BioSystems xxx (2010) xxx–xxx

Contents lists available at ScienceDirect

BioSystems

journal homepage: www.elsevier.com/locate/biosystems

Stabilizing patterning in the *Drosophila* segment polarity network by selecting models *in silico*

Gautier Stoll[a,b,c,*], Mirko Bischofberger[d,e], Jacques Rougemont[d,e], Felix Naef[d,e,**]

[a] *Institut Curie, 26 rue d'Ulm, Paris F-75248, France*
[b] *INSERM, U900, Paris F-75248, France*
[c] *Mines ParisTech, Fontainebleau F-77300, France*
[d] *The Institute of Bioengineering (IBI), School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland*
[e] *Swiss Institute of Bioinformatics (SIB), 1015 Lausanne, Switzerland*

ARTICLE INFO

Article history:
Received 21 May 2010
Accepted 15 July 2010

Keywords:
Segment polarity network
Systems biology
Discrete models
Drosophila development

ABSTRACT

The segmentation of *Drosophila* is a prime model to study spatial patterning during embryogenesis. The spatial expression of segment polarity genes results from a complex network of interacting proteins whose expression products are maintained after successful segmentation. This prompted us to investigate the stability and robustness of this process using a dynamical model for the segmentation network based on Boolean states. The model consists of intra-cellular as well as inter-cellular interactions between adjacent cells in one spatial dimension. We quantify the robustness of the dynamical segmentation process by a systematic analysis of mutations. Our starting point consists in a previous Boolean model for *Drosophila* segmentation. We define mathematically the notion of dynamical robustness and show that the proposed model exhibits limited robustness in gene expression under perturbations. We applied *in silico* evolution (mutation and selection) and discover two classes of modified gene networks that have a more robust spatial expression pattern. We verified that the enhanced robustness of the two new models is maintained in differential equations models. By comparing the predicted model with experiments on mutated flies, we then discuss the two types of enhanced models. *Drosophila* patterning can be explained by modelling the underlying network of interacting genes. Here we demonstrate that simple dynamical considerations and *in silico* evolution can enhance the model to robustly express the expected pattern, helping to elucidate the role of further interactions.

1. Background

The embryonic development in the fruit fly *Drosophila melanogaster* has given tremendous insight into the signaling pathways and interactions involved in morphogenesis and body patterning (Ingham and McMahon, 2001; Sanson, 2001; Hatini and DiNardo, 2001; Sanchez and Thieffry, 2003; Jaeger et al., 2004; Kobayashi et al., 2003). To understand such phenomena, spatial (inter-cellular) and temporal expression patterns of the genes and proteins involved also need to be taken into account, adding a further layer of complexity to models of cell autonomous signaling pathways.

Studying the development of the *Drosophila* embryo has helped to understand how signaling and transcription networks can pattern the syncytial blastoderm prior to cellularization, and how this pre-patterning then leads to the formation of segments. The gene networks involved in the sequential steps of the patterning process have been studied extensively: the expression of gap genes occurs first and establishes morphogen gradients that encode information about the anterior–posterior axis (Jaeger et al., 2004). This activates the pair rule genes which divide the embryo into pairs of segments (Nusslein-Volhard and Wieschaus, 1980), a process that occurs as the cells are formed (Carroll, 1990). These genes in turn regulate the segment polarity network (DiNardo and O'Farrell, 1987) for anterior–posterior patterning within each parasegment. Since segment polarity is completed after less than 10 h, these processes involve mainly transient expression and silencing of gene products. However, the genes do reach stable expression patterns compare to biochemical reactions time scales (van den Heuvel et al., 1993). It is therefore interesting to study what mechanisms produce sufficiently stable and patterned expression levels.

The segment polarity genes interact through a well-studied network: the main genes involved are *wingless* (WG) and *engrailed*

* Corresponding author at: Institut Curie, 26 rue d'Ulm, Paris F-75248, France.
** Corresponding author at: Swiss Institute of Bioinformatics (SIB), 1015 Lausanne, Switzerland. Tel.: +41 21 6931621; fax: +41 21 6931621.
E-mail addresses: gautier.stoll@curie.fr (G. Stoll), Mirko.Bischofberger@epfl.ch (M. Bischofberger), Jacques.Rougemont@epfl.ch (J. Rougemont), Felix.Naef@epfl.ch (F. Naef).

0303-2647/$ – see front matter © 2010 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.biosystems.2010.07.014

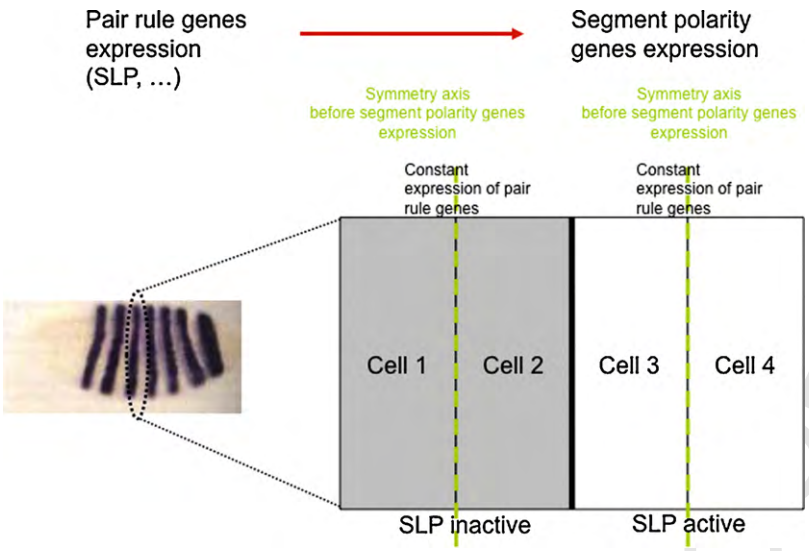


Fig. 1. Relation between expression of segment polarity genes and symmetry breaking. The issue of differentiated expression of segment polarity genes is related to a symmetry breaking problem in this model: the segment polarity genes are expressed differently inside a region where pair rule genes expression is constant. The model chosen in this work consists of four cells (with a periodic condition, i.e. cell 1 next to cell 4), each pair belonging to a region of constant expression of pair rule genes (identical in cells 1, 2 and in cells 3, 4). Therefore the expression of pair rule genes has a symmetry axis (in green). The differential expression of segment polarity genes breaks this symmetry.

