Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Theoretical Biology 235 (2005) 431–449
Journal of Theoretical Biology
www.elsevier.com/locate/yjtbi

Robustness and fragility of Boolean models for genetic regulatory networks

Madalena Chaves[a,*], Réka Albert[b], Eduardo D. Sontag[a]

[a]Department of Mathematics and BioMaPS Institute for Quantitative Biology, Rutgers University, Piscataway, NJ 08854, USA
[b]Department of Physics and Huck Institutes for the Life Sciences, Pennsylvania State University, University Park, PA 16802, USA

Received 8 August 2004; received in revised form 9 December 2004; accepted 26 January 2005
Available online 19 March 2005

Abstract

Interactions between genes and gene products give rise to complex circuits that enable cells to process information and respond to external signals. Theoretical studies often describe these interactions using continuous, stochastic, or logical approaches. We propose a new modeling framework for gene regulatory networks, that combines the intuitive appeal of a qualitative description of gene states with a high flexibility in incorporating stochasticity in the duration of cellular processes. We apply our methods to the regulatory network of the segment polarity genes, thus gaining novel insights into the development of gene expression patterns. For example, we show that very short synthesis and decay times can perturb the wild-type pattern. On the other hand, separation of time-scales between pre- and post-translational processes and a minimal prepattern ensure convergence to the wild-type expression pattern regardless of fluctuations.

Keywords: Gene regulatory networks; Segment polarity genes; Boolean models

1. Introduction

Understanding how genetic information is translated into proteins to produce various cell types remains a major challenge in contemporary biology (Wolpert et al., 1998). Gene products often regulate the synthesis of mRNAs and proteins, forming complex networks of regulatory interactions. Concurrently with experimental progress in gene control networks (Davidson and et al., 2002), several alternative modeling frameworks have been proposed. In the continuous-state approach, the concentrations of cellular components are assumed to be continuous functions of time, governed by differential equations with mass-action (or more general) kinetics (Reinitz and Sharp, 1995; von Dassow et al., 2000; Gursky et al., 2001). Stochastic models address the deviations from population homogeneity by transforming reaction rates into probabilities and concentrations into numbers of molecules (Rao et al., 2002). Finally, in the discrete approach, each component is assumed to have a small number of qualitative states, and the regulatory interactions are described by logical functions (Mendoza et al., 1999; Sánchez and Thieffry, 2001; Yuh et al., 2001; Kauffman et al., 2003; Ghysen and Thomas, 2003; Bodnar, 1997; Albert and Othmer, 2003; Espinosa-Soto et al., 2004).

The kinetic details of protein–protein or protein–DNA interactions are rarely known, but there is increasing evidence that the input–output curves of regulatory relationships are strongly sigmoidal and can be well approximated by step functions (Yuh et al., 2001; Thomas, 1973). Moreover, both models and experiments suggest that regulatory networks are remarkably robust, that is, they maintain their function even when faced with fluctuations in components and reaction rates

*Corresponding author.
E-mail addresses: madalena@math.rutgers.edu (M. Chaves), ralbert@phys.psu.edu (R. Albert), sontag@math.rutgers.edu (E.D. Sontag).

0022-5193/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jtbi.2005.01.023

(von Dassow et al., 2000; Alon et al., 1999; Eldar et al., 2002; Carlson and Doyle, 2002; Conant and Wagner, 2004; Espinosa-Soto et al., 2004). These observations lend support to the assumption of discrete states for genetic network components and of combinatorial rules for the effects of transcription factors (Glass and Kauffman, 1973; de Jong et al., 2004). The extreme of discretization, Boolean models, consider only two states (expressed or not), closely mimicking the inference methods used in genetics (Kauffman et al., 2003; Thomas, 1973; Kauffman, 1993). It is straightforward to study the effect of knock-out mutations or changes in initial conditions in this framework, and the agreement between a real system and a Boolean model of it is a strong indication of the robustness of the system to changes in kinetic details (Albert and Othmer, 2003).

In discrete models the decision whether a network node (component) will be affected by a synthesis or decay process is determined by the state of effector nodes (nodes that interact with it). Typical time-dependent Boolean models use synchronous updating rules (Kauffman et al., 2003; Albert and Othmer, 2003; Bodnar, 1997; Kauffman, 1993), assuming that the time-scales of the processes taking place in the system are similar. In reality the time-scales of transcription, translation, and degradation can vary widely from gene to gene and can be anywhere from minutes to hours. Logical models following the formalism introduced by René Thomas (Thomas, 1973) allow asynchronism by associating two variables to each gene: a state variable describing the level of its protein, and an image variable that is the output of the logical rule whose inputs are the state variables of effector nodes. Whether the future state variable of a gene equals the image or current state variable depends on the update order and, in the absence of temporal information, the Thomas formalism focuses on determining the steady states, where the state and image variables coincide (Mendoza et al., 1999; Sánchez and Thieffry, 2001; Ghysen and Thomas, 2003; Bernot et al., 2004). The effect of asynchronous updates on the dynamics of the system, however, has not been explored yet.

In this paper, we present a methodology for testing the robustness of Boolean models with respect to stochasticity in the order of updates. Through this, we are also probing the system itself: will individual variations lead to unexpected gene expression patterns? In the asynchronous method, the synthesis/decay decision is made at different time-points for each node, allowing individual variability in each process' duration, but more importantly, it allows for decision reversal if the dynamics of effector nodes changes. It becomes possible to reproduce, e.g., the overturning of mRNA decay when its transcriptional activator is synthesized, a process that synchronous update cannot capture. Thus, replacing synchronous with asynchronous updates is not merely a technical detail, but rather a fundamental paradigm shift from pointwise in time to potentially continuous communication between nodes. Indeed, the effective synthesis or decay times for a certain node are determined by the time interval between the latest update of its effector nodes and its current update time, and can be any positive fraction of the unit time interval. We propose three algorithms, with varying freedom in the relative duration of cellular processes, and find that very short transcription or decay times have the potential to derail the wild-type development process.

The steady states of a Boolean model will remain the same regardless of the mechanism of update, but its dynamical behavior can be drastically altered due to the stochastic nature of the updates; for instance, the same initial state may lead to different steady states or limit cycles. Since the duration of synthesis and decay processes is not known, we randomly explore the space of all possible time-scales and update orders, and derive the probability of different outcomes. Our methods offer a systematic way of exploring generic behavior of gene regulatory networks and comparing it to experimentally observed outcomes. To present a concrete example, we generalize a previously introduced Boolean model of the *Drosophila* segment polarity genes (Albert and Othmer, 2003). This model reproduces the wild-type steady state pattern of the segment polarity genes as well as the gene patterns of mutants, but its dynamic behavior is not directly comparable to that of the real system. Here we show that asynchronous update leads to a much more realistic model that gives further insights into the robustness of the gene regulatory network.

2. The segment polarity gene network in *Drosophila*

The *Drosophila melanogaster* segment polarity genes represent the last step in the hierarchical cascade of gene families initiating the segmented body of the fruit fly. While the preceding genes act transiently, the segment polarity genes are expressed throughout the life of the fly, and their periodic spatial pattern is maintained for at least 3 h of embryonic development (Wolpert et al., 1998). The regulatory roles of the previously expressed genes such as the pair-rule genes *fushi tarazu*, *runt*, *even-skipped* are incorporated in the prepattern (initial state) of the segment polarity genes. The stable maintenance of the segment polarity gene expression is due to the interactions between these genes (see Fig. 1), and it is a crucial requirement in the development and stability of the parasegmental furrows. The best characterized segment polarity genes include *engrailed* (*en*), *wingless* (*wg*), *hedgehog* (*hh*), *patched* (*ptc*), *cubitus interruptus* (*ci*) and *sloppy paired* (*slp*), encoding for diverse proteins including transcription factors as well as secreted and receptor proteins.

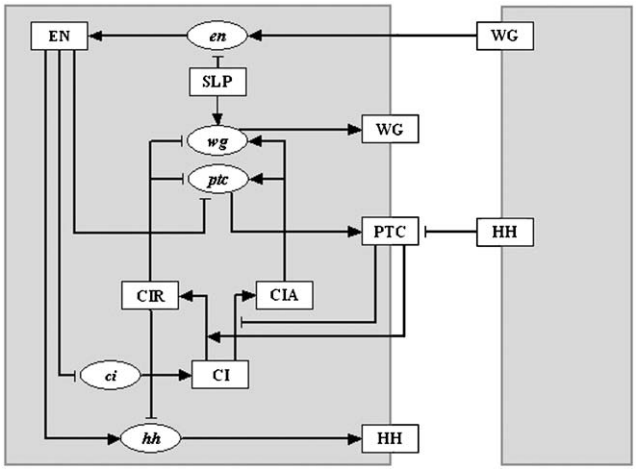

Fig. 1. The network of interactions between the segment polarity genes. The gray background layers illustrate two neighboring cells, indicating that some interactions in this network are inter-cellular. The shape of the nodes indicates whether the corresponding substances are mRNAs (ellipses) or proteins (rectangles). The edges of the network signify either biochemical reactions (e.g. translation, protein interactions) or regulatory interactions (e.g. transcriptional activation). The edges are classified as activating (→) or inhibiting (⊣). Figure adapted from Albert and Othmer (2003).

The pair-rule gene *sloppy paired* (*slp*) is activated before the segment polarity genes and expressed constitutively thereafter (Grossniklaus et al., 1992; Cadigan et al., 1994). *slp* encodes two forkhead domain transcription factors with similar functions that activate *wg* transcription and repress *en* transcription, and since they are co-expressed we designate them both SLP. The *wg* gene encodes a glycoprotein that is secreted from the cells that synthesize it (Hooper and Scott, 1992; Pfeiffer and Vincent, 1999), and can bind to the Frizzled receptor on neighboring cells, initiating a signaling cascade leading to the transcription of *engrailed* (*en*) (Cadigan and Nusse, 1997). EN, the homeodomain-containing product of the *en* gene, promotes the transcription of the *hedgehog* gene (*hh*) (Tabata et al., 1992). In addition to the homeodomain, EN contains a separate repression domain that affects the transcription of *ci* (Eaton and Kornberg, 1990) and possibly *ptc* (Hidalgo and Ingham, 1990; Taylor et al., 1993). The hedgehog protein (HH) is tethered to the cell membrane by a cholesterol linkage that is severed by the dispatched protein, freeing it to bind to the HH receptor PTC on a neighboring cell (Ingham and McMahon, 2001) . The intracellular domain of PTC forms a complex with smoothened (SMO) in which SMO is inactivated by a post-translational conformation change (Ingham, 1998). Binding of HH to PTC removes the inhibition of SMO, and activates a pathway that results in the modification of CI (Ingham, 1998). The CI protein can be converted into one of two transcription factors, depending on the PTC–HH interactions. In the absence of HH signaling CI is cleaved to form CIR, a transcriptional repressor that represses *wg*, *ptc* and *hh* transcription (Aza-Blanc and Kornberg, 1999). When secreted HH binds to PTC and frees SMO, CI is converted to a transcriptional activator, CIA, that promotes the transcription of *wg* and *ptc* (Aza-Blanc and Kornberg, 1999; Ohlmeyer and Kalderon, 1998).

The initial state of the *Drosophila* segment polarity genes includes two-cell-wide SLP stripes followed by two-cell-wide stripes not expressing SLP (Cadigan et al., 1994), single-cell-wide *wg*, *en* and *hh* stripes followed by three cells not expressing them, and three-cell-wide stripes for *ci* and *ptc* (Hooper and Scott, 1992; Wolpert et al., 1998): This pattern is maintained almost unmodified for 3 h[1] (see Fig. 2a), during which time the embryo is divided into 14 parasegments by furrows positioned between the *wg* and *en*-expressing cells (Hooper and Scott, 1992).

