molecular
systems
biology

An atlas of gene regulatory networks reveals multiple three-gene mechanisms for interpreting morphogen gradients

James Cotterell[1,2] and James Sharpe[1,3,*]

[1] EMBL-CRG Systems Biology Research Unit, Centre for Genomic Regulation (CRG), UPF, Barcelona, Spain, [2] MRC Human Genetics Unit, Edinburgh, Scotland, UK and [3] ICREA, Centre for Genomic Regulation (CRG), UPF, Barcelona, Spain
* Corresponding author. EMBL-CRG Systems Biology Research Unit, Center for Genomic Regulation, PRBB, 88 Dr Aiguader, Barcelona 08003, Spain.
Tel.: + 34 93 316 0098; Fax: + 34 93 316 0099; E-mail: james.sharpe@crg.es

The interpretation of morphogen gradients is a pivotal concept in developmental biology, and several mechanisms have been proposed to explain how gene regulatory networks (GRNs) achieve concentration-dependent responses. However, the number of different mechanisms that may exist for cells to interpret morphogens, and the importance of design features such as feedback or local cell–cell communication, is unclear. A complete understanding of such systems will require going beyond a case-by-case analysis of real morphogen interpretation mechanisms and mapping out a complete GRN 'design space.' Here, we generate a first atlas of design space for GRNs capable of patterning a homogeneous field of cells into discrete gene expression domains by interpreting a fixed morphogen gradient. We uncover multiple very distinct mechanisms distributed discretely across the atlas, thereby expanding the repertoire of morphogen interpretation network motifs. Analyzing this diverse collection of mechanisms also allows us to predict that local cell–cell communication will rarely be responsible for the basic dose-dependent response of morphogen interpretation networks.
Molecular Systems Biology **6**: 425; published online 2 November 2010; doi:10.1038/msb.2010.74
Subject Categories: simulation and data analysis; development
Keywords: design space; gene network; morphogen; patterning; systems biology

Introduction

Understanding the relationship between gene regulatory network (GRN) design and function is a central problem in biology. For simple cellular functions (such as bistability and oscillations) identifying network *motifs* has been a productive approach, providing a 'conceptual toolkit' for understanding network design principles (Shen-Orr *et al*, 2002; Mangan and Alon, 2003). However, the relationship between GRN topology (the wiring design) and biological function may not always be straightforward (Chouard, 2008). The ability of many different topologies to encode the same biological function has recently been explored theoretically through the use of genotype–phenotype (GP) maps. These studies revealed that a large number of different topologies could all achieve the same biological function, but intriguingly they could be explained in each case by a common underlying dynamical mechanism (Ma *et al*, 2006; Hornung and Barkai, 2008).

These previous studies applied the GP map approach to highly constrained functions. Such a limited repertoire of dynamical explanations may not be the norm for less constrained functions for which a GP map may be able to uncover a more elaborate mechanism space. To explore this possibility, we applied the GP map approach to mechanisms of morphogen interpretation for which at least several different mechanistic possibilities have been suggested (Lander, 2007). In this study, therefore, we wished to address the following questions: Can we map out the range of mechanistic possibilities that underlie this function? Can such a map serve as a useful theoretical framework in which to explore the general patterning principles for this function?

Results

Exploring design space for a realistic model of development

We chose to explore the mechanisms of morphogen interpretation as multiple mechanistic possibilities have been suggested for this system and is thus a good candidate for

possessing an elaborate mechanism space. Morphogen interpretation is the second step in the two-step process of morphogen-based patterning, the first step being morphogen gradient formation and maintenance that we do not address in this study.

To explore the range of possible morphogen interpretation mechanisms, we sought a biologically verified model of gene regulation. We therefore adapted the continuous mathematical model developed over the last 20 years by Reinitz *et al* (Mjolsness *et al*, 1991; Reinitz and Sharp, 1995; Jaeger *et al*, 2004), which quantitatively captures the spatiotemporal dynamics of gap gene patterning in response to the Bicoid morphogen gradient during *Drosophila* embryogenesis. Our model includes time delays encoded by synchronized updating of gene states in discretized time steps. The gene dynamics depend on the following model parameters: the strength and sign of the interactions between genes, degradation rates and importantly also on cell–cell communication (which is represented by a diffusive process—see Materials and methods for full details). We have also added a new term into the model, to represent stochastic molecular noise. Noise was included in the simulations as the importance of robustness of developmental mechanisms with respect to stochastic noise has been highlighted in the past (Kerszberg, 2004). Our noise term describes temporal fluctuations of molecular concentrations

that generate gene expression variability comparable to that seen in real patterning systems (Supplementary Data section S1). We simulated a one-dimensional spatial system comprising 32 nuclei with a fixed morphogen gradient across the field, and chose a single stripe of expression as the target pattern (Figure 1A), because it represents a particular example of morphogen interpretation and is a simplified version of the well-known and much-studied French Flag problem (Wolpert, 1968). In all, 32 cells were chosen for the simulations because it represents a typical size for a morphogenetic field found in many real patterning systems (Briscoe *et al*, 2001; Wijgerde *et al*, 2002; Bayly *et al*, 2007). As we are searching for the general design principles of stripe-forming networks, our criteria allow stripes of varying widths and positions within the field (see Materials and methods). Furthermore, there are no restrictions on the gene expression time course taken in order to arrive at this final gene expression pattern.

We hypothesized that exploring this system using discrete topologies would serve as a convenient, efficient and meaningful way to represent what is in fact a vast and continuous multidimensional parameter space (Figure 1B–E). The difficulty of using the full continuous space can be considered in the following way. Exploring networks with different numbers of regulatory interactions involves sets of simulations in which certain parameters are kept to zero. In practice, these networks
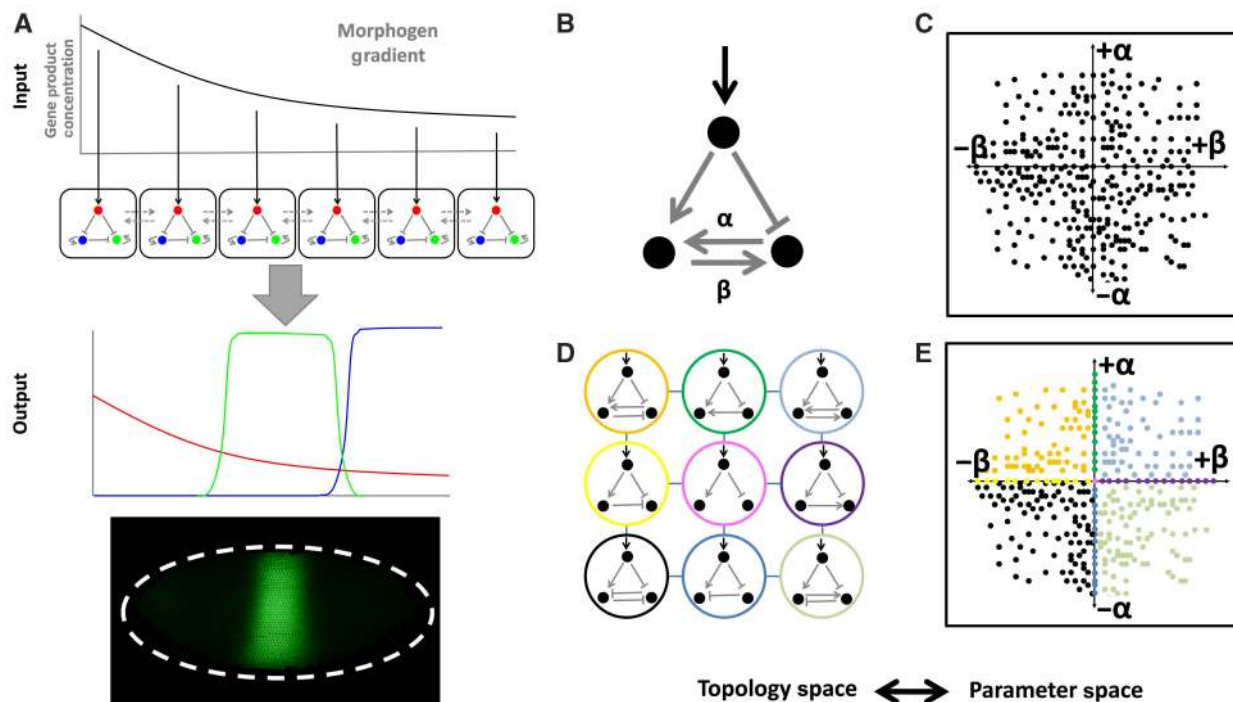

