Douglas Erwin, Andrew Keller, and anonymous referee and the members of the complex system research group, and the epigenetic group for helpful comments. This work was supported by grants of the Generalitat de Catalunya (ISC, FI/98-00342) and the CICYT (PB97-0693 to RVS and PB95-0579 to JGF) and by the Santa Fe Institute (RVS).

## REFERENCES

- ALBERCH, P. (1992). In: *Evolution and Development* (Bonner, J. T., ed.), pp. 313–332. Berlin: Springer-Verlag.
- COLLIER, J. R., MONK, N. A., MAINI, P. K. & LEWIS, J. H. (1996). Pattern formation by lateral inhibition with feedback: a mathematical model of notch-delta intercellular signalling. *J. theor. Biol.* **183**, 429–436.
- CONWAY MORRIS, S. (1998). *The Crucible of Creation*. Oxford: Oxford University Press.
- CORNISH-BOWDEN, A. (1979). *Fundamentals of Enzyme Kinetics*. London: Butterworths.
- DOMANY, E., VAN HEMMEN, J. L. & SCHULTEN, K. (Eds). (1995). *Models of Neural Networks I*. Berlin: Springer-Verlag.
- DUMAIS, J. & HARRISON, L. G. (2000). Whorl genesis in the dasycladalean algae: the pattern formation viewpoint. *Philos. Trans. Roy. Soc. Lond. B. Biol.* **355**, 281–305.
- GLASS, L. & KAUFFMAN, S. A. (1973). The logical analysis of continuous nonlinear biochemical control networks. *J. theor. Biol.* **39**, 103–129.
- GLASS, L. & KAUFFMAN, S. (1972). Cooperative components, spatial localization and oscillatory cellular dynamics. *J. theor. Biol.* **34**, 219–237.
- GONZÁLEZ-GAITÁN, M. & JÄCKLE, H. (1995). Invagination centers within the *Drosophila* stomatogastric nervous system anlage are positioned by Notch-mediated signalling which is spatially controlled through wingless. *Development* **121**, 2313–2325.
- GOODWIN, B. C. (1994). *How the Leopard Changed its Spots*. London: Weidenfeld and Nicholson.

- GOODWIN, B. C. & TRAINOR, L. E. H. (1985). Tip and whorl genesis in Acetabularia by calcium regulated strain fields. *J. theor. Biol.* **117**, 79.
- HUANG, F. (1998) Syntagms in development and evolution. *Int. J. Dev. Biol.* **42**, 487–404.
- HUNDING, A. & ENGELHARDT, E. (1994). Models of pattern formation in early morphogenesis. *J. theor. Biol.* **173**, 401–410.
- HUNDING, A., KAUFFMAN, S. A. & GOODWIN, B. (1990). Drosophila segmentation: supercomputer simulation of prepattern hierarchy. *J. theor. Biol.* **145**, 369–384.
- JACKSON, E. R., JOHNSON, D. & NASH, W. G. (1986). Gene networks in development. *J. theor. Biol.* **119**, 379–396.
- KANEKO, K. & YOMO, T. (1994). Cell division, differentiation and dynamic clustering. *Physica D* **75**, 89–102.
- KANEKO, K. & YOMO, T. (1997) Isologous diversification: a theory of cell differentiation. *Bull. Math. Biol.* **59**, 139–196.
- KAUFFMAN, S. A. (1993). *Origins of Order*. New York: Oxford University Press.
- KIMBLE & SIMPSON (1997). Signalling by the LIN-12/notch pathway. *Annu. Rev. Cell Dev. Biol.* **13**, 333–361.
- LEWIS, J. E. & GLASS, L. (1991). Steady states, limit cycles and chaos in models of complex biological networks. *Int. J. Bifur. Chaos* **1**, 477–483.
- MEINHARDT, H. (1982). *Models of Biological Pattern Formation*. London: Academic Press.
- MEINHARDT, H. & KLINGER, M. (1987). A model for pattern generation on the shell of molluscs. *J. theor. Biol.* **126**, 63–73.
- MIKHAILOV, A. S. (1990). *Foundations of Synergetics I*. Berlin: Springer.
- MIKHAILOV, A. S. & LOSKUTOV, A. YU. (1991). *Foundations of Synergetics II*. Berlin: Springer.
- MJOLSNESS, E., SHARP, D. H. & REINITZ, J. (1991). A connectionist model of development. *J. theor. Biol.* **152**, 429–456.
- NEWMAN, S. A. (1994). Generic physical mechanisms of tissue genesis: a common basis for development and evolution. *J. Evol. Biol.* **7**, 467–488.
- NEWMAN, S. A. & COMPER, W. D. (1990). 'Generic' physical mechanisms of morphogenesis and pattern formation. *Development* **110**, 1–18.
- NICOLIS, G. & PRIGOGINE, I. (1977). *Selforganization in Nonequilibrium Systems*. New York: Wiley.
- OSTER, G., ALBERCH, P., MURRAY, J. D. & SHUBIN, N. (1988). Evolution and genetic rules: the shape of the vertebrate limb in ontogeny and phylogeny. *Evolution* **42**, 862–880.
- OSTER, G. & MURRAY, J. D. (1989). Pattern formation models and developmental constraints. *J. Exp. Zool.* **251**, 186–202.
- OWEN, M. R., SHERRAT, J. R. & MYERS, S. R. (1999) How far can juxtacrine signal travel. *Proc. Roy. Soc. Lond. B, Biol. Sci.* **266**, 579–585.
- REINITZ, J., MJOLSNESS, E. & SHARP, D. H. (1992). Model of cooperative control of positional information in *Drosophila* by Bicoid and maternal Hunchback. *J. Exp. Zool.* **271**, 47–56.
- ROBEY, E. (1997). Notch in vertebrates. *Curr. Opin. Gen. Dev.* **7**, 551–557.
- SKEATH, J. B. & CARROLL, S. B. (1992). Regulation of proneural gene expression and cell fate during neuroblast segregation in the *Drosophila* embryo. *Development* **114**, 939–946.
- SLACK, J. M. W. (1991). *From Egg to Embryo*. Cambridge: Cambridge University Press.
- SOMPOLINSKY, H., CRISANTI, A. & SOMMERS, P. (1988). Chaos in neural networks. *Phys. Rev. Lett.* **61**, 259–262.
- SOLÉ, R. V., BASCOMPTE, J. & VALLS, J. (1992). Stability and complexity in spatially-extended two-species competition. *J. theor. Biol.* **159**, 469–480.
- SOLÉ, R. V. & LUQUE, B. (1995). Phase transitions and antichaos in Kauffman networks. *Phys. Lett. A* **196**, 331–334.
- VAREA, C., ARAGÓN, J. L. & BARRIO, R. A. (1997). Confined turing patterns in growing systems. *Phys. Rev. E* **56**, 1250.
- WAGNER, A. (1999). Redundant gene functions and natural selection. *J. Evol. Biol.* **12**, 1–16.
- WAGNER, A. (1994). Evolution of gene networks by gene duplication: a mathematical model and its implications on genome organization. *Proc. Natl. Acad. Sci. U.S.A.* **91**, 4387–4391.