The first model of the segment polarity gene network was proposed by von Dassow and collaborators (von Dassow et al., 2000), and is a continuous-state model of 13 equations and 48 unknown kinetic parameters. The main conclusion of the (von Dassow et al., 2000) article is that the gene patterns are robust with respect to variations in the kinetic constants in the rate laws, thus the essential feature of this network is its topology, i.e. the existence and signature (activating or inhibiting) of the interactions. The idea of the network topology determining its dynamics was further explored by Albert and Othmer (2003), who used a slightly different network reconstruction and assumed synchronous Boolean regulation among nodes. In the Albert and Othmer (2003) model each mRNA and protein is represented by a node of a network, and the state of each node is 1 or 0, according to whether the corresponding substance is present or not. The states of the nodes are updated synchronously, and the future state of node i is determined by a Boolean function of its current state and the current states of those nodes that have edges incident on it. The updating functions are based on the experimental information and on the following dynamical assumptions: (i) the synthesis of mRNAs/proteins has the duration of one time step; (ii) the effect of transcriptional activators and inhibitors is never additive, but rather, inhibitors are dominant; (iii) mRNAs decay in one time step if not transcribed; (iv) transcription factors and proteins undergoing post-translational modification decay in one time step if their mRNA is not present; (v) protein–protein binding, such as in the formation of the Patched-Hedgehog complex, is assumed to be instantaneous. In summary, the Albert and Othmer (2003) model assumes that gene transcription, protein translation, mRNA and protein decay all happen on a similar time-scale, while protein complex

[1]A notable exception includes the refinement of the *ptc* pattern.

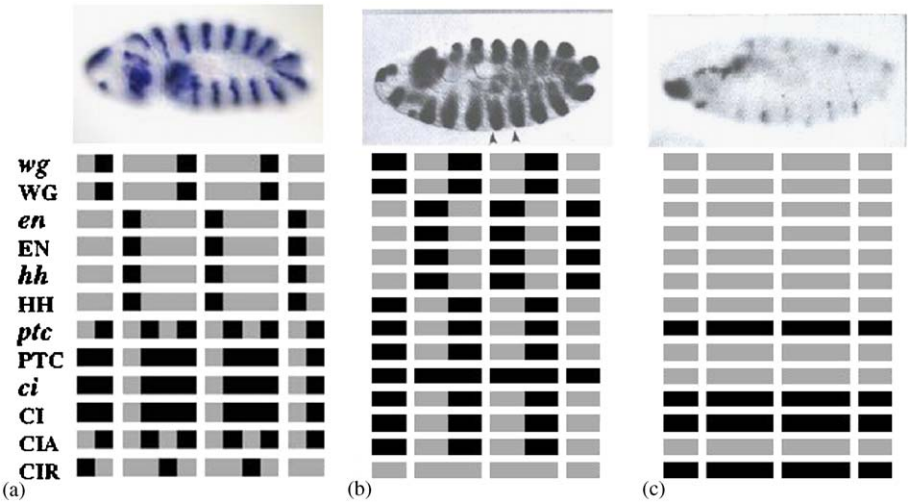

Fig. 2. (a) Top: Illustration of the gene expression pattern of *wingless* on a gastrulating (stage 9) embryo. Other segment polarity genes have similar periodic patterns that are maintained for around 3 h of embryonic development. The parasegmental furrows form at the posterior border of the *wg*-expressing cells (Wolpert et al., 1998). Bottom: Synthesis of the wild-type expression patterns of the segment polarity genes (see also text) (Hooper and Scott, 1992; Wolpert et al., 1998). Left corresponds to anterior and right to posterior in each parasegment. Horizontal rows correspond to the pattern of individual nodes—specified at the left side of the row—over two full and two partial parasegments. Each parasegment is assumed to be four cells wide. A black (gray) box denotes a node that is (is not) expressed. (b) Top: *wingless* expression pattern in a *patched* knock-out mutant embryo at stage 11 (Tabata et al., 1992). The *wingless* stripes broaden, and secondary furrows appear at the middle of the parasegment, indicating a new *en*–*wg* boundary. Bottom: Broad striped steady-state of the Boolean model, obtained when *patched* is kept off (with the change that *ptc* and PTC are not expressed), or when *wg*, *en*, *hh* are initiated in every cell (Albert and Othmer, 2003). This steady-state agrees with all experimental observations on *ptc* mutants and heat-shocked genes (Tabata et al., 1992; Gallet et al., 2000; Martinez-Arias et al., 1988; Schwartz et al., 1995; DiNardo et al., 1988; Ingham et al., 1991; Bejsovec and Wieschaus, 1993). (c) Top: *wingless* expression pattern in an *engrailed* knock-out mutant embryo at stage 11 (Tabata et al., 1992). The initial periodic pattern is disappearing, and gives rise to a non-segmented, embryonic lethal phenotype. Bottom: Non-segmented steady-state of the Boolean model, obtained when *wg*, *en* or *hh* are kept off, or cell-to-cell signaling is disrupted (Albert and Othmer, 2003). This steady-state agrees with all experimental observations on *wg*, *en*, *hh* mutants (Tabata et al., 1992; DiNardo et al., 1988; Schwartz et al., 1995; Hidalgo and Ingham, 1990; Gallet et al., 2000). Gene expression images obtained from http://www.fruitfly.org (a) and Tabata et al. (1992) (b,c).

formation is instantaneous compared to this common time-scale.

The von Dassow et al. (2000) and Albert and Othmer (2003) models agree in their conclusions regarding the robustness of the segment polarity gene network. The simplicity of the Boolean rules in the latter also allows for the exploration of knock-out mutations and changes in the prepattern of the segment polarity genes. Starting from the known initial state of *en*, *wg*, *hh*, *ptc*, *ci* and SLP, and assuming the null(off) state for all other nodes the Albert and Othmer (2003) model leads to a time-invariant spatial pattern (see Fig. 2a) that coincides with the experimentally observed wild-type expression of the segment polarity genes during stages 9–11. Indeed, *wg* and WG are expressed in the most posterior cell of each parasegment, while *en*, EN, *hh* and HH are expressed in the most anterior cell of each parasegment, as is observed experimentally (Ingham, 1998; Tabata et al., 1992), *ptc* is expressed in two stripes of cells, one stripe on each side of the *en*-expressing cells, the anterior one coinciding with the *wg* stripe (Hidalgo and Ingham, 1990; Hooper and Scott, 1992). *ci* is expressed almost ubiquitously, with the exception of the cells expressing *en* (Eaton and Kornberg, 1990). CIA is expressed in the neighbors of the HH-expressing cells, while CIR is expressed far from the HH-expressing cells (Aza-Blanc and Kornberg, 1999). The model indicates that knock-out mutations in *en*, *wg*, *hh* cause the non-segmented gene pattern shown on Fig. 2c, which agrees with experimental observations. Indeed, the *hh* expression in *en* null embryos starts normally, but disappears before stage 10 (Tabata et al., 1992). In *wg* null embryos, *en* is initiated normally but fades away by stage 9, as observed by DiNardo et al. (1988), while *ci* is ubiquitously expressed (Schwartz et al., 1995). In *hh* mutant embryos the *wg* expression disappears by stage 10 (Hidalgo and Ingham, 1990), as does the expression of *ptc*, and there is no segmentation (Gallet et al., 2000). On the other hand, *ptc* knockout mutations or over-expressed initial states lead to the broad-striped pattern of Fig. 2b.[2] Indeed,experimental results indicate broad *en*, *wg* and *hh* stripes (Tabata et al., 1992; Gallet et al., 2000; Martinez-Arias et al., 1988) and Gallet et al. (2000) find that a new ectopic groove forms at the second *en*−*wg* interface at the middle of the

[2]The only difference between the *ptc* mutant and heat-shock pattern is that the former does not express *ptc* and PTC.

parasegment. Also, *ci* is not expressed at this ectopic groove (Schwartz et al., 1995). In heat-shock experiments the *wg* and *ptc* stripes expand anteriorly when *hh* or *en* are ubiquitously induced (Gallet et al., 2000), while narrower *ci* stripes emerge after a transient decay of *ci* (Schwartz et al., 1995). Intriguingly, the Albert and Othmer (2003) model finds that a knock-out mutation of *ci* does not change the *en*, *wg*, *hh* patterns but disrupts *ptc* expression; experiments indicate that the segmental grooves are present and *wg* is expressed until stage 11, but *ptc* expression decays (Gallet et al., 2000). In summary, the simple synchronous Boolean model (Albert and Othmer, 2003) captures perfectly the wild-type and mutant expression patterns of the segment polarity genes, and thus serves as a good starting point for a more realistic model that relaxes the assumption of synchronicity.

We focus our attention on a single parasegment of four cells, thus the total number of nodes we consider is $4 \times 13 = 52$. We use the same interaction topology and logical rules as the synchronous model (Albert and Othmer, 2003), but instead of assuming that the states of all nodes are updated simultaneously, we update the state of each node individually (see Table 1). To maintain the highest generality, we incorporate possible cell-to-cell variations in synthesis and decay processes.[3]

Throughout the text, the notation "wg_1^t" or "$wg_1(t)$" represent the state of *wingless* mRNA in the first cell of the parasegment at time t. Similar notations apply for other mRNAs and proteins. There are 4 cells in each parasegment, and we adopted periodic boundary conditions, meaning that: $node_{4+1} = node_1$ and $node_{1-1} = node_4$. The wild-type initial state corresponds to:

$$wg_4^0 = 1, \quad en_1^0 = 1, \quad hh_1^0 = 1, \quad ptc_{2,3,4}^0 = 1, \quad ci_{2,3,4}^0 = 1 \tag{1}$$

and the remaining nodes are zero. The asynchronous model represented in Table 1 exhibits the same steady states as the synchronous model developed in Albert and Othmer (2003). Note that three of the four main steady states agree perfectly with experimentally observed states corresponding to wild-type, *en*, *wg* or *hh* mutant and *ptc* mutant embryonic patterns (Tabata et al., 1992; DiNardo et al., 1988; Schwartz et al., 1995; Hidalgo and Ingham, 1990; Gallet et al., 2000; Martinez-Arias et al., 1988; Bejsovec and Wieschaus, 1993; Ingham et al., 1991; Hooper and Scott, 1992; Wolpert et al., 1998). A summary is presented in Table 2.

[3]We follow the Albert and Othmer (2003) model in assuming very short time-scales for PTC–HH binding and SMO activation, and consequently in Fig. 1 and in the regulatory rules we connect the CI post-translational modifications to HH signaling. We have verified that this assumption can be relaxed without any qualitative changes in the results.

Table 1
Regulatory functions governing the states of segment polarity gene products in the model

Node	Boolean updating function in the asynchronous algorithm
SLP_i	$SLP_i(t) = \begin{cases} 0 & \text{if } i \in \{1,2\} \\ 1 & \text{if } i \in \{3,4\} \end{cases}$
wg_i	$wg_i(t) = (CIA_i(\tau_{CIA})$ and $SLP_i(\tau_{SLP})$ and not $CIR_i(\tau_{CIR}))$ or $[wg_i(\tau_{wg})$ and $(CIA_i(\tau_{CIA})$ or $SLP_i(\tau_{SLP}))$ and not $CIR_i(\tau_{CIR})]$
WG_i	$WG_i(t) = wg_i(\tau_{wg})$
en_i	$en_i(t) = (WG_{i-1}(\tau_{WG1})$ or $WG_{i+1}(\tau_{WG2}))$ and not $SLP_i(\tau_{SLP})$
EN_i	$EN_i(t) = en_i(\tau_{en})$
hh_i	$hh_i(t) = EN_i(\tau_{EN})$ and not $CIR_i(\tau_{CIR})$
HH_i	$HH_i(t) = hh_i(\tau_{hh})$
ptc_i	$ptc_i(t) = CIA_i(\tau_{CIA})$ and not $EN_i(\tau_{EN})$ and not $CIR_i(\tau_{CIR})$
PTC_i	$PTC_i(t) = ptc_i(\tau_{ptc})$ or $(PTC_i(\tau_{PTC})$ and not $HH_{i-1}(\tau_{HH1})$ and not $HH_{i+1}(\tau_{HH2}))$
ci_i	$ci_i(t) =$ not $EN_i(\tau_{EN})$
CI_i	$CI_i(t) = ci_i(\tau_{ci})$
CIA_i	$CIA_i(t) = CI_i(\tau_{CI})$ and [not $PTC_i(\tau_{PTC})$ or $HH_{i-1}(\tau_{HH1})$ or $HH_{i+1}(\tau_{HH2})$ or $hh_{i-1}(\tau_{hh1})$ or $hh_{i+1}(\tau_{hh2})]$
CIR_i	$CIR_i(t) = CI_i(\tau_{CI})$ and $PTC_i(\tau_{PTC})$ and not $HH_{i-1}(\tau_{HH1})$ and not $HH_{i+1}(\tau_{HH2})$ and not $hh_{i-1}(\tau_{hh1})$ and not $hh_{i+1}(\tau_{hh2})$

Each node is labeled by its biochemical symbol and subscripts signify cell number. The times τ_j signify the last time node j was updated before t.