(EN). Their expression produces the stable pattern necessary for the normal embryonic development (van den Heuvel et al., 1993). The segment polarity network consists of six main genes, plus a pair rule gene *sloppy pair* (SLP), whose expression starts before and essentially acts as a trigger that pre-patterns the segments, though with a higher symmetry than the final segment polarity patterning (Fig. 1). The network is not a simple cascade but contains several feedback loops. Therefore, the spatio-temporal evolution is difficult to grasp intuitively and a mathematical model can be useful. Most modeling approaches (von Dassow et al., 2000; Ingolia and Murray, 2004) to the segment polarity network involve differential equations representing mass-action kinetics and Michaelis-Menten processes. Although such models rely often on parameters that are not accessible to experimentation, they were successfully applied to explain oscillations (Barkai and Leibler, 2000) question stability properties (Becskei and Serrano, 2000) or reproduce mutant phenotypes (Chen et al., 2004). In fact the parameter space in such models can be small enough to make theoretical predictions possible (von Dassow et al., 2000; Ingolia and Murray, 2004).

As an alternative, dynamical evolution can be modeled as Boolean update rules (Sanchez and Thieffry, 2003; Li et al., 2004; Thomas and Kaufman, 2001). Such models are not based on biophysical principles, but have the advantage of keeping the number of unknown parameters within manageable limits. These approaches are particularly suitable in the common situation when the exact biochemical process is not known with sufficient details, because Boolean logic describes gene expression in a language close to that used by experimentalists in cell biology (active/inactive states).

One consequence of patterning is that gene expression patterns are refined as the embryo progresses through the different stages of development. For example, during the transition from the pair rule to the segment polarity patterning, the symmetry axis through the middle of each parasegment is lost (Fig. 1). Our main goal below is to study how the genetic network can robustly establish and maintain this spatial expression pattern. We investigate this problem using existing Boolean models (Albert and Othmer, 2003) which we simulate in the presence of fluctuations following (Stoll et al., 2006). Using *ad hoc* mathematical definitions of robustness, we systematically assess which network modifications contribute most to stabilizing or destabilizing the wild-type patterning. While the original working model has suboptimal robust patterning, we find minimal modifications to the network that can significantly improve the robustness of the correctly patterned state in the presence of cellular fluctuations.

2. Results and Discussion

2.1. The Initial Boolean Model Does Not Lead to Dynamically Robust Patterning

Using our previous stochastic approach (see Section 4) we simulated the segment polarity network from (Albert and Othmer, 2003). These authors computed the stable states of the synchronized Boolean dynamics analytically (without noise) and found not only the expected biological final states, but also lethal gene expression patterns reminiscent of known mutants, e.g. the broad type pattern. In addition the size of the basins of attraction for the wild-type fixed point was estimated and interpreted as measure of stability of the patterning process.

We further investigated the stability properties of the attractor states in this model by introducing noise in the simulation allowing occasional random flips of the states of the genes. The noise strength ("node flipping probability" or NFP = 0.005) is such that the ~20% of the trajectories from the initial condition in (Albert and Othmer, 2003) will reach the final state without perturbation. Using the simulated trajectories to sample the stationary state we compute the distribution of Hamming distances $d(H)$ to the target wild-type pattern. This distribution is concentrated around a non-zero value (Fig. 2, black curve), meaning that the system has a large probability to be in a state with abnormal patterning, namely the "broad type" pattern which is an attractor for the noise free model (Albert and Othmer, 2003). Moreover the probability of the wild-type pattern (zero Hamming distance) is negligible. We therefore conclude that although the correct wild-type pattern is a fixed point of the dynamics, it is rather unstable against small stochastic perturbations.

A possible reason for the lack of robustness against noise lies within the network topology itself: in the two rightmost cells of the parasegment, SLP induces WG but blocks EN, while in the two

Please cite this article in press as: Stoll, G., et al., Stabilizing patterning in the *Drosophila* segment polarity network by selecting models *in silico*. BioSystems (2010), doi:10.1016/j.biosystems.2010.07.014

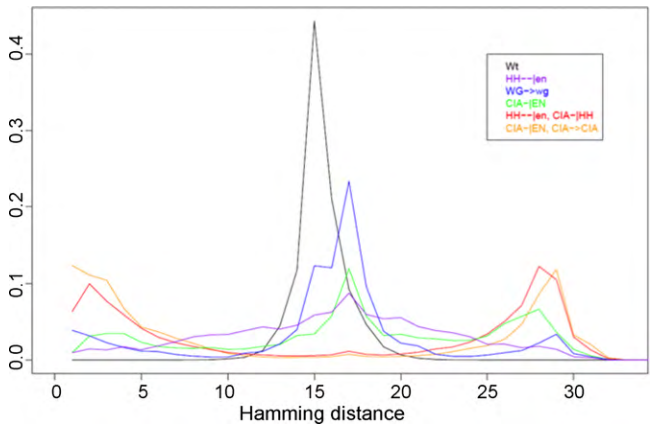


Fig. 2. Hamming distance distribution for different mutations. The Hamming distance distribution compares the biological wild-type final state to the stationary state of the Boolean dynamics. Each color curve is associated with a mutated network model with one new connection constructed *in silico* that increases the probability that the Hamming distance is small. "Wt" means the wild-type original model, the newly added connections are given explicitly, with the following convention: "→" and "-|" are intra-cellular the cell, "–|" is a connection between adjacent cells. The high peak of the "Wt" model corresponds to the "broad type" final state (see Fig. 3), in which no patterning of segment polarity occurs. For mutated models, the Hamming distance distribution has one central peak ("no stripe" state, see Fig. 3) corresponding to a state with no segmentation and two other peaks corresponding to the desired final state of patterning and its symmetrical complement. *Note*: the scale starts at $d = 0$ (leftmost points correspond to $d = 0$).

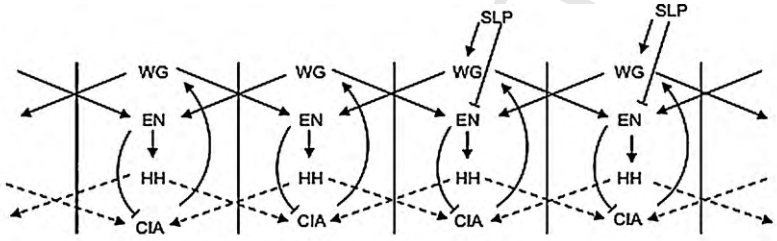
first cells EN is activated by WG from adjacent cells, the latter being activated through SLP (Fig. 3). Therefore, the "broad type" state can occur with a high probability. In order to stabilize the "biologically correct" fixed point, WG and EN should be subject to (direct or indirect) self-inhibition from neighboring cells.

