From Genes to Patterns: A Framework for Modeling the Emergence of Embryonic Development from Transcriptional Regulation

Jimena Garcia-Guillen, Ezzat El-Sherif*

School of Integrative Biological and Chemical Sciences (SIBCS), University of Texas Rio Grande Valley (UTRGV), Edinburg, Texas, USA

* Corresponding author: ezzat.elsherif@utrgv.edu

Abstract

Understanding embryonic patterning, the process by which groups of cells are partitioned into distinct identities defined by gene expression, is a central challenge in developmental biology. This complex phenomenon is driven by precise spatial and temporal regulation of gene expression across many cells, resulting in the emergence of highly organized tissue structures. While similar emergent behavior is well understood in other fields, such as statistical mechanics, the regulation of gene expression in development remains less clear, particularly regarding how molecular-level gene interactions lead to the large-scale patterns observed in embryos. In this study, we present a modeling framework that bridges the gap between molecular gene regulation and tissue-level embryonic patterning. Beginning with basic chemical reaction models of transcription at the single-gene level, we progress to model gene regulatory networks (GRNs) that mediate specific cellular functions. We then introduce phenomenological models of pattern formation, including the French Flag and Temporal Patterning/Speed Regulation models, and integrate them with molecular/GRN realizations. A key feature of our framework is the explicit articulation of underlying assumptions at each level of the model, from transcriptional regulation to tissue patterning. By making these assumptions clear, we provide a foundation for future experimental and theoretical work to critically examine and challenge them, thereby improving the accuracy and relevance of gene regulatory models in developmental biology. As a case study, we explore how different strategies for integrating enhancer activity affect the robustness and evolvability of gene expression patterns in embryonic development. Our simulations suggest that a two-step regulation strategy—enhancer activation followed by competitive integration at the promoter—ensures more standardized integration of new enhancers into developmental GRNs, highlighting the adaptability of eukaryotic transcription. These findings provide new insights into the transcriptional mechanisms underlying embryonic patterning, offering a framework for future experimental and theoretical investigations.

1 Introduction

A fundamental challenge in developmental biology is understanding how a group of cells is partitioned into distinct identities, where each group is defined by the expression of one or a few specific genes (1–4). This process of cell diversification is further subdivided recursively, creating increasingly specialized subgroups, ultimately forming the complex and highly organized structure of the embryo with its diverse cell types. This phenomenon is known as embryonic patterning. Embryonic patterning relies on the precise regulation of gene expression—spatially and temporally—ensuring the right genes

are activated in the right cells at the right time (1,2,5). This regulation controls how cells acquire their identities and organize into complex structures. Embryonic patterning, therefore, is an emergent phenomenon, resulting from the intricate regulation of many interacting genes across numerous cells. This concept of emergence parallels statistical mechanics, where complex behaviors in macroscopic systems, like gases or liquids, arise from interactions among many microscopic particles (6). Although the underlying laws for these particles, such as Newton's laws or quantum mechanics, are relatively simple, their collective behavior leads to emergent large-scale phenomena. Similarly, in embryonic development, gene interactions across many cells produce the complex patterns guiding organism formation (7,8). However, unlike well-understood models in statistical mechanics, our knowledge of gene regulation remains incomplete, particularly regarding how individual genes and their interactions drive these developmental patterns.

In emergent phenomena, not all features of the underlying microscopic elements are critical to the behavior of the larger macroscopic system. Similarly, in embryonic development, not all details of transcription at the single-gene level are essential contributors to the emergent patterns of gene expression that guide development. Therefore, a key step in understanding how gene regulation mediates embryonic patterning is to develop abstract, simplified models (9,10) that focus on the essential mechanisms at the single-gene level while still capturing the emergent gene expression patterns. However, we currently lack such models.

On one hand, reductionist approaches that delve into the molecular details of transcriptional machinery have uncovered vast amounts of information (5,11,12), but they often fail to clarify how these details contribute to the overarching process of pattern formation. On the other hand, existing models of gene regulation that is used in modeling embryonic pattern formation are often based on simplified systems (13), such as those found in prokaryotes. While these models offer insights, prokaryotic gene regulation is much simpler than that of eukaryotes—particularly in metazoans—making it unclear which aspects of eukaryotic transcription are crucial for mediating pattern formation. Identifying these key aspects, even at a theoretical level, would enable more targeted experimental work aimed at understanding the transcriptional mechanisms that drive embryonic patterning. To address this challenge, we need to establish a modeling framework that links gene regulation at the molecular level to the higher-level process of pattern formation across embryonic tissues. Such a framework should not simply aim to reproduce experimental data, but rather expose the underlying assumptions about how gene regulation occurs, providing new perspectives on the problem.

In this paper, we propose such a framework by developing simple toy models of transcription and gene regulatory interactions at the single-gene level. Building on these models, we construct basic pattern formation models that generate realistic yet simplified gene expression patterns. Even if some of these models are based on outdated experiments in prokaryotes, this effort will clarify the assumptions driving our understanding of gene regulation, highlight areas for future investigation, and point to key differences between prokaryotic and eukaryotic transcription. By making these assumptions explicit, we aim to shed light on the connection between the molecular processes of transcription and the large-scale phenomenon of embryonic pattern formation, opening new avenues for exploration in developmental biology and developmental genetics.

We begin our modeling effort by focusing on the transcription process of single genes, using basic chemical reaction models. From there, we model how multiple genes interact to form gene regulatory

networks (GRNs) that mediate specific functions within single cells. Next, we introduce phenomenological models of embryonic pattern formation, including various versions of the French Flag model and the Temporal Patterning/Speed Regulation models. These models are then extended with molecular or GRN realizations, representing a comprehensive modeling approach that connects transcriptional-level processes to tissue-level pattern formation. As a case study, we then explore the impact of different strategies for integrating the activities of multiple enhancers regulating a single gene on the robustness and evolvability of embryonic patterning. Our simulations suggest that the two-step gene regulation strategy observed in eukaryotes—first enhancer activation, followed by competitive integration of enhancer activities at the promoter—provides a more standardized approach for incorporating newly evolved enhancers into developmental GRNs. This insight emphasizes the evolutionary adaptability of eukaryotic transcriptional regulation and its role in shaping robust embryonic development.

2 Materials and methods

2.1 Modeling the regulation of a single gene

We begin our modeling of transcription during embryonic patterning by focusing on the kinetics of a single gene regulated by a single activator. Transcription initiation is typically controlled by a non-coding DNA sequence associated with the gene. In prokaryotes and simple eukaryotes, this regulatory region is a short DNA sequence located directly upstream of the coding sequence, known as the promoter. Transcription factors (TFs) that regulate the gene bind to the promoter, either aiding in the recruitment of RNA polymerase (in the case of activator TFs) or blocking it (in the case of repressor TFs), thereby controlling transcription.

In higher eukaryotes, transcription initiation is more complex and often occurs in two stages. Regulatory TFs first bind to a non-coding region associated with the gene, known as an enhancer (5), which may be distant from the promoter or even located far from the gene itself. These enhancers influence the recruitment of general TFs to the promoter, where transcription is initiated. For simplicity, we will model gene regulation here as the binding of a single activator to a single regulatory region. This region may be considered either the promoter itself or a functional combination of both the promoter and an enhancer, abstracting away the details of their interaction. Later in the paper, we will extend this model to explore the case of interactions between one or more enhancers and the gene promoter, capturing the more complex regulatory dynamics of transcription.

In this simple model of gene activation, an activator TF, denoted as X, binds to the regulatory region of a gene Y (Figure 1A). We represent the unbound regulatory region as D_0 and the region bound to X as D_1. Transcription factors bind and unbind the regulatory region stochastically, with binding and unbinding rates denoted as k_1 and k_{-1}, respectively (Figure 1A), described by the equation:

$$D_0 + X \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} D_1 \tag{1.1}$$

When the regulatory region is unbound (D_0), transcription is inactive, whereas when X is bound (D_1), the gene *Y* is transcribed at a rate k_t, producing transcripts of *Y*, as described by the equation:

$$D_1 \overset{k_t}{\rightarrow} D_1 + Y \tag{1.2}$$

Transcripts are degraded at a rate λ, as modeled by the equation:

$$Y \overset{\lambda}{\rightarrow} \emptyset \tag{1.3}$$

For simplicity, we assume that no transcription occurs in the absence of activator binding (i.e., no basal transcription).

The three reaction equations that describe the transcriptional activation of gene Y by transcription factor (TF) X (Eqs. 1.1–1.3) can be translated into a set of three differential equations using the law of mass action (see Figure S1 and Computer Simulations 1–2 for further details), as follows:

$$\frac{dD_1}{dt} = k_1(1 - D_1)X - k_{-1}D_1 \tag{2.1}$$

$$\frac{dX}{dt} = -\frac{dD_1}{dt} \tag{2.2}$$

$$\frac{dY}{dt} = k_t D_1 - \lambda Y \tag{2.3}$$

Here, D_1 represents the fraction of DNA bound by X (with $D_1 + D_0 = 1$). These three differential equations can be solved to determine $Y(t)$, the concentration of gene Y transcripts over time. However, because TF binding to DNA occurs on the millisecond timescale, while transcription proceeds on the order of seconds, we can apply the quasi-steady-state assumption. This assumption suggests that the steady-state values of X and D_1 are reached much faster than the steady state of Y:

$$\frac{dD_1}{dt} = 0 \tag{3}$$

By substituting Eq. 3 into Eqs. 2.1–2.3, we reduce the system to a single equation (see detailed derivation in Text S1), which takes the form of the Michaelis-Menten equation:

$$\frac{dY}{dt} = \frac{\frac{k_t X}{K_X}}{1 + \frac{X}{K_X}} - \lambda Y \tag{4.1}$$

Here, K_X is the dissociation constant of TF X, representing the rate at which TF X detaches from DNA relative to its binding rate, and serves as a measure of the binding strength of X to DNA (lower values of K_X indicate stronger binding):

$$K_X = \frac{k_{-1}}{k_1} \tag{4.2}$$

As demonstrated by Computer Simulations (Text S2), the Michaelis-Menten equation (Eq. 4.1) yields results that are consistent with the three original equations when $k_1, k_{-1} \gg k_t$ (see Figure S2). To calculate the steady-state concentration of Y as a function of changing X, we set the transcription rate $\frac{dY}{dt} = 0$, yielding the steady-state solution Y_{ss}:

$$Y_{ss} = \frac{k_t}{\lambda} \frac{\frac{X}{K_X}}{1 + \frac{X}{K_X}} \tag{4.3}$$

This results in a monotonic increase in steady-state transcription as the concentration of TF X increases, with saturation occurring at a maximum transcription rate of $\frac{k_t}{\lambda}$ as X exceeds the dissociation constant K_X (see simulation results in Figure 1B).

Thermodynamic State Ensemble (TSE) modelling

As discussed earlier, deriving the Michaelis-Menten relationship from reaction equations can be quite tedious, particularly for more complex scenarios involving multiple transcription factors (TFs) binding to the cis-regulatory region. A more streamlined approach for modeling transcriptional regulation that produces comparable results is the "Thermodynamic State Ensemble" (TSE) modeling (14–16). In this method, changes in transcript levels are influenced by the transcription rate (T_x) and transcript degradation:

$$\frac{dY}{dt} = T_x - \lambda Y \tag{5.1}$$

The transcription rate, T_x, is expressed as the proportion of TF, DNA, and RNA polymerase configurations conducive to transcription relative to all possible configurations. Below, we outline how to use the TSE strategy to model gene regulation; for a more in-depth and rigorous explanation, readers are referred to the detailed exposition in [reference].