Table 2
Complete characterization of the model's steady states

Steady state	Expressed nodes
Wild-type	wg_4, WG_4, en_1, EN_1, hh_1, HH_1, $ptc_{2,4}$, $PTC_{2,3,4}$, $ci_{2,3,4}$, $CI_{2,3,4}$, $CIA_{2,4}$, CIR_3
Broad stripes	$wg_{3,4}$, $WG_{3,4}$, $en_{1,2}$, $EN_{1,2}$, $hh_{1,2}$, $HH_{1,2}$, $ptc_{3,4}$, $PTC_{3,4}$, $ci_{3,4}$, $CI_{3,4}$, $CIA_{3,4}$
No segmentation	$ci_{1,2,3,4}$, $CI_{1,2,3,4}$, $PTC_{1,2,3,4}$, $CIR_{1,2,3,4}$
Wild-type variant	wg_4, WG_4, en_1, EN_1, hh_1, HH_1, $ptc_{2,4}$, $PTC_{1,2,3,4}$, $ci_{2,3,4}$, $CI_{2,3,4}$, $CIA_{2,4}$, CIR_3
Ectopic	wg_3, WG_3, en_2, EN_2, hh_2, HH_2, $ptc_{1,3}$,$PTC_{1,3,4}$, $ci_{1,3,4}$, $CI_{1,3,4}$, $CIA_{1,3}$, CIR_4
Ectopic variant	wg_3, WG_3, en_2, EN_2, hh_2, HH_2, $ptc_{1,3}$, $PTC_{1,2,3,4}$, $ci_{1,3,4}$, $CI_{1,3,4}$, $CIA_{1,3}$, CIR_4

3. Randomly perturbed time-scales

As in the context of parallel computation systems, the fundamental difference between synchronous and asynchronous updates is at the level of task coordination and

data communication among nodes in a network (Bertsekas and Tsitsiklis, 1989). Synchronous algorithms are highly coordinated: at pre-determined instants, all the nodes "stop" and exchange the current information among themselves. For instance, suppose there are N nodes, where each node i "computes" the state of variable x_i, according to a function $f_i(x_1, x_2, \ldots, x_N)$ $(i = 1, \ldots, N)$. When all the N nodes have finished phase k, they exchange their current states, x_i^k, and then proceed to phase $k+1$, that is

$$x_i^{k+1} = f_i(x_1^k, x_2^k, \ldots, x_N^k).$$

Asynchronous algorithms, on the other hand, admit a greater flexibility at the level of process coordination. Each node is allowed to have its own "computation rate", that is, during any time interval $[t_a, t_b]$, node i may be updated only once, while node j may be updated $\ell > 1$ times. In this case, communication delays between nodes may occur, and some possibly outdated information may be used: for instance, node j uses the same value $x_i(t_a)$ throughout its ℓ updates in the interval $[t_a, t_b]$. However, an overall gain in efficiency in achieving the final result may be expected. For instance, in our example, the wild-type steady state is reached in less than 4 steps with the asynchronous algorithms (see Sections 4, 5), while with the synchronous algorithm 6 steps are needed (Albert and Othmer, 2003).

In general, we may say that node i updates its state at times:

$$T_i^1, T_i^2, \ldots, T_i^k, \ldots, k \in \mathbb{N}_0$$

and the local variables, x_i, are updated according to

$$x_i[T_i^k] = f_i(x_1[\tau_{1i}^k], \ldots, x_N[\tau_{Ni}^k]), \tag{2}$$

where τ_{ji}^k is defined as

$$\tau_{ji}^k = \text{the latest available communication to node } i, \text{ from node } j.$$

There is usually a distinction between *totally* or *partially* asynchronous algorithms: the latter impose an updating constraint (every variable is updated at least once in any interval of a fixed length), while the former simply ensure that a variable is updated infinitely many times.

In a first numerical experiment we consider a totally asynchronous algorithm, with the highest degree of individual variability in each process' duration. The time unit of the synchronous model is randomly perturbed, so that the set of updating times for each node i $(1 \leqslant i \leqslant N)$ is of the form

$$T_i^{k+1} = T_i^k + 1 + \varepsilon r_i^k, \quad k \in \mathbb{N},$$

where r_i^k are random numbers generated at each iteration, out of a uniform distribution in the interval $[-1, 1]$. The value $\varepsilon \in [0, 1)$ is the magnitude of the perturbation (the case $\varepsilon = 0$ coincides with the synchronous algorithm). At any given time t, the next node(s) to be updated is(are) j such that $T_j^\ell = \min_{i,k}\{T_i^k \geqslant t\}$, for some ℓ. Since the duration of synthesis and decay processes is not known, through this algorithm one may randomly explore the space of all possible time-scales and update orders, and derive the probability of different outcomes. The set of updating times $\{T_i^k, k \in \mathbb{N}_0\}$ may vary with each execution of the algorithm, so an element of stochasticity is naturally introduced.

Always starting from the wild-type initial condition (1), this experiment was conducted over a wide range of perturbations $(10^{-12} \leqslant \varepsilon \leqslant 0.65)$, and 30 000 trials were executed for each ε. The results (see Fig. 3) show that all of the model's steady states may occur with a certain frequency: the wild-type pattern with only 57%, followed by the broad-striped pattern (24%) observed in heat-shock experiments and *ptc* mutants (Gallet et al., 2000) and by the pattern with no segmentation (15%) observed in *en*, *hh* or *wg* mutants (Tabata et al., 1992), the latter two corresponding to embryonic lethal phenotypes (Gallet et al., 2000).

We observe that each of the steady state patterns *occurs with a frequency which is independent of the value of* ε, *for* $\varepsilon < 0.15$. This may indicate that it is the order in which the protein and mRNA nodes are updated that determines the steady state pattern. In order to test this hypothesis, we designed a second experiment assuming that

(A1) Every node is updated exactly once during each unit time interval $(k, k+1]$ $(k = 0, 1, 2, \ldots)$, according to a given order ϕ^k.

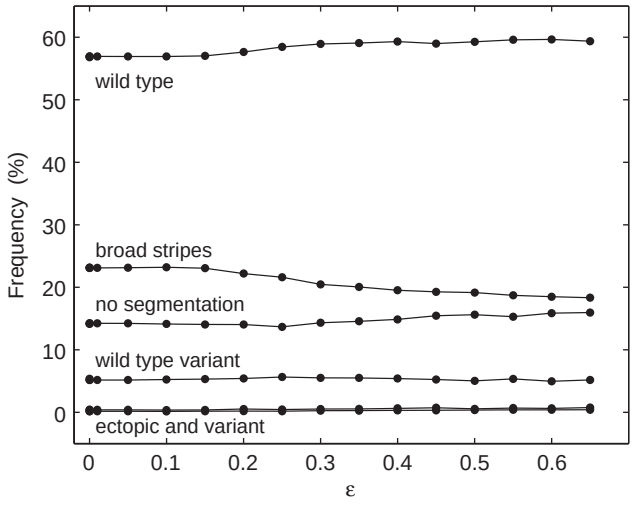

Fig. 3. Fragility of the regulatory network. With the totally asynchronous algorithm the wild-type initial state can lead to one of the six distinct steady states. Each . corresponds to an ε perturbation of the unit time step. Note that the $\varepsilon \to 0$ limit does not give the same results as a synchronous update, demonstrating the fundamental difference between synchronous and asynchronous models.

Table 3
The frequencies of the six steady states observed in the partially asynchronous model confirm those observed for the totally asynchronous model

Steady state	Incidence[%]
Wild-type	56
Broad stripes	24
No segmentation	15
Wild-type variant	4.2
Ectopic	0.98
Ectopic variant	0.68

The frequencies are computed from 30 000 executions.

This order ϕ^k is a permutation of $\{1,\ldots,N\}$, chosen randomly (again out of a uniform distribution over the set of all $N!$ possible permutations) at the beginning of the time unit k. Then we have

$$T_i^k = N(k-1) + \phi^k(i), \quad k \in \mathbb{N},$$

so that $\phi^k(j) < \phi^k(i)$ implies $T_j^k < T_i^k$, and node j is updated before node i. The partially asynchronous algorithm leads to the same patterns, with incidence rates very similar to those observed with the totally asynchronous algorithm (see Table 3).

These results indicate the fragility of the wild-type gene pattern with respect to changes in the time-scales of synthesis and decay processes. While more than half of the random time-scale combinations still lead to the expected outcome, a considerable percentage results in loss of the prepattern and an inviable final state.

3.1. Imbalance between CIA and CIR

Further analysis shows that the divergence from wild-type can be attributed to an imbalance between the two opposing Cubitus Interruptus transcription factors (CIA, CIR) in the posterior half of the parasegment. Indeed, the expression of CIA and CIR in both the broad stripes and the no segmentation patterns is clearly distinct from that in the wild-type pattern. In the next set of numerical experiments, we explore the effects of CIA/CIR expression in the formation of the final pattern.

In wild-type, the two Cubitus Interruptus proteins, CIA and CIR, are expressed in different cells of the posterior part of the parasegments, namely,

$$CIA_3 = 0, \quad CIA_4 = 1,$$

$$CIR_3 = 1, \quad CIR_4 = 0$$

and the maintenance of these complementary ON/OFF states is essential in the wild-type pattern. To investigate the effect of an imbalance between the two Cubitus Interruptus proteins, we considered two disruptive cases: the (transient) overexpression of CIR, or the (transient) overexpression of CIA and absence of CIR in both posterior cells.

More precisely, in the totally asynchronous algorithm (choosing $\varepsilon = 0.1$), we transiently imposed an expression pattern for the Cubitus proteins as follows:

(a) $CIA_{3,4}^t = 1$ and $CIR_{3,4}^t = 0$ for $t \in [3, 3+\tau]$;
(b) $CIR_{3,4}^t = 1$ for $t \in [3, 3+\tau]$,

where τ is the duration of the transient. The overexpression starts after three unit time steps. The duration of the transient was

$$\tau \in \{0, 0.3, 0.75, 1.5, 2.75, 3\},$$

so when $\tau = 0$ the results of the general totally asynchronous algorithm are recovered.

Our results show that even a small transient imbalance between CIA and CIR causes a clear bias towards a mutant state: the broad stripes mutant in case (a), or the no segmentation mutant in case (b) (see Fig. 4a and b, respectively). Thus any perturbation that leads to such an imbalance has as severe effects as a mutation in *ptc* (causing the broad striped pattern) or either of *en*, *wg* or *hh* (causing the non-segmented pattern).

These numerical experiments also open the way to many other questions: are there particular sequences that lead to a given steady-state? How is the evolution from the initial to steady-state? How robust is the asynchronous model with respect to initial conditions?

4. Time-scale separation uncovers robustness of the model

In both of the previous algorithms we assumed no bias towards a preferred protein/mRNA updating sequence and, as a result, an unrealistic divergence from the wild-type pattern is observed, with high incidence of inviable states. Based on the fact that post-translational processes such as protein conformational changes or complex formation usually have shorter durations than transcription, translation or mRNA decay, we introduce a distinct time-scale separation by choosing to update proteins first and mRNAs later. This leads to a model which is very robust, in the sense that the wild-type pattern occurs with a frequency of 87.5% and only one other steady-state is observed, the broad striped pattern, with a frequency of 12.5%. We completely characterize this model by theoretically showing that only two of the six steady states are possible (and occur with well-determined frequencies), and identifying the order of updates that leads to divergence from wild-type. We also show that the wild-type state is really an attractor for the system, while the pathway to the broad stripes state may show oscillatory cycles.