Figure 1 Combining topology space with a realistic model of gene regulation. (**A**) Our model of development is derived from a realistic model of *Drosophila* anterior–posterior patterning (Mjolsness *et al*, 1991; Reinitz and Sharp, 1995; Jaeger *et al*, 2004). This spatial model consists of a one-dimensional row of cells with the GRN repeated in each cell. Cells can signal to one another by means of diffusible gene products (dashed arrows). Specifically, we look for GRNs that have the ability to generate an output with a single stripe of gene expression (green line) by interpreting a morphogen input signal in the form of a gradient (black line). An example of a single stripe of gene expression for the *Krüppel* gene is shown, the data for which was taken from the FlyEx database (Poustelnikova *et al*, 2004; Pisarev *et al*, 2009). (**B**) A GRN topology where two of the gene–gene interactions α and β correspond to the parameter space in (C). (**C**) A parameter space of the two parameters α and β. Dots are random parameter sets from this space. (**D**) A topology space is created if all values of α and β that are positive are considered gene–gene activations, those values of α and β that are negative are considered gene–gene repressions and those values of α and β that are 0 are considered to generate no gene–gene interaction. Regions of parameter space corresponding to the different topologies are indicated by the different colored circles surrounding the topologies and the different colored dots in (**E**). Where topologies differ by a single gene–gene interaction (one Hamming distance) they are linked by a blue line. Such links connect regions of close parameter space.

have fewer parameters than the more complex ones—in other words they reside in a parameter space with fewer dimensions. Representing this full range of complexities (with differing dimensionalities) within a single continuous parameter space is therefore inefficient. In a sense, dimensions of the system 'collapse' as we go from complex networks to simple ones. The discrete topological representation of a continuous parameter space is illustrated in Figure 1. Furthermore, we demonstrated that changes in topology have a greater effect on the resulting phenotype and dynamics than a corresponding simple parameter change (that does not change the topology) (Supplementary Data section S2). Therefore, the topology-focused approach should bias our sampling of parameter space towards a greater diversity of mechanisms.

We chose to simulate all possible topologies of gene network consisting of three genes. The major advantage of choosing three-gene networks is the ability to perform exhaustive analysis on all possible topologies. There is much evidence suggesting that larger biological circuits are comprised of combinations of small network modules linked together (Milo *et al*, 2002; Thieffry and Sanchez, 2003). An example from development is the modules found for endomesoderm specification in the sea urchin (Peter and Davidson, 2009). Furthermore, it has been previously demonstrated that the dynamics of larger networks can be accurately captured using models that reduce the system to include only the more important information processing components (Ingolia and Murray, 2004). Hence, detailed analysis of the behavior of small networks has direct relevance to real patterning systems.

To explore the full design space of three-gene networks, we enumerated every possible unique topology with the condition that one of the genes was activated by the morphogen. This resulted in 9710 topologies when including every permutation by which the morphogen can activate one of the genes (see Materials and methods). As we use a realistic continuous model of gene regulation, topology alone is not enough to define a GRN, and we define a 'genotype' as a specific topology with a specific set of parameters (dots in Figure 1C and E; the strengths of gene–gene interaction and also diffusion rates are variable in this study). We tested 30 000 randomly chosen parameter sets for each topology. Hence, we assessed ∼300 million different genotypes (9710 topologies × 30 000 parameter sets). Each genotype was simulated four times with a different series of random molecular fluctuations and the resulting expression pattern tested against our stripe definition and noise robustness criteria (stability against random stochastic fluctuations—see Materials and methods). From the 300 million genotypes tested, 3379 were able to produce a stripe in a noise-robust manner. These successful 'solutions' mapped to 471 GRN topologies of the fully enumerated set of 9710. Examples of the gene expression time courses of successful solutions are shown in Supplementary Figure S5.

The *complexity atlas* reveals that a variety of distinct dynamical mechanisms can achieve the same function

A major technical challenge is how to explore the dynamical mechanisms of so many (3379) different solutions. If we were

not using discrete topologies but instead were simply exploring the continuous multidimensional parameter space (Figure 1B and C), then one potential approach would be to ignore topological information, and to perform unsupervised clustering on the parameter values for all the 3379 successful solutions. To explore this approach, we measured the Euclidean distances between all pairwise combinations of successful parameter sets that were functional (3379 solutions). For each parameter set pair we measured the Euclidean distance in all possible permutations of the GRN topology, and the lowest Euclidean distance for any of the permutations was taken as the distance for this pairwise combination. The result of agglomerative hierarchical clustering is shown in Supplementary Figure S6, demonstrating that the algorithm failed to find any significant groups or clusters within the dataset.

We therefore explored a new topology-focused approach. Previous studies considered the full collection of functional topologies as a list or *ensemble* (Ma *et al*, 2006; Hornung and Barkai, 2008). By contrast, we hypothesized that by explicitly linking topologies together into a non-directed graph based on topological similarity (similar to neutral networks in the field of evo-devo; Smith, 1970; Schuster *et al*, 1994; Ciliberti *et al*, 2007), the shape of this 'metagraph' (Ciliberti *et al*, 2007) might provide insight into the underlying mechanisms. To link our 471 functional GRN topologies into a connected graph, we defined topologies (nodes) as neighbors (edges) if the addition or removal of a single gene–gene interaction could change one GRN topology into the other (Figure 1D). We found that 448 (95%) of the 471 functional topologies form a single connected graph. We termed the resulting connected graph an 'atlas' as the notion of neighborhood allows the similarity relationships between all topologies to be mapped-out explicitly. We next reorganized the layout of this atlas to generate a 'complexity atlas' (Figure 2), which highlights the regions with more regulatory links (higher complexity) versus fewer links (lower complexity, where minimal topologies will reside). This approach immediately revealed a striking asymmetry of the atlas—whereas the higher levels of complexity show a fairly uniform distribution of successful stripe-forming topologies, less complex networks converge onto multiple discrete core topologies (highlighted by dashed colored lines in Figure 2). The separation of these 'stalactites' suggested that they may each represent a distinctly different way of achieving a stripe—islands of related GRN topologies that share a common mechanism. We identified six-core topologies residing at the bottom of stalactites for further analysis, based on their mutational robustness (see Supplementary Figure S7 and Supplementary Materials and Methods section S3); robustness to topological changes ('width' of the stalactite) and to parameter variation (shading).

If our six-core topologies employ distinct mechanisms, this should be reflected by distinctly different gene expression dynamics over space and time. The gene expression time course for each successful solution can be represented by a space–time plot (Figure 3A), and we therefore performed hierarchical clustering on a simplified version of these plots (taking into account every permutation of the three genes—see Materials and methods section Clustering the space-time plots). We employed a Euclidean distance difference function between the space–time plots for each solution. To control for
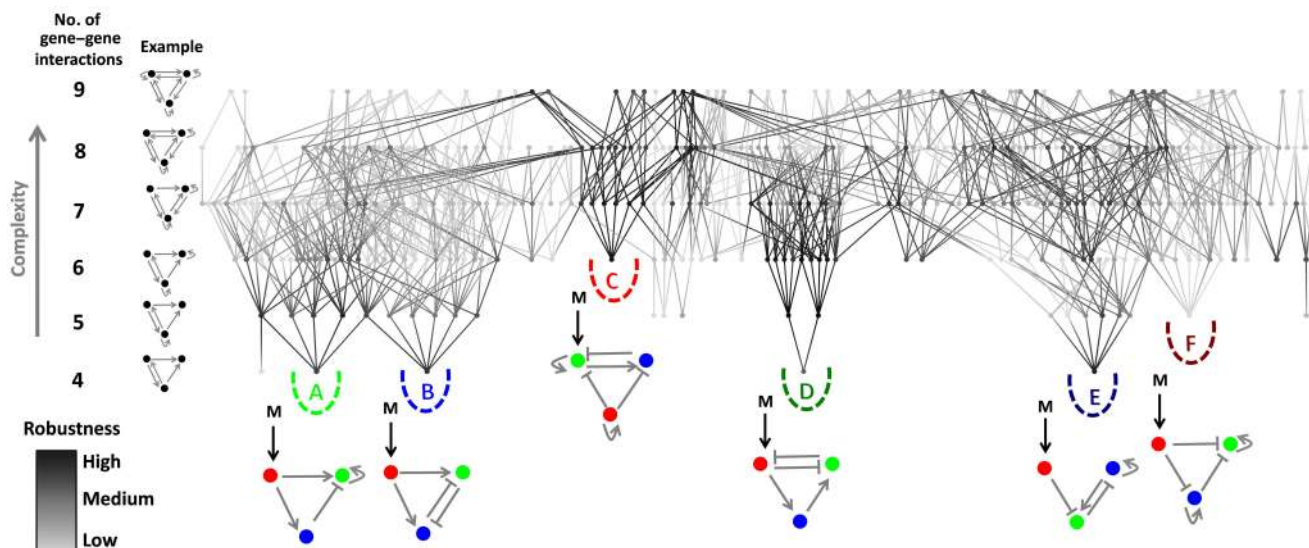
Figure 2 Creating a complexity atlas reveals the core topologies for single-stripe patterning. Nodes are GRN topologies, and edges link those with a single-topological change. The GRN topologies are laid out such that topological complexity increases up the *y* axis (see left-hand key). Spacing along the *x* axis is organized to reduce edge crossing. This reveals a striking structure to the network, in which 'stalactites' are seen protruding from the bottom of the atlas. These stalactites converge downwards to individual 'core' topologies (which are illustrated below the atlas). Mutational robustness, as measured by the fraction of functional parameter space, is shown by the shading (darker topologies are more robust). The six-core topologies chosen for further investigation are shown beneath along with their corresponding color-coded label (A–F). The morphogen input is denoted by 'M.' Genes in each core topology have the following color code: green is the stripe gene, red is the gene receiving the morphogen input (if it is not also the stripe gene) and blue is the default remaining gene. If the gene receiving the morphogen input is also the stripe gene, it is green and the remaining two genes are randomly assigned a red or blue color.

translation and scaling effects in the stripe pattern, we compared the gene expression profiles at just three positions—at the center of each of the low–high–low regions of the stripe pattern. We compared every permutation of the genes for each space–time plot comparison, and the smallest distance of each of these permutation comparisons was taken as the distance between these two space–time plots. This revealed that the spatiotemporal dynamics of the gene expression time courses naturally cluster into six groups (Figure 3B). Furthermore, these groups display a 1-to-1 correspondence with the six-core topologies, strengthening the hypothesis that they represent six different mechanisms.