In the remainder we simulate gain and loss of function mutations to seek a modified, *robust* model of body patterning, i.e. exhibiting a highly probable small Hamming distance to the wild-type patterning at steady-state. Issues related to (transient) trajectories the patterned state will be discussed later.

2.2. In Silico Analysis Proposes Stabilizing Interactions

One way to explain the lack of robustness of the above model is that it may still be incompletely characterized experimentally. All connections included so far in the model are based on biochemical evidence, and may not be exhaustive yet. We attempted to probe whether other connections could stabilize the symmetry breaking event. Given the complex regulatory functions integrating the signals at each node in the network, an exhaustive study of every possible new connection is beyond our means. We thus resorted to a greedy approach: (a) we generated all models with one additional connection compared to the network in Fig. 4; (b) we selected the best models; (c) we iterated the procedure. The score used for selection was the probability $P(H \leq 4)$ of observing a Hamming distance to the wild-type state smaller than 4 (see Section 4). This value of 4 represents one mismatch per cell (see Section 4 for the reason why the model has 4 cells) and we have checked that our conclusions do not depend much on this choice.

The greedy procedure resulted in the tree shown in Fig. 5. Here we only considered biochemically realistic connections, namely transcriptional regulation (a protein links to a mRNA) or post-transcriptional (a protein links to another protein) type. Our first observation was that two new connections, HH–|en and CIA-|EN (see caption of Fig. 2 for the difference between –| and -|) already improved the stability of the wild-type patterning, with the probability $P(H \leq 4)$ being closed to its theoretical maximum of 1/2 (the value of 1/2 is due to the wild-type patterning being non-symmetric, the symmetric version of this state shares the same



"Broad type" final state

	cell1	cell2	cell3	cell4
wg	0	0	1	1
WG	0	0	1	1
en	1	1	0	0
EN	1	1	0	0
hh	1	1	0	0
HH	1	1	0	0
ptc	0	0	1	1
PTC	0	0	1	1
PH	0	0	1	1
SMO	1	1	1	1
CI	0	0	1	1
CI	0	0	1	1
CIA	0	0	1	1
CIR	0	0	0	0

"Biological" final state

	cell1	cell2	cell3	cell4
wg	0	0	0	1
WG	0	0	0	1
en	1	0	0	0
EN	1	0	0	0
hh	1	0	0	0
HH	1	0	0	0
ptc	0	1	0	1
PTC	0	1	1	1
PH	0	1	0	1
SMO	1	1	0	1
CI	0	1	1	1
CI	0	1	1	1
CIA	0	1	0	1
CIR	0	0	1	0

"No stripe" final state

	cell1	cell2	cell3	cell4
wg	0	0	0	0
WG	0	0	0	0
en	0	0	0	0
EN	0	0	0	0
hh	0	0	0	0
HH	0	0	0	0
ptc	0	0	0	0
PTC	1	1	1	1
PH	0	0	0	0
SMO	0	0	0	0
CI	1	1	1	1
CI	1	1	1	1
CIA	0	0	0	0
CIR	1	1	1	1

Fig. 3. Simplified representation of network models and their main final states. The top part is a fusion of a part of the network of Fig. 4 and the one-dimensional model of patterning from Fig. 1. The bottom part represents the Boolean final states, each line being a gene and each column a cell. The left state is the broad type (lethal) pattern and the center state is the biological one; the right state ("no stripe") corresponds to expression of mutated *Drosophila*, when wg, en or hh is blocked. It is easy to understand why the right state appears naturally with a high probability.

Please cite this article in press as: Stoll, G., et al., Stabilizing patterning in the *Drosophila* segment polarity network by selecting models *in silico*. BioSystems (2010), doi:10.1016/j.biosystems.2010.07.014

$$SLP_i^{t+1} = SLP_i^t = \begin{cases} 0 \text{ if } (i \bmod 4) = 1 \text{ or } (t \bmod 4) = 2 \\ 1 \text{ if } (i \bmod 4) = 3 \text{ or } (t \bmod 4) = 0 \end{cases}$$

$$wg_i^{t+1} = \left(CIA_i^t \text{ and } SLP_i^t \text{ and not } CIR_i^t\right) \text{ or } \left[wg_i^t \text{ and } \left(CIA_i^t \text{ or } SLP_i^t\right) \text{ and not } CIR_i^t\right]$$

$$WG_i^{t+1} = wg_i^t$$

$$en_i^{t+1} = \left(WG_{i-1}^t \text{ and } WG_{i+1}^t\right) \text{ and not } SLP_i^t$$

$$EN_i^{t+1} = en_i^t$$

$$hh_i^{t+1} = EN_i^t \text{ and not } CIR_i^t$$

$$HH_i^{t+1} = hh_i^t$$

$$ptc_i^{t+1} = CIA_i^t \text{ and not } EN_i^t \text{ and not } CIR_i^t$$

$$PTC_i^{t+1} = ptc_i^t \text{ or } \left(PTC_i^t \text{ and not } HH_{i-1}^t \text{and not } HH_{i+1}^t\right)$$

$$PH_i^t = PTC_i^t \text{ and } \left(HH_{i-1}^t \text{or } HH_{i+1}^t\right)$$

$$SMO_i^t = \text{not } PTC_i^t \text{ or } HH_{i-1}^t \text{or } HH_{i+1}^t$$

$$ci_i^{t+1} = \text{not } EN_i^t$$

$$CI_i^{t+1} = ci_i^t$$

$$CIA_i^{t+1} = CI_i^t \text{ and } \left(SMO_i^t \text{ or } hh_{i-1}^t \text{or } hh_{i+1}^t\right)$$

$$CIR_i^{t+1} = CI_i^t \text{ and } SMO_i^t \text{ and not } hh_{i-1}^t \text{ and not } hh_{i+1}^t$$

Fig. 4. Boolean model of segment polarity genes (Albert and Othmer, 2003). The left panel illustrates the network of segment polarity genes that is identical in all four cells, with intra-cellular and inter-cellular connections. Upper case and lower case names refer to protein and mRNA respectively. The right panel shows the detailed logic applied for updating gene expression in a discrete time Boolean model.

probability). Therefore, we can conclude that the model proposed by Albert et al. based on observed connections is nearly complete. It should be noticed that our approach cannot predict the biochemical mechanism behind the new connections. It is not excluded that these new regulations occur through intermediate genes.