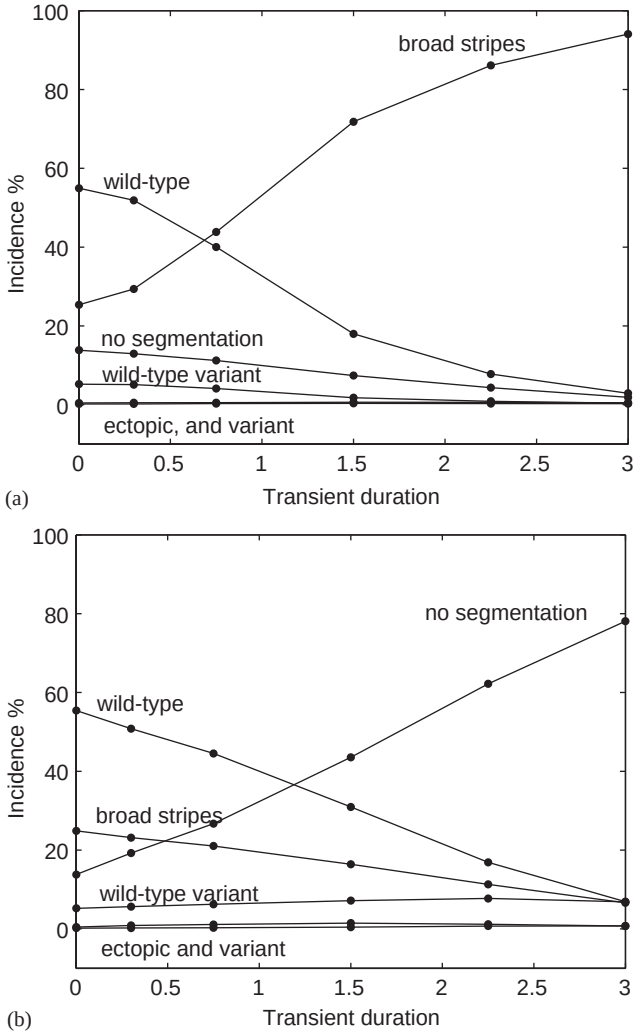

Fig. 4. Bias towards mutant states. The x-axis represents the duration of the transient, τ (in unit time steps). The incidence probabilities were computed over 20 000 trials. (a) The case $CIA^t_{3,4}=1$ and $CIR^t_{3,4}=0$ leads to the broad striped pattern. (b) The case $CIR^t_{3,4}=1$ leads to the no segmentation pattern.

Assuming that

(A2) All the proteins are updated before all the genes,

the kth iteration of the two-time-scale algorithm proceeds as follows:

(A3) At the beginning of the kth time unit, generate a random permutation, ϕ^k_{Prot} of $\{1,\ldots,L\}$, and a random permutation, ϕ^k_{mRNA} of $\{L+1,\ldots,N\}$ (using a uniform distribution over, respectively, the sets of $L!$ and $(N-L+1)!$ possible permutations). Then the N nodes are updated in the order given by $\phi^k=(\phi^k_{Prot},\phi^k_{mRNA})$, according to Eq. (2), with

$$\tau^k_{ji}=\begin{cases} T^{k-1}_j, & \phi^k(j)\leqslant\phi^k(i),\\ T^k_j, & \phi^k(j)>\phi^k(i).\end{cases}$$

As an example, suppose that

$$N=5,\quad L=3,\quad \phi^1_{Prot}=\{2,1,3\},\quad \phi^1_{mRNA}=\{5,4\}.$$

Then, $\phi^1=\{2,1,3,5,4\}$, and $T^1_1=2$, $T^1_2=1$, $T^1_3=3$, $T^1_4=5$, $T^1_5=4$. The nodes are updated as follows: (for simplicity of notation, we will write $x^k_i:=x_i[T^k_i]$):

$$x^1_2=f_2(x^0_1,x^0_2,x^0_3,x^0_4,x^0_5),$$

$$x^1_1=f_1(x^0_1,x^1_2,x^0_3,x^0_4,x^0_5),$$

$$x^1_3=f_3(x^1_1,x^1_2,x^0_3,x^0_4,x^0_5),$$

$$x^1_5=f_5(x^1_1,x^1_2,x^1_3,x^0_4,x^0_5),$$

$$x^1_4=f_4(x^1_1,x^1_2,x^1_3,x^0_4,x^1_5).$$

Some general inferences about the updating rules can be made. For example, the translation process only depends on the presence of the transcript, which is decided in the previous time unit, thus $Prot^t=mRNA^{t-1}$. The beginning of a transcription process depends on the presence of transcription factors, and since mRNAs are updated after proteins, $mRNA^t=Prot^t$. The outcome of post-translational processes depends on the order of updates, for example the rule for a binding process will be $Complex^t=Prot^{t_1}_1$ and $Prot^{t_2}_2$, where t_1 and t_2 can be either $t-1$ or t (see Table 4).

4.1. Two steady states

The trajectory of the system is thus defined by a sequence of permutations (obtained as described in (A3)), and the corresponding sequence of states:

$$\{\phi^k\},\ \{x^k\}\quad \text{for } k=0,1,2,\ldots. \tag{3}$$

We will show that for pre-patterns that satisfy $wg^0_4=1$, $ci^0_1=0$ and $ptc^0_1=0$ (which include the pattern observed in the wild-type at stage 8), the only possible steady states for system (3) are the wild-type pattern experimentally observed at stages 9–11, and the pattern with broad wg stripes. We assume that all the proteins are absent initially (at $T=0$), and that the *sloppy pair* gene is maintained at a constant value: $SLP_{1,2}=0$ and $SLP_{3,4}=1$. This pattern for SLP is responsible for permanent absence (or expression) of some of the segment polarity genes, and corresponding proteins, in certain cells of the parasegment. By direct inspection of the model, it follows that

$$wg^0_{1,2}=0\Rightarrow wg^T_{1,2}=0,\ WG^T_{1,2}=0\quad \text{for } T\geqslant 0 \tag{4}$$

$$en^T_{3,4}=0,\ EN^T_{3,4}=0\quad \text{for } T\geqslant 0, \tag{5}$$

$$hh^T_{3,4}=0,\ HH^T_{3,4}=0\quad \text{for } T\geqslant 0 \tag{6}$$

Table 4
Regulatory functions governing the states of segment polarity gene products in the two-time-scale asynchronous algorithm

Node	Boolean updating function in the two-time-scale algorithm
wg_i	$wg_i^t = (CIA_i^t$ and SLP_i^t and not $CIR_i^t)$ or $[wg_i^{t-1}$ and $(CIA_i^t$ or $SLP_i^t)$ and not $CIR_i^t]$
WG_i	$WG_i^t = wg_i^{t-1}$
en_i	$en_i^t = (WG_{i-1}^t$ or $WG_{i+1}^t)$ and not SLP_i^t
EN_i	$EN_i^t = en_i^{t-1}$
hh_i	$hh_i^t = EN_i^t$ and not CIR_i^t
HH_i	$HH_i^t = hh_i^{t-1}$
ptc_i	$ptc_i^t = CIA_i^t$ and not EN_i^t and not CIR_i^t
PTC_i	$PTC_i^t = ptc_i^{t-1}$ or $(PTC_i^{t-1}$ and not $HH_{i-1}^{t_1}$and not $HH_{i+1}^{t_2})$
ci_i	$ci_i^t =$ not EN_i^t
CI_i	$CI_i^t = ci_i^{t-1}$
CIA_i	$CIA_i^t = CI_i^{t_3}$ and (not $PTC_i^{t_4}$ or $HH_{i-1}^{t_5}$ or $HH_{i+1}^{t_6}$ or $hh_{i\pm1}^{t-1})$
CIR_i	$CIR_i^t = CI_i^{t_7}$ and $PTC_i^{t_8}$ and not $HH_{i-1}^{t_9}$ and not $HH_{i+1}^{t_{10}}$ and not $hh_{i\pm1}^{t-1}$

Each node is labeled by its biochemical symbol, subscripts signify cell number and superscripts signify time step. Although the updating time of each node varies, each function can be written by using the states of effector nodes at the previous or current time steps. The individual times $t_1 \ldots t_{10}$ can take the values $\{t-1, t\}$.

and

$$ci_{3,4}^0 = 1 \Rightarrow ci_{3,4}^T = 1 \text{ for } T \geqslant 0 \text{ and } CI_{3,4}^T = 1 \text{ for } T \geqslant 1, \quad (7)$$

$$ci_{3,4}^0 = 0 \Rightarrow ci_{3,4}^T = 1 \text{ for } T \geqslant 1 \text{ and } CI_{3,4}^T = 1 \text{ for } T \geqslant 2. \quad (8)$$

The next statement reflects the fact that the effect of wg_4 activating en_1 propagates to inhibit ci_1 which then eliminates all forms of CI from the first cell.

Fact 1. *Assume that* $wg_4^0 = 1$, $ci_1^0 = 0$ *and* $ptc_1^0 = 0$. *For any* $T \geqslant 0$, *if* $wg_4^t = 1$ *for all* $0 \leqslant t \leqslant T$, *then* $CI_1^t = 0$ *for all* $3 \leqslant t \leqslant T+3$ *and* $CIR_1^t = 0$ *for all* $0 \leqslant t \leqslant T+3$.

Proposition 4.1. *Assume* $wg_4^0 = 1$, $ci_1^0 = 0$ *and* $ptc_1^0 = 0$. *Under assumptions* (A1)–(A2), $wg_4^T = 1$, *for all* $T \geqslant 0$.

Proof. We will argue by contradiction. Suppose that there do exist times $t \geqslant 1$ with $wg_4^t = 0$, and let T be the minimum of such times, that is,

$$wg_4^T = 0 \quad \text{and} \quad wg_4^t = 1 \quad \text{for all } 0 \leqslant t < T.$$

From the model's equations, together with assumptions (A1)–(A2):

$$WG_4^t = wg_4^{t-1},$$

$$wg_4^t = (CIA_4^t \text{ and not } CIR_4^t) \text{ or } (wg_4^{t-1} \text{ and not } CIR_4^t)$$

for all $t \geqslant 1$. So, it follows that

$$WG_4^t = 1 \quad \text{for all } 0 \leqslant t \leqslant T, \quad (9)$$

$$CIR_4^t = 0 \quad \text{for all } 0 \leqslant t < T \quad \text{and} \quad CIR_4^T = 1. \quad (10)$$

Now, from Fact 1 it also follows that

$$CIR_1^t = 0 \quad \text{for all } 0 \leqslant t \leqslant T+2. \quad (11)$$

The equation for CIR_4 is

$$CIR_4^t = CI_4^{t_d} \text{ and not } [\text{not } PTC_4^{t_a} \text{ or } HH_3^{t_c} \text{ or } HH_1^{t_b} \text{ or } hh_3^{t-1} \text{ or } hh_1^{t-1}] \quad (12)$$

(where $t_a, \ldots, t_d \in \{t, t-1\}$ depend on the permutation ϕ^t). Recall also that

$$hh_1^t = EN_1^t \text{ and not } CIR_1^t, \quad (13)$$

$$EN_1^t = en_1^{t-1}, \quad (14)$$

$$en_1^{t-1} = WG_4^{t-1} \text{or} WG_2^{t-1}. \quad (15)$$

From (12):

$$CIR_4^T = 1 \Rightarrow hh_1^{T-1} = 0$$

and then from (11) and (13):

$$hh_1^{T-1} = 0 \Rightarrow EN_1^{T-1} = 0.$$

Now by Eqs. (14) and (15):

$$EN_1^{T-1} = 0 \Rightarrow en_1^{T-2} = 0 \Rightarrow WG_4^{T-2} = 0 \text{ and } WG_2^{T-2} = 0,$$

which contradicts Eq. (9). Thus, it must be that $wg_4^T = 1$ for all times T, as we wanted to show. □

The following are now immediate conclusions from the model:

Corollary 4.2. $CIR_4^T = 0$ *for all* $T \geqslant 0$, $en_1^T = 1$ *and* $WG_4^T = 1$ *for all* $T \geqslant 1$. $EN_1^T = 1$, $ci_1^T = 0$ *and* $hh_1^T = 1$ *for all* $T \geqslant 2$. $CI_1^T = 0$ *and* $HH_1^T = 1$ *for all* $T \geqslant 3$. *And finally*, $CIA_1^T = CIR_1^T = 0$ *for all* $T \geqslant 4$.

Corollary 4.3. $ptc_1^T = 0$ *and* $PTC_1^T = 0$ *for all* $T \geqslant 0$, *and* $CIR_2^T = 0$ *for all* $T \geqslant 3$.

In conclusion, from Proposition 4.1 it is clear that neither the no segmentation nor the two ectopic patterns are steady states of system (3) under assumptions (A1)–(A2), because all of these states imply $wg_4 = 0$. In addition, Corollary 4.3 shows that the wild-type variant, where *PTC* is ubiquitous, cannot be a steady-state. Also, any of the states with $wg_{1,2} = 1$ is immediately prevented by the initial condition (4). This leaves only the "regular" wild-type or the mutant with broad *wg* stripes.

4.2. Divergence from wild-type

Under assumptions (A1)–(A2), divergence from the wild-type pattern occurs if and only if the first permutation (in particular ϕ^1_{Prot}) is of a particular form. Thus, convergence (or divergence) to the wild-type pattern is decided at the first iterate ($T = 1$).

