

Gene Network Organization, Mutation, and Selection Collectively Drive Developmental Pattern Evolvability and Predictability

Harry Booth ¹ and Z. Hadjivasiliou ^{1,2,3,*}

¹*The Francis Crick Institute, Mathematical and Physical Biology Laboratory, 1 Midland Road, London NW1 1AT, United Kingdom*

²*University College London, Department of Physics and Astronomy, Gower Street, London WC1E 6BT, United Kingdom*

³*London Centre for Nanotechnology, 19 Gordon Street, London WC1H 0AH, United Kingdom*



(Received 16 November 2024; accepted 11 August 2025; published 16 September 2025; corrected 15 December 2025)

A hallmark of development is the generation of spatial patterns driven by morphogen gradients and gene regulatory networks (GRNs). Although the mechanistic basis by which GRNs orchestrate cellular responses and tissue patterning during development is well understood, their evolutionary dynamics remain less clear. Still, mutations in regulatory elements that govern GRN-driven patterning are a key mechanism for patterning evolution. In this study, we use the *de novo* evolution of a stripe phenotype as a model framework to investigate the evolutionary dynamics associated with the emergence of new spatial gene expression boundaries and the adjustment of existing boundaries. To probe general principles of GRN-driven pattern evolution we introduce a new high-throughput theoretical framework that rapidly produces a comprehensive dataset of evolutionary trajectories. We leverage this large dataset to investigate the types of mutations that drive different phenotypic shifts in spatial patterning. Our findings suggest that the order in which mutations in gene-gene interactions appear and the specific combination of gene-gene interactions that mutate together determine the evolvability of novel spatial gene expression patterns. We interpret our results in the context of epistatic effects that naturally arise in networks of interconnected genes and show how contingencies and constraints emerge in our system. Our results elucidate the interplay between mutation and gene network organization, revealing how historical contingencies arise and impact the evolvability of GRNs and the predictability of their evolutionary outcomes.

DOI: [10.1103/q79n-k93j](https://doi.org/10.1103/q79n-k93j)

I. INTRODUCTION

A central challenge in evolutionary developmental biology is to elucidate how phenotypic variation arises from underlying developmental mechanisms. At the core of development lies the spatial patterning of tissues into different cell types, driven by gene regulatory networks (GRNs) that interpret morphogen gradients [1]. While the mechanistic basis of how GRNs influence spatial patterning during development is well established, the interplay of these networks with evolutionary forces such as mutation, selection, and historical contingency remains poorly understood. Yet several studies have revealed the critical role of GRNs in generating morphological variation both within and across species [2–6], underscoring the importance of investigating their evolutionary dynamics in this developmental context.

Understanding how mutations at the sequence level impact GRN function and their evolutionary dynamics is a key milestone for the study of GRN-driven pattern evolution. The current consensus is that GRN evolution advances through mutations not in the regulatory genes or transcription factors

(TFs) themselves, but in *cis*-regulatory elements (CREs) or enhancers that drive TF expression levels [6–8]. In this way pleiotropic effects associated with TF mutations are bypassed, while mutations in regulatory regions are able to generate variation for selection to act upon. Therefore, the ways in which mutations in CREs reorganize gene interactions plays a central role in understanding the evolution of GRNs and the spatial patterns they coordinate during development.

Recent work has investigated the role of point mutations in native CREs and the insertion of *de novo* CREs into a GRN controlling the patterning of fruit fly embryos. This revealed that the impact of mutations on the spatial pattern depends on whether mutations affect existing CREs or introduce new regulatory elements in the system [9]. New assays that allow targeted genetic manipulation of GRNs *in vivo* can therefore generate large datasets of mutations which link changes in regulatory and TF sequences to phenotypic shifts [10,11]. Furthermore, synthetic patterning networks offer promising tools for investigating the mechanistic role of genes by highlighting how the regulatory interactions give rise to novel patterns and facilitate evolutionary exploration [12,13]. However, it is currently unfeasible to characterize the GRN reorganization which may result from high-throughput experimental mutation assays at a scale large enough to extract general principles of network rewiring in these systems. Furthermore, the difficulty of imposing a selection pressure for novel patterns in multicellular development means that the role of GRN rewiring within an evolutionary process is challenging to

*Contact author: zena.hadjivasiliou@ucl.ac.uk

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

analyze. Hence, although new experimental tools promise to advance our understanding of the evolution of development, the complexity of the genotype-phenotype map specified by GRN-driven patterning coupled with the timescales over which evolutionary change occurs, present challenges that experimental tools alone may struggle to overcome.