As an example, we will use the TSE strategy to re-derive the rate equation governing the binding of a single activator X to the cis-regulatory region of gene Y. In this case, the denominator of T_x accounts for two possible states: unbound DNA (represented by 1) and DNA bound by X (represented by $\frac{X}{K_X}$, where K_X is the dissociation constant of X). The numerator includes only the states conducive to transcription—DNA bound by X, represented by $\frac{X}{K_X}$. The overall expression is then multiplied by the transcription rate k_t, which occurs once RNA polymerase binds to the DNA. This yields:

$$T_x = k_t \frac{\frac{X}{K_X}}{1 + \frac{X}{K_X}} \tag{5.2}$$

Substituting this into the differential equation for transcript dynamics gives the familiar Michaelis-Menten expression for a single activator (Eq. 4.1):

$$\frac{dY}{dt} = k_t \frac{\frac{X}{K_X}}{1 + \frac{X}{K_X}} - \lambda Y \tag{5.3}$$

Next, we apply the TSE strategy to model gene regulation by a single repressor X (Figure 1C). In this case, the denominator of T_x remains similar to the activation case, but the numerator reflects only the state conducive to transcription—unbound DNA, represented by 1 (i.e., transcription occurs in the absence of repressor X). This yields the following equation for the transcription rate:

$$\frac{dY}{dt} = k_t \frac{1}{1 + \frac{X}{K_X}} - \lambda Y \tag{6}$$

When plotting the steady-state value of Y as a function of X, the transcription rate reaches a maximum at zero concentration of the repressor X, then progressively decreases as X concentration increases (see computer simulation result in Figure 1D).

Modeling cooperativity

Often, multiple copies of the same transcription factor (TF) are required to bind cooperatively in order to effectively recruit RNA polymerases. In this section, we will use the Thermodynamic State Ensemble (TSE) approach to model the cooperative binding of *n* copies of activator TF X for the activation of gene Y. We will assume a case of extreme cooperativity, meaning that all *n* copies of X must bind simultaneously to the regulatory region of Y, whereas fewer than *n* copies are unable to bind.

In this scenario, the denominator of T_x represents all possible binding events, which include unbound DNA and the binding of *n* copies of X to the cis-regulatory elements. The numerator, however, includes only the state conducive to gene activation, which is the simultaneous binding of *n* copies of X to the regulatory region of Y (see Figure 1E). This yields the rate equation:

$$T_x = \frac{k_t \left(\frac{X}{K_X}\right)^n}{1 + \left(\frac{X}{K_X}\right)^n} \tag{7.1}$$

This type of equation is commonly referred to as the Hill equation.

Cooperativity introduces a nonlinear, switch-like behavior in the gene's response to activator binding. The higher the value of *n*, the more pronounced this switch-like behavior becomes, with a sharper transition from low to high input. As *n* increases further, the gene's response approaches digital behavior, as demonstrated by comparing the cases of $n = 2$ and $n = 5$ in Figures 1F and 1G, respectively. The point at which the gene switches from low to high output is determined by the dissociation constant K_X, which characterizes the binding affinity of TF X (as shown in Eq. 7.1).

Similarly, the Hill equation for the cooperative binding of *n* repressors follows a comparable form:

$$T_x = \frac{k_t}{1 + \left(\frac{X}{K_X}\right)^n} \tag{7.2}$$

In both cases, the cooperative binding of TFs leads to a sharp, switch-like regulation of gene expression, reflecting the impact of cooperativity on transcriptional control.

Modeling multiple TFs binding cis-regulatory module

So far, we have considered the binding of one or multiple copies of the same transcription factor (either activator or repressor) to the regulatory region of a gene. In this section, we will extend this by considering the binding of multiple different TFs. Specifically, we will examine the binding of two TFs, A and B, to the regulatory region of gene Y. This can be extended to cases involving a greater number of different TFs with straightforward modifications.

The effect of binding multiple TFs depends on the logic that governs their regulation of the gene, often referred to as the "transcription logic" of the regulatory module. Below, we will explore several common transcriptional logic models, though more elaborate scenarios are certainly possible.

Two ORed Activators

In this model, the binding of either A, B, or both activates the gene (Figure 2A). Using the Thermodynamic State Ensemble (TSE) approach, and assuming there is no cooperative binding between A and B (i.e., the binding of A does not influence the binding of B and vice versa), we obtain the following equation:

$$\frac{dY}{dt} = \frac{\frac{A}{K_A} + \frac{B}{K_B} + \frac{A}{K_A}\frac{B}{K_B}}{1 + \frac{A}{K_A} + \frac{B}{K_B} + \frac{A}{K_A}\frac{B}{K_B}} - \lambda Y \tag{8.1}$$

In cases where A and B bind cooperatively, with cooperativity values n_A and n_B, we derive the following Hill equation:

$$\frac{dY}{dt} = \frac{\left(\frac{A}{K_A}\right)^{n_A}}{1 + \left(\frac{A}{K_A}\right)^{n_A}} + \frac{\left(\frac{B}{K_B}\right)^{n_B}}{1 + \left(\frac{B}{K_B}\right)^{n_B}} - \lambda Y \tag{8.2}$$

As shown in Figure 2B, simulation result of this relationship matches the OR logic as expected. Moving forward, we will focus on the binding of single copies of A and B, as the extension to cooperative binding is straightforward.

Two Competitive ORed Activators

Here, we consider the case where both A and B are activators that bind the same site on the regulatory region of gene Y, leading to competitive binding (Figure 2C). In this case, the denominator includes unbound DNA, DNA-bound A, and DNA-bound B. However, the simultaneous binding of both A and B

is excluded because they compete for the same binding site and cannot bind at the same time. The numerator includes the states where A or B binds independently, resulting in the reaction equation:

$$\frac{dY}{dt} = \frac{\frac{A}{K_A} + \frac{B}{K_B}}{1 + \frac{A}{K_A} + \frac{B}{K_B}} - \lambda Y \tag{9}$$

As expected, simulation result of the Hill-function extension of this relationship matches the competitive OR logic as described above (Figure 2D).

Two ANDed Activators

In this model, both A and B must bind simultaneously to activate gene expression (Figure 2E). This logic results in the following equation:

$$\frac{dY}{dt} = \frac{\frac{A}{K_A} \cdot \frac{B}{K_B}}{1 + \frac{A}{K_A} + \frac{B}{K_B} + \frac{A}{K_A}\frac{B}{K_B}} - \lambda Y \tag{10}$$

As expected, simulation result of the Hill-function extension of this relationship matches the competitive AND logic (Figure 2F).

Two ANDed Repressors

So far, we have considered activators, but the extension to models involving both activators and repressors is straightforward. As an example, we will examine the case of two ANDed repressors. In this case, only unbound DNA allows for transcription, and the transcription rate T_x is the product of the individual repression effects of A and B:

$$\frac{dY}{dt} = \frac{1}{1 + \frac{A}{K_A} + \frac{B}{K_B} + \frac{A}{K_A}\frac{B}{K_B}} - \lambda Y \tag{11}$$

2.2 Modeling gene regulatory networks

So far, we have considered the regulation of single genes by one or more transcription factors (TFs). During development, terminal and housekeeping genes may be passively regulated by other transcription factors. However, many TFs regulate other TFs, forming gene regulatory networks (GRNs) (Figure 3A). In this section, we will explore examples of GRNs that act within a single cell and how to model them. Later, we will extend this to GRNs that function across multiple cells within embryonic tissues.

Positive Feedback Loop

We begin by considering feedback loops involving two genes. If activation is positive and repression is negative, a positive feedback loop results when the product of all regulatory links' signs is positive. Positive feedback loops can function as toggle switches or memory devices—if one of the genes is activated, it remains active.

A positive feedback loop can be realized as either two genes activating each other (Eqs. 12.1, 12.2; Figure 3B) or two genes repressing each other (Eqs. 13.1, 13.2; Figure 3C). In a loop of two activators, both genes can be either high ("on") or low ("off") (Figures 3B', 3B"). In a loop of two repressors, the GRN stabilizes in a state where one gene is active and the other is repressed (Figures 3C', 3C"). In these equations, K_{XY} denotes the dissociation constant of TF Y binding to the regulatory region of gene X.

$$\frac{dX}{dt} = \frac{\left(\frac{Y}{K_{XY}}\right)^n}{1 + \left(\frac{Y}{K_{XY}}\right)^n} - \lambda X \tag{12.1}$$

$$\frac{dY}{dt} = \frac{\left(\frac{X}{K_{YX}}\right)^n}{1 + \left(\frac{X}{K_{YX}}\right)^n} - \lambda Y \tag{12.2}$$

For a repression-based positive feedback loop, the equations become:

$$\frac{dX}{dt} = \frac{1}{1 + \left(\frac{Y}{K_{XY}}\right)^n} - \lambda X \tag{13.1}$$

$$\frac{dY}{dt} = \frac{1}{1 + \left(\frac{X}{K_{YX}}\right)^n} - \lambda Y \tag{13.2}$$

Negative Feedback Loop

A negative feedback loop consists of two genes where one activates and the other represses (Figure 3D). Negative feedback loops help maintain homeostasis, ensuring that gene expression remains stable. For instance, the GRN modeled by Eqs. 14.1–14.2 maintains gene Y's transcriptional activity around the activation threshold for TF X binding to its regulatory region (K_{YX}), and gene X's activity around the deactivation threshold for TF Y binding to its regulatory region (K_{XY}) (Figure 3D').

$$\frac{dX}{dt} = \frac{\left(\frac{Y}{K_{XY}}\right)^n}{1 + \left(\frac{Y}{K_{XY}}\right)^n} - \lambda X \tag{14.1}$$

$$\frac{dY}{dt} = \frac{\left(\frac{X}{K_{YX}}\right)^n}{1 + \left(\frac{X}{K_{YX}}\right)^n} - \lambda Y \tag{14.2}$$

Oscillatory GRNs

While two-gene negative feedback loops typically maintain homeostasis, more complex three-gene negative feedback loops can result in oscillations, depending on network parameters. For example, a GRN with two activators and one repressor (Figure 3E, E') can behave as an oscillator (Eqs. 15.1–

15.3). A GRN composed of three repressors, also known as a repressilator (17), can also function as an oscillator (Figure 3F, F').

$$\frac{dX}{dt} = \frac{\left(\frac{Z}{K_{XZ}}\right)^n}{1 + \left(\frac{Z}{K_{XZ}}\right)^n} - \lambda X \tag{15.1}$$

$$\frac{dY}{dt} = \frac{\left(\frac{X}{K_{YX}}\right)^n}{1 + \left(\frac{X}{K_{YX}}\right)^n} - \lambda Y \tag{15.2}$$

$$\frac{dZ}{dt} = \frac{\left(\frac{Y}{K_{ZY}}\right)^n}{1 + \left(\frac{Y}{K_{ZY}}\right)^n} - \lambda Z \tag{15.3}$$

Developmentally inspired GRNs

So far, we have presented example GRNs with specific wiring capable of executing various functions within the cell. The structures we've discussed are not the only possible configurations, as many different GRN architectures with varying parameter values can accomplish the same function. However, GRNs involved in development, particularly those involved in embryonic pattern formation, have distinct characteristics. One notable feature is the minimal reliance on activators; most transcription factors (TFs) involved in embryonic patterning are repressors, where activation is achieved through the repression of a repressor, a process called de-repression.

We will now explore some common GRN architectures that perform specific tasks during embryonic pattern formation, relying on repression and de-repression to mediate regulatory relationships between genes. Another important feature of these GRNs is the extensive cross-regulation between the genes, where all constituent genes repress each other. Despite this complexity, variations in the strength of repression among the genes can produce meaningful and distinct functions, as illustrated by the GRNs discussed below.

Multi-stable GRNs

A multi-stable GRN functions similarly to the positive feedback loop or genetic toggle switch discussed earlier (Fig. _). However, developmental GRNs typically consist of more than two genes. The simplest extension of the toggle switch to more than two genes is a multi-stable GRN, composed of M mutually repressing genes. This repression is characterized by small dissociation constants (θ_s), indicating strong repression (Eq. 16; Fig. 3G). In this configuration, depending on the initial conditions, one gene remains active while the others are inactive (Fig. 3G'). Other variations of this realization are possible, where subsets of genes are co-expressed due to weakened repressive links between them.

$$G_m = \frac{1}{1 + \left(\frac{G_i}{\theta_s}\right)^n} - \lambda G_m \tag{16}$$

Genetic Cascade

Another developmentally inspired GRN is the genetic cascade, where a sequence of genes is activated one after another. A simple realization of this cascade involves one gene activating the next, which in turn represses the previous gene, and so on. In a GRN based on de-repression logic, all genes are naturally active but kept repressed by other genes in the cascade (with small dissociation constants, θ_s, indicating strong repression; Eqs. 17.1, 17.2). The only exception is that a gene weakly represses the gene preceding it in the cascade (with larger dissociation constants, θ_w, indicating weaker repression). When the cascade is initialized by activating the first gene, a domino effect occurs where the next gene represses the previous one, leading to the de-repression of the subsequent gene, and so forth (Fig. 3H').