Recall that the wild-type pattern requires *wingless* not to be expressed in the third cell ($wg_3 = 0$). The next Fact (proved in the Appendix) essentially says that a stable $wg_3 = 0$ induces the absence of both *engrailed*, *hedgehog* in the second cell, as well as the absence of CIA_3, and maintains the expression of PTC_3.

Fact 2. *Assume* $ptc_3^0 = 1$ *and* $en_2^0 = 0$.

(a) *Let* $T \geqslant 1$. *If* $wg_3^t = 0$ *for all* $0 \leqslant t \leqslant T$, *then*

$$en_2^t = 0, \quad EN_2^t = 0, \quad 0 \leqslant t \leqslant T+2,$$

$$hh_2^t = 0, \quad 1 \leqslant t \leqslant T+2,$$

$$HH_2^t = 0, \quad 1 \leqslant t \leqslant T+3,$$

$$PTC_3^t = 1, \quad 1 \leqslant t \leqslant T+3, \text{ and}$$

$$CIA_3^t = 0, \quad CIR_3^t = 1, \quad 2 \leqslant t \leqslant T+3.$$

(b) *Furthermore, if* $ci_3^0 = 0$, *then also* $CIA_3^1 = 0$ *and part* (a) *holds for any* $T \geqslant 0$.

With the help of this fact, we establish that wg_3 may become expressed only at the first iterate or else it is never expressed. Thus the two time-scale model provides a strong natural restriction on the formation of an inviable state: if $wg_3^1 = 0$, then $wg_3^T = 0$ for all $T \geqslant 0$, implying that such trajectories will never converge to the broad striped pattern.

Proposition 4.4. *Assume that the initial condition satisfies* $wg_3^0 = 0$, $ptc_3^0 = 1$, $hh_{2,4}^0 = 0$, *and* $ci_3^0 = 1$. *Then* $wg_3^{T_1} = 1$ *and* $wg_3^T = 0$ *for all* $0 \leqslant T < T_1$, *only if* $T_1 = 1$.

Proof. To obtain $wg_3^T = 1$ with $wg_3^{T-1} = 0$ it is necessary that

$$wg_3^T = CIA_3^T \text{ and not } CIR_3^T \Rightarrow CIA_3^T = 1 \text{ and } CIR_3^T = 0.$$

But, if $wg_3^T = 0$ for $T = 1$, then, by Fact 2, the activator CIA_3 is zero for $T = 2, 3, 4$. Then (by induction on T) expression of wg_3 is prevented at any later time. □

In addition, it is possible to completely characterize the updating permutation (ϕ^1) that leads to $wg_3^1 = 1$ and, as a consequence, exactly compute the probability of divergence from (or convergence to) the wild-type steady state (Section 4.3).

Proposition 4.5. *Assume that assumptions* (A1) *and* (A2) *hold. Assume that the initial condition satisfies* $wg_3^0 = 0$, $ptc_3^0 = 1$ *and* $hh_{2,4}^0 = 0$.

(a) *If* $ci_3^0 = 0$, *then* $wg_3^1 = 0$.
(b) *If* $ci_3^0 = 1$, *then* $wg_3^1 = 1$ *if and only if the permutation* ϕ^1 *satisfies the following sequence among the proteins CI, CIA, CIR and PTC*:

$$\begin{array}{llll} CIR_3 & CI_3 & CIA_3 & PTC_3, \\ CI_3 & CIR_3 & CIA_3 & PTC_3, \\ CI_3 & CIA_3 & CIR_3 & PTC_3, \end{array} \tag{16}$$

while the other proteins may appear in any of the remaining slots.

Proof. Part (a) follows immediately from Fact 2(b). To prove part (b), we start by noticing that, because $SLP_3 = 1$ and $wg_3^0 = 0$,

$$wg_3^1 = CIA_3^1 \text{ and not } CIR_3^1,$$

so that

$$wg_3^1 = 1 \Leftrightarrow CIA_3^1 = 1 \text{ and } CIR_3^1 = 0.$$

Following assumptions (A1)–(A2), the model's equations for CIA_3^1 and CIR_3^1 are given by

$$CIA_3^1 = CI_3^{t_a} \text{ and [not } PTC_3^{t_b} \text{ or } HH_2^{t_c} \text{ or } HH_4^{t_d} \text{ or } hh_2^0 \text{ or } hh_4^0],$$

$$CIR_3^1 = CI_3^{s_a} \text{ and not [not } PTC_3^{s_b} \text{ or } HH_2^{s_c} \text{ or } HH_4^{s_d} \text{ or } hh_2^0 \text{ or } hh_4^0],$$

where $t_a, \ldots, s_a \in \{0, 1\}$ and depend on the permutation ϕ^1. These expressions may be simplified by observing that: (a) $hh_{2,4}^0 = 0$, and thus also (b) $HH_{2,4}^{0,1} = 0$. Therefore,

$$CIA_3^1 = CI_3^{t_a} \text{ and not } PTC_3^{t_b},$$

$$CIR_3^1 = CI_3^{s_a} \text{ and } PTC_3^{s_b}.$$

The values for $CI_3^{0,1}$ and $PTC_3^{0,1}$ are determined by

1. $CI_3^0 = 0$ and $CI_3^1 = ci_3^0 = 1$,
2. $PTC_3^0 = 0$ and $PTC_3^1 = ptc_3^0$ or $[\cdots] = 1$, since $ptc_3^0 = 1$

and recall that both $CIA_i^0 = 0$ and $CIR_i^0 = 0$. Therefore, *it is necessary that* CI_3 *is updated before* CIA_3 *and* PTC_3 *is updated after* CIA_3, because otherwise $CIA_3^1 = 0$. Finally, CIR_3 *must be updated before* PTC_3, because otherwise $CIR_3^1 = 1$. In other words:

$$t_a = 1, \quad t_b = 0, \quad s_a \in \{0, 1\}, \quad s_b = 0.$$

It is easy to see that any of sequences (16) is also sufficient to obtain $wg_3^1 = 1$. □

Finally, we will show that whenever wg_3 becomes expressed at time $T = 1$, it is afterwards periodically expressed, every third step. Such trajectories cannot converge to the wild-type pattern. In other words, initial permutations of form (16) are not included in the basin of attraction of the wild-type pattern.

Proposition 4.6. *Assume* $wg_4^0 = 1$, $ci_1^0 = 0$ *and* $ptc_1^0 = 0$. *For any* $T \geqslant 1$, *if* $wg_3^T = 1$, *then* $wg_3^{T+3} = 1$.

Proof. Recall that, by Proposition 4.1, $wg_4^t = 1$ for all $t \geqslant 0$. By Corollary 4.3, $CIR_2^t = 0$ for all $t \geqslant 3$.

Now, pick any $T \geqslant 1$ and assume that $wg_3^T = 1$. Then

$$WG_3^{T+1} = 1 \Rightarrow en_2^{T+1} = 1$$
$$\Rightarrow EN_2^{T+2} = 1 \Rightarrow hh_2^{T+2} = wg_3^T = 1,$$

where the last implication follows from Fact 4(in Appendix A). Then (using either (7) or (8))

$$CIA_3^{T+3} = \text{not } PTC_3^{t_a} \text{ or } HH_2^{t_b} \text{ or } hh_2^{T+2},$$

$$CIR_3^{T+3} = PTC_3^{s_a} \text{ and not } HH_2^{s_b} \text{ and not } hh_2^{T+2}$$

(where $t_a, t_b, s_a, s_b \in \{T+2, T+3\}$, and depend on the permutation ϕ^{T+3}). So $hh_2^{T+2} = 1$ implies

$$CIA_3^{T+3} = 1 \quad \text{and} \quad CIR_3^{T+3} = 0$$

and therefore $wg_3^{T+3} = 1$, as we wanted to show. □

Whenever wg_3 is not expressed, some other nodes also stabilize, after the appropriate number of iterations. These are summarized next.

Corollary 4.7. *Assume that* $wg_3^t = 0$ *for all* $t \geqslant 0$. *Then* $WG_3^t = 0$, $en_2^t = 0$ *for all* $t \geqslant 1$. $EN_2^t = 0$, $hh_2^t = 0$ *and* $ci_2^t = 1$ *for all* $t \geqslant 2$. *Finally*, $HH_2^t = 0$ *and* $CI_2^t = 1$ *for all* $t \geqslant 3$.

4.3. Probability of convergence to wild-type

The wild-type pattern is in fact an attractor for the asynchronous model: every trajectory which is not of form (16) converges to the wild-type pattern (see Appendix B). The probability that this happens is therefore determined by counting all the possible states of form (16):

$$\text{Prob(wild type)} = 1 - \frac{\#\text{ permutations as in Eq. 16}}{\text{Total }\#\text{ permutations}}.$$

Let L be the number of protein nodes to be updated at each iterate (there are 7 proteins in each of the four cells so $L = 28$). Out of the L proteins, only 4 need to satisfy one of particular sequences (16) in their relative positions. So let M_L be the number of possible permutations satisfying any of sequences (16). Then

$$\text{Prob(wild type)} = 1 - \frac{M_L\,(L-4)!}{L!}.$$

The next proposition is proved in the Appendix B and shows that, in fact, this is a constant number, independent of L. That is,

$$\text{Prob(wild type)} = 0.875.$$

Proposition 4.8. *For any* $L \geqslant 4$,

$$M_L = \frac{3!}{2}\sum_{P=4}^{L}\binom{P-1}{3} = \frac{1}{2}\sum_{j=1}^{L-3} j(j+1)(j+2)$$

and

$$\frac{M_L\,(L-4)!}{L!} = \frac{1}{8}$$

5. A Markov chain process

As a Boolean model, there are only a finite set, say $\mathscr{S}$, of distinct states (in the total state space $\{0,1\}^N$) reachable by the system. Starting from any state $\mathscr{S}_a \in \mathscr{S}$, each permutation ϕ of $\{1,\ldots,N\}$ takes the system to some other state $\mathscr{S}_b \in \mathscr{S}$. It is possible to theoretically identify all the distinct intermediate and final states of the system as well as all the possible transitions after one iteration. Thus the asynchronous algorithm consisting of the N node functions (2) together with assumptions (A1)–(A3) may be characterized as a Markov chain process, by identifying the d distinct states

$$\mathscr{S} = \{\mathscr{S}_1, \mathscr{S}_2, \ldots, \mathscr{S}_d\}$$

and the $d \times d$ transition matrix P, where each entry P_{ij} denotes the probability of a transition from state S_i to state S_j. The probabilities P_{ij} are simply the fraction of the total number of permutations that (in one iterate) transform the state $\mathscr{S}_i$ into state $\mathscr{S}_j$. The matrix P is a stochastic matrix, since all its rows add up to 1. A state $\mathscr{S}_a$ with the property that all permutations leave the state unchanged (that is, $P_{aa} = 1$ and $P_{aj} = 0$ for $j \neq 0$) is called an *absorption state* of the Markov chain, and it is also a steady-state of system (2). In the asynchronous model there are only two absorption states, corresponding to the wild-type and broad *wingless* stripe mutant patterns, as described above. Isolating the two rows and columns that correspond to these absorption states, the transition matrix may be partitioned as

$$P = \begin{bmatrix} P_a & 0 \\ R_a & \bar{P} \end{bmatrix},$$

where $\bar{P}$ is of size $(d-2) \times (d-2)$. It is a well know result(see any standard book on probability theory, for instance (Feller, 1970)) that $I - \bar{P}$ is an invertible

matrix, and

$$\begin{bmatrix} \bar{T}_1 \\ \bar{T}_2 \\ \vdots \\ \bar{T}_{d-2} \end{bmatrix} = (I - \bar{P})^{-1} \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix},$$

or $\bar{T}_i = 1 + \sum_{j=1}^{d-2} p_{ij}\, \bar{T}_j$. The values $\bar{T}_i$ provide an estimated time for absorption when the chain starts from state $\mathscr{S}_i$ (if a is an absorption state, then $\bar{T}_a = 0$). In Fig. 5, a schematic diagram of transitions is shown, together with probabilities and estimated times for absorption. This diagram was obtained from a simulation starting with the initial wild-type pattern (observed at stage 8 of the embrionic development), and following assumptions (A1)–(A3), as well as the additional

(A4) Each protein or gene is updated simultaneously in the four cells,

meaning that there is no cell-to-cell variation in the duration of molecular processes. In this case, the total number of possible permutations is a manageable $7! \times 5! = 604800$:

$$\Phi = \{\phi = (\phi_{Prot}, \phi_{mRNA}):$$
$$\phi_{Prot} \text{ is a permutation of } \{1,2,3,4,5,6,7\},$$
$$\phi_{mRNA} \text{ is a permutation of } \{1,2,3,4,5\}\}.$$

The total number of distinct states of the Markov chain (under assumptions (A1)–(A4)) is $d = 48$. The transition probabilities matrix was computed exactly, by counting all the $7! \times 5!$ transitions from each of the 48 states.