A related question is the relative role of stochasticity and constraints in determining evolutionary outcomes. Historical contingency during evolution describes a sensitivity to past events that are often inconsequential at the time they occur but that nonetheless impact subsequent evolution. This sensitivity becomes heightened when adaptation takes place on “rugged” fitness landscapes [14], a property which emerges through the complex, nonlinear nature of many biological genotype-phenotype mappings [15]. While historical contingency has been investigated using evolutionary simulations in protein evolution [16], and in GRNs with simple evolved functions [17], these investigations have yet to be extended to genotype-phenotype maps defined by multicellular development processes such as patterning. For example, although it is known that spatial pattern accessibility is dependent on GRN topology [13,18], we do not yet understand when constraints imposed by existing network topology might become influential in the evolution of new developmental patterns or how these constraints interact with mutation and the strength of selection. For instance, neutral evolution—whereby mutations with little or no impact on the phenotype accumulate—has been shown to increase the accessibility of novel spatial patterns [12]. Still, the ways in which neutral mutations impact the diversification of GRN topology and spatial patterns when placed in an evolutionary process is not yet understood.

Theoretical models of GRN-driven patterning capture the spatiotemporal patterning dynamics observed in various model systems, from the *Drosophila* embryo [19–21] to the vertebrate neural tube [22,23]. While these models do not explicitly track genetic information at a molecular scale, they have provided crucial mechanistic insights into the fundamental characteristics and emergent properties of GRN patterning [24–26] and offer a promising and tractable framework for exploring GRN-driven pattern evolution. Previous work integrating mechanistic GRN patterning models into evolutionary simulations has yielded valuable insights into the constraints and design principles that optimize specific phenotypic outcomes or “fitness” [27–30]. For example, simulations have identified parsimonious paths in GRN space linking Dipteran species and their homologous relationships [31], elucidated the design principles underlying body plan segmentation in flies [32], and explored potential routes toward the *de novo* evolution of body segmentation [33,34]. This mechanism-centered approach highlights the central role of gene interactions and network topology—rather than individual genes—as key drivers of evolutionary processes. A common limitation of these approaches to date has been the efficiency at which they can generate evolutionary trajectories. As a result, previous studies have been limited to datasets of evolutionary repeats that were not large enough to investigate general principles of mutation and selection in a developmental context.

In this work we explore the interplay among mutation, gene network organization, and natural selection in shaping

the evolution of GRNs and developmental patterns. We use stripe-forming networks as a framework to investigate how new gene expression boundaries are formed and how existing gene expression boundaries diversify. We develop a high-throughput computational framework that integrates a mechanistic description of GRN patterning with an evolutionary process. By leveraging the framework’s capacity to rapidly evolve target phenotypes, we produce an extensive dataset of independent evolutionary trajectories. This allows us to employ statistical tools to investigate general principles of GRN-driven pattern evolution. The formation and adjustment of gene expression boundaries is a hallmark of developmental patterning and its evolution [1,35] and so this approach can yield general insights about the evolution of development. Our framework is easily adaptable and offers a platform to study developmental pattern evolution, which has the potential to link physical constraints that emerge from genetic architecture to evolutionary change.

II. METHODS

High-throughput framework for GRN evolution

We model a minimal GRN downstream of a morphogen and introduce mutation of the regulatory interactions between genes [Fig. 1(a)]. Our theoretical framework includes three key components (Fig. 1): (i) GRN-mediated spatial patterning, (ii) mutation and fitness evaluation, and (iii) a Monte Carlo scheme which simulates an evolutionary trajectory.

We consider a one-dimensional tissue composed of a discrete number of cells, over which a morphogen gradient is established Fig. 1(a). Each cell has an identical GRN consisting of three genes $\{a, b, c\}$, with gene a activated by the morphogen, m Figs. 1(a) and 1(b). The network is specified by a weight matrix W whose entries w_{ij} encode the pairwise regulatory interactions between genes Fig. 1(b). We then assume that the concentration dynamics of gene product j in cell $i \in \{1, \dots, N_{\text{cells}}\}$ satisfy

$$\frac{dg_j^{(i)}}{dt} = \sigma \left(\sum_{k \in \{a,b,c\}} w_{kj} g_k^{(i)} + w_{mj} m^{(i)} \right) - \gamma_g g_j^{(i)}, \quad (1)$$

$$g_j^{(i)}(0) = g_0 \quad \forall j \in \{a, b, c\}, \quad (2)$$

where $\sigma(I) = \frac{1}{1 + \exp(-\theta I)}$ is a sigmoidal activation function with corresponding steepness parameter $\theta > 0$ and γ_g is the gene product degradation rate. This description of gene regulation is widely used in both theoretical and experimental studies of spatial patterning in biological systems [21,37,38].