The possibility of six mechanisms to achieve the same function was intriguing because previous studies with more restrictive functions found a common dynamical mechanism to explain them (Ma *et al*, 2006; Hornung and Barkai, 2008). Hence, to confirm whether our six stalactites are genuinely distinct dynamical mechanisms, or just modified versions of each other, we directly studied the dynamics and mechanistic requirements of each one in detail (Figure 3C; Box 1). The distinct nature of these five is highlighted by the variety of patterning strategies that they employ. The core topologies of two of these mechanisms are entirely feed-forward networks (A and F). One of these is the well-known incoherent type 1 feed-forward motif in dose-response mode (Kim *et al*, 2008), whereas the other utilizes the two upstream genes as inhibitors defining the anterior and posterior boundaries of stripe gene expression, respectively. Another mechanism (C) depends on an intrinsic oscillator, which is rapidly 'frozen' to produce the spatial stripe. The core topology of this mechanism has a modular structure allowing 'roles' to be assigned to each of the modules. One of the modules acts as the oscillator, whereas the

other acts as the repressor that freezes the oscillator after a period of time. Mechanism (B) employs mutual inhibition where the two mutually repressing genes have opposite expression profiles. Finally, mechanism (E) contains a two-gene bistable module that is modulated by a gradient of repression to generate a stripe gene expression profile (all these dynamics are explained in more detail in Box 1).

Strikingly, although cell–cell communication was allowed for all the genotype assessments, five of the six mechanisms make no use of this option—they interpret the morphogen gradient in a completely cell-autonomous manner acting like a band-pass filter. The prevalence of cell-autonomous mechanisms is heightened by the discovery of one mechanism D, which by contrast is absolutely dependent on local cell–cell communication. It involves overlapping activator and repressor expression domains to create a zone of net activation for the stripe gene. Importantly, it does not use cell–cell communication just to sharpen the stripe or to make it more robust with respect to stochastic noise—the absence of communication between neighboring cells renders it completely unable to respond in a concentration-dependant manner.

Design features of stripe-forming GRNs

Which design features do the stripe-forming GRNs tend to employ? Can this tell us anything about the functional significance of design features observed in real morphogen interpretation systems? In general, the core GRNs rely on three particular design features, which are over represented in the complexity landscape compared with what would be expected by chance; feed-forward, autopositive feedback and negative feedback (these design features are discussed in more detail in
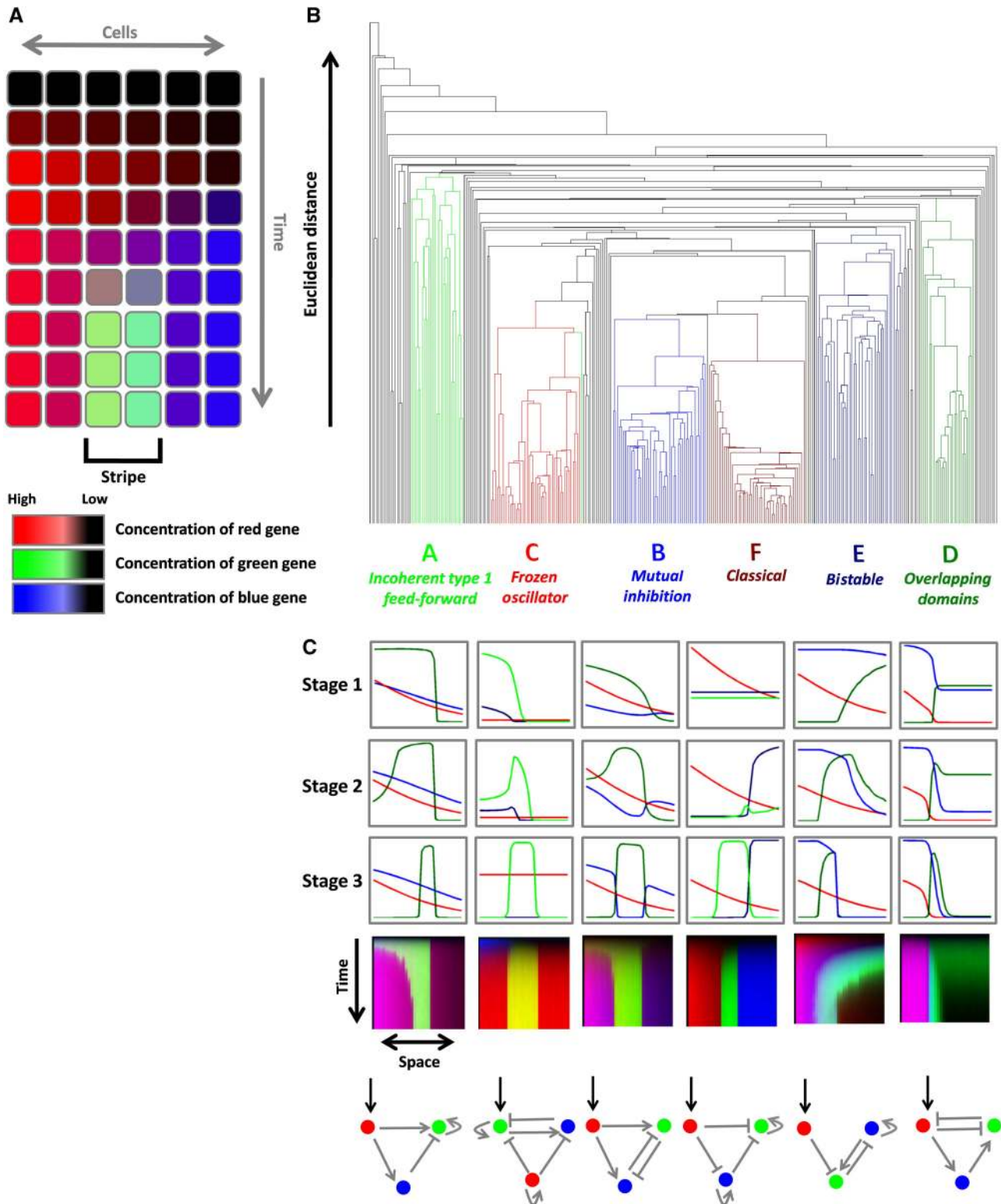
Figure 3 The six mechanisms are distinct and (**A**) How to construct a space–time plot. The concentration of the three genes in each cell (*x* axis) at each time point (*y* axis) is indicated by the intensity of the red, green or blue, respectively. (**B**) A hierarchical clustering of 300 solution space–time plots, 50 of which were randomly generated for each core topology. This tree shows that the six mechanisms are distinct as the space–time plots naturally cluster into six separate categories. Branches of the six main groups are colored according to the core topology from which the solution derives. (**C**) Gene expression graphs of each mechanism at different stages are shown below the tree along with a single space–time plot that captures these stages in a single image. A narrative description of each of the mechanisms can be found in Box 1. The core topologies for each of the mechanisms are also shown beneath the space–time plots.

Box 1 Comparison of the six mechanisms

See Figure 3C for corresponding gene expression profiles of the different stages.

(A) Incoherent type 1 feed-forward
Description: (Stage 1) The red gene is activated by the morphogen and hence starts to mimic its gradient pattern. The red gene activates the green gene, and switches on the green genes positive feedback above a certain threshold. (Stage 2) A combination of time and dose leads to the concentration of the repressing blue gene product building up to a high enough concentration for it to force the green gene product down, but only on the left-hand side where the concentration is higher. (Stage 3) In the central region, the repression from the blue gene product is lower allowing the green gene product to reach a high steady state.

(B) Mutual inhibition
Description: One of the real mechanisms found to be involved in *Drosophila* anterior–posterior patterning. (Stage 1) The red gene product mimics the expression pattern of the morphogen. The green gene is activated more strongly than the blue gene and hence switches on earlier on the anterior side where the red activating gene product is higher. (Stage 2) Eventually the blue gene product builds up to a high enough concentration to start forcing the green gene product down on the very anterior side where the concentration is highest. (Stage 3) The result is a final gene expression pattern where the blue and green gene products form two mutually exclusive expression zones.

