# Spatial pattern formation during morphogenesis

Sungsoo Lim

May 13, 2018

**Abstract**

Various biological systems and processes are understood to be emergent properties of the collective interactions of networks. Morphogenesis, the organized spatial distribution of cells to establish different body parts during an organism's development, is an interesting example. Several of general structural and dynamical properties of biological networks have been established, with which the molecular details of spatial pattern formation during morphogenesis can be understood. To this end, this essay describes some of the general properties of biological networks. The reaction-diffusion model is also introduced, and in this context, the left-right asymmetry and the segmentation of embryos are discussed.

# 1 Introduction

Systems analysis, based on interactions between the components, provides interesting examples of emergent phenomena in various levels of biology. Such analysis can be used to understand the properties of specific interaction networks, ranging from the microscopic interactions in protein-folding mechanisms, interactions in molecular networks within a cell, and intercellular interactions in populations [1]. In particular, developments in experimental procedures and the resulting accumulation of data have made such analysis in molecular networks not only possible, but also a mainstream of thought. The advent of high-throughput molecular data and the increasing availability of such data make studies of biological processes more amenable to quantitative, network-based analysis, where the structural and dynamical properties of the networks are discussed [2]. This shift in focus of biological research allows us to understand biological processes as interacting systems and their emergent properties.

Morphogenesis during embryonic development, the precise spatial patterning of cells to produce body parts, provides an interesting example of an emergent behavior based on the interactions between the network components. A brief overview of molecular networks is presented, focusing on the structural properties that allow the mechanistic descriptions of various spatial pattern formations during morphogenesis. Then the reaction-diffusion model of morphogenesis is introduced, and two experimentally considered examples of the model are discussed.

# 2 Molecular networks

## 2.1 Overview

Molecular networks are defined on the interaction maps of DNA, RNA, proteins, and complexes of these molecules. Focusing on the structural and dynamical properties, the network descriptions provide mechanistic representations of various biological processes. Known molecular networks include [1], [2]:

- **Protein-protein interaction networks**: Interactions between proteins such as formation of protein complexes and the activation/inhibition of one protein by another.
- **Genetic regulatory networks**: Activation/inhibition of genes, and can be used to understand spatiotemporal information of cellular activity.

## 2.2 Structural properties

Independent of the level of description, graph theory has been useful in structural descriptions; individual components are identified as nodes and the resulting interactions as edges. The directionality of the edges provides further information on the structure of the networks.

Directed edges can specify a regulator and a target, providing information on the regulatory relationships in the networks. Topological parameters are established, several of the commonly used ones being [2], [3]:

(i) **Degree**: The number of edges connected to one node. A node with higher than average degree is known as a hub. In the case of a directed graph, the degree of a node can be further divided into the number of edges directed in and out.

(ii) **Average clustering coefficient**: For each node, the fraction of existing edges based on the total possible edges from its closest neighboring nodes is calculated. These fractions are averaged over all nodes.

(iii) **Characteristic path length**: The number of edges in the shortest path between two nodes, averaged over all pairs of nodes.

One example of the information that these parameters provide is the "small-world" structure [3]. "Small-world" networks have the topological properties of small characteristic path lengths and high clustering coefficients. Additionally, many of the biological networks considered have also been shown to exhibit the properties of "scale-free" networks [4]. An important feature of such networks has been identified as the existence of hubs and the power-law degree distribution: the fraction $P(k)$ of nodes in the network having $k$ connections to other nodes has the distribution:

$$P(k) \sim k^{-\gamma}$$

with $\gamma$ in the range of $2 < \gamma < 3$ [4]. The minimal model describing these properties incorporates two mechanisms: (i) growth in the number of nodes and edges over time and (ii) preferential attachment, where the probability of new edge acquisition is proportional to the existing degree of a node [4]:

$$p_i = \frac{k_i}{\sum_j k_j},$$

where the probability of a new node attaching to an existing node ($i$) is proportional to its relative degree ($k_i$) in the network. In this model, the clustering-degree function $C(k)$, the clustering coefficients of nodes as the function of the degree $k$, is constant [5].