The evolutionary starting point in our simulations is given by an initial weight matrix, W_0 , with a minimal number of nonzero interactions [Fig. 1(c)], representing an initial state for network evolution. Subsequent mutation adds new regulatory interactions between genes, or adjusts the magnitude of existing ones. Thus, the weight matrix W represents the genotype in our model, and mutation acts directly on its entries to produce a mutant GRN.

Previous work has suggested that multiplicative weight changes most accurately represent the effect of point mutations on binding affinities [39]. We model this as $w_{ij}^{n+1} = w_{ij}^n \exp(k)$, where k is normally distributed with $k \sim$

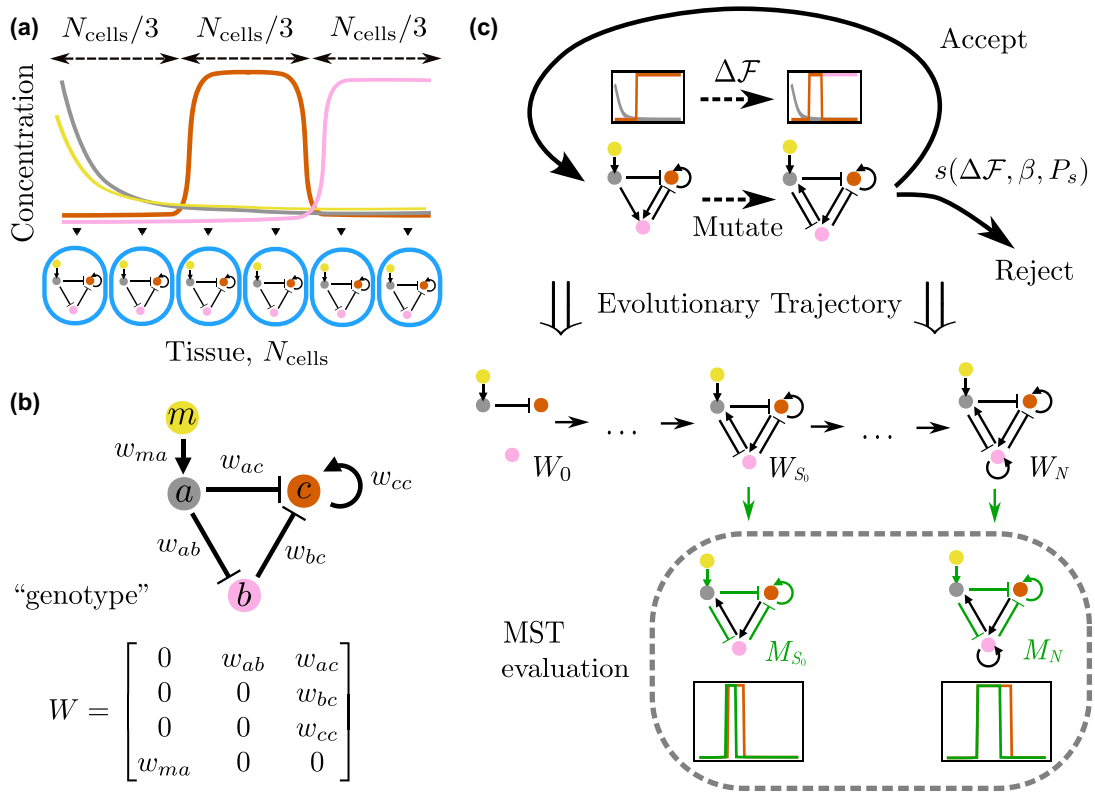


FIG. 1. Method summary. (a) Model of developmental patterning and target stripe. A tissue is made up of a discrete number of cells (blue ovals), each with an identical three-node gene regulatory network (GRN). A morphogen gradient (yellow) is assumed to be at steady state across the tissue. The target stripe profile (orange) has boundaries which define three equal size regions across the tissue. The gray and pink circles represent the two other genes in the network. Lines in the plot are a schematic of the putative concentration at steady state for the morphogen (yellow) and three genes shown in the diagram below (gray, orange, and pink). (b) An example GRN. The morphogen, m (yellow) activates gene a (gray), which then regulates the output gene c (orange) directly and through an intermediary gene b (pink). The regulatory organization is specified by a matrix W that holds the weights that define an effective strength for activation (positive weights) and inhibition (negative weights) arrows and represents the genotype in our analysis. (c) A schematic outlining the Monte Carlo scheme for simulating evolutionary trajectories. The resident genotype associated with the current phenotype is mutated by introducing or editing existing interactions (weight changes). The new phenotype defines a fitness differential $\Delta\mathcal{F}$, which then determines the probability $s(\Delta\mathcal{F}, \beta, P_s)$ of fixation for the mutant based on the selection strength β and population size P_s . This process then repeats until a predefined fitness threshold is reached. The sequence of accepted mutants is collected to form an “evolutionary trajectory.” The *minimal stripe topology* (MST) associated with the first stripe forming mutant W_{S_0} and the final mutant W_N are then evaluated. The MST is the topology associated with the minimal set of regulatory interactions able to maintain a stripe phenotype. Details of parameters used in the evolutionary algorithm are provided in Ref. [36].