It is clear from this diagram that the decision between convergence to the wild-type or the mutant patterns is indeed decided at the first iteration, in agreement with Propositions 4.1 and 4.4. Furthermore, the diagram shows the possibility of periodic oscillations (of period at least three) in the mutant branch (see also Proposition 4.6). Although in practice the probability of a limit cycle is very small, this prevents (theoretical) convergence to the mutant state, and considerably increases the absorption times to the mutant state. The robustness of the two-time-scale model is illustrated by the fast convergence to the wild-type pattern (expected time to convergence is 4 steps), contrasted with the long and oscillation-strewn path toward the broad striped pattern (expected time to convergence is 15 steps).

6. Identifying minimal pre-patterns

A necessary condition for convergence to the wild-type is that $ptc_3^0 = 1$. Otherwise the trajectory immediately fails to enter the basin of attraction of the wild-type state:

Fact 3. *Assume* $ptc_3^0 = 0$. *Then* $wg_3^T = 1$, *for some* $T \in \{1,2,3\}$.

Proof. Note that $ptc_3^0 = 0$ implies $PTC_3^1 = 0$. Using Eqs. (6) and (7), (8), the equations for ptc_3 and CIA_3 simplify to

$$CIA_3^t = \text{not } CIR_3^t = \text{not } PTC_3^{t_a} \text{or } HH_2^{t_b} \text{ or } hh_2^{t-1},\ t \geqslant \tau,$$

$$ptc_3^t = CIA_3^t \text{ and not } CIR_3^t = CIA_3^t,\ t \geqslant 1,$$

$$PTC_3^t = ptc_3^{t-1} \text{ or } [PTC_3^{t-1} \text{ and not } HH_2^{t_a} \text{ and not } hh_2^{t-1}],\ t \geqslant 1,$$

where $\tau = 2$ (respectively, $\tau = 3$), if $ci_3^0 = 1$ (respectively, $ci_3^0 = 0$). Consider first the case $ci_3^0 = 1$. The activator protein will be turned on either at the first or second iterations: CIA_3 may become activated at $t = 1$ because $PTC_3^1 = 0$, (if the permutation ϕ^1 is such that CI_3 is updated before CIA_3). If CIA_3 is not actived at $t = 1$, then it certainly is actived at $t = 2$ (because $CIA_3^1 = 0$ implies $ptc_3^1 = 0$ and $PTC_3^2 = 0$). A similar argument shows that $CIR_3^{1,2} = 0$.

Consider next the case $ci_3^0 = 0$: we have $CIA_3^1 = CIR_3^1 = 0$, and CIA_3 is turned on at the second or third iterations, by a similar argument as before. Therefore, the *wingless* gene is also expressed after CIA_3, at $T = 1, 2$ or 3. □

Another necessary condition for convergence to wild-type is that

$$wg_4^0 = 1 \text{ or } en_1^0 = 1 \text{ or } ci_4^0 = 1.$$

Otherwise, the trajectories cannot converge to the wild-type nor to the mutant steady states, In this case, the only possible steady-state is a "lethal" state, where expression of *PTC*, *ci*, *CIA* and *CIR* is ubiqitous and all others are absent.

Proposition 6.1. *Assume* $wg_{3,4}^0 = 0$, $en_1^0 = 0$, $ptc_3^0 = 1$ *and* $ci_{3,4}^0 = 0$. *Then*

(a) $wg_3 = 0$ *for all* $t \geqslant 0$.
(b) $PTC_1^t = 1$ *for all* $t \geqslant 4$.
(c) $wg_4 = 0$ *for infinitely many* t.

Proof. Part (a) follows from Fact 2(b), for any $T \geqslant 0$:

$$wg_3^t = 0, \quad t \leqslant T \Rightarrow CIA_3^t = 0, \quad 0 < t \leqslant T + 3.$$

Thus, by induction, $wg_3^t = 0$ for $t \geqslant 0$.

Part (a) and Fact 4(Appendix 4) imply $hh_2^t = 0$ and $HH_2^t = 0$ for all $t > 0$, so that

$$ptc_1^t = CIA_1^t \text{ and not } CIR_1^t \text{ and not } EN_1^t,$$

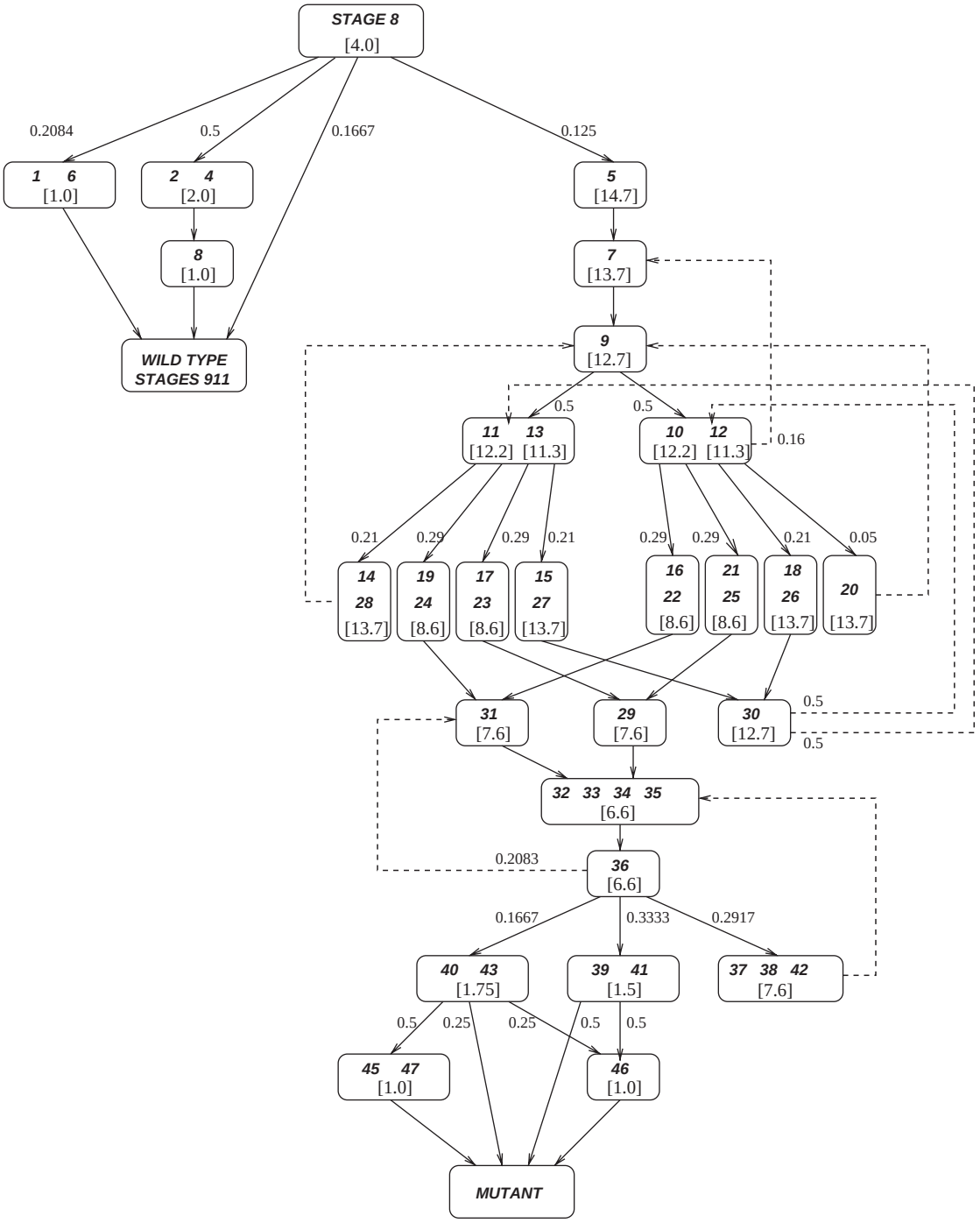

Fig. 5. Robustness of the regulatory network modeled with the two-time-scale algorithm. There are 48 states reachable from the wild-type initial state. The arrows are labeled by the transition probabilities between states (if unlabeled, the probabilities are 1), and the expected times to absorption into the corresponding steady-state are indicated between square brackets.

$PTC_1^t = ptc_1^{t-1}$ or PTC_1^{t-1}, $\quad t \geqslant 1$,

$CIA_1^t = CI_1^{t_a}$ and not $PTC_1^{t_b}$, $\quad t \geqslant 1$.

If $ptc_1^0 = 1$, then it is clear that $PTC_1^t = 1$ for all $t \geqslant 1$. Consider the case $ptc_1^0 = 0$. Then $PTC_1 = 0$ as long as $ptc_1 = 0$. Note that $ci_4^0 = 0$ implies $CI_4^1 = 0$, $CIA_4^1 = 0$ and also $wg_4^1 = 0$. Then $wg_4^{0,1} = 0$ and $en_1^{0,1} = 0$ imply $EN_1^{1,2,3} = 0$ and $CI_1^{2,3,4} = 1$. So, either at $T = 2$ or $T = 3$, we will have $CIA_1^T = 1$, and therefore also $ptc_1^T = 1$ and $PTC_1^{T+1} = 1$. Thus, $PTC_4^t = 1$ for $t \geqslant 4$, proving part (b).

Finally, to argue by contradiction, suppose that $wg_4^t = 1$ for $t \geqslant T_a$. Then by Corollary 4.3, $PTC_1^t = 0$

for $t \geqslant T_b > T_a$, which contradicts part (b). Hence, wg_4 cannot become permanently expressed. □

By Proposition 4.1, together with Fact 2(b), a *sufficient* condition for convergence to wild-type is

$$wg_4^0 = 1, \quad ptc_3^0 = 1, \quad ptc_1^0 = 0, \quad ci_{1,3}^0 = 0.$$

Another *sufficient* condition (which allows the presence of *cubitus* in the third cell) is

$$wg_4^0 = 1, \quad ptc_3^0 = 1, \quad PTC_3^0 = 1.$$

The argument in the proof of Proposition 4.5, shows that, if $PTC_3^0 = 1$ then $wg_3^1 = 0$. Then, by Proposition 4.4, it follows that $wg_3^t = 0$ for all times.

In conclusion, while the wild-type initial state allows for an ambiguity in the final states, we find that a remarkably minimal prepattern, consisting of wg_4 and ptc_3, is sufficient to guarantee the convergence to the wild-type steady-state. In other words, the initiation of two genes in two cells is enough to compensate for initiation delays in any and all other genes, irrespectively of the variations in individual synthesis and decay processes. This suggests a remarkable error correcting ability of the segment polarity gene control network.

7. Conclusions

In summary, we proposed an intuitive and practical way of introducing stochasticity in qualitative models of gene regulation. We explored three possible ways of incorporating the variability of transcription, translation, post-translational modification and decay processes (see Table 5 for a comparison between the synchronous and three asynchronous algorithms). Applying our methods on a previously introduced model of the *Drosophila* segment polarity genes gave us new insights into the dynamics and function of the interactions among the segment polarity genes and, through it, into the robustness of the embryonic segmentation process. Our results suggest that unrestricted variability in synthesis/decay/transformation time-scales can lead to a divergence from the wild-type development process, with an expected divergence probability of 45%. On the other hand, if the duration of post-translational transformations is consistently less than the duration of transcription, translation and mRNA/protein half-lives, the wild-type steady-state will be achieved with a high probability, despite significant variability in individual process durations. We find that a remarkably sparse prepattern is sufficient to ensure the convergence to the wild-type steady state of these genes. This dual behavior, robustness to changes in the initial state but fragility with respect to temporal variability, is reminiscent of Highly Optimized Tolerance, a feature of highly structured, non-generic complex systems with robust, yet fragile external dynamics (Carlson and Doyle, 2002). Similar robust-yet-fragile features have also been found in the of structure of diverse networks (Jeong et al., 2000, 2001; Albert et al., 2000).