For $m = 1$:

$$G_m = \frac{1}{1 + \left(\frac{G_i}{\theta_s}\right)^n} - \lambda G_m \tag{17.1}$$

For $m \neq 1$:

$$G_m = \frac{1}{1 + \left(\frac{G_m - 1}{\theta_w}\right)^n} \cdot \frac{1}{1 + \left(\frac{G_i}{\theta_s}\right)^n} - \lambda G_m \tag{17.2}$$

Oscillators

The basic logic of an oscillator, which is based on mutually repressive links and de-repressive interactions, is similar to that of a genetic cascade. However, in this case, the last gene in the sequence does not repress the first gene, allowing the system to cycle continuously (Eqs. 18.1, 18.2; Fig. 3I, I').

For $m = 1$:

$$G_m = \frac{1}{1 + \left(\frac{G_i}{\theta_s}\right)^n} - \lambda G_m \tag{18.1}$$

For $m \neq 1$:

$$G_m = \frac{1}{1 + \left(\frac{G_m - 1}{\theta_w}\right)^n} \cdot \frac{1}{1 + \left(\frac{G_i}{\theta_s}\right)^n} - \lambda G_m \tag{18.2}$$

2.3 Cellular differentiation of GRNs: Space enters the scene

So far, we have discussed GRNs that function within a single cell. In development, although cells within a tissue share the same genes and GRNs (due to having the same genome), they are regulated differently, allowing each cell to express a unique set of transcription factors (TFs) and follow distinct

developmental trajectories. This manifests as gene expression patterns across the tissue, where different subsets of cells express different genes.

While such expression patterns could, in theory, take any form, they usually adopt specific, stereotyped patterns. Two major classes of patterns typically observed during early developmental stages are periodic and non-periodic patterns (Fig. 4A, A'; shown for simplicity as a one-dimensional row of cells). Periodic patterns mediate the division of tissue into serial structures, such as demarcating the segmented structures along the anterior-posterior (AP) axis in arthropods and vertebrates. Non-periodic patterns, which are more common, mediate the division of tissue into distinct cell fates (regionalization). This differential expression of genes across a tissue results from regulating the same GRN differently in various cells (Figure 4B).

To regulate cells differently across a tissue, a morphogen gradient is often employed. An organizer (usually positioned at one extremity of the tissue to be patterned) secretes a ligand that diffuses and forms a morphogen gradient (Figure 4C). Different concentrations of the gradient set distinct developmental paths for cells along the tissue. However, tissue elongation—whether due to growth or convergent extension—often accompanies embryonic patterning. In some cases, tissue elongation occurs after patterning (Figure 4C'), while in others, it takes place concurrently with patterning.

Patterning a tissue during elongation is mechanistically distinct from patterning a non-elongating tissue (or a tissue patterned after elongation), as it often involves a dynamically retracting morphogen gradient. As a result, cells are exposed to changing concentrations of the morphogen over time (Figure 4D). In summary, embryonic patterning can be categorized based on the nature of the pattern (periodic vs. non-periodic) and whether the tissue is elongating or non-elongating. These developmental processes are explored further in the Results section.

The mechanisms that drive morphogen gradient formation range from passive diffusion to highly regulated ligand transport across a tissue, which are covered extensively in other reviews and are beyond the scope of this study. Here, we use a simplified formulation for the gradient, independent of the mechanism that generates it. For the case of a non-retracting gradient in a non-elongating tissue, we use the following sigmoid function, where x_0 specifies the inflection point of the gradient, and k denotes the gradient slope:

$$M(x) = \frac{1}{1 + e^{-k(x - x_0)}} \tag{19.1}$$

For the case of a retracting gradient, we employ the following moving sigmoid function, where v is the velocity of gradient retraction:

$$M(x, t) = \frac{1}{1 + e^{-k[x - (x_0(0) - vt)]}} \tag{19.2}$$

3 Results

In this section, we will discuss two of the most prominent models of embryonic pattern formation: the French Flag model (2,3) and the Temporal Patterning (or Speed Regulation) model (18,19). These models are phenomenological in nature, focusing on descriptive aspects and leaving out the molecular details of their genetic or molecular realization. We will explore four variations of each model, addressing the four common scenarios of embryonic patterning: (1) a non-periodic pattern in a non-elongating tissue, (2) a non-periodic pattern in an elongating tissue, (3) a periodic pattern in a non-elongating tissue, and (4) a periodic pattern in an elongating tissue.

Following this, we will apply the gene regulation models previously discussed to investigate different molecular and genetic realizations of these two phenomenological models of embryonic pattern formation. Although the reaction-diffusion model is another significant framework for understanding pattern formation, little is known about its regulation at the molecular level, and it will not be covered in this study.

3.1 The French Flag Model

The core mechanism of the French Flag model (2,3,18) is that different concentrations of a morphogen gradient activate different genes or cellular states (Fig. 5A). This mechanism is ideal for mediating non-periodic patterns in non-elongating tissue (Fig. 5A, left panel), where certain ranges of the morphogen gradient activate specific genes. However, a more challenging scenario arises when applying the French Flag model to pattern an elongating tissue during elongation. The model assumes a stable gradient to set precise boundaries between gene expression domains. To address this, a retracting gradient (M'; black in Fig. 5A, right panel) activates a gene (M; grey in Fig. 5A, right panel) with a slow decay rate (ideally $\lambda = 0$), forming a stable long-range morphogen gradient (grey) from a retracting short-range gradient (black). This stable gradient then mediates the division of the tissue into distinct gene expression domains. Extending the French Flag model to periodic patterns (Fig. 5B) is conceptually straightforward: alternating ranges of morphogen gradient concentrations turn a gene on and off. However, achieving this at the molecular level is more complex and will be discussed later.

It is important to note that the French Flag model is a purely phenomenological, descriptive model and does not specify how it is realized at the molecular or genetic level. Here, we will explore potential molecular and GRN (gene regulatory network) realizations of this model. Despite the simplicity of the French Flag model, realizing it at the GRN level is challenging, as some genes must be both activated and repressed by the same morphogen gradient. For instance, in Fig. 5A, the red gene is activated by moderately high levels of the morphogen gradient (grey), but repressed at very high concentrations. This implies that the red gene responds to two different thresholds, T1 and T2, of the same morphogen gradient. None of the models discussed so far can accommodate this feature. While this phenomenon has been experimentally observed, it has not yet been fully elucidated and will be revisited in the Discussion. For now, we will limit ourselves to simpler models where each gene responds to a single threshold of the morphogen.

We will start by reviewing potential GRN realizations of the French Flag model through a basic gene expression pattern where genes are expressed in bands extending to the tissue ends (Fig. 5C, C'). This

simple pattern can be realized by adjusting the activation and repression thresholds of patterning genes.

$$M'^{(x,t)} = \frac{1}{1 + e^{-k[x-(x_0(0)-vt)]}} \tag{20.1}$$

$$M(x,t) = \frac{M'}{1 + M'} \tag{20.2}$$

For example, in Fig. 5C, the blue gene (G1) is strongly activated by the morphogen gradient (M, grey in Fig. 5C, C'), extending its expression towards the lower end of the gradient. The red gene (G2), on the other hand, is weakly activated by M. This can be achieved by setting G1 to have a low activation threshold (low K_{G1M}, indicating strong binding sites for M at G1's CRM) and G2 to have a high activation threshold (high K_{G2M}, indicating weaker binding sites for M at G2's CRM). The expression domains of the green and orange genes (G3 and G4, respectively) are formed by having them repressed by M—strongly in the case of G3 (low K_{G3M}) and weakly in the case of G4 (high K_{G4M}).

To create gene expression domains that do not extend to the boundaries of the morphogen gradient, two morphogen gradients can be employed. An example of this is shown in Fig. 5D, D', where two gradients (M1 and M2, represented by light and dark grey, respectively) together define the boundaries of the blue and red genes (G1 and G2). In this scenario, G1 is strongly repressed by M1 but weakly repressed by M2 (i.e., small K_{G1M1} and large K_{G1M2}), while G2 is weakly repressed by M1 and strongly repressed by M2 (i.e., large K_{G2M1} and small K_{G2M2}). This approach can be easily generalized to multiple genes by adjusting the repression thresholds of the two gradients for each gene.

$$\frac{dG_i}{dt} = \frac{\left(\frac{M}{K_{GiM}}\right)^n}{1 + \left(\frac{M}{K_{GiM}}\right)^n} - \lambda G_i \tag{22}$$

$$\frac{dG_i}{dt} = \frac{\left(\frac{M1}{K_{GiM1}}\right)^n}{1 + \left(\frac{M1}{K_{GiM1}}\right)^n} \cdot \frac{\left(\frac{M2}{K_{GiM2}}\right)^n}{1 + \left(\frac{M2}{K_{GiM2}}\right)^n} - \lambda G_i \tag{23}$$

To generate multiple gene expression domains from a single morphogen gradient, cross-regulatory interactions between genes are required. The specific GRN wiring can vary, but one straightforward realization is shown in Fig. 5E. In this configuration, the morphogen gradient M activates patterning genes (G1, G2, and G3) at different thresholds. Without cross-regulatory interactions, this would result in broad expression domains, extending from the high concentration of M to the respective activation thresholds. Repression from other genes is needed to refine these domains. For instance, G1 is repressed by both G2 and G3, restricting its expression to the lower end of the gradient.

$$G_i = \frac{1}{1 + \left(\frac{M}{K_{GiM}}\right)^n} \cdot \frac{1}{1 + \left(\frac{G_{i+1}}{\theta_w}\right)^n} \cdot \prod_{j=i+2}^{N} \frac{1}{1 + \left(\frac{G_j}{\theta_s}\right)^n} - \lambda G_i \tag{24.1}$$

For $i = N - 1$:

$$G_i = \frac{1}{1+\left(\frac{M}{K_{GiM}}\right)^n} \cdot \frac{1}{1+\left(\frac{G_{i+1}}{\theta_w}\right)^n} - \lambda G_i \tag{24.2}$$

For $i = N$:

$$G_i = \frac{1}{1+\left(\frac{M}{K_{GiM}}\right)^n} - \lambda G_i \tag{24.3}$$

Other GRN configurations are also possible. For instance, in the AC-DC GRN motif (20–22), the morphogen gradient thresholds emerge from cross-regulation between genes, rather than being dictated by the gradient itself.

So far, we have discussed the formation of non-periodic patterns, noting that their GRN realization is more challenging than the simplicity of the phenomenological French Flag model suggests. This difficulty primarily stems from the challenge of making a single gene respond to two different thresholds of the same morphogen gradient, which we addressed using either two morphogen gradients or cross-regulatory interactions. Applying the French Flag model to periodic patterns is even more complex, as genes must respond to multiple thresholds, with the morphogen gradient alternately activating and repressing the same gene. One solution is similar to the non-periodic case: decomposing the periodic pattern into several non-periodic patterns, each mediated separately by two morphogen gradients. The final periodic pattern is the sum of all sub-patterns.

$$G_i = \frac{1}{1+\left(\frac{M}{K_{GiM}}\right)^n} \cdot \frac{1}{1+\left(\frac{G_{i+1}}{\theta_w}\right)^n} \cdot \prod_{j=i+2}^{N} \frac{1}{1+\left(\frac{G_j}{\theta_s}\right)^n} - \lambda G_i \tag{25.1}$$

For $i = N - 1$:

$$G_i = \frac{1}{1+\left(\frac{M}{K_{GiM}}\right)^n} \cdot \frac{1}{1+\left(\frac{G_{i+1}}{\theta_w}\right)^n} - \lambda G_i \tag{25.2}$$

For $i = N$:

$$G_i = \frac{1}{1+\left(\frac{M}{K_{GiM}}\right)^n} - \lambda G_i \tag{25.3}$$

This method of generating periodic patterns has been studied extensively in the context of anterior-posterior (AP) patterning in the early Drosophila embryo (Fig. 5F, G). It offers insights into how cis-regulatory modules are structured during development. A prominent example is the expression of the *even-skipped* (*eve*) gene in the early *Drosophila* embryo, which forms seven stripes (23,24). Rather than arising from a periodic process like an oscillator or reaction-diffusion mechanism, this pattern results from decomposing the periodic pattern into several non-periodic sub-patterns, each composed of a single stripe or pair of stripes. Each sub-pattern is mediated by a dedicated enhancer, and the final expression of eve is the sum of all enhancer activities. The regulatory logic of each enhancer is primarily determined by the multiplicative (AND) interactions of repressing transcription factors (TFs) known as

gap genes, which form morphogen gradients in the early *Drosophila* embryo. Fig. 5F, G illustrates the regulatory logic of two enhancers (25): the 3+7 enhancer (which mediates the formation of the 3rd and 7th stripes) and the 4+6 enhancer (which mediates the formation of the 4th and 6th stripes). The positioning of these stripes is determined by the binding strengths of the gap gene morphogens *hunchback* (*hb*) and *knirps* (*kni*) to the 3+7 and 4+6 enhancers along the AP axis.