$N(\mu_M, \sigma_M^2)$ and w_{ij}^n and w_{ij}^{n+1} represent the regulatory effect of gene i on j before and after mutation, respectively. Since point mutations typically decrease affinity [39,40], we set $\mu_M < 0$. While point mutations can generate *de novo* regulatory interactions [41], GRN evolution is thought to occur more efficiently through alternative mechanisms such as duplication [42,43] or co-option [6,8]. Therefore, to enable mutations of larger effects we also consider additive weight changes such that $w_{ij}^{n+1} = w_{ij}^n + k$, where k is normally distributed with $k \sim N(0, \sigma_A^2)$. This allows the distribution of weight change magnitudes $|w_{ij}^{n+1} - w_{ij}^n|$ to be independent of the initial magnitude $|w_{ij}^n|$. Additive mutations enable the introduction of novel interactions (i.e., $w_{ij}^n = 0 \rightarrow w_{ij}^{n+1} \gg 0$). Furthermore, additive mutations enable the sign of a regulatory interaction to change from activating to inhibiting or vice versa. Finally, to investigate the impact of mutations that change a single versus multiple gene interactions in the network, we assume

that a single mutant may carry multiple weight changes simultaneously. We implement this by allowing each weight in W to change independently with probability p . This generates a distribution of weight changes per mutant dictated by p (Ref. [36], Sec. 1.2.1).

Note that p could be interpreted to reflect constraints which emerge from the architecture of CREs and TF interactions and is hence separate from the mutation rate. For instance, multiple weight changes could occur simultaneously when single mutations on the same CRE affect the binding affinity of multiple TFs, or when TFs themselves mutate. In this way, our framework is set up to investigate how constraints emerging from genetic architecture may impact evolutionary trajectories and evolvability.

We define a target stripe by a specific width and position Fig. 1(a), and compute the fitness differential $\Delta\mathcal{F}$ of mutants relative to the resident phenotype Fig. 1(c). The phenotype ϕ upon which selection acts is taken to be the steady-state

expression of gene c Fig. 1(a). Fitness is evaluated in a two-stage process employing two distinct fitness functions, $\mathcal{F}_S$ and $\mathcal{F}_R$. $\mathcal{F}_S$ is an indicator function which recognizes the transition to a stripe by evaluating phenotypes against a set of general stripe criteria (see Ref. [36], Sec. 1.3). These conditions ensure a stripelike pattern, but they do not enforce any particular boundary locations. $\mathcal{F}_R$ measures the average distance of the steady-state expression of gene c to a target stripe pattern with specific boundaries. The target pattern is a stripe that splits the tissue into three equally sized parts Fig. 1(a). Selection for this phenotype can be imposed through the following fitness functional:

$$\mathcal{F}_R(\phi) = \frac{3\gamma_g}{L} \left(\int_{L/3}^{2L/3} \phi(x) dx - \frac{1}{2} \left(\int_0^{L/3} \phi(x) dx + \int_{2L/3}^L \phi(x) dx \right) \right), \quad (3)$$

where L is the tissue length and γ_g is the gene product degradation rate. The first integral rewards high expression of the gene product in the central region of the tissue. This is then penalized by the average expression of the gene product in the neighboring regions. Finally, this is normalized into the range $[-1, 1]$ by dividing by the upper bound on the difference between the first and second terms. The fitness differential $\Delta\mathcal{F}$ between the resident phenotype ϕ and mutant phenotype ϕ' is then evaluated as $\mathcal{F}_S(\phi') - \mathcal{F}_S(\phi)$ if $\mathcal{F}_S(\phi') \neq \mathcal{F}_S(\phi)$ and $\mathcal{F}_R(\phi') - \mathcal{F}_R(\phi)$ otherwise. This choice is motivated by the fact a phenotype may display a stripelike pattern while not necessarily having a close match in expression space to a stripe with specific boundaries. In these instances we assume its fitness to be higher than a phenotype with no stripelike characteristics. This combination of selection functions is designed to generate conditions that prioritizes the generation of new gene expression boundaries before the adjustment of existing boundary positions to a specific target pattern. As a consequence, comparison of expression profiles to a target pattern only takes place between phenotypes with the same number of gene expression boundaries. While the conditions that favor the evolution of spatial patterning are in general unknown, our choice of fitness function captures the advantage of increasing the number of differentiated cell types [44].