It is interesting to understand the mechanism induced by the new hypothetical connections. Referring to previous remarks and to Fig. 3, the network needs to block, or at least control, the direct and indirect inter-cellular connections between EN and WG. The most straightforward solution is to introduce a blocking connection (typed -|) between cells: HH -| EN (or possibly HH -| FZ) as found in the left branch of Fig. 5. This possibility, shown in Fig. 6, is however not sufficient. Namely the activation of WG by CIA should also be controlled, which can be achieved by the new repression HH -| CIA. The other main branch of the tree contains the new link CIA -| EN and refines the activation of CIA. In that case, CIA -| EN has a similar action as HH -| EN but acts intra-cellularly instead of across the membrane, through an indirect induction HH → CIA (Fig. 3). Our results thus suggest new connections that were not found in most of recent models based on differential equations (von Dassow et al., 2000; Ingolia and Murray, 2004; Ma et al., 2006), except by Dayarian et al. (2009). This recent work also suggests the positive feedback of WG, and the connection CIA-| EN. This is likely due to a different formulation of the problem: in these studies, a model is robust if it produces the proper trajectory starting from an initial condition that already has the correct EN and WG patterning. In contrast we require that the wild-type patterning remains globally stable, i.e. for any state reachable by perturbation, the system has a high probability of returning to the wild-type patterning. In other words, we test a more global robustness property than Albert et al. (see Fig. 7). It could be argued that a pathway is missing in our model, namely the feedback on SLP (EN -| SLP -| EN) as in Ingolia and Murray (2004). However, if this was necessary to obtain the right pattern, our analysis would have selected the effective loop EN → en, because EN -| SLP -| EN results in a positive feedback on EN protein expression, that can be replaced by EN → en → EN.

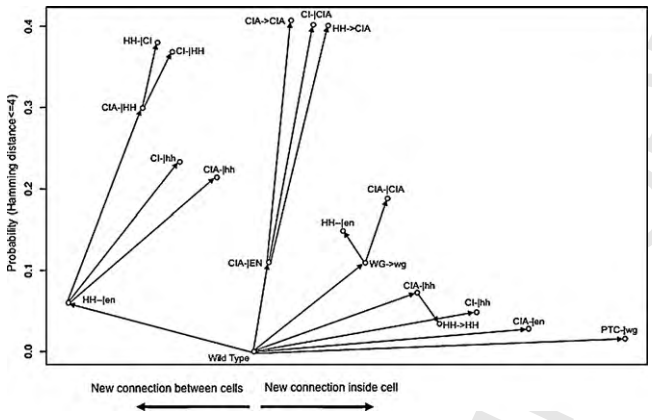


Fig. 5. Tree obtained from the greedy evolutionary algorithm. Each point represents a network of interacting genes. These models were chosen among perturbations that increase the probability of having a small Hamming distance between stationary state and wild-type pattern (see Fig. 2), represented by the *y*-axis. The new connections are given explicitly, with the same convention as in Fig. 2. Two main branches can be observed, the left one has a new inter-cellular connection and the right one only has new intra-cellular connections. The top of the highest branch propose the new connections CIA → CIA and CI-|CIA. These two links essentially refine the Boolean description of both the expression of CIA and the influence of CI onto CIA.

2.3. Proposed Connections Are Model Independent

While useful for rapidly exploring modified connectivities, the Boolean approach remains a highly simplified approximation of the microscopic reality. Therefore, it is useful to test the results obtained above with a more detailed dynamical model to determine if our theoretical predictions prevail. In von Dassow et al. (2000), Ingolia and Murray (2004), and Ma et al. (2006), differential equation models were developed to study the same biological network. In these studies the authors defined the level of robustness as the percentage of parameter values for which the correct pattern is reached from an initially patterned WG and EN expression.

For optimal comparison we use the model of von Dassow et al. (2000) rather than that of Ingolia and Murray (2004) because the

Please cite this article in press as: Stoll, G., et al., Stabilizing patterning in the *Drosophila* segment polarity network by selecting models *in silico*. BioSystems (2010), doi:10.1016/j.biosystems.2010.07.014

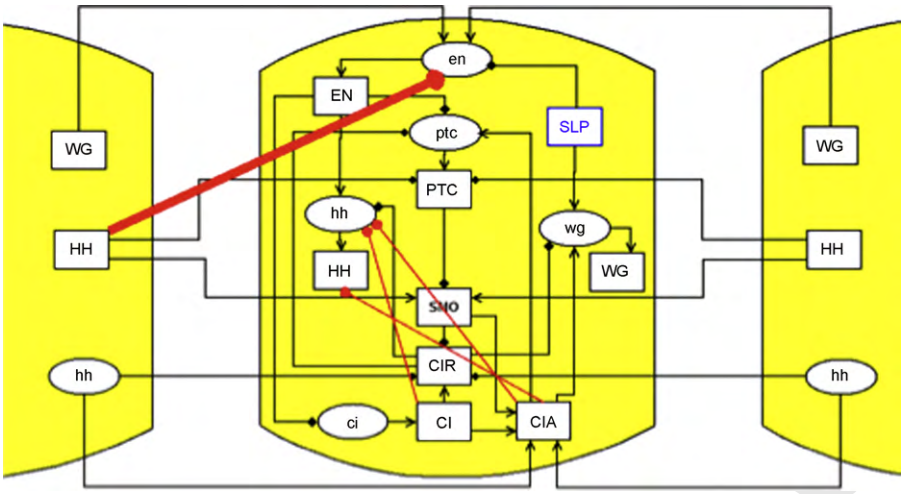


Fig. 6. Models of hypothetical connections with inter-cellular HH-|en. The network of segment polarity genes is increased with the new connections that stabilize the desired final state (big red arrow). Along with the new inter-cellular link, the model requires one other new intra-cellular connection among the three possibilities represented by red (inhibiting) arrows.

latter did not allow SLP to have a fixed expression as it is involved in the positive feedback loop EN-|SLP-|EN (although the link EN-|SLP has been observed; Kobayashi et al., 2003), our approach to robustness does not seem to require it (cf. above). In order to test if the connections identified through the Boolean approach also lead to a consistent phenotype in the ODE models, we slightly modified the model in von Dassow et al. (2000) using the Java tool (http://rusty.fhl.washington.edu/ingeneue/) to reflect the model studied here (Fig. 4 and Albert and Othmer, 2003). Specifically we modified the topology to consider four consecutive cells with periodic boundaries, as in Fig. 1. Secondly we considered SLP expression as fixed.

Using the same measure of robustness as in von Dassow et al. (2000), we found that the connections identified through the Boolean approach (Figs. 6 and 8) also increased the robustness of the differential equation model albeit with different quantitative changes (Table 1). The main difference is that the positive feedback WG → wg is necessary in a differential equation approach (as in other ODE models; von Dassow et al., 2000; Ingolia and Murray, 2004) whereas WG → wg is one possibility for increasing robustness of the Boolean model. This is likely due to the difference in the detailed description of WG activation: in the Boolean approach, its mRNA expression is itself subject to positive feedback (see the logical expression of wg in Fig. 4). In summary, the main new connections proposed by our Boolean approach induce a consistent change when implemented in ODE based models, indicating that they reflect relevant topological changes.