This example, along with others, demonstrates that an additive OR logic is mediated by using separate enhancers spaced far enough to avoid interference between their activities, while multiplicative AND logic is mediated by placing TF binding sites close to each other within the same enhancer.

A simple example of this is shown in Fig. 5H, H', where gene G is expressed in two domains, each mediated by the additive effects of two separate enhancers: Enhancer 1 (E1) and Enhancer 2 (E2). E1 is strongly repressed by morphogen 1 (M1) and weakly repressed by morphogen 2 (M2), while E2 is weakly repressed by M1 and strongly repressed by M2.

$$E_i = \frac{\left(\frac{M1}{K_{EiM1}}\right)^n}{1+\left(\frac{M1}{K_{EiM1}}\right)^n} \cdot \frac{\left(\frac{M2}{K_{EiM2}}\right)^n}{1+\left(\frac{M2}{K_{EiM2}}\right)^n} - \lambda G_i \tag{26.1}$$

$$G = E_1 + E_2 - \lambda G \tag{26.2}$$

3.2 Temporal Patterning or the Speed Regulation Model

In temporal patterning, a temporal sequence is translated into a spatial pattern rather than directly mediating a spatial pattern (18,26,27). This mode of patterning, although less intuitive, is a standard mechanism in the development of neurogenesis (28,29) in many animals and at various stages of development. Recent studies have shown it is involved in many other tissues as well, such as segmentation and regionalization of the anterior-posterior (AP) axis in short-germ insects (30,31) and vertebrates (32), ventral neural tube patterning (33), and limb bud development . This raises the possibility that temporal patterning may be the default mechanism for embryonic pattern formation. While the reason for this preference is not fully understood, it has been suggested that this mechanism may offer greater robustness compared to alternatives, such as the French Flag model. It has also been hypothesized that since animals evolved from single-celled organisms, many evolutionarily conserved gene regulatory networks (GRNs) are temporal in nature, and multicellularity may have evolved by translating these conserved temporal patterns into spatial ones.

The Speed Regulation (SR) model (19) was recently proposed as a unifying mechanism for various types of temporal patterning observed in different tissues. The SR model synthesizes several patterning schemes that all share the core feature of a morphogen gradient modulating the speed of a temporal sequence, whether periodic or non-periodic. In the non-periodic version of the SR model, each cell within a tissue progresses through successive states (depicted in different colors in Fig. 6A), with each state defined by the expression of one or more genes. The speed of transitions between these states is controlled by a molecular factor (referred to as the "speed regulator," shown in grey at the top of Fig. 6A). At very low or zero concentrations of the speed regulator, the transitions slow down to the point where the states are indefinitely stabilized. When a group of cells experiences a gradient of the speed

regulator (Fig. 6A'), all cells transition through successive states, but at progressively slower speeds as they encounter lower concentrations of the gradient. This creates the illusion of cellular states propagating as waves from high to low gradient concentrations. These waves, known as "kinematic" or "pseudo-waves," do not rely on diffusion or cell-to-cell communication (34,35). We refer to this version of the model as "gradient-based speed regulation," which is particularly suited for patterning non-elongating tissues (Fig. 6A', left). The model can also pattern elongating tissues if the gradient retracts as a wavefront, a process we call "wavefront-based speed regulation" (Fig. 6A', right). The SR model can also generate periodic structures if the sequential gene activation process is driven by a biological clock instead of a sequential transition of states (Fig. S_). Notably, in the wavefront-based version of SR, if the wavefront is in the form of a tapered gradient (a superposition of the gradient-based and wavefront-based SR models), kinematic waves will propagate from high to low concentrations of the gradient, in the opposite direction of wavefront retraction, as the tissue elongates.

Like the French Flag (FF) model, the SR model is phenomenological and does not propose a molecular mechanism or GRN realization that can transform temporal sequences (or oscillations) into patterns. A straightforward realization of the model is for a morphogen gradient (the speed regulator) to modulate the overall transcription rate, including mRNA decay rates (Fig. 6B).

For $j = 1$:

$$G_j = \frac{M}{1+M} \times \prod_{i=2}^{J} \frac{1}{1+\left(\frac{G_i}{\theta_s}\right)^n} - \lambda G_j \tag{27.1}$$

For $j \neq 1$:

$$G_j = \frac{M}{1+M} \times \left(\frac{1}{1+\left(\frac{G_{m-1}}{\theta_w}\right)^n} \prod_{i=1, i\neq j, j-1}^{J} \frac{1}{1+\left(\frac{G_i}{\theta_s}\right)^n} \right) - \lambda G_j \tag{27.2}$$

While this approach works well, it requires that the decay rates of all genes involved in patterning be regulated by the morphogen gradient in proportion to the regulation of their transcription. Achieving this at the molecular level may be unfeasible.

An alternative mechanism, recently proposed to modulate the timing of GRNs, is the Enhancer Switching Model. This model posits that each patterning gene is simultaneously wired into two GRNs (Fig. 1C): (i) a dynamic GRN that drives periodic or sequential gene activities, and (ii) a static GRN that stabilizes gene expression patterns. The concentration of the speed regulator (shown in grey in Fig. 1C) activates the dynamic GRN while repressing the static GRN, thus balancing the contribution of each GRN to the overall dynamics and consequently regulating the speed of gene expression. At high concentrations of the speed regulator, the dynamic GRN is dominant, facilitating rapid oscillations or sequential gene activities. At low concentrations, the static GRN dominates, resulting in slower oscillations or sequential gene activities. Each gene is connected to these two GRNs through two enhancers: (i) a dynamic enhancer that encodes the wiring of the gene within the dynamic GRN, and (ii) a static enhancer that encodes the wiring of the gene within the static GRN (Fig. 1D).

$$G_j = k_{td}E_{dj} + k_{ts}E_{sj} - \lambda G_j \tag{28.1}$$

For the dynamic enhancer, E_{dj}:

$$E_{dj} = \frac{M}{1+M} \times \prod_{i=2}^{J} \frac{1}{1+\left(\frac{G_i}{\theta_s}\right)^n} \tag{28.2}$$

For the static enhancer, E_{sj}:

$$E_{sj} = \frac{M}{1+M} \times \frac{1}{1+\left(\frac{G_{m-1}}{\theta_w}\right)^n} \prod_{i=1, i\neq j, j-1}^{J} \frac{1}{1+\left(\frac{G_i}{\theta_s}\right)^n} \tag{28.3}$$

$$E_{sj} = \frac{M}{1+M} \times \prod_{i=1, i\neq j}^{J} \frac{1}{1+\left(\frac{G_i}{\theta_s}\right)^n} \tag{28.4}$$

3.3 Case study: Impact of enhancer integration mechanisms on embryonic patterning

So far, we have discussed various embryonic pattern formation mechanisms using basic models of transcription. To that end, we employed simple toy models at each level of the modeling process. One of the main advantages of this approach is that it allows us to examine the effect of fundamental transcriptional and gene regulatory mechanisms on the overall performance of embryonic patterning systems. In this section, we will explore a specific example to demonstrate this. We will examine how the integration of multiple enhancer activities influences the performance of the enhancer switching model.

In our modeling of the enhancer switching mechanism (Fig. 6C, and Eqs. 27.1–27.3), we used a weighted sum of the two enhancer activities (Eq. 27.1). The physical interpretation of this relationship is that each enhancer can transcribe the gene independently (Fig. 7A). In other models of enhancer function, however, enhancers compete for binding to the promoter, and the enhancer that wins the competition sets the transcriptional state of the promoter (Eq. 28; Fig. 7B) (36). In the first scenario (Fig. 7A; Eq. 27.1), each enhancer determines its own final transcriptional rate (denoted as k_{td} and k_{ts} in Eq. 27.1), whereas in the second scenario, the normalized activities of competing enhancers set the transcriptional state of the promoter, but the maximum final transcriptional rate (denoted as k_t in Eq. 28) is set by the promoter itself (Fig. 7B).

The normalization effect of the enhancer competition scenario may be a crucial mechanism for integrating enhancer activities. The overall transcription level of a gene is a key parameter in any GRN, as the activation thresholds of genes within a GRN depend on the average transcription rates of the

genes regulating them. Enhancers are rapidly evolving genetic elements and are the primary drivers of novelty in animal evolution. However, this also introduces a risk, as allowing a newly evolved enhancer to take control of a key transcriptional parameter, such as the overall average transcription rate, could disrupt normal function. Alternatively, in the enhancer competition model, enhancers first compete to activate the promoter, with the promoter acting as the final arbiter for the transcription rate.

To investigate this further, we simulated our enhancer switching model for both the additive case (Fig. 7A, A') and the enhancer competition case (Fig. 7B, B'). Both scenarios resulted in comparable performance (Fig. 7A' and 7B'). We then examined the effect of an overactive enhancer (Fig. 7A" and 7B"). In the additive case, the overactive enhancer disrupted the overall function of the GRN (Fig. 7A"), whereas in the enhancer competition case, the normalization effect mitigated the impact of the overactive enhancer, preserving the overall performance of the GRN (Fig. 7B").

$$G_j = k_t \left(\frac{k_{ed} E_{dj} + k_{es} E_{sj}}{1 + k_{ed} E_{dj} + k_{es} E_{sj}} \right) - \lambda G_j \tag{29}$$

4 Discussion

In this paper, we developed a comprehensive framework to model the emergence of embryonic patterning, linking molecular gene regulation to tissue-level organization. We began by modeling transcription at the single-gene level using basic chemical reaction models and extended this to model gene regulatory networks (GRNs) that govern specific cellular functions. We then introduced phenomenological models of embryonic pattern formation, such as the French Flag model and Temporal Patterning models, integrating these with molecular and GRN realizations. Through our case study on enhancer integration, we demonstrated that a two-step gene regulation strategy—enhancer activation followed by competitive integration at the promoter—ensures robustness and evolvability in gene expression patterns, emphasizing the adaptability of eukaryotic transcriptional regulation. Our results provide new insights into the mechanisms that underpin embryonic patterning and the robustness of developmental processes.

Modeling gene regulation during embryonic development has traditionally followed two distinct strategies. The first involves using finely tuned gene regulatory network (GRN) models to fit large datasets, which we refer to as the "simulation approach." The second relies on simplified, abstract models that capture the overall behavior of the observed data, which we call the "toy model approach." Both approaches have limitations. The simulation approach, though data-driven, and the toy model approach, though more conceptual, often employ gene regulatory models inspired by prokaryotic gene regulation systems. These models may not accurately reflect the complexity of eukaryotic transcriptional machinery, particularly in animals. In this paper, we adopt the toy model approach, acknowledging its inherent limitations. However, our aim is to explicitly articulate the various assumptions made when modeling gene regulation, from the molecular scale to the tissue level. This transparency allows us to critically examine and challenge these assumptions throughout our modeling process. The simplicity and low computational cost of the toy model approach—unlike the simulation approach, which requires extensive data collection and optimization—facilitates this exploration. In fact, we present our model not as a definitive framework, but as a basis for questioning and refining the assumptions at each step. Our hope is that future work by our group and others will further challenge and improve upon these ideas.