In most observed metabolic and protein-protein interaction networks from model organisms, such as yeast and *Drosophila*, it has been shown that the "small-world" characteristics are satisfied with short characteristic path lengths and high average clustering coefficients [6], [7]. Additionally, the protein-protein interaction networks in these organisms display a power-law degree distribution [6], [7]. Similar "small-world" and scale-free properties have been shown in the genetic regulatory network in yeast [8]. However, measurements of $C(k)$ in the yeast network follow $C(k) \sim k^{-\beta}$ $(1 < \beta < 2)$. This provides an interesting observation that lower-degree nodes have more connected neighborhoods than higher-degree nodes [6].

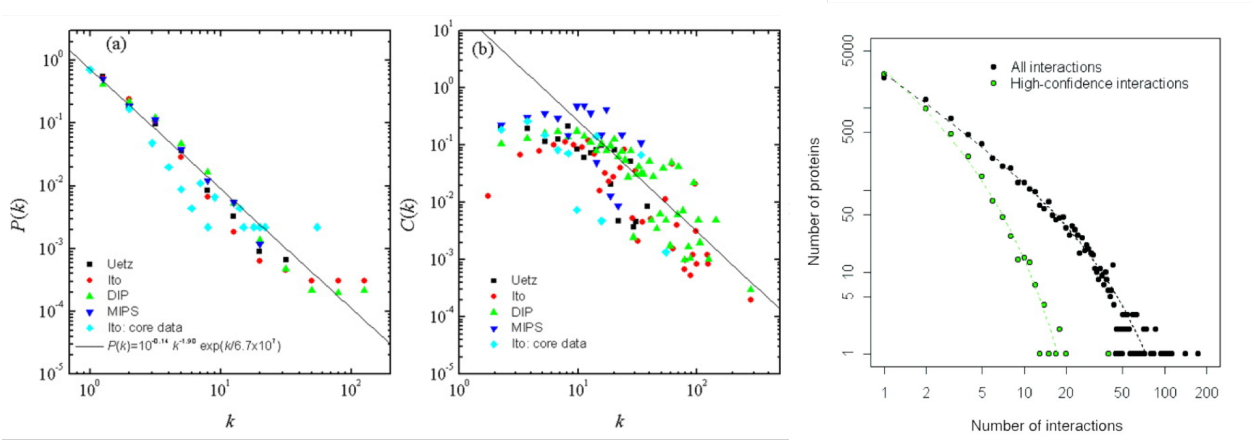


**Figure 1:** Degree distribution, $P(k)$, and average clustering coefficient, $C(k)$, in yeast (left, middle), and the number of interactions per protein in *Drosophila* (right) are shown. The degree distribution can be approximated by a power law with the exponent $\gamma = 2.5$. $C(k) = B/k^{\beta}$ is approximated by the exponent $\beta$ of 2. Similar to yeast, *Drosophila* degree distribution can also be approximated by a power-law distribution. Figure adapted from [7] and [6]

## 2.3 Genetic regulatory networks

A fundamental response of a cell to internal and external signals is the regulation of its genetic activity, the activation/inactivation of specific genes depending on the present stimuli [1]. Signal transduction pathways, which describe the interlinked chain of biochemical reactions within a cell that transmits the perceived signals, and gene regulatory networks provide the mechanistic descriptions of the cell's coordinated response. Combined with other environmental molecules, a genetic regulatory network is an interaction map of molecular regulators that collectively determines the turning on or off of genes at specific location and time. In these networks, nodes represent the transcription factors (proteins that bind to promoter regions to regulate the activation/inhibition of genes) and the target genes, and directed edges represent the regulation of the target genes by their transcription factors. In turn, these transcription factors are regulated by the upstream protein-protein and signaling networks [2].