(C) Frozen oscillator
Description: (Stage 1) The morphogen sets the green gene and the blue gene products oscillating because they are in a negative feedback loop. The phase of the oscillation is different in different cells because of the difference in the strength of the morphogen activation. (Stage 2) The red gene product starts to build up everywhere in a uniform distribution because of its positive feedback. (Stage 3) It represses both the green and blue gene and stops the oscillations forcing both genes off, except in the central region, which are in a phase of the oscillation allowing the concentrations to reach a high steady state.

(D) Overlapping domains
Description: This mechanism is completely dependent upon diffusion. The stripe gene is activated and inhibited by the two genes. Because of the activator having a higher diffusion constant than the inhibitor, the expression domain of the activator extends further than that of the inhibitor allowing a region for a stripe to form. (Stage 1) The red gene product starts to form a gradient as it is activated by the morphogen. The red gene strongly activates the blue gene giving it a similar expression profile though at a higher concentration. The blue gene product activates the green gene, and the red gene inhibits the green gene meaning the green gene product can only start to increase in concentration on the right-hand side. The green gene inhibits the red gene, which causes the sharp threshold break in the red gradient. (Stage 2) The green gene forces the red gene completely off on the right-hand side, in turn leading to decay of the blue gene product and then the green gene product. (Stage 3) Only in the overlap region where the blue gene is activating the green gene and where there is no repression from the red gene can a stripe form.

(E) Bistable
Description: (Stage 1) The red gene is activated by the morphogen, and thus its product also forms a gradient. The blue gene activates itself and thus starts to switch on everywhere. The blue gene also activates the green gene whose product will thus start to increase in concentration, but only on the right-hand side as it is repressed by the red gene on the left-hand side. (Stage 2) The green gene product can build up to a high enough concentration on the right-hand side to start to force down the blue gene product. (Stage 3) The green gene, however, is also dependent upon the blue gene for activation and thus after a delay its product concentration also starts to fall. The result is a single stripe of gene expression.

(F) Classical
Description: One of the real mechanisms found to be involved in *Drosophila* anterior–posterior patterning. (Stage 1) The red gene is activated by the morphogen, and thus its product also forms a gradient. (Stage 2) The blue and the green genes activate themselves and they start to switch on where the repression from the red gene is lowest. (Stage 3) The blue gene also represses the green gene meaning it is forced off at the very right. Only in the central zone where the repression from the red and the blue gene is lowest can the green gene product reach a high steady state.

Supplementary Data section S4). The feed-forward motif is of particular interest as it is employed by four of the core topologies. The prevalence of such mechanisms suggests why in many morphogen patterning systems there is often more than one morphogen input that feeds into many genes of the system by both activation and inhibition. For example, *Drosophila* blastoderm patterning, vertebrate hindbrain patterning and neural tube patterning all employ multiple morphogen inputs (Jaeger *et al*, 2004; Ulloa and Briscoe, 2007; Tümpel *et al*, 2009).

Mechanisms show a discrete distribution within the underlying parameter space

To further explore the distinction between the six mechanisms, we asked whether a simple topological description is enough to specify each mechanism, or whether instead certain parameter restrictions are also necessary. Extensive analysis of correlations between parameter values confirmed the latter to be true, and more importantly that these parameter restrictions are different for each mechanism (Supplementary Figure S8). We thus wished to go further and explore the distribution of mechanisms within underlying continuous parameter space (Figure 1B–E). Are there continuous transitions between mechanisms or are the mechanisms discretely different? To address this question, we tested whether it was possible to smoothly interpolate through parameter space between all pairs of mechanisms without losing functionality (illustrated in Figure 4A and B). We specifically looked at the transition between the solution parameter sets of the core topologies. We explored all pairwise combinations of transitions between these topologies as shown in Figure 4C. For certain pairs of core topologies, a smooth transition would be impossible, as there are topological clashes. In other words,
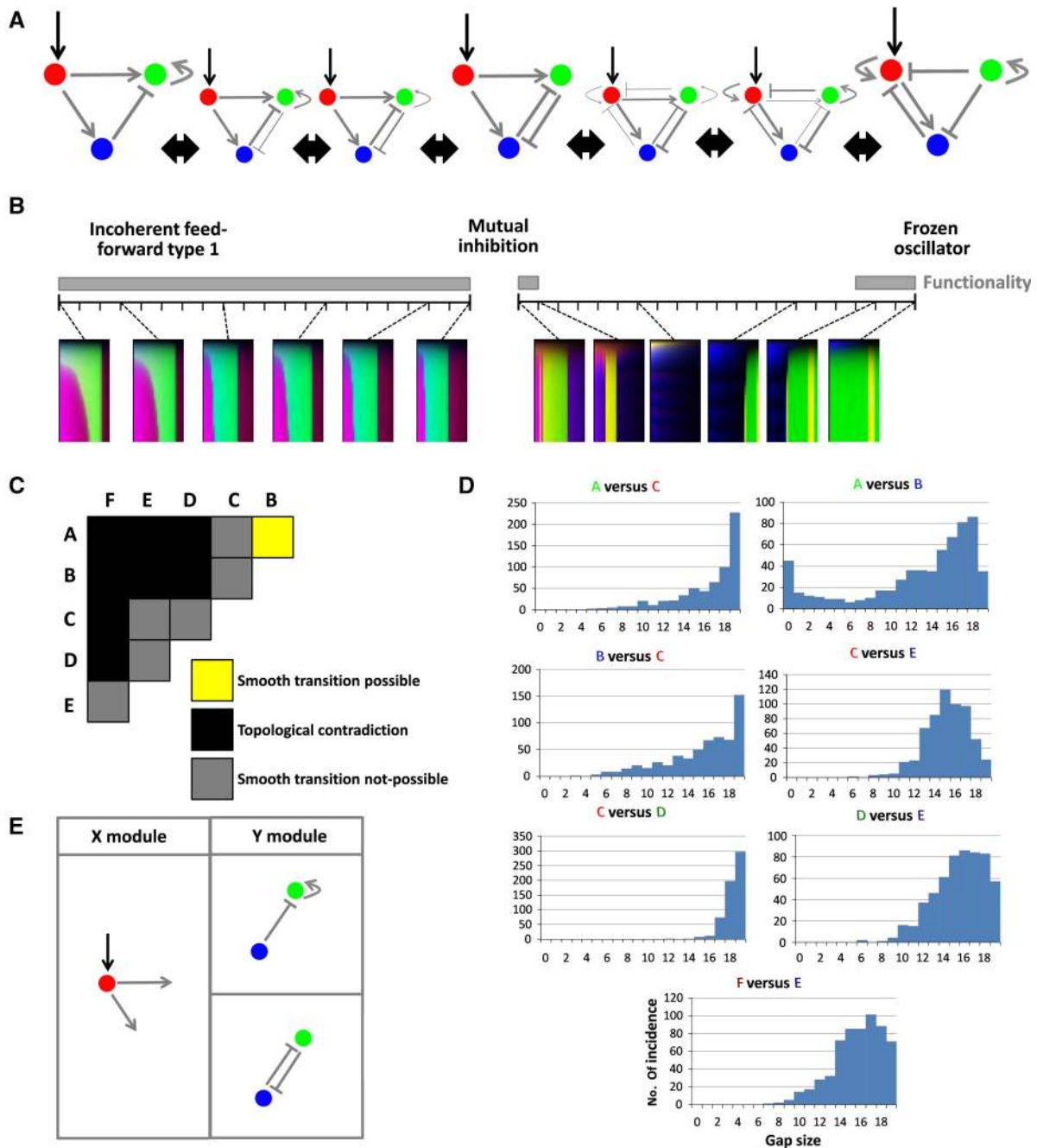
Figure 4 Interpolating between parameter sets reveals the mechanisms to be discrete. (**A**) Illustration of the method of interpolating directly through parameter space using functional parameter sets from core topologies. Gene–gene interactions that are present in one-core topology and not the other are reduced to 0 as the interpolation moves away from that topology. (**B**) Showing two interpolations through continuous parameter space from A-to-B and B-to-C (whose positions on the complexity atlas are indicated). The horizontal black line represents the linear interpolation between the parameter combinations. They gray bar above illustrates where an interpolation was functional. Example of space–time plots from various points of the interpolation are shown beneath. (Left) In the first example (A-to-B), some interpolations are possible without losing the stripe-forming functionality, as illustrated by the continuous gray bar and space–time plots with a stripe at each stage of the interpolation. (Right) By contrast, the interpolation from B-to-C passes through a large non-functional region of parameter space (broken gray bar). (**C**) A triangular matrix describing whether it is possible to interpolate between the core topologies for each mechanism. Where there is a topological contradiction (one-core topology requires a repression, whereas the other requires an activation) squares are colored black, and an interpolation was never tested. For each case where an interpolation was possible, we attempted 625 interpolations between randomly selected parameter sets, the results of which are shown in (D). If a single interpolation was successful, squares are colored yellow. When every interpolation was not successful, the square was colored gray. (**D**) Histograms of 625 attempted interpolations between functional parameter sets from non-topologically contradictory core topologies. For each interpolation, 20 direct equidistant steps were taken. The gap size is the number of non-functional steps within that interpolation. Only the transition A-to-B is possible as indicated by gap sizes of 0. (**E**) The modular nature of mechanisms A and B. The two mechanisms share an X module and only differ depending on which Y module they utilize.

where one core topology requires an activation, the other requires a repression. As a core GRN topology contains only the essential gene–gene interactions for a particular mechanism, a change from activation to a repression or *vice versa* will by definition break the mechanism. Hence, there is a guaranteed non-functional gap between the two functional parameter domains of these core GRN topologies. In situations where there were no topological clashes, we chose to explore what happened as we interpolated between 25 randomly chosen parameter sets responsible for one mechanism to 25 of those responsible for the other. By doing this, we could measure the 'gap size' between the solutions. The gap size is the length of the non-functional region along the interpolation. We could then plot a histogram of the gap size for each of the 625 pairwise parameter interpolations. The results are shown in Figure 4D.