One specific aspect of gene regulation we examine in this paper, which relates to the complex transcriptional machinery of animals, is the impact of multiple enhancer integration on the performance of gene regulatory networks (GRNs) during development—an aspect largely overlooked in previous modeling efforts. This represents a small departure from the prokaryote-inspired models traditionally used in gene regulation studies, moving instead toward the more intricate transcriptional mechanisms found in animals, particularly during development. However, this is neither the only nor the most critical aspect to reconsider in efforts to model gene regulation more realistically. A widely used assumption, which we also adopted in our modeling, is the simplistic regulatory logic of individual genes within the GRN, typically modeled with basic OR and/or AND relationships. Under this assumption, the complexity of GRN performance stems from the network's structure rather than from the computational capabilities of individual genes. Yet, it has been observed that single genes can exhibit far more complex regulatory logic. For instance, a single transcription factor (TF) may act as an activator or a repressor, depending on the concentration of co-regulating TFs. These intricate regulatory interactions are likely evolutionarily conserved and could serve as modular building blocks, akin to Lego pieces, which can be used as is or slightly modified to construct more extensive GRNs.

Another key aspect of gene regulation that warrants further investigation using our modeling framework is how the timing of GRNs is modulated. Recent findings suggest that many phenomena in embryonic patterning arise from the modulation of GRN timing by morphogen gradients, supporting a temporal/speed regulation model. In this paper, we model this effect using two distinct hypotheses. However, given the largely unexplored computational complexity of transcriptional machinery in developmental genes, it is conceivable that additional mechanisms for timing regulation exist. Moreover, recent experimental findings suggest other assumptions in current models may need re-evaluation. For example, the effects of transcriptional bursting and the formation of transcription factor (TF) clusters or condensates on GRN performance—particularly in embryonic patterning—remain unclear. These newly discovered phenomena could significantly influence how GRNs function. We believe that these, along with other assumptions, should and can be rigorously tested using the modeling framework presented here.

Figure Legends

Figure 1. Modeling the regulation of a single gene. (A) Schematic representation of gene activation by a single activator transcription factor. The transcription factor X binds to the unbound regulatory region (D_0) of gene Y, transitioning it to the bound state (D_1). Binding and unbinding occur at rates k_1 and k_{-1}, respectively. Transcripts are degraded at rate λ. As an illustration, transcript degradation is shown here but will be omitted in subsequent schematics. This model demonstrates the fundamental biochemical reactions involved in transcription initiation. **(B)** Simulation of steady-state transcription levels as a function of activator concentration. The graph depicts the steady-state concentration of gene Y transcripts (Y_{ss}) increasing monotonically with activator X concentration. Saturation is observed when the concentration of X surpasses the dissociation constant (K_x), resulting in a maximum transcription rate of k_t/λ. **(C)** Schematic representation of gene repression by a single repressor transcription factor. Repressor X binds to the regulatory region of gene Y, inhibiting transcription. The unbound state D_0 transitions to the bound state D_1 upon binding of X, with rates k_1 and k_{-1} indicating binding and unbinding dynamics, respectively. **(D)** Simulation of steady-state transcription levels as a function of repressor concentration. The graph depicts the steady-state concentration of gene Y transcripts (Y_{ss}) decreasing as the concentration of repressor X increases. Maximum transcription occurs when X is zero, diminishing progressively with higher X concentrations. **(E)** Modeling cooperativity in gene activation. Schematic representation of the cooperative binding of n copies of activator X to the regulatory region of gene Y. All n copies must bind simultaneously to activate transcription, representing extreme cooperativity. **(F)** Effect of cooperative binding on gene activation (n = 2). The simulation demonstrates how cooperative binding of two activator molecules leads to a switch-like increase in gene expression. The response curve shows a sharper transition compared to non-cooperative binding. **(G)** Effect of cooperative binding on gene activation (n = 5). The simulation illustrates that with higher cooperativity (n = 5), the gene's response to activator concentration becomes more digital, exhibiting an even sharper switch-like behavior.

Figure 2. Modeling multiple transcription factors binding to a cis-regulatory module. (A) Schematic of gene activation by two ORed activators. Transcription factors A and B independently bind to the regulatory region of gene Y. The presence of either A or B is sufficient to activate transcription, illustrating an OR logic gate. **(B)** Simulation of transcription levels with two ORed activators. The graph shows how the steady-state concentration of gene Y transcripts varies with different concentrations of A and B, reflecting the OR logic. **(C)** Schematic of gene activation by two competitive ORed activators. Activators A and B compete for the same binding site on gene Y's regulatory region. Binding of either A or B activates transcription, but they cannot bind simultaneously due to competition. **(D)** Simulation of transcription levels with competitive ORed activators. The graph illustrates how the competition between A and B affects gene Y's expression, with transcription levels depending on the relative concentrations of A and B. **(E)** Schematic of gene activation by two ANDed activators. Both transcription factors A and B must simultaneously bind to gene Y's regulatory region to initiate transcription, representing an AND logic gate. **(F)** Simulation of transcription levels with ANDed activators. The graph demonstrates the switch-like activation of gene Y only when both A and B are present, showcasing the cooperative effect of AND logic on gene expression.

Figure 3. Modeling gene regulatory networks (GRNs). (A) Schematic representation of a gene regulatory network (GRN). Multiple genes interact to regulate each other's expression, forming complex networks that mediate specific cellular functions. Shown a positive feedback loop between two genes as an example. **(B)** Positive feedback loop between two activator genes. Genes X and Y activate each other's expression. This mutual activation can lead to bistable states where both genes are either on or off. **(B', B'')** Simulation of gene expression states in the positive feedback loop. The graph shows stable states achieved in the system, with both genes being not expressed (B') or highly expressed (B''), depending on initial conditions. **(C)** Positive feedback loop between two repressor genes. Genes X and Y repress each other's expression. This mutual repression results in a system where one gene is active while the other is repressed. **(C', C'')** Simulation of gene expression states in the mutual repression loop. The graph demonstrates the system stabilizing with one gene active and the other inactive, based on initial expression levels. **(D)** Negative feedback loop between an activator and a repressor gene. Gene X activates gene Y, while gene Y represses gene X. This configuration promotes homeostasis, maintaining stable gene expression levels. **(D')** Simulation of gene expression in the negative feedback loop. The graph shows how the system maintains stable expression levels of genes X and Y over time. **(E)** Oscillatory GRN with two activators and one repressor. Gene X activates gene Y, gene Y activates gene Z, and gene Z represses gene X. This network can produce oscillations in gene expression over time. **(E')** Simulation of oscillatory behavior in the gene network. The graph displays cyclic fluctuations in the expression levels of genes X, Y, and Z, characteristic of an oscillatory system. **(F)** Repressilator network composed of three repressor genes. Genes X, Y, and Z repress one another in a cyclic manner, forming an oscillatory system known as a repressilator. **(F')** Simulation of oscillations in the repressilator network. The graph shows the periodic expression patterns of genes X, Y, and Z over time. **(G)** Multi-stable gene regulatory network with mutually repressing genes. Multiple genes repress each other strongly, leading to stable states where only one gene is expressed while the others are repressed. **(G')** Simulation of gene expression states in the multi-stable network. The graph shows that depending on initial conditions, one gene remains active while others are inactive. **(H)** Genetic cascade illustrating sequential gene activation. Each gene in the cascade represses the previous gene and weakly represses the next one. Activation of the first gene triggers a domino effect of gene expression. **(H')** Simulation of gene expression in the genetic cascade. The graph depicts the sequential activation and repression of genes over time, following the cascade logic. **(I)** Oscillatory gene network in which a series of genes repress each other in sequence without the last gene repressing the first, resulting in continuous oscillations of gene expression. **(I')** Simulation of oscillations in the gene network. The graph shows the cyclic expression patterns of the genes over time.

Figure 4. Introduction of spatial patterns in gene regulatory networks. (A, A') Illustration of non-periodic and periodic patterns in a tissue. The schematic shows a linear arrangement of cells exhibiting non-periodic (distinct regions) and periodic (repeating units) gene expression patterns. **(B)** Conceptual representation of cells within a tissue sharing the same GRN but being regulated differently. Different regulatory inputs lead to distinct expression patterns despite identical GRNs in each cell. **(C)** Formation of a morphogen gradient by an organizer. An organizer at one end of the tissue secretes a ligand that diffuses to form a gradient, providing positional information to cells. **(C')** Patterning followed by tissue elongation. The diagram illustrates how a tissue is patterned by a morphogen gradient before undergoing elongation, resulting in the expansion of patterned regions. **(D)** Concurrent tissue

elongation and patterning with a retracting gradient. A moving morphogen gradient, exposing cells to changing morphogen concentrations over time, patterns the tissue as it elongates.

Figure 5. The French Flag Model and its gene regulatory network realizations. (A) *French Flag model for non-periodic patterning.* Left panel: A stable morphogen gradient divides the tissue into distinct gene expression domains (blue, red, green). Right panel: In an elongating tissue, a retracting gradient (black) induces a stable gradient with slow decay (gray) to pattern the tissue. **(B)** French Flag model applied to periodic patterning. Alternating concentrations of a morphogen gradient lead to the activation and repression of genes, creating repeating patterns across the tissue. **(C)** GRN realization using activation and repression thresholds. Genes G1-G4 are activated or repressed by different thresholds of the morphogen gradient M (gray), resulting in specific expression domains. **(C')** Simulation of GRN realization shown in (C). **(D)** Gene expression boundaries defined by two morphogen gradients. Gradients M1 (dark gray) and M2 (light gray) overlap to establish precise expression domains for genes G1 (blue) and G2 (red). **(D')** Simulation of GRN realization shown in (D). **(E)** GRN configuration involving cross-regulatory interactions. The morphogen gradient M activates genes G1, G2, and G3 at different thresholds, while cross-repressions between them refine their expression domains. **(E')** Simulation of GRN realization shown in (E). **(F)** The periodic expression of the gene eve in Drosophila is the result of the additive effect of multiple stripe-specific enhancers, each mediating one or two stripes. The logic of each enhancer is mediated by the multiplicative effect of several TF factors. Shown is the regulatory logic of enhancer 3+7 (mediating the expression of eve stripes 3 and 7), and enhancer 4+6 (mediating the expression of eve stripes 4 and 6). Both enhancers mediate the formation of stripes by integrating inputs from gap gene morphogens (mainly *hb* and *kni*). **(G)** A closer look at the regulatory logic of the eve 3+7 and 4+6 enhancers in *Drosophila* given the gene expression patterns of gap genes *hb* and *kni*. Harboring strong binding sites (thick line) for kni (green) but weak bind sites (thin line) for hb (blue), enhancer 3+7 place stripes 3 and 7 away from kni expression and close to abutting hb expressions. Harboring weak binding sites for kni but strong bind sites for hb, enhancer 4+6 place stripes 4 and 6 closer to central kni expression and away from abutting hb expressions. **(H)** *Additive effect of separate enhancers on gene expression.* Gene G is expressed in two domains, each regulated by a distinct enhancer (E1 and E2). The overall expression pattern is the sum of both enhancer activities. The same basic principle can be extended to form periodic pattern of arbitrary repetitions. **(H')** Simulation of GRN realization shown in (H).

Figure 6. Temporal Patterning or the Speed Regulation Model. (A) Schematic of the temporal sequence of cellular states in temporal patterning. Cells progress through successive states (depicted in different colors), with transitions regulated by a speed regulator (gray). **(A')** Gradient-based and wavefront-based speed regulation in tissue patterning. Left: In a non-elongating tissue, a morphogen gradient modulates the speed of transitions between cellular states, resulting in the induction of non-periodic waves that eventually stabilize into a non-periodic pattern. Right: In an elongating tissue, a retracting wavefront regulates the speed of state transitions, resulting in the formation of a non-periodic pattern.

(B) Realization of the Speed Regulation model by modulating transcription and decay rates. The morphogen gradient jointly affects the overall transcription and mRNA degradation rates to control the

timing of gene expression. **(B')** Simulation of the GRN realization in (B). **(C)** Enhancer switching model as a GRN realization of the Speed Regulation model. Each patterning gene is regulated by the additive sum of a dynamic GRN (driving sequential activities) and a static GRN (stabilizing expression) through two separate enhancers. The speed regulator modulates the balance between these GRNs. **(C')** Simulation of the GRN realization in (C).