Based on the directionality of the interactions, interesting observations are obtained on the functionality of these networks. In yeast, the genetic regulatory network has a scale-free out-degree distribution, with a few transcription factors regulating many different target genes. The in-degree distribution for given target genes is represented as an exponential function [9]. Such distributions point to the fact that combinatorial regulation of a target gene by different transcription factors is less frequent than regulation of several targets by the same transcription factor [9]. Additionally, a key difference in the structural aspect is that genetic regulatory networks show hierarchical organization. When the binding targets of *E. coli* transcription factors are analyzed, a few key regulators regulate downstream units

for specific functions [10].

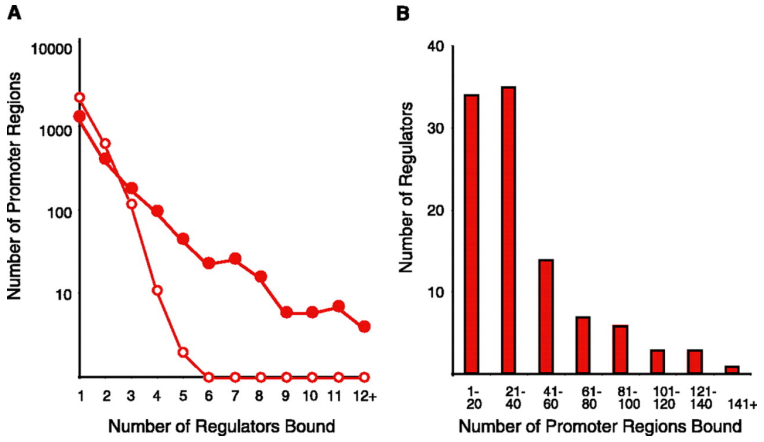


**Figure 2:** (A) The in-degree distribution of the regulated genes (number of transcription factors bound per promoter region; solid symbols) as compared to the in-degree distribution of a randomized network (open symbols). (B) The out-degree distribution of transcription factors with the number of promoter regions bound per regulator. These data are from yeast. Figure from [9]

Genetic regulatory networks are an interesting example, as these interactions provide the framework in which the process of morphogenesis in developmental biology is studied [17], [20]. In the case of pattern formation during morphogenesis, morphogen (a type of signaling molecule) gradients across the cell population are established and used, by which the genetic regulatory networks in specific cells respond by regulating the set of genes that generate the developmental patterns. Thus, in conjunction with these morphogen gradients, these networks provide the regulatory networks for specific spatial patterns in the genetic activities of the cells [17], [20].

## 2.4 Network motifs

The high clustering coefficients of various molecular networks have indicated that these networks are locally represented with modules with different functionalities [5]. Although the modularity concept had been previously applied, these "network motifs" were first considered in the genetic regulatory network of *E. coli*, where specific interaction motifs of different feedback structures have been shown [11]. Further studies have also indicated the presence of these common motifs in other networks and organisms, such as the genetic regulatory networks in the human embryonic stem cells [13].

The motif representations can help elucidate the dynamical properties of the networks. Based on the interaction maps of these motifs, simple dynamical models can be constructed, from which it can be shown how each network motif can carry out specific information-processing functions, such as threshold-dependent behaviors [12]. Some of the network motifs that are common and specific to developmental genetic networks are [12]:

i **Auto-regulation motifs**: Auto-regulations can be either positive, where a node can enhance its own production, and negative, where a node represses its production.

ii **Toggle switches**: These consist of double-negative or double-positive interactions between two nodes, and they can mutually inhibit or enhance each other's activities/productions.

iii **Feedforward loops**: These consist of a cascade of three nodes with directional regulation, with the added feature of the first upstream regulator also interacting with the third downstream effector.

Auto-regulation motifs can be used to explain the propagation of molecular perturbations, while toggle switches are used to generate spatially alternating gene activities. Based on the directionality of information flow from signaling molecules to effector genes, feedforward loops can also provide a further layer of regulation to fine-tune the effector activities.