The majority of attempted interpolations have a gap. For six out of the seven mechanisms pairs, we find no interpolations from one mechanism to the other that preserve functionality all the way. The exception to this finding is the interpolation between the incoherent type 1 feed-forward and mutual inhibition mechanisms (A and B), where there are a substantial proportion of interpolations with no gap. As a control, we performed interpolations between parameter sets from topologies of the same mechanism categories each of which had at least some interpolations with no gap (Supplementary Data section S6). Hence, this result strongly suggests that, in general, the nature of mechanism space is discrete as illustrated in Supplementary Figure S10. To our knowledge, this is the first time that the discrete nature of mechanism space has been demonstrated by a concrete example.

This exception of (A-to-B) is due to the inherent modularity of the incoherent type 1 feed-forward and mutual inhibition mechanisms (as illustrated in Figure 4E). Both motifs can be considered as the combination of two modules: a feed-forward module (X), which is the same for both motifs, and a positive-feedback on C (Y), which is either a direct positive feedback (for mechanism A) or an indirect double-negative feedback (for mechanism B). This is reminiscent of the result found in Ma *et al* (2006) for the segment polarity GRN, whereby inherent modularity allowed for 'combinatorial variability' in the GRN topologies able to achieve the 'stripe-sharpening' function.

Mapping design space for morphogen interpretation

To generate a first explicit map of mechanisms of morphogen interpretation, we explored the distribution of the mechanisms across the range of more complex topologies in the full complexity atlas. By comparing the spatiotemporal dynamics of gene expression for each genotype, we were able to assign the majority of solutions to one of the six mechanisms. This was confirmed by an automatic method that resulted in the assignment of 93% of the 3379 solutions (see Materials and methods). In all, 76% of the GRN topologies were found to work via a single mechanism and 2% were found to have the potential to work via two different mechanisms depending on the exact parameter values (Figure 5A; in these cases, the

topology contains the core topology of both mechanisms). This result demonstrates the value of the spatial nature of our atlas—it reveals the explicit mapping of the positions, domains and topological overlaps of the mechanisms relative to each other in 'design space' (Kitano, 2007). Our comparison with real known systems shows that the atlas can serve as a new conceptual framework within which to explore and understand the possible mechanisms of metazoan pattern formation and their relationships to each other.

The atlas uncovers known motifs for morphogen interpretation and also predicts new ones

We next explored whether our atlas of dynamical mechanisms had revealed any real known morphogen interpretation systems (Figure 5B). Within the currently accepted gap gene network (which interprets the Bicoid morphogen in *Drosophila*), we find motifs corresponding to two of our core GRN topologies (Jäckle *et al*, 1986; Kraut and Levine, 1991; Clyde *et al*, 2003). The mutual inhibition in B reflects either the 'alternating cushions' or 'mutually exclusive domains' model of gap gene expression that occurs between *kni* and *hb* and between *gt* and *Kr* (Kraut and Levine, 1991; Clyde *et al*, 2003; Vakulenko *et al*, 2009). In *Drosophila*, it is believed that this model is also complemented by a more hierarchical mechanism like F, in which different genes define the anterior and posterior boundaries of each gap gene, for example *hb* and *kni* defining the anterior and posterior boundary of *Kr* (Clyde *et al*, 2003). The core topology of A is found within the GRN that controls the mesoderm inducer *Xenopus* Brachyury (XBra) expression (Green, 2002). XBra is strongly expressed only at intermediate levels of activin signaling (Latinkić *et al*, 1997).

This correspondence between three of our motifs and real developmental networks in both vertebrates and invertebrates suggests that the remaining topologies are also likely to be biologically realistic (such as C, D and E), and can now be added to the repertoire of possible stripe-forming networks. The identification of real developmental motifs emphasizes the power of our technique to complement the motif approach of previous studies (Shen-Orr *et al*, 2002; Mangan and Alon, 2003). Rather than searching for overrepresented motifs in real network data and then studying what dynamical behavior they are capable of, here we perform the reverse—we start by identifying theoretically functional motifs and then search for their occurrence within real networks.

Performance features of the six mechanisms

We next explored the dynamical performance of the six mechanisms to investigate why certain motifs are observed in real biological systems and the likelihood of employment of the non-observed mechanisms in real biological systems. A summary of performance features of the six mechanisms is shown in Table I, and a more extensive discussion can be found in Supplementary Data section S8. We analyzed features including mutational and noise robustness, how fast equilibrium is reached, the number of cell types generated and the typical features of stripes produced by the six-core mechanisms. Two mechanisms deserve special attention following
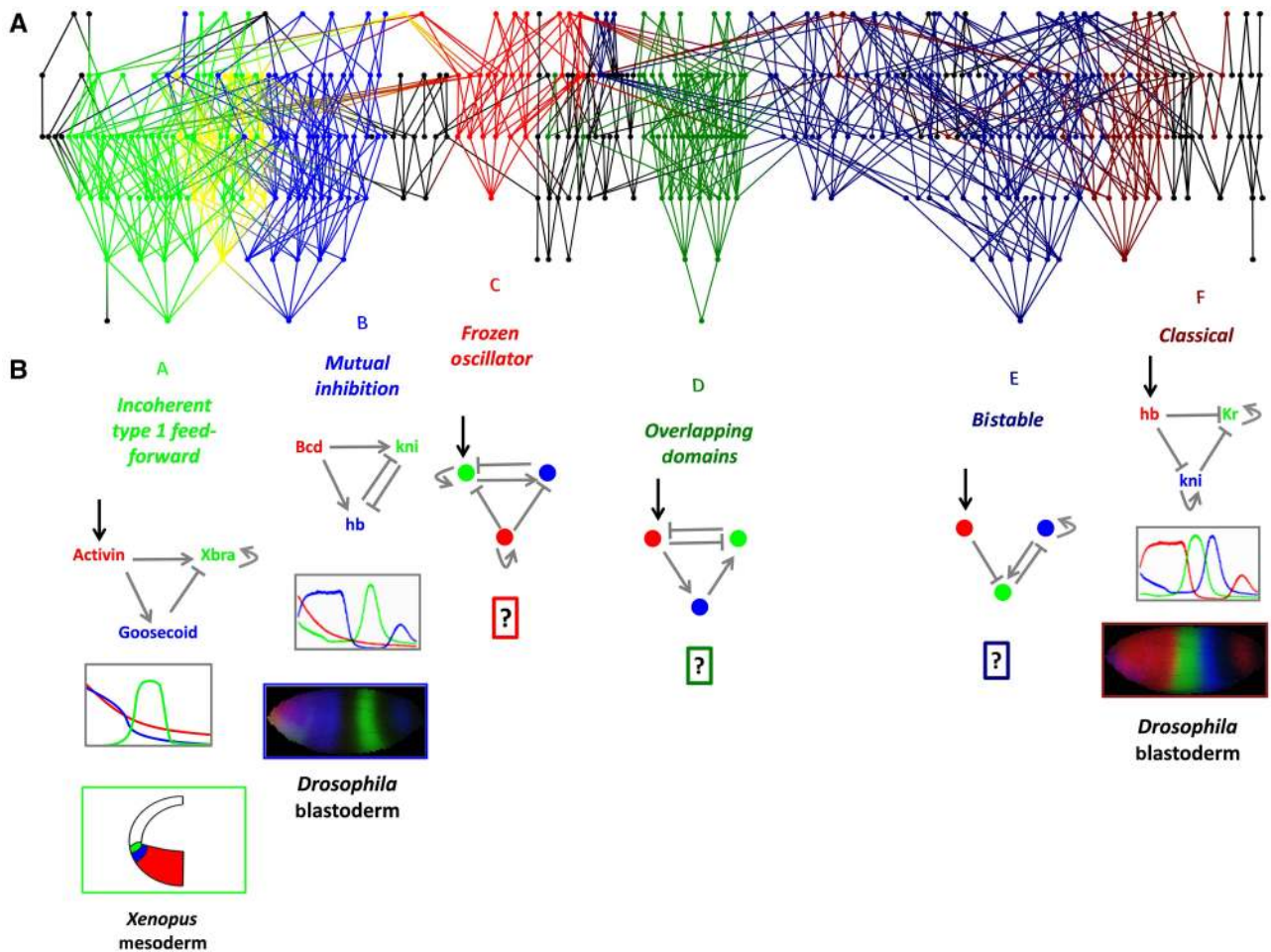
Figure 5 Mapping the six mechanisms to the complexity atlas provides the first explicit map of mechanisms for morphogen interpretation. (**A**) The topologies of the complexity atlas are colored according to the mechanism by which they produce the single stripe of gene expression (as shown in Figures 2 and 3). Topologies that are capable of performing multiple mechanisms are shown in yellow. Mechanisms occupy locally connected regions of the complexity atlas, and together cover 78% of the topologies. (**B**) The known biological systems that each morphogen interpretation mechanism is associated with are illustrated by the core topologies and the images beneath. Morphogen interpretation mechanisms from diverse contexts including *Drosophila* and *Xenopus* can be seen. Two of the mechanisms (B and F) are involved in *Drosophila* axial patterning. The gap gene subnetworks that correspond to our mechanisms are shown in the bounded region along with the real quantified gene expression patterns for the corresponding genes (from the FlyEx database; Poustelnikova *et al*, 2004; Pisarev *et al*, 2009). The core topology of A is found within the GRN that controls the mesoderm inducer *Xenopus* Brachyury (XBra) expression (Green, 2002). Schematic versions of the gene expression pattern of the three genes in the XBra control network are shown beneath. Those mechanisms labeled with a '?' are those that have not been observed in real biological contexts. These can be added to our repertoire of morphogen interpreting mechanisms.