Figure 7. Case study: Impact of enhancer integration mechanisms on embryonic patterning. (A) *Additive enhancer integration model.* Multiple enhancers independently contribute to the transcription rate of a gene. The total transcription is the weighted sum of the individual enhancer activities. **(A')** Simulation of gene expression with additive enhancer integration under normal conditions. The performance of the enhancer switching model is shown, demonstrating effective temporal patterning. **(A")** Impact of an overactive enhancer in the additive model. The simulation illustrates how an overactive enhancer disrupts gene expression, leading to aberrant patterning and loss of GRN function. **(B)** Enhancer competition integration model. Enhancers compete for binding to the promoter of a gene. The promoter's transcriptional state is determined by the enhancer that successfully binds, normalizing the transcriptional output. **(B')** Simulation of gene expression with enhancer competition integration under normal conditions. The model shows comparable performance to the additive case, maintaining proper temporal patterning. **(B")** Mitigation of overactive enhancer effects in the competition model. The simulation demonstrates that the competitive mechanism normalizes transcription levels, preserving GRN function despite the presence of an overactive enhancer.

References

1. Negrete J, Oates AC. Towards a physical understanding of developmental patterning. Nat Rev Genet. 2021 Aug;22(8):518–31.

2. Wolpert L. Positional information and the spatial pattern of cellular differentiation. J Theor Biol. 1969 Oct;25(1):1–47.

3. Green JBA, Sharpe J. Positional information and reaction-diffusion: two big ideas in developmental biology combine. Development. 2015 Apr 1;142(7):1203–11.

4. Müller P, El-Sherif E. A systems-level view of pattern formation mechanisms in development. Dev Biol. 2020 Apr 1;460(1):1.

5. Spitz F, Furlong EEM. Transcription factors: from enhancer binding to developmental control. Nat Rev Genet. 2012 Sep;13(9):613–26.

6. Anderson PW. More is different. Science. 1972 Aug 4;177(4047):393–6.

7. Davidson EH. Emerging properties of animal gene regulatory networks. Nature. 2010 Dec 16;468(7326):911–20.

8. Levine M, Davidson EH. Gene regulatory networks for development. Proc Natl Acad Sci USA. 2005 Apr 5;102(14):4936–42.

9. Alon U. An Introduction to Systems Biology: Design Principles of Biological Circuits. illustrated ed. CRC Press; 2006.

10. Alon U. Network motifs: theory and experimental approaches. Nat Rev Genet. 2007 Jun;8(6):450–61.

11. Cramer P. Organization and regulation of gene transcription. Nature. 2019 Sep;573(7772):45–54.

12. Furlong EEM, Levine M. Developmental enhancers and chromosome topology. Science. 2018 Sep 28;361(6409):1341–5.

13. Kim YJ, Rhee K, Liu J, Jeammet S, Turner MA, Small SJ, et al. Predictive modeling reveals that higher-order cooperativity drives transcriptional repression in a synthetic developmental enhancer. eLife. 2022 Dec 12;11.

14. Bintu L, Buchler NE, Garcia HG, Gerland U, Hwa T, Kondev J, et al. Transcriptional regulation by the numbers: applications. Curr Opin Genet Dev. 2005 Apr;15(2):125–35.

15. Bintu L, Buchler NE, Garcia HG, Gerland U, Hwa T, Kondev J, et al. Transcriptional regulation by the numbers: models. Curr Opin Genet Dev. 2005 Apr;15(2):116–24.

16. Sherman MS, Cohen BA. Thermodynamic state ensemble models of cis-regulation. PLoS Comput Biol. 2012 Mar 29;8(3):e1002407.

17. Elowitz MB, Leibler S. A synthetic oscillatory network of transcriptional regulators. Nature. 2000 Jan 20;403(6767):335–8.

18. Diaz-Cuadros M, Pourquié O, El-Sherif E. Patterning with clocks and genetic cascades: Segmentation and regionalization of vertebrate versus insect body plans. PLoS Genet. 2021 Oct 14;17(10):e1009812.

19. Zhu X, Rudolf H, Healey L, François P, Brown SJ, Klingler M, et al. Speed regulation of genetic cascades allows for evolvability in the body plan specification of insects. Proc Natl Acad Sci USA. 2017 Sep 25;114(41).

20. Balaskas N, Ribeiro A, Panovska J, Dessaud E, Sasai N, Page KM, et al. Gene regulatory logic for reading the Sonic Hedgehog signaling gradient in the vertebrate neural tube. Cell. 2012 Jan 20;148(1–2):273–84.

21. Perez-Carrasco R, Barnes CP, Schaerli Y, Isalan M, Briscoe J, Page KM. Combining a Toggle Switch and a Repressilator within the AC-DC Circuit Generates Distinct Dynamical Behaviors. Cell Syst. 2018 Apr 25;6(4):521-530.e3.

22. Panovska-Griffiths J, Page KM, Briscoe J. A gene regulatory motif that generates oscillatory or multiway switch outputs. J R Soc Interface. 2013 Feb;10(79):20120826.

23. Goto T, Macdonald P, Maniatis T. Early and late periodic patterns of even skipped expression are controlled by distinct regulatory elements that respond to different spatial cues. Cell. 1989 May 5;57(3):413–22.

24. Schroeder MD, Greer C, Gaul U. How to make stripes: deciphering the transition from non-periodic to periodic patterns in Drosophila segmentation. Development. 2011 Jul;138(14):3067–78.

25. Clyde DE, Corado MSG, Wu X, Paré A, Papatsenko D, Small S. A self-organizing system of repressor gradients establishes segmental complexity in Drosophila. Nature. 2003 Dec 18;426(6968):849–53.

26. Durston A, Wacker S, Bardine N, Jansen H. Time space translation: a hox mechanism for vertebrate a-p patterning. Curr Genomics. 2012 Jun;13(4):300–7.

27. Palmeirim I, Henrique D, Ish-Horowicz D, Pourquié O. Avian hairy gene expression identifies a molecular clock linked to vertebrate segmentation and somitogenesis. Cell. 1997 Nov 28;91(5):639–48.

28. Kohwi M, Doe CQ. Temporal fate specification and neural progenitor competence during development. Nat Rev Neurosci. 2013 Nov 20;14(12):823–38.

29. Naidu VG, Zhang Y, Lowe S, Ray A, Zhu H, Li X. Temporal progression of Drosophila medulla neuroblasts generates the transcription factor combination to control T1 neuron morphogenesis. Dev Biol. 2020 May 19;

30. El-Sherif E, Averof M, Brown SJ. A segmentation clock operating in blastoderm and germband stages of Tribolium development. Development. 2012 Dec 1;139(23):4341–6.

31. Sarrazin AF, Peel AD, Averof M. A segmentation clock with two-segment periodicity in insects. Science. 2012 Apr 20;336(6079):338–41.

32. Oates AC, Morelli LG, Ares S. Patterning embryos with oscillations: structure, function and dynamics of the vertebrate segmentation clock. Development. 2012 Feb;139(4):625–39.

33. Dessaud E, Yang LL, Hill K, Cox B, Ulloa F, Ribeiro A, et al. Interpretation of the sonic hedgehog morphogen gradient by a temporal adaptation mechanism. Nature. 2007 Nov 29;450(7170):717–20.

34. BECK MT, VÁRADI ZB. One, Two and Three-dimensional Spatially Periodic Chemical Reactions. Nature. 1972 Jan 3;

35. Rohde LA, Bercowsky-Rama A, Valentin G, Naganathan SR, Desai RA, Strnad P, et al. Cell-autonomous timing drives the vertebrate segmentation clock's wave pattern. 2024 Feb 13;

36. Bothma JP, Garcia HG, Ng S, Perry MW, Gregor T, Levine M. Enhancer additivity and non-additivity are determined by enhancer strength in the Drosophila embryo. eLife. 2015 Aug 12;4.

Figure 1

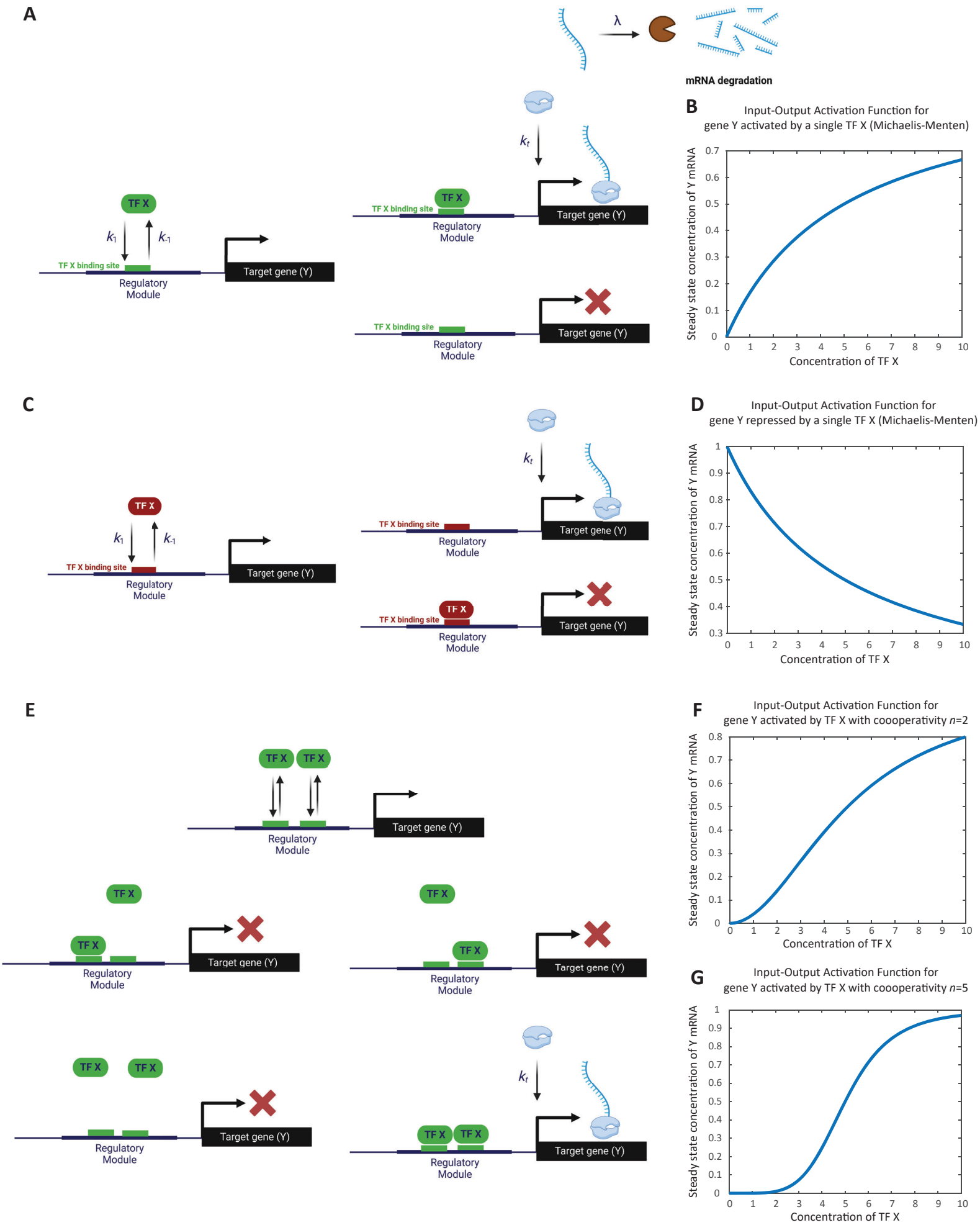

A
λ
mRNA degradation
TF X
k_1
k_{-1}
TF X binding site
Regulatory Module
Target gene (Y)
k_t
B
Input-Output Activation Function for gene Y activated by a single TF X (Michaelis-Menten)
Steady state concentration of Y mRNA
Concentration of TF X
C
D
Input-Output Activation Function for gene Y repressed by a single TF X (Michaelis-Menten)
E
F
Input-Output Activation Function for gene Y activated by TF X with coooperativity n=2
G
Input-Output Activation Function for gene Y activated by TF X with coooperativity n=5