We model monomorphic evolutionary dynamics with a sequential fixation model, by assuming that the rate at which mutants appear is lower than the time it takes for them to fix [45–47]. The Monte Carlo scheme for an evolutionary trajectory is defined by the probability that a mutant fixes in a population of size P_s :

$$s = \frac{1 - \exp(-2\beta\Delta\mathcal{F})}{1 - \exp(-2P_s\beta\Delta\mathcal{F})}, \quad (4)$$

where β represents the strength of selection [48]. Refer to Sec. 1 in Ref. [36] for further details on the dynamical equations, evolutionary algorithm and parameters used.

III. RESULTS

A. Multiple GRN topologies evolve following selection for stripe phenotype

We use the algorithm defined in the previous section to generate $\sim 10^5$ independent evolutionary trajectories from the

initial conditions to the target phenotype. We found that 88.2% of our simulations converged to the target phenotype within 250 000 generations (Sec. 4 in Ref. [36]). We recorded the evolutionary trajectory of accepted genotypes, $(W_n^{(i)})_{n=0}^{N_i}$, where N_i is the total number of accepted mutants leading to the target phenotype for the i th trajectory Fig. 1(c). We defined four key stages across simulations; the first accepted mutant ($n = 1$), a period of prestripe formation for $n \in (1, S_0)$, stripe inception at $n = S_0$ and poststripe formation for $n \in (S_0, N_i]$. Therefore, S_0 is the mutant number for which a stripe first manifests Fig. 2(a).

To classify the mechanistic outcomes of individual trajectories, we identified the minimal subset of regulatory interactions that maintain a basic stripe phenotype, which we term the *minimal stripe topology* (MST). We determine the MST for the evolved networks at two key time points during evolution: the point of stripe formation, $n = S_0$, and the point at which the target stripe is achieved, $n = N_i$, and denote the MSTs at these two time points by $M_{S_0}^{(i)}$ and $M_{N_i}^{(i)}$, respectively (Fig. 1(c) and Sec. 1.6 in Ref. [36]). Using this information we infer a distribution of unique evolutionary outcomes relating to the core mechanisms of stripe formation Fig. 2(b). We identified a total of 732 unique MSTs, with the eight most frequent (MST:1-8) accounting for $\sim 75\%$ of outcomes.

MST:{1, 2, 5, 6, 7} feature the well-studied feed-forward incoherent type-2 (I2-FFL) mechanism, found in the gap gene network responsible for embryo segmentation in fruit flies [38,50]. MST:4 is the I3-FFL mechanism [51], which has not been observed in natural stripe-forming regulatory networks but has been investigated synthetically [13]. MST:{3, 8} feature a negative feedback loop with gene b serving as a controller to determine different responses across inputs, a mechanism which has been shown to enable perfect biochemical adaption [52]. Therefore, the dominant MSTs that have evolved in our simulations are linked to networks that have been identified in developmental patterning contexts. The remaining MST outcomes occurred with a mean frequency of less than 0.05%. We discuss and interpret these rare MSTs and their relationship to MST:1-8 in the supplemental material (Sec. 2 in Ref. [36]).

To illustrate that our MST classification captures essential differences in transcriptional dynamics, we solved each of the final networks to steady state and collected concentration profiles at intermediary times [38] (see Sec. 1.7 in Ref. [36]). These spatial snapshots of gene expression levels as steady state is approached generate a high-dimensional “fingerprint” for each evolved network and its dynamics (for an example see Movie 1 [36]). A two-dimensional projection of these dynamical profiles indicates that networks are in general grouped dynamically according to their MST identity Fig. 2(c). This suggests a correspondence between the essential gene interactions for basic stripe formation and the temporal dynamics of gene expression, driven by the complete and (often) more complex network.

Overall, the range of stripe-forming mechanisms identified in our evolution experiments highlights the many-to-one aspect of the genotype-to-phenotype map for stripe-forming mechanisms as determined before [38]. A question that follows is what determines evolutionary paths towards a specific mechanism. In what follows we investigate how the muta-