Table I A summary of the performance features of the six-core mechanisms

Mechanism	Robustness			Patterning speed (min)	Number of cell types generated	Irreversibility
	To parameter changes (mutational) rank	To noise rank	To dynamic morphogen			
IFF	3	1	Fragile	125	2/3	Yes (without cell–cell communication)
MI	4	4	Robust	75	2/3	No
FO	1	2	Robust	45	2	Yes
OD	5	5	Robust	215	3	No
BI	2	6	Robust	80	3	No
Classical	6	3	Robust	110	3	Yes (without cell–cell communication)

Details of each of the individual analysis can be found in Supplementary Data section 8. Noise robustness rank is that measured at the 1% per time step noise level. Mutational robustness is measured as the number of solutions found for each of the core topologies. Patterning speed was converted to minutes from number of iterations using the assumption of 25 s per iteration as described in Supplementary Data section 1.

the analysis of these features; the frozen oscillator and mutual inhibition mechanisms.

The undescribed frozen oscillator mechanism is a strong candidate for employment in real biological systems given its robustness to noise and mutation, and its invariance with respect to changes in the morphogen input. It is also of modular design and can pattern fast to produce a stripe almost anywhere in a field of cells. Indeed in some respects, it resembles the clock and wavefront mechanism proposed and favored for somitogenesis (Pourquié, 2003), as it involves the spatial manifestation of a temporal phenomenon. It is still unclear how the clock of the clock and wavefront mechanism is stopped at the molecular level (Kulesa *et al*, 2007). The frozen oscillator mechanism suggests one way by which such oscillations can be stopped; by means of a repressor that acts like a timer and shuts down oscillations after a particular amount of time has passed.

The mutual inhibition type core topology is observed in a number of different patterning contexts including *Drosophila* blastoderm and dorsal–ventral patterning of the vertebrate neural tube, which both involve activating gradients of morphogen and cross-repressions between the downstream genes. Why then is this type of configuration often observed in real patterning systems? The answer could lie in the fact that the mutual inhibition mechanism allows equilibrium to be reached particularly fast and is thus suitable to contexts where fast patterning is required.

Another key feature of the mutual inhibition mechanism could endow a major advantage in terms of the evolvability of mechanisms is the symmetrical nature of the mutual inhibition motif. Such a motif can be used multiple times to produce multiple stripes with minor pleiotropic constraints. This final point is an important consideration when such small networks appear in the context of a larger gene network, as they may have other functions to perform. Thus, certain mechanisms may be favored over others in larger network contexts. The mechanism of network evolution also has an important bearing on the utilization of different motifs in larger gene networks.

In summary then, different mechanisms have different features that will be of use in different patterning contexts. Exploring the design space of possible mechanisms can be used to suggest why some mechanisms are observed more often than others in real biological contexts and give insights into specific patterning processes.

Role of cell–cell communication in morphogen interpretation

Can the design space tell us anything about the general features of morphogen interpretation? Recently, it has been demonstrated in the Wingless system of the *Drosophila* wing imaginal disc that local cell–cell communication is essential for proper morphogen interpretation (Piddini and Vincent, 2009). This observation highlights the question of the role of cell–cell communication in morphogen interpretation. Our results so far from the complexity atlas strongly suggest that cell–cell communication will not generally be necessary for the underlying morphogen interpretation mechanisms. Does cell–cell communication thus have some secondary role such as

providing robustness to noise or 'fine tuning' the features of the gene expression domains such as stripe position or width? In order to explore this question, we performed a series of simulations using the core topologies of the six mechanisms with and without the presence of noise in the upstream morphogen input and with and without diffusion (Supplementary Data section S9). These results show that cell–cell communication can provide a small but significant increase in the robustness to noise in the morphogen input.

We next explored the role of diffusion in fine tuning the morphogen interpretation process. We analyzed the effects on stripe position and width generated by the core topologies with and without diffusion. Furthermore, we analyzed how the features of the stripes changed when we adjusted the diffusion parameter for solution parameter sets of the core topologies. These results revealed that diffusion can serve as a fine-tuning parameter allowing the adjustment of stripe width and position. For example, increasing diffusion results in thinner stripes for the incoherent feed-forward type 1 and mutual inhibition mechanisms but instead results in a larger stripe for the frozen oscillator, bistable and classical mechanisms. Increasing diffusion for the overlapping domains mechanisms has the effect of moving the stripe towards the source of the morphogen with no significant size change.

Generality of the model

Would the same results hold if we changed the underlying modeling formalism? In order to explore this question, we performed extensive further tests changing various aspects of the model (see Supplementary Data section S10 for full details). These parameters included the amount of noise, the size of the spatial field, the functional criteria of the stripe of gene expression, shape of the morphogen gradient and the initial conditions. In each case, a whole new complexity landscape was created by resampling many thousand parameter combinations per topology, and the results show that the structure of the complexity landscape changes very little when each of these features are changed.

However, the result changes more markedly when we expand the model space by changing the gene regulatory input function for the model. In particular, we changed the Michaelis–Menten input function described in Equation (3) to a sigmoid input function described in Equation (4). Most of the main stalactites can be observed though some mechanisms can function with less gene–gene interactions. By expanding the model even further and including features such as explicit delays, varying decay or more complex regulatory input–output functions (our study utilizes simple addition of inputs), we would expect the results to change further. However, these changes would only be expected to increase the repertoire of morphogen interpretation mechanisms further thus not changing the general message that for some biological functions a variety of distinct underlying strategies exists.

Discussion

In this study, we have combined the approaches from two previously separate lines of research. The neutral network

concept has been used to address evolutionary questions such as the relationship between mutational robustness and innovation (Smith, 1970; Schuster *et al*, 1994; Ciliberti *et al*, 2007). Here, we have adapted the approach to focus on a different question: the relationship between network design (motif topology) and dynamical mechanism (Ma *et al*, 2006; Hornung and Barkai, 2008). Rather than considering topologies as a list or ensemble, we created a complexity atlas by utilizing the notion of neighborhood and the concept that minimal topologies represent core mechanisms (Salazar-Ciudad *et al*, 2000; Munteanu and Solé, 2008). Stalactites naturally emerge from this approach, and we have shown that although core topology alone is insufficient to define a dynamical mechanism (Supplementary Figure S8), nevertheless stalactites can correspond to a classification of mechanisms. Although the different mechanisms map to separate regions of the underlying parameter space, more conventional methods failed to find these mechanistic classes (Supplementary Figure S6). We therefore believe the complexity atlas to be a powerful concept, which will be applicable to many other studies of network mechanism.