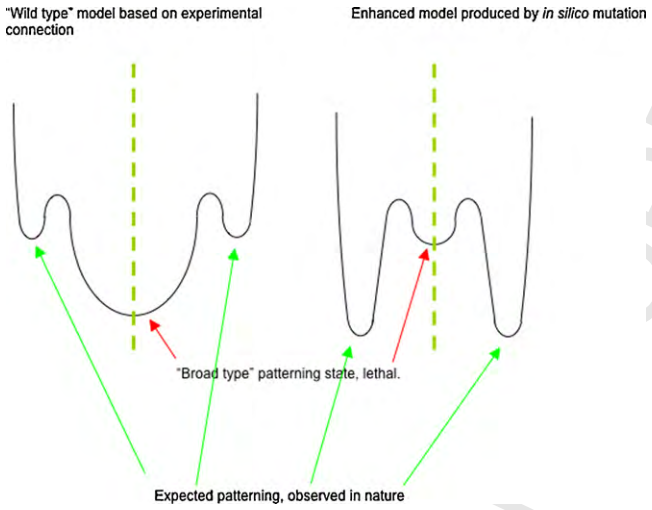


Fig. 7. Schematic representation of final states stability as a potential energy landscape; *x*-axis represent the abstract position of states, the *y*-axis the associate potential energy. Such representation is often used to illustrate the notion of *dynamical symmetry breaking*: a dynamical system having a given symmetry, but with non-symmetric stable states. The dashed green axis reflects the symmetry of the gene expression network and SLP expression. The red arrows represent the lethal broad type patterning and the green the biological non-symmetric patterning. The effect of new connections is simply to enhance the probability of correct patterning compared to the broad type one.

2.4. Trajectories to Robust Patterning

So far we focused on the stability of the final patterning, as a steady-state property. Therefore our criteria for robustness did not depend on the initial condition of the dynamics and we ignored the trajectory leading to the wild-type patterning. As expected, the symmetrical version of the patterning (symmetry according to the axis of Fig. 1) has exactly the same stability properties as the original pattern (Fig. 2).

Even though we do not expect trajectories in Boolean models to accurately reflect temporal sequences of biochemical events (better descriptions are provided by finer dynamical models such as ODE models), we used these Boolean trajectories to study how the final state is reached form the initial state (as defined by Albert and Othmer, 2003) and thereafter maintained. For this, we compute the distribution of the number of Boolean steps (1) leading from the initial pattern (described in Albert and Othmer, 2003) to the final pattern, and (2) from the final pattern back to itself. Figs. 9 and 10

Table 1

Robustness analysis applied to a one-dimensional ordinary differential equation (ODE) model of segment polarity genes: the connections are the same than those in the Boolean model (Fig. 4) but the dynamics is simulated using the modelling tool developed in von Dassow et al. (2000). The measure of robustness is identical to that in von Dassow et al. (2000), i.e. the probability (in %) reaching the correct patterning for random values of unknown rate parameters. The hypothetical new connections, like those in Figs. 6 and 8, enhance the robustness, with the maximal robustness for the model with three added connections. Note that ODE models require the positive feedback loop WG → wg.

Please cite this article in press as: Stoll, G., et al., Stabilizing patterning in the *Drosophila* segment polarity network by selecting models *in silico*. BioSystems (2010), doi:10.1016/j.biosystems.2010.07.014

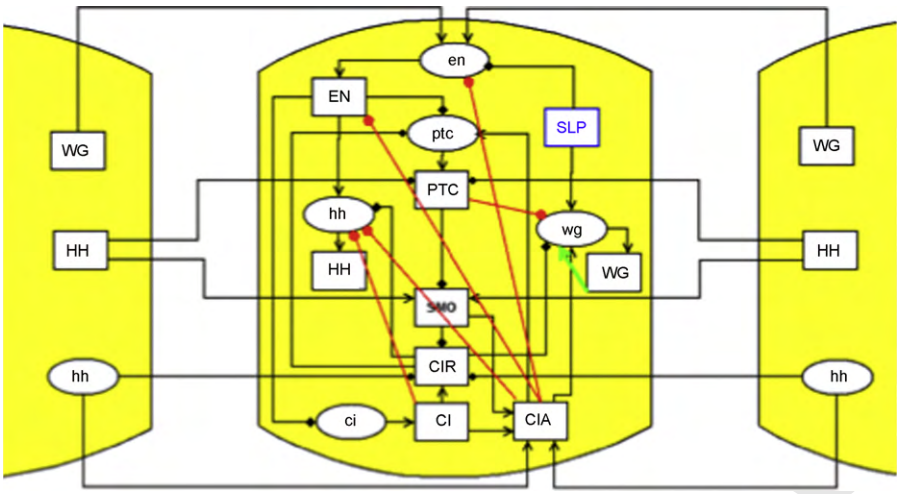


Fig. 8. Models with hypothetical intra-cellular connections. Similar to Fig. 6, but without any new inter-cellular connection (green arrows are activating, red are inhibiting).

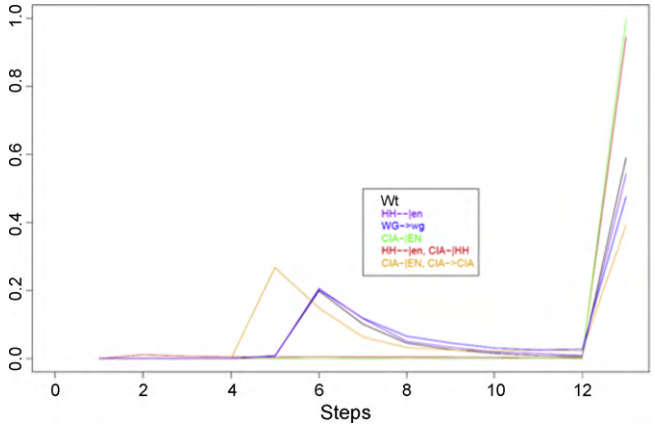


Fig. 9. Probability distribution for the number of Boolean steps (using noisy updates, cf. Section 4) taken from the initial segment polarity genes expression to wild-type stable patterning, colors correspond to models in Fig. 2. The value at step 13 is the probability that the number of steps is larger than 12.

illustrate these two distributions, with a NFP of 0.005 as before. From these two figures and Fig. 2, we can interpret the dynamical behavior of the time trajectories producing the wild-type patterning in the different models:

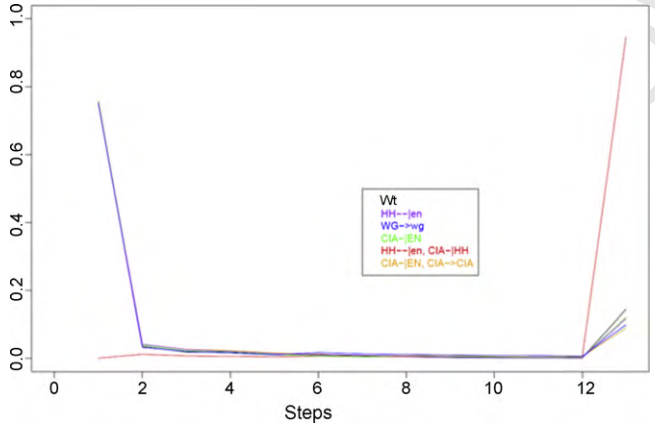


Fig. 10. Probability distribution of return times (in number of Boolean steps) to the wild-type patterning; colors correspond to models in Fig. 2. The value at step 13 is the probability that the number of steps is larger than 12.

- For the wild-type model, the trajectory from initial to final state typically takes six steps, as expected from the deterministic dynamics. Comparing with Fig. 2, we conclude that after the system has left the deterministic trajectory, it almost never comes back. Therefore, after a sufficiently long time, the noise will irreversibly drive the system towards the broad type pattern.
- For models with added (HH-|en) and (WG → wg) interactions, the step distribution is almost identical to the wild-type, except for the slightly lower probability that reaching the final patterning would take more than 12 steps. The main change is seen in Fig. 2: the wild-type pattern stability is higher (when compare to wild-type pattern stability in the wild-type model). This indicates that these models have the capability of returning to the wild-type patterning even under random perturbation of gene expressions.
- For models with added CIA-|EN and (HH-|en, CIA-|HH), the deterministic trajectory is completely lost. In the second model, the wild-type itself is not a stable state of the deterministic dynamics. In such cases, the noise itself drives the system towards the wild-type patterning (in the second case, the system oscillates around this patterning, as illustrated in Fig. 2 with a higher probability of Hamming distance 2 than 1).
- For the model with new (CIA-|EN, CIA → CIA), the deterministic trajectory is shorter and the wild-type patterning is stable under noise.

Accordingly, the model with (CIA-|EN, CIA → CIA) has the best performance, though it is obtained by adding one connection to an unfavorable model (CIA-|EN). Nevertheless, the interpretation of trajectories from synchronized Boolean dynamics should be taken with care: this is why we prefer to concentrate on the stability properties of such models at steady-state. The detailed investigation of trajectories should be more reliable with continuous models such as stochastic differential equations.

2.5. Comparison With Phenotypes in Mutants Helps to Segregate Between Enhanced Models

Here we analyze model predictions for networks that correspond to studied mutants which interfere with segment polarity (Tabata et al., 1992; DiNardo et al., 1988; Hidalgo and Ingham, 1990). For instance, it was shown that removal of hh, wg or en impairs segmentation. Albert et al. implemented these mutations in their model and noticed that the fixed point corresponded to the "no stripe" phenotype (Fig. 3). The effect of the three mutants on the Hamming distance in our enhanced model is shown in Fig. 11.

Please cite this article in press as: Stoll, G., et al., Stabilizing patterning in the *Drosophila* segment polarity network by selecting models *in silico*. BioSystems (2010), doi:10.1016/j.biosystems.2010.07.014

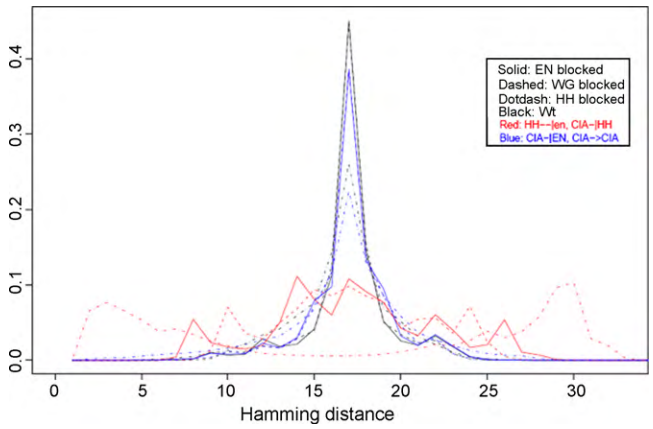


Fig. 11. Hamming distance distributions of biological mutations (blocking wg, en or hh) on the three different models: (1) wild-type model, (2) model with inter-cells connection HH–|en and intra-cell connection CIA-|HH, (3) model with intra cell connections CIA-|EN and CIA → CIA. The central peak (around a Hamming distance of 17) corresponds to a state with not segmentation ("no stripe" state, see Fig. 3).

The peak around $d = 17$ corresponds to the "no stripe" state. Fig. 11 shows that an enhanced model with inter-cell HH –| en link fails to reproduce this "no stripe" state. Therefore, we believe that the observed mutant phenotypes exclude the left branch of the tree in Fig. 4 as it is not consistent with experimental data.

2.6. Dynamical and Structural Robustness Are Interdependent

Biological systems are expected to exhibit robustness to fluctuations or parameter variations, but there it is not easy to give a general definition of this concept (Kitano, 2007). Most robustness studies consider noisy variations on the dynamics, as it is done for example in the works on the same biological system, with ODE (von Dassow et al., 2000; Ingolia and Murray, 2004). Because most systems biology approaches based on simulating network models separate the long time behavior (qualitative influence, mainly encoded in the genome) and short time behavior (detailed interaction between molecules), we propose that robustness should both be studied through topological and dynamical modifications, the latter being more reminiscent of the short time properties. In the present work, the notion of robustness is formalized as the probability of reaching a stationary state close to the wild-type pattern. This is a definition for a *dynamical* robustness, considering dynamical (or short time) variations. In a straightforward manner, we can compute the *structural* (or topological) robustness by systematically modifying all logic functions defining the dynamical updates (Section 4) and estimating $P(H \leq 4)$.