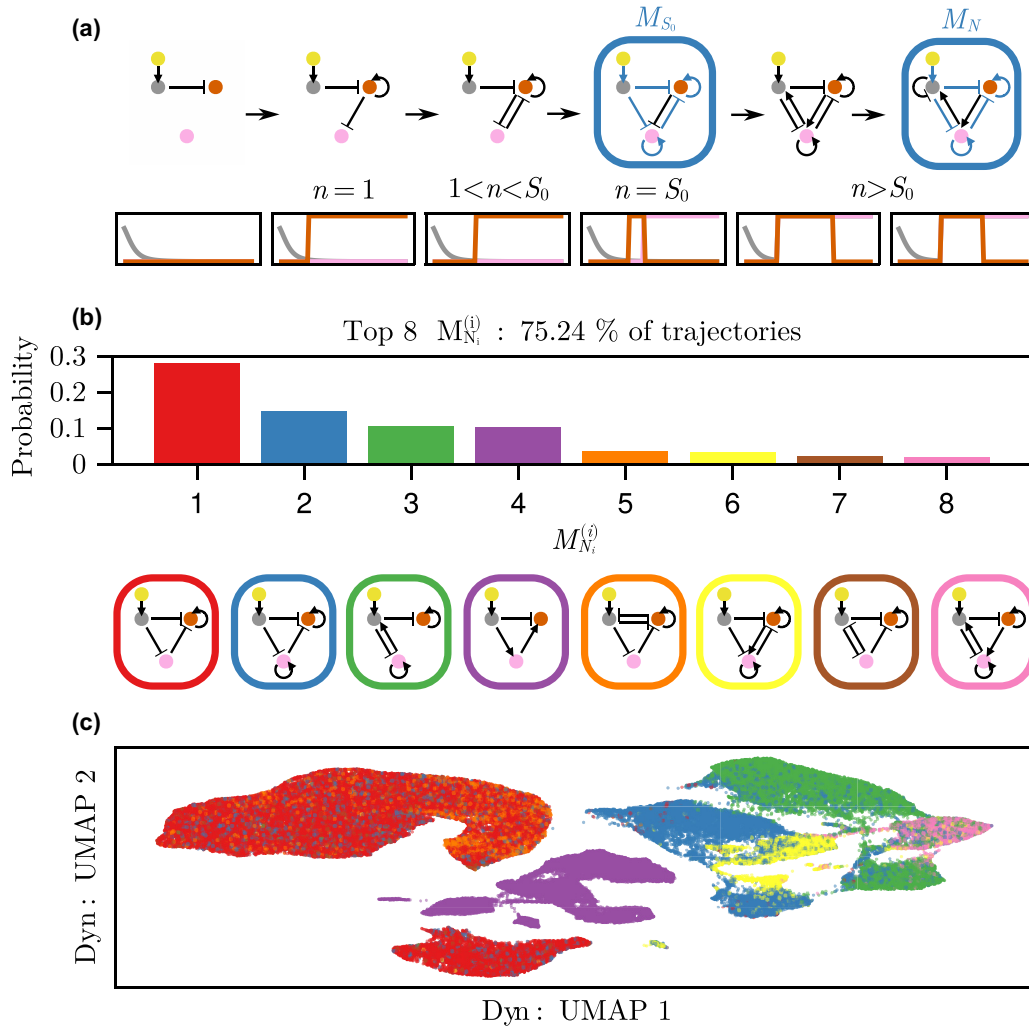


FIG. 2. The distribution of adaptive outcomes relating to the core mechanisms of stripe formation. (a). An example evolutionary trajectory showing topological changes in the underlying network along with the associated phenotypes. Each step indicates an accepted mutant. The change in the network topology (genotype) is shown on top with the corresponding gene expression profile (phenotype) below. The value of n is the index of the shown accepted mutant. The interactions which form the MST at $n = S_0$ and $n = N$ are highlighted and denoted as M_{S_0} and M_N , respectively. (b) The distribution of top eight MST outcomes $M_{N_i}^{(i)}$ across evolutionary simulations. Colors correspond to the MSTs visualized below the distribution. (c) A UMAP [49] embedding of the dynamics associated with final evolved networks $W_{N_i}^{(i)}$, with color coding indicating MST classification shown in (b).

tional process interacts with GRN organization to drive and constrain network evolution and pattern innovation.

B. New gene expression boundaries require multiple weight changes per mutant

To understand how mutation interacts with GRN organization and selection for novel patterns, we asked whether specific types of mutations were typically associated with the formation of new gene expression boundaries or the adjustment of existing ones. To this end, we characterized the phenotypic transition associated with each accepted mutant.

We reasoned that there are three “routes” to a stripe pattern Fig. 3(a). Starting from the shallow proximal boundary associated with the initial genotype, evolution can first establish a sharp proximal boundary followed by a sharp distal boundary through sequential mutants, do this in reverse, or establish both proximal and distal boundaries simultaneously

[colored arrows in Fig. 3(a)]. In between establishing new boundaries, mutation and selection can continue to alter the phenotype, e.g., by adjusting the location or amplitude of existing boundaries [self-loops in Fig. 3(a)]. Finally, mutants can be phenotypically neutral, creating no phenotypic changes and no fitness change compared to the resident phenotype.