Figure 2

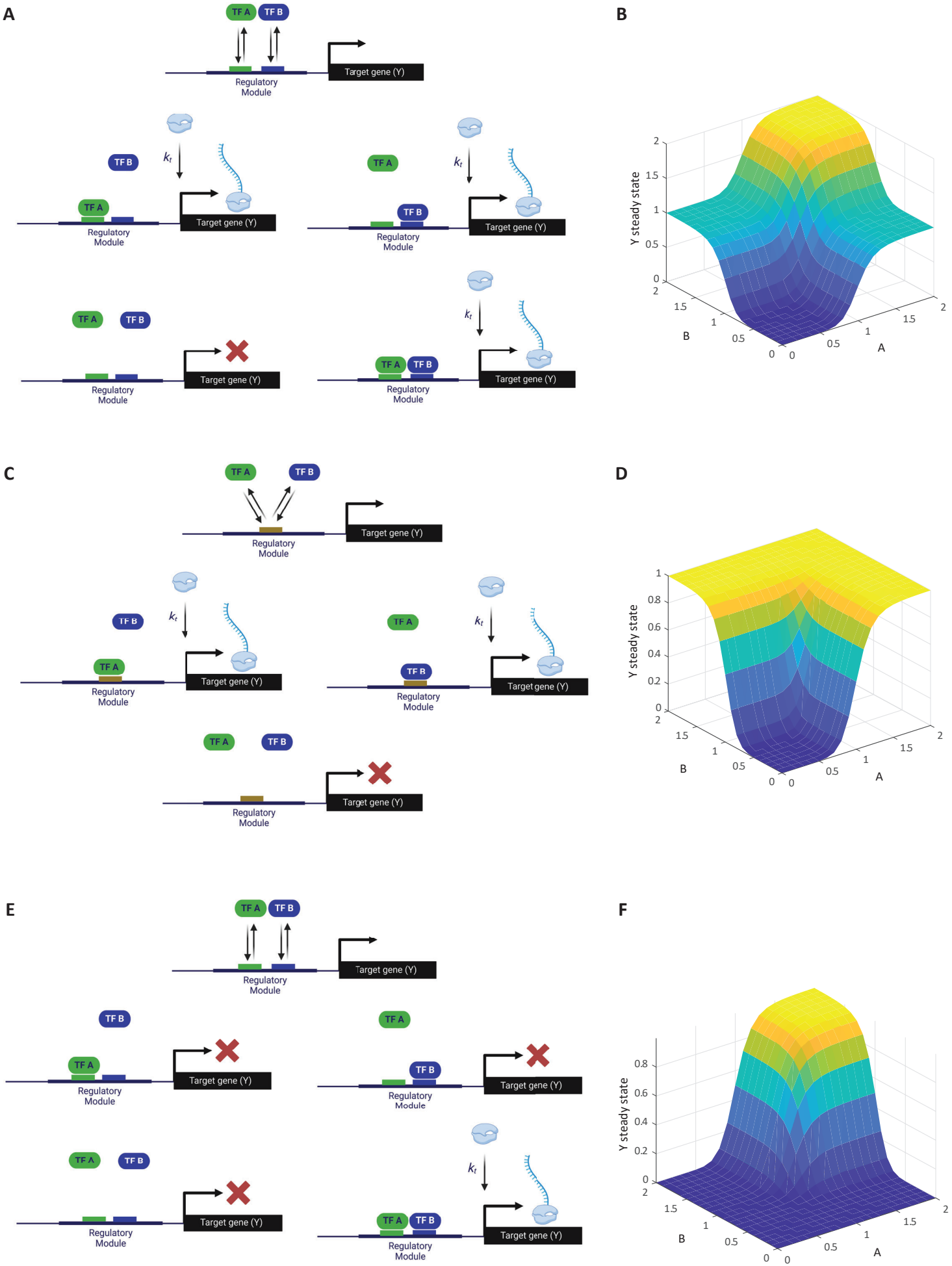

A
TF A
TF B
Regulatory Module
Target gene (Y)
k_t
B
Y steady state
B
A
C
D
E
F

Figure 3

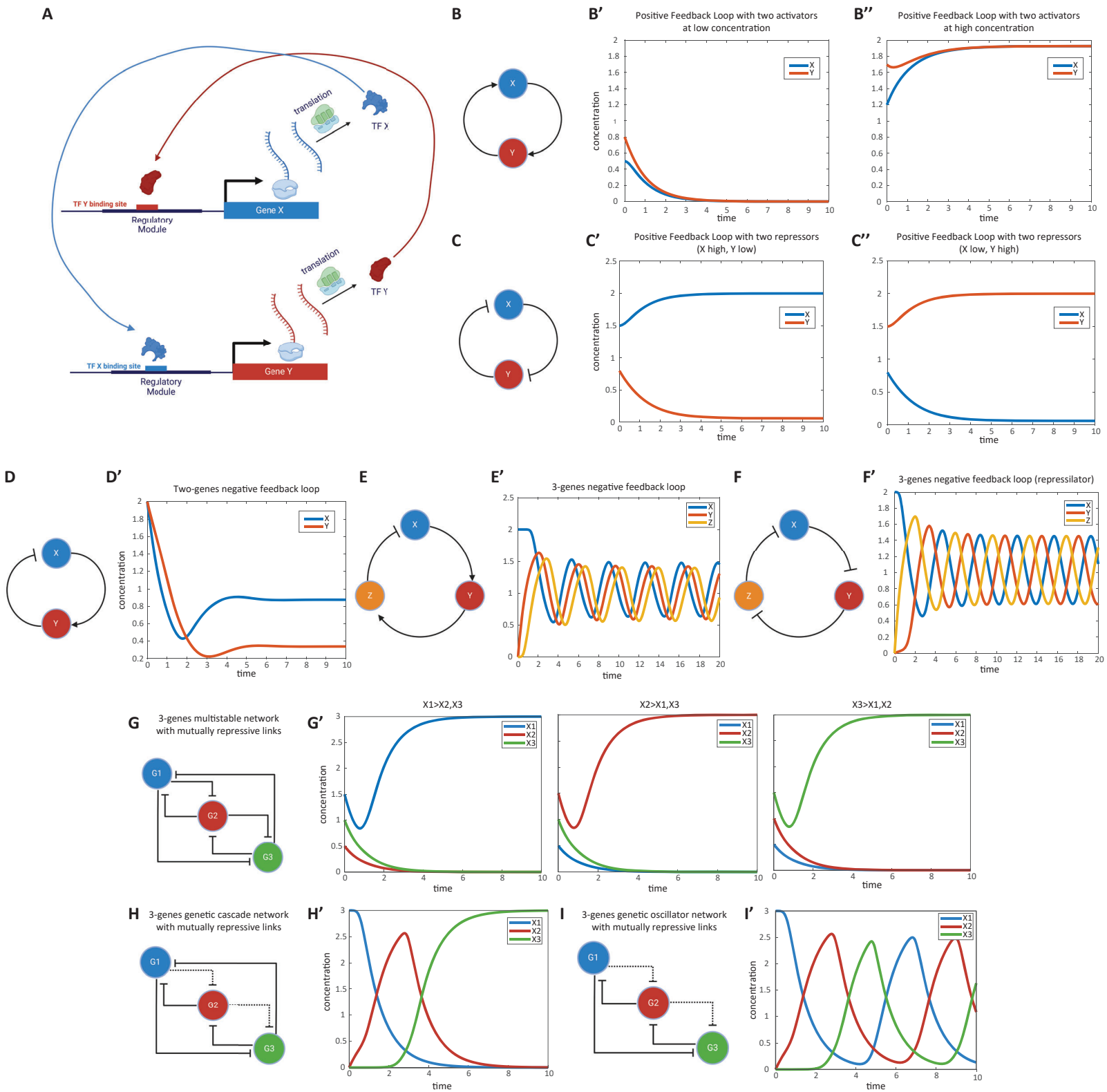

A
translation
TF X
TF Y binding site
Regulatory Module
Gene X
translation
TF Y
TF X binding site
Regulatory Module
Gene Y
B
X
Y
B'
Positive Feedback Loop with two activators at low concentration
B''
Positive Feedback Loop with two activators at high concentration
C
C'
Positive Feedback Loop with two repressors (X high, Y low)
C''
Positive Feedback Loop with two repressors (X low, Y high)
concentration
time
D
D'
Two-genes negative feedback loop
E
E'
3-genes negative feedback loop
F
F'
3-genes negative feedback loop (repressilator)
Z
G
3-genes multistable network with mutually repressive links
G1
G2
G3
G'
X1>X2,X3
X2>X1,X3
X3>X1,X2
X1
X2
X3
H
3-genes genetic cascade network with mutually repressive links
H'
I
3-genes genetic oscillator network with mutually repressive links
I'

Figure 4

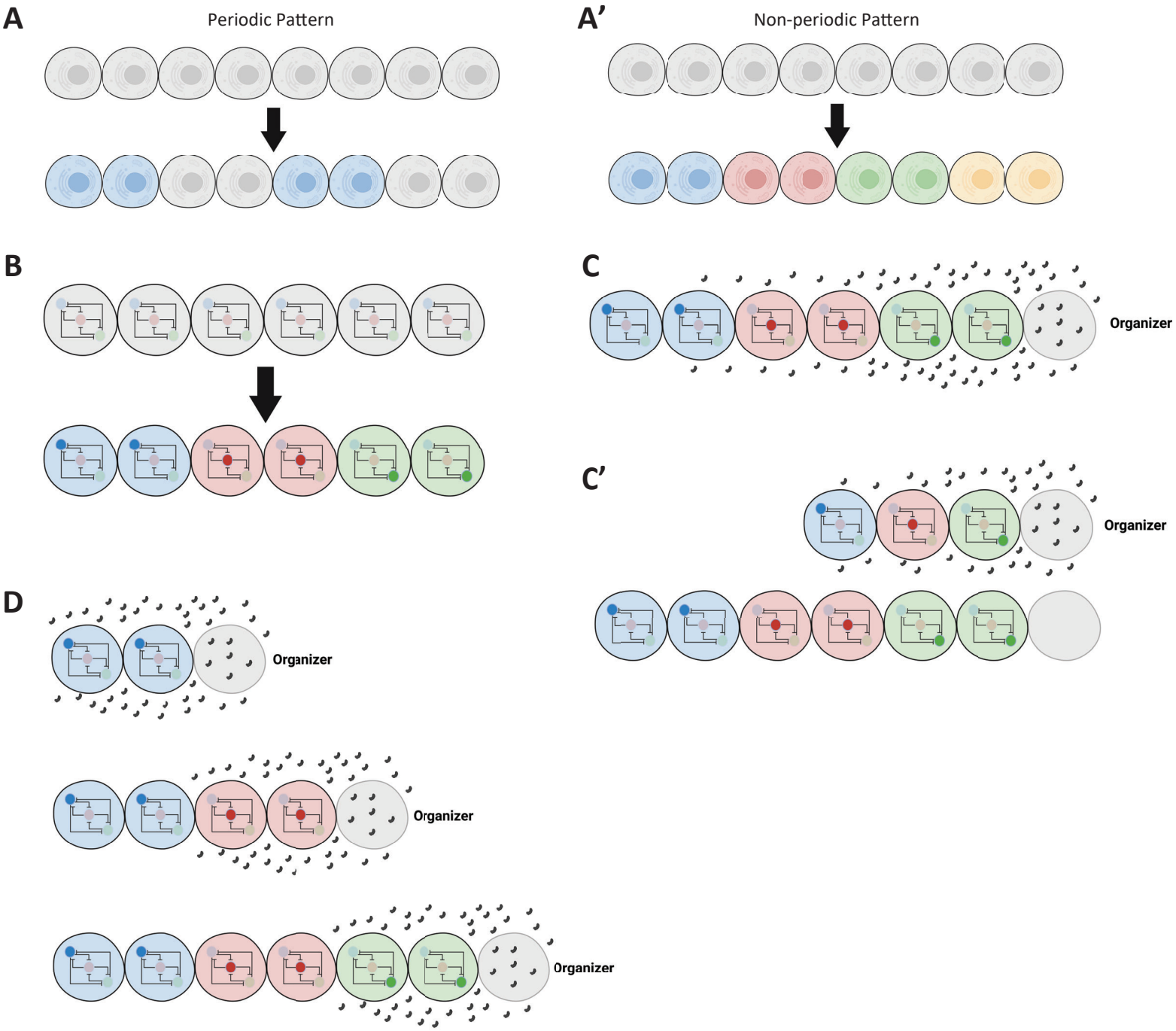

A
Periodic Pattern
A'
Non-periodic Pattern
B
C
Organizer
C'
Organizer
D
Organizer
Organizer
Organizer