# 3 Pattern formation in morphogenesis

Morphogenesis is described by the organized spatial distribution of cells during the development of an organism. With the positional information provided by morphogens across the cell population, the distribution of cells is in turn regulated by their genetic activities. Although it had preceded much of the experimental work done on the molecular underpinnings of the processes, the reaction-diffusion model first proposed by Alan Turing in 1952 can be and are used to describe the general principles behind some of these processes. This model provides a mechanism by which the initial spatial symmetry in the morphogen concentration is broken, leading to spatial genetic patterning in the cells.

However, in experimentally validated morphogenic processes, the asymmetric sources of patterning are already present. Thus, in most instances, morphogenic pattern formations do not arise from a homogeneous initial state, a point that Turing had considered while introducing the model. For example, in the molecular morphogenesis in *Drosophila* embryos, the pre-determined asymmetry is inherited from the mother; fertilized eggs contain maternally determined morphogen gradients, which initiates the patterning of the embryo. Similarly, the left-right asymmetry initiation is controlled by a reaction-diffusion mechanism, although the pre-patterns of the activator and inhibitor are present at the onset, which are also maternally determined. These two examples are discussed in a later section, focusing on the mechanistic details that support the reaction-diffusion model and the general biological network properties previously discussed.

## 3.1 Reaction-Diffusion

The reaction-diffusion (RD) model considers pattern formation based on mutual interactions between chemicals and their differential diffusion [14]. In the context of morphogenesis in developmental biology, a main starting point may be to understand where the spatial

symmetry-breaking mechanism comes from. What's really interesting in the RD model is that the spatially uniform concentrations of morphogens can be shown to be unstable with respect to perturbations because of diffusion, which can in general lead to a wide variety of self-regulated pattern formations.

In the original model, two interacting and diffusing morphogens are considered, where the $N$ individual cells provide the discrete positional information in an one dimensional circle. With a continuous position variable, the model considers two morphogens ($u_1$ and $u_2$) [14], [15]:

$$\partial_t u_1 = f_1(u_1, u_2) + D_1 \nabla^2 u_1,$$
$$\partial_t u_2 = f_2(u_1, u_2) + D_2 \nabla^2 u_2.$$

where the dynamical system of two morphogen concentrations $u_1$ and $u_2$ is described by the nonlinear functions of production, degradation, and interaction terms ($f_i(u_1, u_2)$) and the corresponding constant diffusion coefficients ($D_i$). An important assumption in the model is that there is no prior spatiotemporal asymmetry in the system, indicating that the functions $f_i$ and the diffusion coefficients $D_i$ do not depend on time and position. Then, linearization about the uniform solutions shows that the perturbations $(u_{p1}, u_{p2})$ in one dimension satisfy [15]:

$$\partial_t u_{p1} = a_{11} u_{p1} + a_{12} u_{p2} + D_1 \partial_x^2 u_{p1},$$
$$\partial_t u_{p2} = a_{21} u_{p1} + a_{22} u_{p2} + D_1 \partial_x^2 u_{p2},$$

and the $a_{ij}$'s are the Jacobians of the $f_i$'s with respect to $u_i$'s at the uniform solutions. The particular solution $u_p(x,t)$ can be shown to have an exponential spatial factor $e^{iqx}$ with the wave number $q$ and an exponential factor in time $e^{k_q t}$ with the corresponding growth factor $k_q$ [15]. Then, eigenanalysis of the growth factor $k_q$ gives the stability conditions for the uniform state in the absence of diffusion:

$$a_{11} + a_{22} < 0,$$
$$a_{11} a_{22} - a_{12} a_{21} > 0.$$

Combined with non-negative diffusion constants, it can be shown that the following holds for diffusion-induced instability of the uniform state:

$$D_1 a_{22} + D_2 a_{11} > 0.$$

Based on these conditions for the existence of Turing instability in the presence of diffusion, the coefficients $a_{11}$ and $a_{22}$ must have opposite signs, and the same condition holds for $a_{12}$ and $a_{21}$ [15]. These conditions are generally understood that one morphogen has a positive auto-regulation (for instance, $u_1$, an activator) and the other a negative auto-regulation ($u_2$, an inhibitor). Furthermore, $u_1$ can either serve as an activator or inhibitor of $u_2$ production, whereas $u_2$ shows the opposite pattern with respect to $u_1$. Thus, we can

understand the necessary network motif to be a directed activation/inhibition loop between the two components with nested auto-regulations. Additionally, the decay lengths ($l_i$) of the activator and inhibitor are defined on the diffusion constants, where $l_i = (\frac{D_i}{a_{ii}})^{1/2}$. In one dimensional simplification, this regime of parameter values leads to a time-independent state the following wave number [15], [16]:

$$q = \left[\frac{1}{2}\left[\frac{1}{l_1} - \frac{1}{l_2}\right]\right]^{1/2}.$$

The condition for the existence of a finite wave number instability is $l_2 > l_1$, indicating that the relatively faster diffusibility of the inhibitor can lead to the mechanism by which fluctuations spontaneously break spatial symmetry in the uniform state. Considering the case that the auto-activator activates the production of the inhibitor (Figure 3), fluctuations will increase the level of the activator across a cell, which leads to increased local levels of both the activator and the inhibitor ("local activation"). Then, the diffusibility condition indicates that there are regions of "lateral inhibition" where the increase in the inhibitor levels in neighboring cells prevents the level of activator from increasing. These patterns are then spatially propagated.

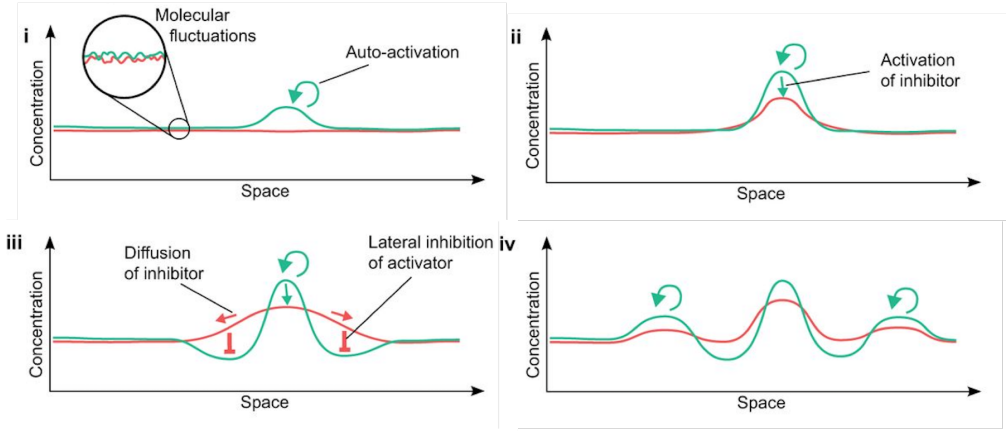


**Figure 3:** Schematic diagram (i-iv) of the diffusion-induced instability in the activator-inhibitor system. Green and red lines represent the concentration of activator and inhibitor respectively. The case of $a_{12} > 0$ and $a_{21} < 0$ is considered. Figure adapted from [17].

There are other instabilities that can arise depending on the parameter values. For instance, when there is a cross-coupling between the activator and inhibitor, where

$$a_{12}a_{21} > \frac{1}{4}(a_{11} + a_{22})^2,$$

and the production rate of the activator is greater than the degradation rate of the inhibitor, there can be a time-periodic state with spatial uniformity [15], [16].