All our algorithms concur in suggesting that the divergence from wild-type can be attributed to an imbalance between the two opposing Cubitus

Table 5
Comparison of synchronous and asynchronous algorithms

	Synchronous	Totally asynchronous	Random order	Two-time-scale
Assume	Nodes are updated at multiples of the unit time interval	The time between updates is perturbed in a range $\pm\varepsilon$	Each node is updated at a randomly selected point of the unit time interval	In each time interval proteins are updated first, then mRNAs
Update	$T^k = k$	$T^k = T^{k-1} + 1 + \varepsilon r^k$	$T^k = N(k-1) + \phi^k$ ϕ^k-node permutation	$T^k = N(k-1) + \phi_{tt}^k$ $\phi_{tt}^k \in (\phi_{Prot}^k, \phi_{mRNA}^k)$
Pros	Correctly identifies all steady states Can be solved analytically	Allows for unlimited variability in process durations	Does not depend on any perturbation parameter ε	Allows separation of post-translational and pre-translational processes
Cons	Dynamics is unrealistic	Can have unrealistically short transcription and translation times		Only useful when process durations can be separated
Results	Prepattern errors can be corrected	Divergence from the wild-type process is possible Cause: imbalance between two transcription factors		Development is stable if PTC is prepatterned

Interruptus transcription factors (CIA, CIR) in the posterior half of the parasegment. Thus the complementary regulation and pattern of these opposing transcription factors (Aza-Blanc and Kornberg, 1999) is a vital requirement for the correct functioning of the segment polarity gene network. The totally asynchronous algorithm predicts that perturbations to the post-translational modification of Cubitus Interruptus can have effects as severe as mutations: a transient over-expression of CIR leads to the pattern with no segmentation, while transient expression of CIA and not CIR leads to the broad striped pattern. With the two-time-scale algorithm we find that the condition for the divergence from the wild-type pattern is that, in the third cell of the parasegment, the post-translational modification of CI precedes the synthesis of the Patched protein. The biological realization of this condition appears unlikely, since PTC is documented as being ubiquitously expressed during cellularization (stage 5) (Taylor et al., 1993), while the post-translational modification of CI requires SMO that is only weakly expressed until stage 8 (Alcedo et al., 2000). Our model predicts that if for any reason the PTC protein is absent in the period when the pair-rule proteins decay and the regulation between the segment polarity genes starts, the wild-type expression pattern is unreachable.

Our methods combine the benefits of discrete-state models with a continuum in time-scales. In the absence of quantitative information, we considered every possible time-scale or update order, but as the two-time-scale model demonstrates, existing information can be easily incorporated. We were able to describe the system in a rigorous mathematical way, to identify the relatively few types of behavior possible in the system (the attractors in state space) and to theoretically prove the convergence toward these states. Our results underscore that predictive mathematical modeling is possible despite the scarcity of quantitative information on gene regulatory processes.

Acknowledgements

The work of M.C. was supported in part by NIH Grants P20 GM64375 and Sanofi-aventis. R.A. gratefully acknowledges an Alfred P. Sloan Research Fellowship. The work of E.D.S. was supported in part by NSF Grant CCR-0206789 and NIH Grants P20 GM64375 and R01 GM46383.

Appendix A. Additional proofs

Proof of Fact 1. For $T=0$, the statement follows directly from the model's equations and by using the assumptions (A1)–(A2) repeatedly:

$$WG_4^1=1,\quad en_1^1=1,\quad EN_1^2=1,$$

as well as

$$ci_1^{0,2}=0,\quad CI_1^{0,1,3}=0,\quad CIA_1^{0,1}=0,\quad CIR_1^{0,1}=0,$$
$$PTC_1^{0,1}=0.$$

We have (using (6))

$$ptc_1^t=CIA_1^t \text{ and not } CIR_1^t \text{ and not } EN_1^t,$$

$$PTC_1^t=ptc_1^{t-1} \text{ or } [PTC_1^{t_b} \text{ and not } HH_2^{t_c}],$$

$$CIR_1^t=CI_1^{t_a} \text{ and } PTC_1^{t_b} \text{ and not } HH_2^{t_c} \text{ and not } hh_2^{t-1},$$

so that we conclude

$$ptc_1^{0,1,2}=0,\quad PTC_1^{2,3}=0,\quad CIR_1^{2,3}=0.$$

We next prove the fact by induction. First note that, for any $t\geqslant 0$:

$$wg_4^t=1\Rightarrow WG_4^{t+1}=1\Rightarrow en_1^{t+1}=1\Rightarrow EN_1^{t+2}=1$$

and this implies

$$ci_1^{t+2}=0\Rightarrow CI_1^{t+3}=0. \tag{17}$$

Now assume that the fact holds for some $T\geqslant 1$ and that

$$wg_4^t=1,\quad 0\leqslant t\leqslant T.$$

By the induction hypothesis, we know that

$$CI_1^t=0,\ 3\leqslant t\leqslant T+2,\ \text{and } CIR_1^t=0,\ 0\leqslant t\leqslant T+2,$$

By (17), $wg_4^T=1$ implies $CI_1^{T+3}=0$, and this together with $CI_1^{T+2}=0$ also guarantees that $CIR_1^{T+3}=0$, as we wanted to show. □

Proof of Fact 2. To prove part (a), assume that $wg_3^t=0$ for all $0\leqslant t<T$, with $T\geqslant 1$. Since $en_2^0=0$, then $EN_2^{0,1}=0$ and $hh_2^1=0$, $HH_2^{1,2}=0$. For $t\geqslant 3$ apply Fact 4 to obtain the desired value for hh_2 and HH_2. Note that $ptc_3^0=1$ implies $PTC_3^1=1$ and, together with $HH_2^{1,2}=0$, also $PTC_3^2=1$; then the value of PTC_3 follows from Fact 5. For $T\geqslant 2$, and using (6) and (8), we have

$$CIA_3^t=\text{not } PTC_3^{t_b} \text{ or } HH_2^{t_c} \text{ or } hh_2^{t-1},$$

$$CIR_3^t=PTC_3^{s_b} \text{ and not } HH_2^{s_c} \text{ and not } hh_2^{t-1}$$

so, the values for hh_2, HH_2 and PTC_3, indeed imply that $CIA_3^t=0$ and $CIR_3^t=1$, for $2\leqslant t\leqslant T+3$.

To prove part (b), note that if $ci_3^0=0$, then also $ci_3^{0,1}=0$ and hence $CIA_3^1=0$. But now $CIA_3^1=0$ together with $wg_3^0=0$ immediately imply that $wg_3^1=0$, and therefore, the results in part (a) are valid for all $T\geqslant 0$. □

Fact 4.

$hh_2^{T+2}=HH_2^{T+3}=wg_3^T$ and not CIR_2^{T+2} for all $T\geqslant 0$.

In particular, if $CIR_2^t=0$ for all $t\geqslant 3$, then $hh_2^{T+2}=HH_2^{T+3}=wg_3^T$, for all $T\geqslant 1$.

Proof. Given any $t \geqslant 0$ it is easy to see that

$WG_3^{t+1} = wg_3^t$

$en_2^{t+1} = WG_1^{t+1}$ or $WG_3^{t+1} \equiv WG_3^{t+1}$,

$EN_2^{t+2} = en_2^{t+1}$,

$hh_2^{t+2} = EN_2^{t+2}$ and not CIR_2^{t+2},

$HH_2^{t+3} = hh_2^{t+2}$,

where the equation for en_2^{t+1} follows from (4). □

Fact 5. *$PTC_3^T = 1$ and $PTC_3^{T+1} = 0$, for some $T > 0$, only if $wg_3^t = 1$ for some $t \in \{T-3, T-2\}$ (chosen according to the permutation ϕ^{T-2}).*

Proof. To see this, simply notice that

$$\begin{aligned} PTC_3^{T+1} &= ptc_3^T \text{ or } [PTC_3^T \text{ and not } HH_2^{t_a} \text{ and not } HH_4^{t_b}] \\ &= ptc_3^T \text{ or } [PTC_3^T \text{ and not } HH_2^{t_a}] \\ &= ptc_3^T \text{ or } [PTC_3^T (\text{and not } wg_3^{t_a-3} \text{ or } CIR_2^{t_a-1})] \end{aligned}$$

because from (6) $HH_4^T = 0$, and by Fact 4. Note that $t_a \in \{T, T+1\}$, depending on the permutation ϕ^{T+1}. So, for PTC_3 to vanish it is necessary that both $ptc_3^T = 0$ and $wg_3^{t_a-3} = 1$. □

Appendix B. Attractiveness of the wild-type pattern

Assuming that the trajectory is not of form (16), the only accessible steady-state is the wild-type. In this case, to establish convergence of the trajectory, it is enough to show that each node attains a constant value after a finite number of iterates. And in fact, from Propositions 4.1, 4.4 and Corollaries 4.2, 4.3, all the nodes become fixed after at most t iterates, as indicated:

$wg_{1,2} = 0, \quad WG_{1,2} = 0, \quad t \geqslant 0, \quad (4)$

$wg_4 = 1, \quad WG_4 = 1, \quad t \geqslant 1, \quad$ Proposition 4.1

$wg_3 = 0, \quad WG_3 = 0, \quad t \geqslant 1, \quad$ Proposition 4.4

$en_1 = 1, \quad EN_1 = 1, \quad t \geqslant 2, \quad$ Corollary 4.2

$en_2 = 0, \quad EN_2 = 0, \quad t \geqslant 2, \quad$ Corollary 4.7

$en_{3,4} = 0, \quad EN_{3,4} = 0, \quad t \geqslant 0, \quad (5)$

$hh_1 = 1, \quad HH_1 = 1, \quad t \geqslant 3, \quad$ Corollary 4.2

$hh_2 = 0, \quad HH_2 = 0, \quad t \geqslant 3, \quad$ Corollary 4.7

$hh_{3,4} = 0, \quad HH_{3,4} = 0, \quad t \geqslant 0, \quad (6)$

$ci_1 = 0, \quad CI_1 = 0, \quad t \geqslant 3, \quad$ Corollary 4.2

$ci_2 = 1, \quad CI_2 = 1, \quad t \geqslant 3, \quad$ Corollary 4.7

$ci_{3,4} = 1, \quad CI_{3,4} = 1, \quad t \geqslant 2, \quad (7), (8)$

$ptc_1 = 0, \quad PTC_1 = 0, \quad t \geqslant 0, \quad$ Corollary 4.3

$CIA_1 = 0, \quad CIR_1 = 0, \quad t \geqslant 4, \quad$ Corollary 4.2

$CIA_3 = 0, \quad CIR_3 = 1, \quad t \geqslant 2, \quad$ Fact 2

$ptc_3 = 0, \quad PTC_3 = 1, \quad t \geqslant 1, \quad$ Fact 2

$CIA_2 = 1, \quad CIR_2 = 0, \quad t \geqslant 4, \text{ from } CI_2 = 1, hh_n = 1$

$ptc_2 = 1, \quad PTC_2 = 1, \quad t \geqslant 5,$

from $CIA_2 = 1, \quad CIR_2 = 0, \quad EN_2 = 0$

$CIA_2 = 1, \quad CIR_2 = 0, \quad EN_2 = 0$

$CIA_4 = 1, \quad CIR_4 = 0, \quad t \geqslant 3, \quad wg_4 = 1$

$ptc_4 = 1, \quad PTC_4 = 1, \quad t \geqslant 5,$

from $CIA_4 = 1, \quad CIR_4 = 0.$

$CIA_4 = 1, \quad CIR_4 = 0.$

This is indeed a complete characterization of the wild-type steady-state.