Using this new tool, we have discovered the first case of a well-defined biological function for which at least six different three-gene mechanisms exist. These likely represent major classes of network design (dynamical systems) important to the problem of morphogen gradient interpretation (Figure 5), and can be added to the repertoire of possible stripe-forming mechanisms, to aid our understanding of real model systems (Lander, 2007). We predict that those undescribed mechanisms for morphogen interpretation namely (C, D and E) will also underlie real biological systems. The discovery that five of the six mechanisms are cell autonomous strengthens the emerging view that networks of cross-regulatory transcription factors may be at the heart of many real systems, as has been proposed for SHH (Dessaud *et al*, 2008). Interestingly, it has recently been demonstrated in the Wingless system of the *Drosophila* wing imaginal disc that local cell–cell communication is essential for proper morphogen interpretation (Piddini and Vincent, 2009). Our study predicts that local cell–cell communication will rarely be responsible for the dose-dependent response of morphogen interpretation networks. We demonstrate that it is more likely to be involved in providing noise robustness or precision to the system, rather than its underlying functionality.

Finally, the power of such function-topology maps to discern the underlying network topology and mechanism for a function was previously demonstrated (Ma *et al*, 2009). However, such maps thus far have only demonstrated the existence of one or two mechanisms that are capable of generating a given function. In contrast, the results presented here caution that some well-defined biological functions may instead possess a wide range of alternative dynamical explanations (even for relatively simple three-gene networks and a limited modeling formalism) as demonstrated in Figure 5. This predicted feature has potential benefits as well as caveats. It suggests that in synthetic biology, there will exist a range of different ways to design a circuit to perform a desired function—thereby giving a greater range of options from which to choose the most practical design.

Materials and methods

Enumerating all GRN topologies

A topology can be represented in the form of a matrix w_{ij}, where i and j represent the position in those matrices and values 1, −1 and 0 represent activation, repression and no interaction, respectively. We generated all possible matrices that correspond to unlabeled topologies and then removed isometric equivalents by comparing them in all possible permutations. There are 19 683 gene network matrices before non-isometric topologies have been removed, and this is reduced to 3284 topologies in the fully enumerated set.

The morphogen gene is a gene that activates one of the genes of the GRN but is not affected by the GRN. Each GRN topology is represented multiple times with the morphogen feeding into the different genes (exact number depends on the amount of symmetry in the GRN topology). The morphogen is taken account of in the topology generation by extending the GRN matrix ($i = i + 1$) to include the input from morphogen (which is permutated independently). When the morphogen is included, the number of isometric topologies increases from 3284 to 9710.

Creating an Atlas of GRNs by including explicit neighbor definitions

Two GRN topologies are considered neighbors in the atlas if the two GRN topologies are one Hamming distance apart (a single gene–gene interaction change). The Hamming distance can be measured by the following equation where

$$D(w, w') = \sum_{i,j} \left| \text{sgn}(w_{ij}) - \text{sgn}(w'_{ij}) \right|, \qquad (1)$$

D is the Hamming distance between the matrices of two GRN topologies w and w' and i and j represent the position in those matrices. The matrices are compared in every permutation and the lowest D of those permutations is taken as the Hamming distance. Hence, two GRN topologies are neighbors if the gain or removal of any one interaction can transform one of the GRN topologies into the other.

The gene regulation model

We employed a biologically verified model of gene regulation for this problem, and therefore adapted the continuous mathematical model developed over the last 20 years by Reinitz *et al* (Mjolsness *et al*, 1991; Reinitz and Sharp, 1995; Jaeger *et al*, 2004), which quantitatively captures the spatiotemporal dynamics of gap gene patterning in response to the Bicoid morphogen gradient during *Drosophila* embryogenesis. The model is described by

$$\frac{dg_{ij}}{dt} = \chi \left[\Phi \left[\sum_{l=1}^{N_g} W^{li} g_{lj} + M \right] \right] + D_i \nabla^2 g_{ij} - \lambda g_{ij} + \eta(t) g_{ij}, \qquad (2)$$

where g_{ij} is the concentration of the ith gene in the jth cell, $\Phi(x)$ is a function defining the interaction among genes (which can take the form of a Michaelis–Menten, sigmoid or other non-linear input function), W^{li} is a matrix containing the strength of gene-to-gene regulation parameters, M is the morphogen input described in more detail in the section 'configuration of the spatial domain' below, $\chi(x)$ is the Heaviside function (to prevent negative gene product production rates), D_i is the diffusion constant for the ith gene, which we use to represent local cell–cell signaling, λ is the decay rate (set to 0.05) and $\eta(t)$ is a noise term, which adds uniformly distributed fluctuations ($\pm 1\%$) to the concentration of every gene in every cell at every time step. There is zero autocorrelation in the noise term. The parameters that could vary in the model were regulation W^{il} and diffusion D_i. Full details of the model are described in Supplementary Materials and Methods section S1. The input function describes the relationship between the activation and inhibition of a gene and its actual expression. The main input function used in this work took the form of a Michaelis–Menten function, which is defined by

$$O = \frac{1}{(1+I)}, \qquad (3)$$

where I is the total input into the gene and O is the output of the function. The alternative function used to test the dependence of the results on the exact gene regulation model used was a sigmoid function.

$$O = \frac{1}{(1 + (e^{(5-5I)}))}, \tag{4}$$

where I is the total input into the gene and O is the output of the function.

Parameter range distributions

For each GRN topology, 30 000 different parameter sets were tested (a GRN topology with a specific set of parameters we called a genotype). There are up to 12 variable parameters for a three-gene network; diffusion for each individual gene and then the strengths of the interaction values between the genes. The parameters are chosen randomly though biased towards lower numbers through a logarithmic probability distribution. The logarithmic probability distribution was implemented in order to take account of the fact that a small change in a small parameter value will have a greater effect on a network's behavior than a small change in a larger parameter value. The logarithmic probability distribution is described by

$$V = 0.9995^{i}R, \tag{5}$$

where i is a random number between 0 and 10 000 and R is the parameter range and V is the resulting parameter value. Parameter ranges are as follows: regulation 0–10 and diffusion 0–0.05.

Configuration of the spatial domain

The simulations take place on a theoretical one-dimensional row of 32 cells. Zero-flux boundary conditions are used throughout this work. The simulation starts with every gene in every cell set to have a concentration of 0.1. This was necessary because the noise term used is a percentage noise term and thus if the concentration was always 0 at the start of the simulation, then the products of any genes with positive feedbacks without any other input would remain at 0. The simulation is also initiated by the positive input from the morphogen gradient that does not change throughout the simulation.

The morphogen was chosen to give an approximate input range to the gene that it affects 10–50% of the maximal activation. The morphogen input is defined by

$$M = Id^{c}, \tag{6}$$

where M is the morphogen input, I is the morphogen concentration in the left-most cell of the field, d is the reduction of morphogen concentration in each subsequent cell of the morphogen gradient and c is the cell position. For the 10–50% input range, $I=1$ and $d=0.93$ was used for the Michaelis–Menten function and $I=1$ and $d=0.98$ was used for the sigmoid function.

Stripe-forming functional definition

For a genotype (GRN topology with a specific parameter set) to be considered functional, it had to produce a stripe of gene expression for at least one of the genes. For each gene, we measured an abstraction of its gene expression of the one-dimensional field, where each cell was defined as low or high. We defined a cell as low if the gene expression was below 10% of the maximum possible allowed by the model. We defined a cell as high if the gene expression was above 10% of the maximum gene expression allowed by the model. A gene was considered to have a stripe pattern if it had a single region of low for two consecutive cells followed by a single region of high for a maximum of 16 consecutive cells followed by a single region of low for at least two consecutive cells. The two low regions must occur at the extremities of the field. The definition is intentionally loose in the sense that the single stripe can be of any width up to 16 cells and be in any position in the spatial domain. This is because we are interested in the basic design principles of the system, not the details of how to control a specific width. Functional parameter sets that can produce the single stripe of gene expression are termed 'solutions.' Hence, a

single topology has multiple genotypes and can have multiple solutions. The number of solutions that each topology has is a measure of its mutational robustness. A GRN topology must have at least one solution to be considered functional. Furthermore, solutions had to be robust with respect to developmental noise and have reached equilibrium to be considered functional (full details are in Supplementary Materials and Methods section S2).

Clustering the space–time plots

In order to cluster the space–time plots, we needed a suitable difference function. The difference function used is a measure of the Euclidean distance between the gene expression values of any two plots over space and time. In order to control for differences in the position and size of the stripe, we only compare three cells from each space–time plot. These three cells are the cell in the middle of the defined stripe (high) region and the two cells directly in the middle of the defined low regions. Thus, the difference function is

$$D_{i,j} = \frac{\Sigma_1^C \Sigma_0^T \Sigma_1^G \sqrt{(Y_{i,c,t,g} - Y_{j,c,t,g})^2}}{CTG}, \tag{7}$$

where D is the Euclidean distance, i and j are the indices of the respective space–time plots. Y is the array containing the values of the space–time plot, where c is the cell, t is the time step and g is the gene. C is the number of cells being compared, in this case three for the low–high–low regions. T is the number of time points being compared. If the two space–time plots have different numbers of time points, then they are compared up to the length of the shortest time of the two plots. G is the number of genes. Plots are compared for all permutations of the genes, and the smallest difference of all of the permutations is taken as the final difference.