The key aspects of the RD model in pattern formation is the absence of pre-pattern in the regulatory molecules and the interactions between self-regulating morphogens with differential diffusibility. Thus, a patterning system combines a "a short-range positive feedback with a long-range negative feedback" [15]. There are many postulated examples in the literature that represent these details, although the absence of pre-patterns has been a point of contention in the developmental biology community, such that the RD model had not been extensively considered [17]. Gierer and Meinhardt generalized the key ideas in the RD model to embryological processes by considering the sources of activator and inhibitor morphogens [18]. By incorporating pre-determined sources, it was shown that patterns that are independent of the properties of the source distributions can arise, patterns that can sufficiently explain various experimental patterns considered [18]. In this model, it was still assumed that the relative diffusibilities of the activator and inhibitor have to be different, such that the inhibitor concentration is effectively a function of the activator concentration [18].

## 3.2 Left-right asymmetry

These general interaction motifs consisting of auto-regulation and directional/mutual inhibition is a key feature of developmental genetic networks [12]. Thus, the main experimental concern is to determine whether an activator-inhibitor system shows differential diffusibility, which has not been established for many of the model systems. One interesting example where these mechanisms apply is the establishment of left-right asymmetry in zebrafish [19]. The Nodal-Lefty activator/inhibitor system is responsible for these patterns, where Nodal, a morphogen responsible for the right specification, enhances both its own production and that of Lefty, and Lefty inhibits the activity of Nodal. It was also established that the normalized diffusivity ratio of Lefty and Nodal measured was $\approx 14$, indicating that the inhibitor Lefty has a longer spatial range than the activator Nodal [19]. Interestingly, it is postulated that the pre-patterns in Nodal and Lefty provide a positive feedback on the establishment of the patterns, such that the left and right specification pathways are activated [19].

## 3.3 Establishment of body parts in *Drosophila*

During the patterning process in *Drosophila* embryos, cells show discrete segmentations along the one-dimensional anterior-posterior axis (Figure 4) [17], [20]. These segmentations then become different body parts over the course of the organism's development. This process is orchestrated by a cascade of input morphogens and effector genes. Since the morphogens (Bicoid and Caudal) are the only diffusible components within this network, the components relevant to the RD model are the maternally inherited sources for the two morphogens, where Bicoid is originally present at the anterior point and Caudal at the posterior point of the embryo. There are no auto-regulations present for these morphogens, although it is known that Bicoid inhibits Caudal production from its source [20]. Additionally, due to experimental difficulties, there is no consensus on the relative diffusibilities of the two

morphogens [21]. However, what is known is that the diffusion-dependent Bicoid gradient and the anti-parallel Caudal gradient are established, which act as threshold-dependent activating transcription factors of the genetic regulatory network within the cells [17], [20]. Due to these differences between the RD model and the experimental details for this process, other conceptual models, such as the "positional information" model, have been used to explain the morphogen-based segmentation in *Drosophila* embryos [17].