We found that 98.8% of trajectories introduced the proximal boundary first followed by the distal boundary [Fig. 3(b)]. The large majority of these establish a sharp proximal boundary with the first accepted mutant ($n = 1$), followed by adjustment of the newly formed proximal boundary to align it with that of the target stripe ($1 < n < S_0$), as illustrated in Fig. 2(a). The distal boundary is established at $n = S_0$, with its location then similarly adjusted to match the target pattern during $n > S_0$. The prominence of this phenotypic path is consistent with the initial network topology, in which

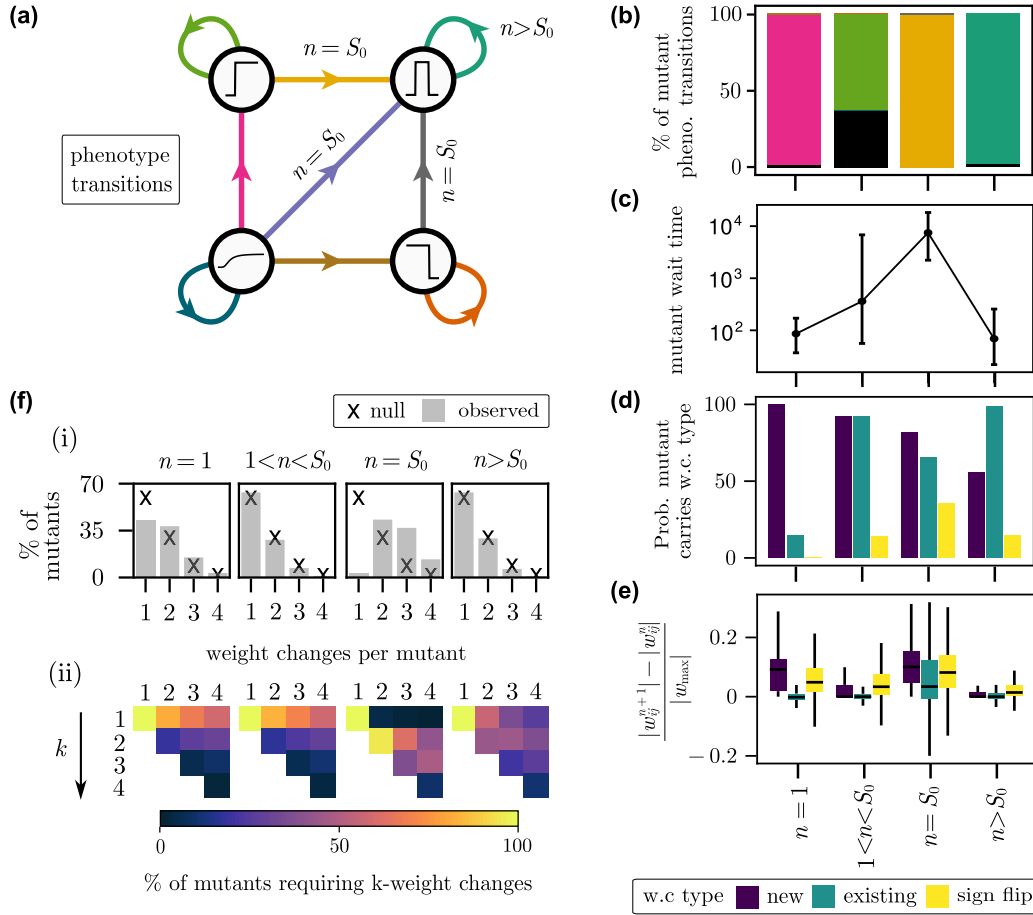


FIG. 3. Key components of stripe evolution. (a) Phenotypic transitions during evolution. The initial phenotype is that of a shallow proximal boundary (bottom left). The first accepted mutant at $n = 1$ may lead to a sharp proximal boundary (top left, pink arrow), a stripe (top right, purple arrow), a sharp distal boundary (bottom right, brown arrow) or adjust the existing phenotype without introducing new boundaries (self-arrow). Arrows between these states indicate the acceptance of a mutant driving the respective phenotypic changes. Self-arrows indicate accepted mutants that adjust the respective phenotype without adding or removing boundaries. Arrows for transitions that lead to stripe inception are marked by $n = S_0$. (b) Percentage of accepted mutants in different evolutionary periods associated with specific phenotypic transitions. Colors correspond to arrows in (a). The black bar for $1 < n < S_0$ corresponds to neutral mutants that do not induce a phenotypic shift or fitness change. (c) Median wait times associated with accepted mutants in each evolutionary period (dots). Range bars represent the 25th to 75th percentiles of the wait time distribution. (d) Classification of the weight change (w.c.) types carried by accepted mutants in different evolutionary periods as follows: introduction of a new weight (purple) and adjustment of existing weight while preserving sign (green) adjustment of existing weight with sign change (yellow). Solid bars indicate the probability of mutants carrying at least one weight change of the type corresponding to the color. (e) The distribution of accepted weight change magnitudes expressed relative to the maximum regulatory effect, defined as $|w_{ij}^{n+1}| - |w_{ij}^n| / |w_{\max}|$, associated with mutants in different evolutionary periods. Color indicates weight change category as in (d). (f) Percentage of accepted mutants carrying one to four weight changes, across different evolutionary periods. Crosses indicate likelihood according to null sampling distribution (preselection), and bars indicate postselection percentages. For each mutant carrying multiple weight changes we evaluate the minimum number of weight changes, k , required to produce the observed fitness gain. This information is visualized in the heatmap, which presents the distribution of required weight changes k against the original number carried by the mutants. Yellow along diagonal entries indicates that all weight changes are necessary for the fitness gain.