Proof of Proposition 4.8. Let P, A, C and R ($\leqslant L$) denote the positions of PTC_3, CIA_3, CI_3 and CIR_3, respectively. Then, from Eq. (16) it is easy to see that

$$P \in \{4, 5, 6, \ldots, L\}, \quad A \in \{2, 3, 4, \ldots, P-1\},$$
$$C \in \{1, 2, 3, \ldots, A-1\}, \quad R \in \{1, \ldots, P-1\} \backslash \{P, A, C\}.$$

To derive a formula for M_L, we note that, for each pair of values P, A, the number of possible combinations of C and R is:

$$\text{— — — —A —R — —P :} \quad (A-1)(P-1-A),$$

$$\text{—R — —A — — — —P :} \quad (A-1)(A-2),$$

respectively, for sequences of the form $CARP$ (top), or $CRAP$ and $RCAP$ (bottom). Therefore, summing over all possible A and P:

$$\begin{aligned} M_L &= \sum_{P=4}^{L} \sum_{A=2}^{P-1} [(A-1)(P-1-A) + (A-1)(A-2)] \\ &= \sum_{P=4}^{L} \sum_{A=2}^{P-1} (A-1)(P-3) \\ &= \frac{1}{2} \sum_{P=4}^{L} (P-3)(P-2)(P-1) \\ &= \frac{1}{2} \sum_{j=1}^{L-3} j(j+1)(j+2). \end{aligned}$$

Now, for $L = 4$,

$$\frac{M_L\,(L-4)!}{L!} = \frac{1}{2}\,3!\,\frac{0!}{4!} = \frac{1}{2}\frac{1}{4} = \frac{1}{8}.$$

Assume now that the equality is true for $L-1$:

$$\frac{M_{L-1}\,(L-5)!}{(L-1)!} = \frac{1}{8}.$$

Then

$$\begin{aligned} M_L &= M_{L-1} + \frac{1}{2}(L-3)(L-2)(L-1) \\ &= \frac{1}{8}\frac{(L-1)!}{(L-5)!} + \left(\frac{1}{2}L - 3\right)(L-2)(L-1) \\ &= \frac{1}{8}(L-1)(L-2)(L-3)(L-4) \\ &\quad + \left(\frac{1}{2}L - 3\right)(L-2)(L-1) \\ &= (L-1)(L-2)(L-3)\left[\frac{1}{8}(L-4) + \frac{1}{2}\right] \\ &= (L-1)(L-2)(L-3)\left[\frac{1}{8}L\right] = \frac{1}{8}\frac{L!}{(L-4)!}, \end{aligned}$$

just as we wanted to show. □

Appendix C. State aggregation in the Markov chain

The tables below show the complete transition probabilities p_{ij} (when not indicated, the transition probabilities are equal to 1). The states numbered *3* and *44* denote, respectively, the wild-type and the mutant state. The initial condition was numbered *48*. The first table shows the complete transition probabilities at step 1, from the wild-type initial condition.

Thus the probability shown in the diagram for the transition from the initial state to the aggregated state {*1 6*} was obtained by adding $p_{48,1} + p_{48,6} = 0.1667 + 0.0417 = 0.2084$. A more complex aggregation formula was used for the transition

$$\begin{aligned} &\{10\ 12\} \rightarrow \{16\ 22\}: \\ &\frac{1}{2}(p_{10,16} + p_{10,22}) + \frac{1}{2}(p_{12,16} + p_{12,22}) \\ &= \frac{1}{2}(0.2083 + 0.0417 + 0.3 + 0.0333) \\ &= 0.29165 \approx 0.29. \end{aligned}$$

The normalization by $1/2$ is justified by the fact that the transition from *9* and *30* to either *10* or *12* is the same.

From initial wild-type state to:					
1	*2*	*3*	*4*	*5*	*6*
0.1667	0.357	0.1667	0.125	0.125	0.0417

From *9* to:			
10	*11*	*12*	*13*
0.25	0.25	0.25	0.25

	7	*16*	*18*	*20*	*21*	*22*	*25*	*26*
From *10* to:	0.2083	0.2083	0.125	0.0417	0.125	0.0417	0.125	0.125
From *12* to:	0.1167	0.3	0.1167	0.05	0.1333	0.0333	0.2	0.05

	14	*15*	*17*	*19*	*23*	*24*	*27*	*28*
From *11* to:	0.2083	0.125	0.125	0.2083	0.125	0.0417	0.125	0.0417
From *13* to:	0.1167	0.1167	0.1333	0.3	0.2	0.0333	0.05	0.05

From *29* or *31* to:			
32	*33*	*34*	*35*
0.25	0.25	0.25	0.25

From *30* to:			
10	*11*	*12*	*13*
0.25	0.25	0.25	0.25

From *36* to:							
31	*37*	*38*	*39*	*40*	*41*	*42*	*43*
0.2083	0.125	0.125	0.2083	0.125	0.125	0.0417	0.0417

From *37* or *38* or *42* to:			
32	*33*	*34*	*35*
0.25	0.25	0.25	0.25

From *39* or *41* to:	
44	*46*
0.5	0.5

From *40* or *43* to:			
44	*45*	*46*	*47*
0.25	0.25	0.25	0.25

References

Albert, R., Othmer, H.G., 2003. The topology of the regulatory interactions predicts the expression pattern of the *Drosophila* segment polarity genes. J. Theor. Biol. 223, 1–18.

Albert, R., Jeong, H., Barabási, A.-L., 2000. Error and attack tolerance in complex networks. Nature 406, 378–382.

Alcedo, J., Zou, Y., Noll, M., 2000. Posttranscriptional regulation of smoothened is part of a self-correcting mechanism in the hedgehog signaling System. Mol. Cell 6, 457–465.

Alon, U., Surette, M., Barkai, N., Leibler, S., 1999. Robustness in bacterial chemotaxis. Nature 397, 168–171.

Aza-Blanc, P., Kornberg, T.B., 1999. Ci a complex transducer of the Hedgehog signal. Trends Genet. 15, 458–462.

Bejsovec, A., Wieschaus, E., 1993. Segment polarity gene interactions modulate epidermal patterning in *Drosophila* embryos. Development 119, 501–517.

Bernot, G., Comet, J.-P., Richard, A., Guespin, J., 2004. Application of formal methods to biological regulatory networks: extending Thomas' asynchronous logical approach with temporal logic. J. Theor. Biol. 229, 339–347.

Bertsekas, D.P., Tsitsiklis, J.N., 1989. Parallel and Distributed Computation, Numerical Methods. Prentice-Hall, Englewood Cliffs, NJ.

Bodnar, J.W., 1997. Programming the *Drosophila* Embryo. J. Theor. Biol. 188, 391–445.

Cadigan, K.M., Nusse, R., 1997. Wnt signaling: a common theme in animal development. Genes Dev. 11, 3286–3305.

Cadigan, K.M., Grossniklaus, U., Gehring, W.J., 1994. Localized expression of *sloppy paired* protein maintains the polarity of *Drosophila* parasegments. Genes Dev. 8, 899–913.

Carlson, J.M., Doyle, J., 2002. Complexity and robustness. Proc. Natl Acad. Sci. 99, 2538–2545.

Conant, G.C., Wagner, A., 2004. Duplicate genes and robustness to transient gene knock-downs in Caenorhabditis elegans. Proc. R. Soc. Lond. B Biol. Sci. 271, 89–96.

Davidson, E.H., et al., 2002. A genomic regulatory network for development. Science 295, 1669–1678.

de Jong, H., et al., 2004. Qualitative simulation of genetic regulatory networks using piecewise-linear models. Bull. Math. Biol. 66, 301–340.

DiNardo, S., Sher, E., Heemskerk-Jongens, J., Kassis, J.A., O'Farrell, P.H., 1988. Two-tiered regulation of spatially patterned engrailed gene expression during *Drosophila* embryogenesis. Nature 332, 45–53.

Eaton, S., Kornberg, T.B., 1990. Repression of ci-D in posterior compartments of *Drosophila* by *engrailed*. Genes. Dev. 4, 1068–1077.

Eldar, A., Dorfman, R., Weiss, D., Ashe, H., Shilo, B.-Z., Barkai, N., 2002. Robustness of the BMP morphogen gradient in *Drosophila* embryonic patterning. Nature 419, 304–308.

Espinosa-Soto C., Padilla-Longoria P., Alvarez-Buylla E.R., 2004. A gene regulatory network model for cell-fate determination during Arabidopsis thaliana flower development that is robust and recovers experimental gene expression profiles. Plant Cell 16, 2923–2939.

Feller, W., 1970. An Introduction to Probability Theory and its Applications, Third ed. Wiley, New York.

Gallet, A., Angelats, C., Kerridge, S., Thérond, P.P., 2000. Cubitus interruptus-independent transduction of the Hedgehog signal in *Drosophila*. Development 127, 5509–5522.

Ghysen, A., Thomas, R., 2003. The formation of sense organs in *Drosophila*: a logical approach. BioEssays 25, 802–807.

Glass, L., Kauffman, S.A., 1973. The logical analysis of continuous, nonlinear biochemical control networks. J. Theor. Biol. 39, 103–129.

Grossniklaus, U., Pearson, R.K., Gehring, W.J., 1992. The *Drosophila sloppy* paired locus encodes two proteins involved in segmentation that show homology with mammalian transcription factors. Genes Dev. 6, 1030–1051.

Gursky, V.V., Reinitz, J., Samsonov, A.M., 2001. How gap genes make their domains: an analytical study based on data driven approximations. Chaos 11, 132–141.

Hidalgo, A., Ingham, P., 1990. Cell Patterning in the *Drosophila* segment: spatial regulation of the segment polarity gene *patched*. Development 110, 291–301.

Hooper, J.E., Scott, M.P., 1992. The Molecular genetic basis of positional information in insect segments. In: Hennig, W. (Ed.), Early Embryonic Development of Animals. Springer, Berlin, pp. 1–49.

Ingham, P.W., 1998. Transducing hedgehog: the story so far. EMBO J 17, 3505–3511.

Ingham, P.W., McMahon, A.P., 2001. Hedgehog signaling in animal development: paradigms and principles. Genes Dev. 15, 3059–3087.

Ingham, P.W., Taylor, A.M., Nakano, Y., 1991. Role of the *Drosophila patched* gene in positional signaling. Nature 353, 184–187.

Jeong, H., Tombor, B., Albert, R., Oltvai, Z.N., Barabási, A.-L., 2000. The large-scale organization of metabolic networks. Nature 407, 651–654.

Jeong, H., Mason, S., Barabási, A.-L., Oltvai, Z.N., 2001. Lethality and centrality in protein networks. Nature 411, 41–42.

Kauffman, S.A., 1993. The origins of Order. Oxford University Press, New York.

Kauffman, S., Peterson, C., Samuelsson, B., Troein, C., 2003. Random Boolean network models and the yeast transcriptional network. Proc. Natl Acad. Sci. USA 100, 14796–14799.

Martinez-Arias, A., Baker, N., Ingham, P.W., 1988. Role of segment polarity genes in the definition and maintenance of cell states in the *Drosophila* embryo. Development 103, 157–170.

Mendoza, L., Thieffry, D., Alvarez-Buylla, E.R., 1999. Genetic control of flower morphogenesis in *Arabidopsis thaliana*: a logical analysis. Bioinformatics 15, 593–606.

Ohlmeyer, J.T., Kalderon, D., 1998. Hedgehog stimulates maturation of Cubitus interruptus into a labile transcriptional activator. Nature 396, 749–753.

Pfeiffer, S., Vincent, J.-P., 1999. Signaling at a distance: transport of Wingless in the embryonic epidermis of *Drosophila*. Cell Dev. Biol. 10, 303–309.

Rao, C.V., Wolf, D.M., Arkin, A.P., 2002. Control exploitation and tolerance of intracellular noise. Nature 420, 231–237.

Reinitz, J., Sharp, D.H., 1995. Mechanism of eve stripe formation. Mech. Dev. 49, 133–158.

Sánchez, L., Thieffry, D., 2001. A logical analysis of the *Drosophila* gap-gene system. J. Theor. Biol. 211, 115–141.

Schwartz, C., Locke, J., Nishida, C., Kornberg, T.B., 1995. Analysis of *cubitus interruptus* regulation in *Drosophila* embryos and imaginal disks. Development 121, 1625–1635.

Tabata, T., Eaton, S., Kornberg, T.B., 1992. The *Drosophila hedgehog* gene is expressed specifically in posterior compartment cells and is a target of *engrailed* regulation. Genes Dev. 6, 2635–2645.

Taylor, A.M., Nakano, Y., Mohler, J., Ingham, P.W., 1993. Contrasting distributions of patched and hedgehog proteins in the *Drosophila embryo*. Mech. Dev. 42, 89–96.

Thomas, R., 1973. Boolean formalization of genetic control circuits. J. Theor. Biol. 42, 563–585.

von Dassow, E., Munro, E.M., Odell, G.M., 2000. The segment polarity network is a robust developmental module. Nature 406, 188–192.

Wolpert, L., Beddington, R., Brockes, J., Jessell, T., Lawrence, P., Meyerowitz, E., 1998. Principles of Development. Current Biology Ltd., London.

Yuh, C.H., Bolouri, H., Bower, J.M., Davidson, E.H., 2001. A logical model of cis-regulatory control in a eukaryotic system. In: Bower, J.M., Bolouri, H. (Eds.), Computational Modeling of Genetic and Biochemical Networks. MIT Press, Cambridge, MA, pp. 73–100.