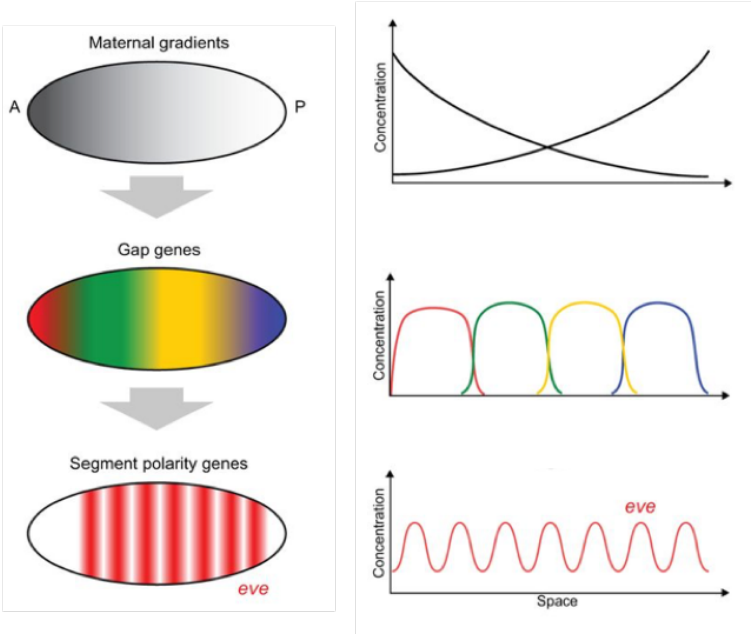


**Figure 4:** The anterior-posterior patterning system in *Drosophila* embryos. The Bicoid and Caudal sources are asymmetrically inherited from the mother, after which diffusion and inhibition of Bicoid on Caudal production produce the gradients. The spatial information provided by the concentration gradients are reinterpreted as spatially alternating genetic activities of gap-genes. The segment polarity genes, such as *eve*, are then regulated to be produced with a specific spatial wave number. Figure adapted from [17].

Thus, focus has been on the formulation of the mechanistic details in the genetic regulatory networks that produce specific spatial patterns from broadly presented concentration gradients. The previously discussed toggle switches and feedforward loops have been shown to be involved in the gap-gene genetic networks and the segmentation polarity gene networks [20]. In this view, the morphogens Bicoid and Caudal are transcription factors that directly activate several of the mutually inhibiting toggle switches of the gap-genes, such as *hb*, *kni*, *kr*, and *gt* (Figure 5). Specific double-negative feedback loops between pairs of the gap-genes translate the spatial information provided by the morphogen gradients into alternating expression patterns. To describe the experimental patterns, it is also proposed that the relative inhibitory strengths between the gap genes are asymmetric across the anterior-posterior axis, with posterior interaction partners exerting stronger inhibition on their more anterior counterparts [22]. Additionally, the specification of the spatial wave number of the

segmentation polarity genes should be considered, although there are currently no available experimental data. One study presents a feedforward loop structure in the morphogen - gap-gene - segmentation gene cascade, where the morphogen further activates the downstream segmentation gene [23].

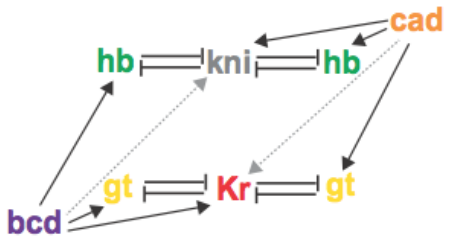


**Figure 5:** The gap-gene genetic regulatory network in *Drosophila* embryo segmentation. Morphogens provide positional cues by activating the gap-genes by a threshold-dependent manner. Pairs of gap genes form mutually inhibiting switches to form spatially alternating patterns. Figure from [20].

# 4 Summary

Pattern formation during morphogenesis gives us an interesting example of an emergent phenomenon in biology. The reaction-diffusion model provides a mechanism by which spatial uniformity can be broken, with resulting spatial pattern formations. In some of the morphogenic processes considered experimentally, however, it is established that the spatial patterning cues are pre-determined with maternal inputs. Additionally, experimental data concerning these processes are also dependent on the organism in which they are tested. Thus, current views on the established models of pattern formation in morphogenesis combine the general ideas of the reaction-diffusion model with molecular details that are specific to the process considered. To this end, more data are accumulated and analysis performed to determine the structural properties of biological networks across different organisms. Several interesting concepts have been established such that they could be considered to be universal. With this, it is hopeful that we can gain better understanding.

# References

[1] Eric Alm and Adam P Arkin. *Biological networks.* Current Opinion in Structural Biology, 13: 193-202, 2003.

[2] Xiaowei Zhu, Mark Gerstein, and Michael Snyder *Getting connected: analysis and principles of biological networks.* Genes and Dev., 21: 1010-1024, 2007.

[3] D.J. Watts and S.H. Strogatz. *Collective dynamics of "small-world" networks*. Nature, 393: 440-2, 1998.

[4] Albert-László Barabási and Réka Albert. *Emergence of scaling in random networks*. Science, 286: 509-12, 1999.