Figure 5

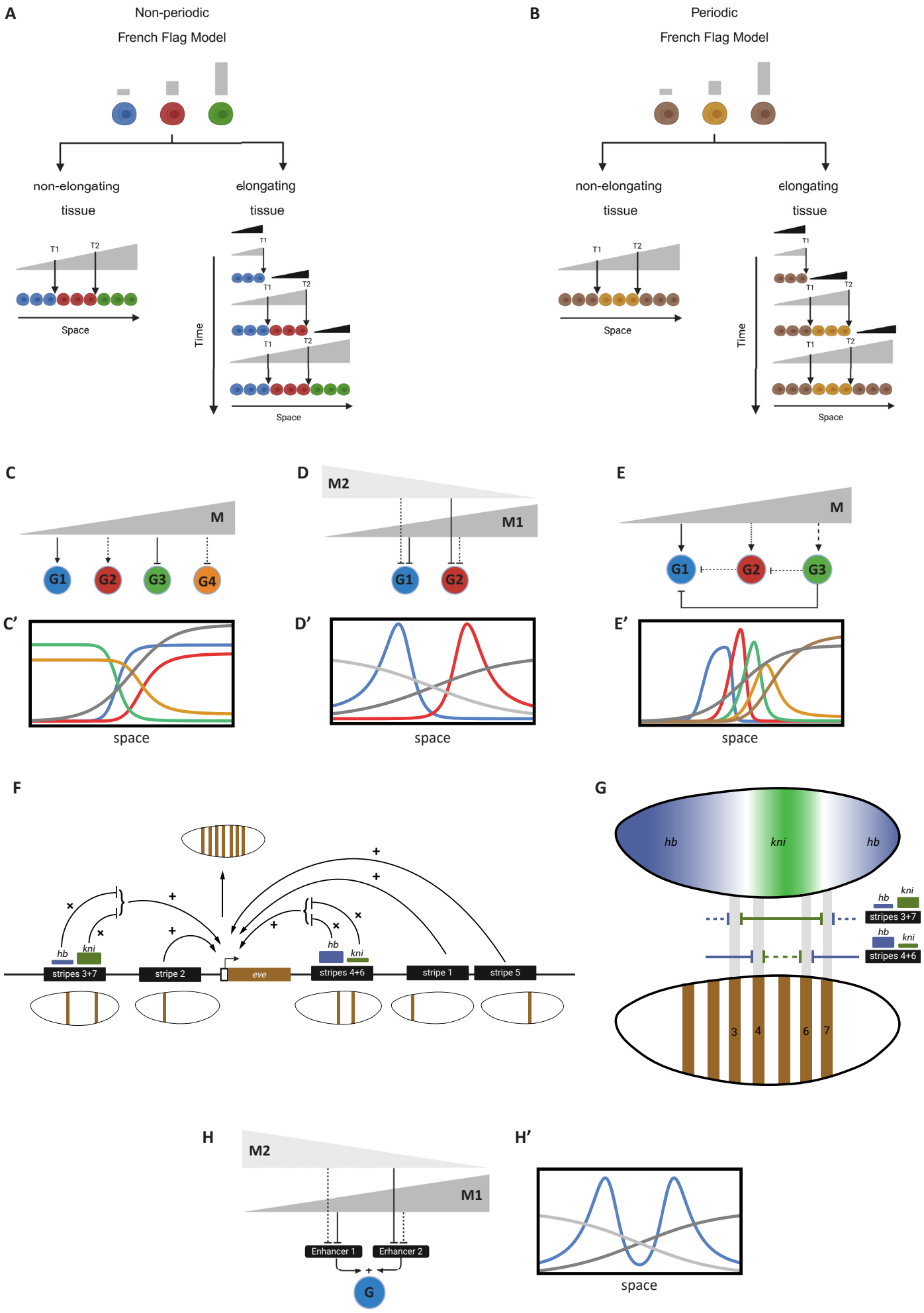

A
Non-periodic
French Flag Model
non-elongating
tissue
elongating
tissue
T1
T2
Space
Time
B
Periodic
French Flag Model
C
M
G1
G2
G3
G4
C'
space
D
M2
M1
D'
E
E'
F
hb
kni
stripes 3+7
stripe 2
eve
stripes 4+6
stripe 1
stripe 5
G
3
4
6
7
H
Enhancer 1
Erhancer 2
H'

Figure 6

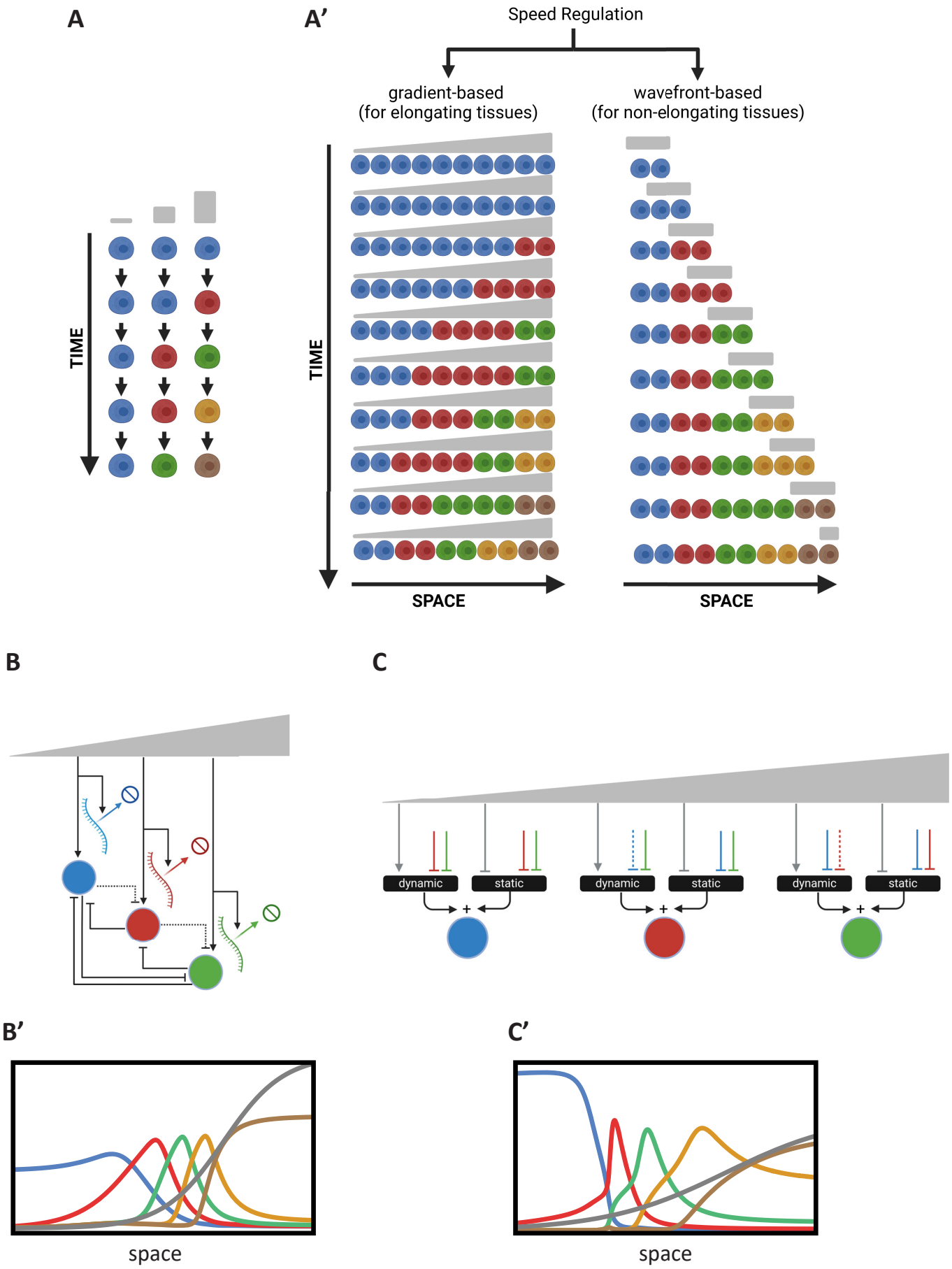
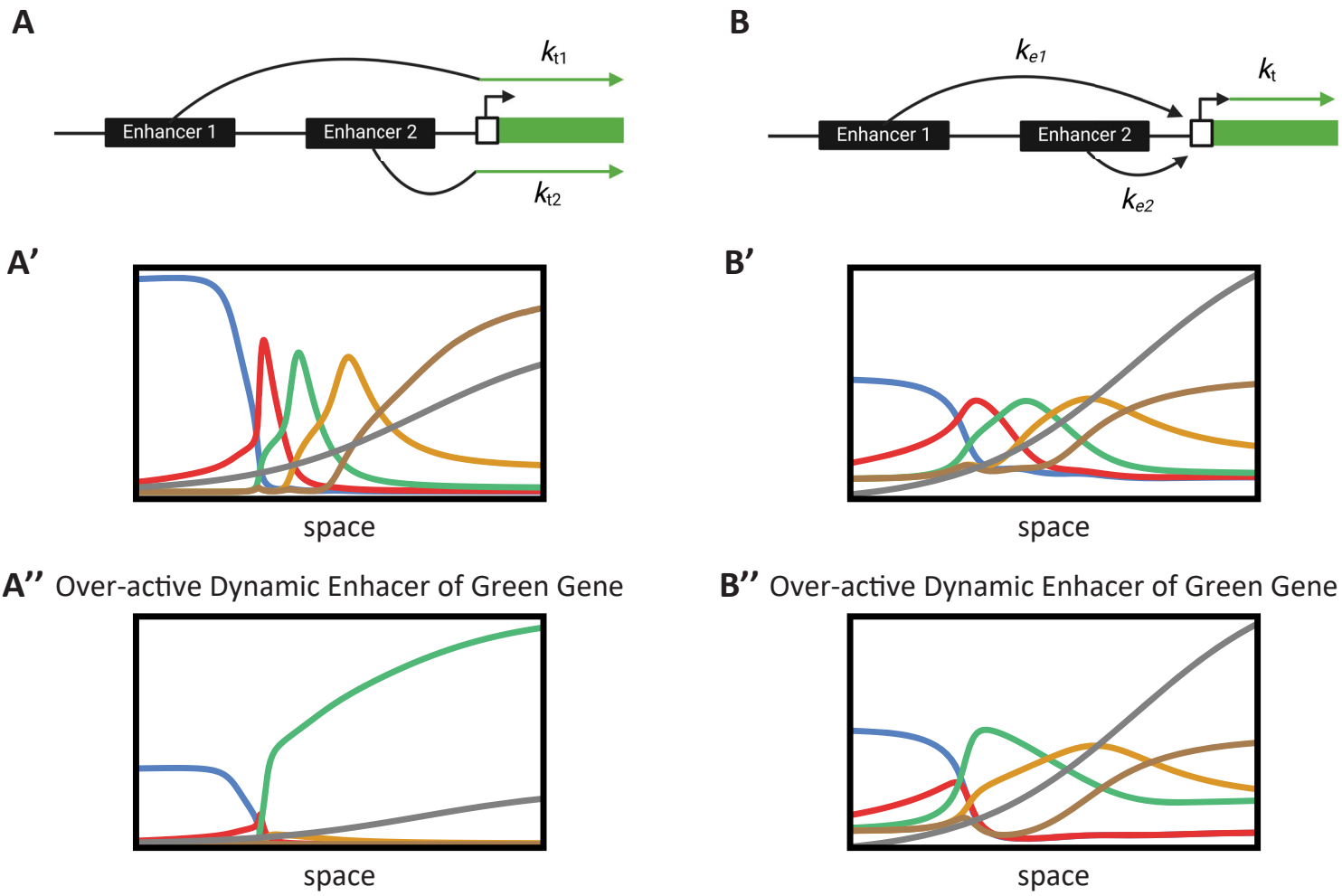

A
A'
Speed Regulation
gradient-based
(for elongating tissues)
wavefront-based
(for non-elongating tissues)
TIME
TIME
SPACE
SPACE
B
C
dynamic
static
dynamic
static
dynamic
static
B'
C'
space
space

Figure 7


A
k_{t1}
Enhancer 1
Enhancer 2
k_{t2}
B
k_{e1}
k_t
Enhancer 1
Enhancer 2
k_{e2}
A'
space
B'
space
A'' Over-active Dynamic Enhacer of Green Gene
space
B'' Over-active Dynamic Enhacer of Green Gene
space