A hierarchical clustering algorithm then generates a tree in an agglomerative manner, starting with the most similar pair of space–time plots increasing the difference value until all space–time plots have been assigned to the tree. The y distance on the tree corresponds to the minimal difference between any two space–time plots each from a different branch.

Interpolating between randomly chosen parameter sets for each mechanism

For each topologically compatible pairwise combination of core topologies from Figure 4C, we performed the following analysis. We randomly collected 25 solutions from each core topology. For each pairwise combination of solutions ($25 \times 25=625$), we interpolated the parameter values (in all viable permutations). We used a linear diagonal interpolation through parameter space, consisting of 20 interpolation steps. For each step, we resimulated the interpolated topology and asked if it could produce the single stripe of gene expression.

Mapping mechanisms to the complexity atlas

Individual solutions for generating a single stripe of gene expression were assigned to mechanism categories based on a combination of topological and gene expression dynamic properties. For a particular solution to be assigned to a particular mechanism category, the GRN topology responsible for the solution must contain the core GRN topology of that mechanism. If a particular solution does not meet the topological criteria for any of the mechanism categories, then it is not assigned a category. If the solution meets the topological criteria of a single-mechanism category, then it is assigned to that category. If, however, it meets the topological criteria of multiple mechanism categories, then gene expression dynamic information is taken into account to decide which category it belongs to. The space–time plot of the solution is compared with all of the solution space–time plots of the 'core GRN topologies' (when the core topology is simulated with a million parameter sets).

The three landmark cells are compared between the solution and each of the space–time plots in the 'core GRN topologies' dataset

for every permutation of the genes. The Euclidean distance is measured between the solution and each core GRN topology space–time plot. The solution is placed in the mechanism category that has a core GRN topology space–time plot with the smallest Euclidean distance to the solution. Using this approach, 3147 solutions out of the total 3379 (93%) were assigned to at least one of these mechanism groups.

Supplementary information

Supplementary information is available at the *Molecular Systems Biology* website (http://www.nature.com/msb).

Acknowledgements

Conflict of interest

The authors declare that they have no conflict of interest.

References

Bayly RD, Ngo M, Aglyamova GV, Agarwala S (2007) Regulation of ventral midbrain patterning by Hedgehog signaling. *Development* **134:** 2115–2124

Briscoe J, Chen Y, Jessell TM, Struhl G (2001) A hedgehog-insensitive form of patched provides evidence for direct long-range morphogen activity of sonic hedgehog in the neural tube. *Mol Cell* **7:** 1279–1291

Chouard T (2008) Beneath the surface. *Nature* **456:** 300–303

Ciliberti S, Martin OC, Wagner A (2007) Robustness can evolve gradually in complex regulatory gene networks with varying topology. *PLoS Comput Biol* **3:** 164–173

Clyde DE, Corado MS, Wu X, Paré A, Papatsenko D, Small S (2003) A self-organizing system of repressor gradients establishes segmental complexity in Drosophila. *Nature* **18:** 849–853

Dessaud E, McMahon AP, Briscoe J (2008) Pattern formation in the vertebrate neural tube: a sonic hedgehog morphogen-regulated transcriptional network. *Development* **135:** 2489–2503

Green J (2002) Morphogen gradients, positional information, and Xenopus: interplay of theory and experiment. *Dev Dyn* **225:** 392–408

Hornung G, Barkai N (2008) Noise propogation and signalling sensitivity in biological networks: a role of positive feedback. *PLoS Comput Biol* **4:** e8

Ingolia NT, Murray AW (2004) The ups and downs of modeling the cell cycle. *Curr Biol* **14:** R771–R777

Jäckle H, Tautz D, Schuh R, Seifert E, Lehmann R (1986) Cross-regulatory interactions among the gap genes of Drosophila. *Nature* **324:** 668–670

Jaeger J, Surkova S, Blagov M, Janssens H, Kosma D, Kozlov KN, Manu, Myasnikova E, Vanario-Alonso CE, Samsonova M, Sharp DH, Reinitz J (2004) Dynamic control of positional information in the early Drosophila embryo. *Nature* **430:** 368–371

Kerszberg M (2004) Noise, delays, robustness, canalization and all that. *Curr Opin Genet Dev* **14:** 440–445

Kim D, Yung-Keun K, Kwang-Hyun C (2008) The biphasic behavior of incoherent feedforward loops in biomolecular regulatory networks. *BioEssays* **30:** 1204–1211

Kitano H (2007) Towards a theory of biological robustness. *Mol Syst Biol* **3:** 137

Kraut R, Levine M (1991) Mutually repressive interactions between the gap genes giant and Krüppel define middle body regions of the Drosophila embryo. *Development* **111:** 611–621

Kulesa PM, Schnell S, Rudloff S, Baker RE, Maini PK (2007) From segment to somite: segmentation to epithelialization analyzed within quantitative frameworks. *Dev Dyn* **236:** 1392–1402

Lander AD (2007) Morpheus unbound: reimagining the morphogen gradient. *Cell* **128:** 245–256

Latinkić BV, Umbhauer M, Neal KA, Lerchner W, Smith JC, Cunliffe V (1997) The Xenopus Brachyury promoter is activated by FGF and low concentrations of activin and suppressed by high concentrations of activin and by paired-type homeodomain proteins. *Genes Dev* **11:** 3265–3276

Ma W, Lai L, Ouyang Q, Tang C (2006) Robustness and modular design of the Drosophila segment polarity network. *Mol Syst Biol* **2:** 70

Ma W, Trusina A, El-Samad H, Lim WA, Tang C (2009) Defining network topologies that can achieve biochemical adaptation. *Cell* **138:** 760–773

Mangan S, Alon U (2003) Structure and function of the feed-forward loop network motif. *Proc Natl Acad Sci USA* **100:** 11980–11985

Milo R, Shen-Orr S, Itzkovitz S, Kashtan N, Chklovskii D, Alon U (2002) Network motifs: simple building blocks of complex networks. *Science* **298:** 824–827

Mjolsness E, Sharp DH, Reinitz J (1991) A connectionist model of development. *J Theor Biol* **152:** 429–453

Munteanu A, Solé RV (2008) Neutrality and robustness in evo-devo: emergence of lateral inhibition. *PLoS Comput Biol* **4:** 1–14

Peter IS, Davidson EH (2009) Genomic control of patterning. *Int J Dev Biol* **53:** 707–716

Piddini E, Vincent JP (2009) Interpretation of the wingless gradient requires signaling-induced self-inhibition. *Cell* **136:** 296–307

Pisarev A, Poustelnikova E, Samsonova M, Reinitz J (2009) FlyEx, the quantitative atlas on segmentation gene expression at cellular resolution. *Nucleic Acids Res* **37:** D560–D566

Pourquié O (2003) Vertebrate somitogenesis: a novel paradigm for animal segmentation? *Int J Dev Biol* **47:** 597–603

Poustelnikova E, Pisarev A, Blagov M, Samsonova M, Reinitz J (2004) A database for management of gene expression data *in situ*. *Bioinformatics* **20:** 2212–2221

Reinitz J, Sharp DH (1995) Mechanism of eve stripe formation. *Mech Dev* **49:** 133–158

Salazar-Ciudad I, Garcia-Fernández J, Solé RV (2000) Gene networks capable of pattern formation: from induction to reaction-diffusion. *J Theor Biol* **205:** 587–603

Schuster P, Fontana W, Stadler PF, Hofacker IL (1994) From sequences to shapes and back: a casestudy in RNA secondary structures. *Proc Biol Sci* **255:** 279

Shen-Orr SS, Milo R, Mangan S, Alon U (2002) Network motifs in the transcriptional regulation network of Escherichia coli. *Nat Genet* **31:** 64–68

Smith JM (1970) Natural selection and the concept of a protein space. *Nature* **225:** 563–564

Thieffry D, Sanchez L (2003) Dynamical modelling of pattern formation during embryonic development. *Curr Opin Genet Dev* **13:** 326–330

Tümpel S, Wiedemann LM, Krumlauf R (2009) Hox genes and segmentation of the vertebrate hindbrain. *Curr Top Dev Biol* **88:** 103–137

Ulloa F, Briscoe J (2007) Morphogens and the control of cell proliferation and patterning in the spinal cord. *Cell Cycle* **6:** 2640–2649

Vakulenko S, Manu, Reinitz J, Radulescu O (2009) Size regulation in the segmentation of Drosophila: interacting interfaces between localized domains of gene expression ensure robust spatial patterning. *Phys Rev Lett* **103:** 16

Wijgerde M, McMahon JA, Rule M, McMahon AP (2002) A direct requirement for Hedgehog signaling for normal specification of all

ventral progenitor domains in the presumptive mammalian spinal cord. *Genes Dev* **16:** 2849–2864

Wolpert L (1968) The French Flag problem: a contribution to the discussion on pattern development and regulation. In *Towards a Theoretical Biology*, Waddington CH (ed), Vol. 1, pp 125–133. Edinburgh: Edinburgh University Press