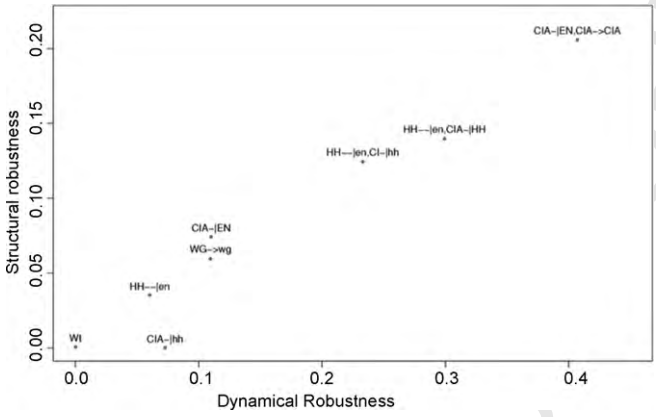


Fig. 12. Dynamical vs. structural robustness for mutated models *in silico*. The diagram represents the relation between dynamical and structural robustness, for enhanced models by adding connections *in silico* (the notation is the same as in Fig. 2). The dynamical robustness is given by $P(H \leq 4)$ (see Section 2) structural robustness is the average dynamical robustness over all possible new connections. These two types of robustness seem quite correlated, excepted for one model (CIA-|hh) for which *in silico* evolution reaches a dead end: no new connections can improve the dynamical robustness.

We compared measures of robustness in Fig. 12. Although they probe different properties of the segment polarity network, their values are strongly correlated. The only outlier is the network where we add CIA -| hh. Inspection of Fig. 5 shows that this link is a dead end in the evolution tree: CIA -| hh increases the dynamical robustness, but no further link starting from this new model can improve the stability.

The correlation of these two definitions of robustness is not obvious for any models of dynamical evolution applied on a network. We think that Fig. 12 has this strong correlation property because the models considered are close approximations to a biological system that has evolved to accomplish a specific task robustly.

3. Conclusion

Using a discrete dynamical model by Albert et al., we assessed to what extent the wild-type patterning in the segment polarity genes is robust against fluctuations in gene expression. We found that a few additions to the original model could lead to robust models of essentially two types, those with added intra-cellular and those with new inter-cellular interactions.

The comparison with mutant phenotypes favors the solution with added intra-cellular connections. We thus invoked dynamical considerations to propose a modified gene network with much improved robustness performances, which held both in discrete dynamical modeling as well as in models of ordinary differential equations.

The segment polarity network being among the best-studied genetic circuits, it naturally triggered interest from theoreticians. Recently, an exhaustive search over network topologies for patterned states showed that the experimentally mapped network is among those showing the most robust gene expressions states (Ma et al., 2006). Our systematic study showed that we can even enhance the robustness properties by slight modifications to the original models.

We have departed from the deterministic model by introducing stochastic flips of gene activity states. Other studies have investigated the effect of desynchronizing the discrete dynamical steps, by introducing variability in the delays between gene interactions (Chaves et al., 2005). For the model studied here the trajectory leading to the patterning was stable against this type of variation.

More generally, our work exemplifies how a spatially patterned gene expression state can be the result of a dynamical symmetry breaking. In other words, although the network is identical in two adjacent cells, the combined effect of inter- and intra-cellular links on the dynamical evolution leads to stable patterns in which genes are not expressed similarly in adjacent cells. Here, this was the case for example for WG, EN and HH. Noticing that current models had limited stability properties against fluctuations, we developed a greedy search for modified networks and were able to modify existing models by proposing new interactions that stabilize the dynamical symmetry breaking process.

4. Methods

We start our analysis following the Boolean model of the segment polarity network by Albert and Othmer (2003). This model describes both intra- and inter-cellular interactions amongst 10 core genes in the network. To limit the number of cells to the minimum the model describes one parasegment which contains four adjacent cells repeated periodically (Fig. 1). The patterning is taken constant along the dorsal-ventral axis and the model is therefore one-dimensional. The network of

Please cite this article in press as: Stoll, G., et al., Stabilizing patterning in the *Drosophila* segment polarity network by selecting models *in silico*. BioSystems (2010), doi:10.1016/j.biosystems.2010.07.014

interacting genes is the same in each of the four cells and inter-cellular connections are restricted to direct neighbors, so that the model is invariant upon translations and reflections (swapping left and right). The genetic network model from Albert et al. is recapitulated in Fig. 4. In Albert and Othmer (2003), discrete time (synchronized) Boolean dynamics was applied to this model: the state of each node can only be active or inactive, and the states of all nodes at time $t+1$ are obtained from the states at time t by applying logical and strictly deterministic update rules (Fig. 4, right panel) reflecting network connections. The pair rule patterning is fed into the network through the state of the SLP gene which is taken as static in this model: it is active ('on') in the first part of the parasegment and inactive ('off') in the second. Notice that some models of the segment polarity network take SLP as a dynamic node (Ingolia and Murray, 2004; Ma et al., 2006).

With the given SLP pre-patterning, the modeled parasegment has two reflection symmetry axes (Fig. 1). The output pattern returned after the SLP pattern is processed by the segment polarity network has less symmetry, even though the network model is entirely symmetric with respect to the two axes. For example, the engrailed pattern is 1, 0, 0, 0 in the four cells and lacks the axial symmetries. This is a typical case of *dynamical symmetry breaking*: because the network is identical in all four cells, this symmetry can only be broken dynamically when the segment polarity genes are activated. Loss of symmetry is accompanied with a multiplication of the possible stable patterns, e.g. for engrailed 0, 1, 0, 0 is an equally likely pattern, obtained from 1, 0, 0, 0 by the reflection about the two symmetry axes. In noisy simulations, the dynamics may stabilize one or the other state depending on initial conditions.

In order to study the dynamical robustness of the symmetry breaking (see below for the exact definition), we simulated the model in the presence of noise: at each time step, the state of a node was allowed to either flip randomly with a fixed probability (node flipping probability: NFP) or to change according to the Boolean function in the model (see Stoll et al., 2006 for a detailed description of this type of noisy simulation). This time evolution generates an ergodic Markov process with a unique stationary state: the time average along a trajectory converges to a unique probability distribution. The consequence is that this asymptotic distribution does not depend on the initial state of the dynamics. We sampled the stationary states, using BNetDyn (http://www.vital-it.ch/Software/) (Stoll et al., 2006) by simulating 200,000 time-steps with NFP = 0.005 on the Boolean network of Albert. We checked that small variations of NFP do not affect qualitatively the results shown below. The choice of 200,000 time-steps is necessary to reach convergence, i.e. quantities which are averaged on trajectories do not depend on the initial condition. With a smaller NFP, longer trajectories would be necessary. Nevertheless, our value for NFP is appropriate since the orbit from the initial condition as in Albert and Othmer (2003) to the wild-type pattern has a probability of 20% to occur unperturbed. Hence we are probing a regime where the network model strongly influences the dynamics.