[5] Réka Albert. *Scale-free networks in cell biology*. Journal of Cell Science, 118: 4947-57, 2005.

[6] Giot, L., Bader, J. S., Brouwer, C., Chaudhuri, A., Kuang, B., Li, Y., Hao, Y. L., Ooi, C. E., Godwin, B., Vitols, E. *et al. A protein interaction map of Drosophila melanogaster* Science, 302: 1727-36, 2003.

[7] Yook, S. H., Oltvai, Z. N. and Barabsi, A. L. *Functional and topological characterization of protein interaction networks*. Proteomics 4, 928-942, 2004.

[8] Tong, A. H., Lesage, G., Bader, G. D., Ding, H., Xu, H., Xin, X., Young, J., Berriz, G. F., Brost, R. L., Chang, M. *et al. Global mapping of the yeast genetic interaction network*. Science, 303: 808-813, 2004.

[9] Lee, T. I., Rinaldi, N. J., Robert, F., Odom, D. T., Bar-Joseph, Z., Gerber, G. K., Hannett, N. M., Harbison, C. T., Thompson, C. M., Simon, I. *et al. Transcriptional regulatory networks in Saccharomyces cerevisiae*. Science, 298: 799-804, 2002.

[10] Haiyuan Yu and Mark Gerstein. *Genomic analysis of the hierarchical structure of regulatory networks*. PNAS, 103(40): 14724-31, 2006.

[11] Shen-Orr, S. S., Milo, R., Mangan, S. and Alon, U. *Network motifs in the transcriptional regulation network of Escherichia coli*. Nat. Genet. 31: 64-8, 2002.

[12] Uri Alon. *Network motifs: theory and experimental approaches*. Nat. Genet. Rev., 8: 450-61, 2007.

[13] Boyer, L. A. *et al. Core transcriptional regulatory circuitry in human embryonic stem cells*. Cell, 122: 947-956, 2005.

[14] A.M. Turing. *The chemical basis of morphogenesis*. Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences, 237: 37-72, 1952.

[15] Henry Greenside and Micheal Cross. *Pattern Formation and Dynamics in Nonequilibrium Systems*. Cambridge University Press, 2009.

[16] M.C. Cross and P.C. Hohenberg. *Pattern formation outside of equilibrium*. Rev. Mod. Phys., 65(3), 1993.

[17] Jeremy B.A. Green and James Sharpe. *Positional information and reaction-diffusion: two big ideas in developmental biology combine*. Development, 142: 1203-1211, 2015.

[18] A. Gierer and H. Meinhardt. *A theory of biological pattern formation.* Kybernetik, 12(1): 30-9, 1972.

[19] Patrick Müller, Katherine W. Rogers, Ben M. Jordan, Joon S. Lee, Drew Robson, Sharad Ramanathan, Alexander F. Schier. *Differential Diffusivity of Nodal and Lefty Underlies a Reaction-Diffusion Patterning System.* Science, 336: 721-724, 2012.

[20] James Briscoe and Stephen Small. *Morphogen rules: design principles of gradient-mediated embryo patterning.* Development, 142: 3996-4009, 2015.

[21] Aude Porcher and Nathalie Dostatni. *The Bicoid Morphogen System.* Current Biology, 20: R249-R254, 2010.

[22] Manu, Surkova, S., Spirov, A. V., Gursky, V. V., Janssens, H., Kim, A.-R., Radulescu, O., Vanario-Alonso, C. E., Sharp, D. H., Samsonova, M. *et al. Canalization of gene expression and domain shifts in the Drosophila blastoderm by dynamical attractors.* PLoS Comput. Biol., 5, e1000303, 2009.

[23] Xu, Z., Chen, H., Ling, J., Yu, D., Struffi, P. and Small, S. *Impacts of the ubiquitous factor Zelda on Bicoid-dependent DNA binding and transcription in Drosophila.* Genes Dev., 28: 608-621, 2014.