To evaluate how strongly a trajectory is attracted to the wild-type pattern, we estimate the dissimilarity between the stationary state and the wild-type pattern. Specifically, we measure their *Hamming distance*, namely the number of genes with different expression in the two patterns. Accordingly the set of all possible final states and their respective probabilities induces a distribution of Hamming distances. Mathematically, suppose that S is a given Boolean state and R is the reference Boolean state (the wild-type pattern of segment polarity gene expression), then the Hamming distance is given by $H(R, S) = \sum_i |R_i - S_i|$. Given a stationary state, i.e. a set of Boolean states $B^{(j)}$ with their respective probability p_j, the Hamming distance distribution is given by the formula $d(H) = \sum_j p_j \delta_{H,H(R,B^{(j)})}$ (where δ is the kronecker symbol).

To validate the proposed interactions in a continuous model we adapted the ODE model in von Dassow et al. (2000) to our topology using the Java software (http://rusty.fhl.washington.edu/ingeneue/). The main modifications consisted in changing the topology to a one-dimensional chain of four cells with periodic boundary conditions, adding the constant SLP input, and implementing the proposed interactions. All modified files for the Ingeneue program (models, affectors and iterators) are provided at http://wiki.epfl.ch/naeflab/patterning. Every model was run at least 100,000 times to assess the percentage of correct patterns.

The Boolean steps distribution was evaluated by a slight modification of BNetDyn. NFP was set at 0.005, sample was done on 5000 trajectories. We checked that the errors are beyond visualization precisions in Figs. 9 and 10, by varying the seed of the random numbers generator.

Authors' contributions

GS carried out the Boolean analysis of the network model, participated in the software development and drafted the manuscript. MB carried out the ODE analysis of the network model and helped to draft the manuscript. JR participated in the software development and helped to draft the manuscript. FN coordinated the work and helped to draft the manuscript.

Acknowledgements

GS, MB and FN were supported by the NCCR molecular Oncology program. GS is member of the "Equipe Biologie des Systèmes du Cancer" from the U900 department, équipe labellisée par La Ligue Nationale Contre le Cancer. We thank Christian Iseli and Eric Viara for their help in data and software management.

References

Albert, R., Othmer, H.G., 2003. The topology of the regulatory interactions predicts the expression pattern of the segment polarity genes in *Drosophila melanogaster*. J. Theor. Biol. 223, 1–18.
Barkai, N., Leibler, S., 2000. Circadian clocks limited by noise. Nature 403, 267–268.
Becskei, A., Serrano, L., 2000. Engineering stability in gene networks by autoregulation. Nature 405, 590–593.
Carroll, S.B., 1990. Zebra patterns in fly embryos: activation of stripes or repression of interstripes? Cell 60, 9–16.
Chaves, M., Albert, R., Sontag, E.D., 2005. Robustness and fragility of Boolean models for genetic regulatory networks. J. Theor. Biol. 235, 431–449.
Chen, K.C., Calzone, L., Csikasz-Nagy, A., Cross, F.R., Novak, B., Tyson, J.J., 2004. Integrative analysis of cell cycle control in budding yeast. Mol. Biol. Cell 15, 3841–3862.
Dayarian, A., Chaves, M., Sontag, E.D., Sengupta, A.M., 2009. Shape, size, and robustness: feasible regions in the parameter space of biochemical networks. PLoS Comput. Biol. 5, e1000256.
DiNardo, S., O'Farrell, P.H., 1987. Establishment and refinement of segmental pattern in the *Drosophila* embryo: spatial control of engrailed expression by pair-rule genes. Genes Dev. 1, 1212–1225.
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, 604–609.
Hatini, V., DiNardo, S., 2001. Divide and conquer: pattern formation in *Drosophila* embryonic epidermis. Trends Genet. 17, 574–579.
Hidalgo, A., Ingham, P., 1990. Cell patterning in the *Drosophila* segment: spatial regulation of the segment polarity gene patched. Development 110, 291–301.
Ingham, P.W., McMahon, A.P., 2001. Hedgehog signaling in animal development: paradigms and principles. Genes Dev. 15, 3059–3087.
Ingolia, N.T., Murray, A.W., 2004. The ups and downs of modeling the cell cycle. Curr. Biol. 14, R771–777.
Jaeger, J., Surkova, S., Blagov, M., Janssens, H., Kosman, D., Kozlov, K.N., Manu Myasnikova, E., Vanario-Alonso, C.E., Samsonova, M., et al., 2004. Dynamic control of positional information in the early *Drosophila* embryo. Nature 430, 368–371.
Kitano, H., 2007. Towards a theory of biological robustness. Mol. Syst. Biol. 3, 137.
Kobayashi, M., Fujioka, M., Tolkunova, E.N., Deka, D., Abu-Shaar, M., Mann, R.S., Jaynes, J.B., 2003. Engrailed cooperates with extradenticle and homothorax to repress target genes in *Drosophila*. Development 130, 741–751.
Li, F., Long, T., Lu, Y., Ouyang, Q., Tang, C., 2004. The yeast cell-cycle network is robustly designed. Proc. Natl. Acad. Sci. U.S.A. 101, 4781–4786 (Epub 2004 March 4722).
Ma, W., Lai, L., Ouyang, Q., Tang, C., 2006. Robustness and modular design of the *Drosophila* segment polarity network. Mol. Syst. Biol. 2, 70.
Nusslein-Volhard, C., Wieschaus, E., 1980. Mutations affecting segment number and polarity in *Drosophila*. Nature 287, 795–801.
Sanchez, L., Thieffry, D., 2003. Segmenting the fly embryo: a logical analysis of the pair-rule cross-regulatory module. J. Theor. Biol. 224, 517–537.
Sanson, B., 2001. Generating patterns from fields of cells. Examples from *Drosophila* segmentation. EMBO Rep. 2, 1083–1088.
Stoll, G., Rougemont, J., Naef, F., 2006. Few crucial links assure checkpoint efficiency in the yeast cell-cycle network. Bioinformatics 22, 2539–2546.
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.
Thomas, R., Kaufman, M., 2001. Multistationarity, the basis of cell differentiation and memory. II. Logical analysis of regulatory networks in terms of feedback circuits. Chaos 11, 180–195.
van den Heuvel, M., Klingensmith, J., Perrimon, N., Nusse, R., 1993. Cell patterning in the *Drosophila* segment: engrailed and wingless antigen distributions in segment polarity mutant embryos. Dev. Suppl., 105–114.
von Dassow, G., Meir, E., Munro, E.M., Odell, G.M., 2000. The segment polarity network is a robust developmental module. Nature 406, 188–192.