the morphogen inhibits output gene c (orange) through the intermediary gene a (gray) [Fig. 1(c)]. This initial constraint means that creating the proximal boundary simply requires increasing positive feedback on gene c . Accordingly, $\sim 96\%$ of accepted mutations at $n = 1$ introduce a self-activation interaction on the output gene c (orange). The wait times associated with the acceptance of mutants in each evolutionary period indicate that the proximal boundary is typically established quickly [Fig. 3(c)]. In contrast, establishing the distal boundary at $n = S_0$ takes orders of magnitude longer [Fig. 3(c)].

Therefore despite the degeneracy in the genotype that underlies stripe formation identified in the previous section, phenotypic trajectories are highly stereotyped. These phenotypic trajectories reflect the accessibility of new gene expression boundaries imposed by gene interactions already present in the system at the onset of selection for a stripelike pattern. Consistently, altering the initial network configuration impacts the typical sequence of phenotypic transitions (Fig. S1 in Ref. [36]).

To understand what type of network changes typically produced each phenotypic transition, we characterized the types

of weight changes carried by mutants accepted within each evolutionary stage [Fig. 3(d)]. Mutants accepted in periods $1 < n < S_0$ and $n > S_0$ primarily carry adjustments to existing weights, and weight changes during these periods are typically of small magnitude [Figs. 3(d) and 3(e)]. These periods are associated with the movement of existing boundaries, indicating that the refinement of existing gene expression boundaries can be achieved through small adjustments to an existing network. In contrast, formation of new boundaries at $n = 1$ and $n = S_0$ typically required the addition of novel regulatory interactions [i.e., new edges in the GRN; Fig. 3(d)] and was often associated with weight changes of larger magnitude [Fig. 3(e)]. Additionally, distal boundary formation was associated with a change in sign of an existing regulatory interaction, from inhibition to activation or vice versa [Fig. 3(d)]. Therefore our model suggests that generating novel gene expression boundaries typically requires more extensive network reorganization than the adjustment of existing boundaries.

To investigate whether the number of weight changes in individual mutants impacts fitness, we next characterized the number of weight changes per accepted mutant in each evolutionary period [Fig. 3(f)]. Preselection, the probability of a single weight change, p , specifies a null distribution for the number of weight changes per mutant (crosses, Fig. 3(f)(i); for derivation see Sec. 1.2.1 in Ref. [36]). Postselection this null distribution may be “rebalanced”, representing the number of weight changes carried by accepted mutants, and may be suggestive of a selective bias towards more or fewer weight changes [bars, Fig. 3(f)(i)].

For $n = 1$ and $n = S_0$ we noted a bias for more weight changes in accepted mutants in the postselection distribution compared to the null distribution. The probability that a stripe-forming mutant carries only a single weight change is $\sim 3\%$, compared to the null expectation of $\sim 60\%$ [Fig. 3(f)(i)]. A similar bias, albeit less pronounced, is observed for $n = 1$ [Fig. 3(f)(i)]. Conversely, the number of weight changes carried by mutants in periods $1 < n < S_0$ and $n > S_0$ follow the null distribution. Taken together, these results suggest that the generation of new gene expression boundaries requires interactions between multiple genes in the network to change simultaneously. Inversely, adjustment of existing boundaries is more readily achieved through single weight changes.

To investigate whether our results depend on how rarely two weight changes appear together, we repeated our analysis for smaller values of p , reducing the frequency at which multiple weight changes appear in the same mutant. Notably, the postselection distribution of weight changes per mutant across the different evolutionary periods was invariant for smaller p , but the wait times for stripe inception increased as expected given the increasing rarity of the multiple mutation for smaller p (Figs. S2(a) and S2(b) in Ref. [36]). This finding suggests that the capacity to generate multiple weight changes at once may impact the evolvability of GRNs and the spatial patterns they regulate during development.

For each mutant carrying multiple weight changes we additionally evaluated the minimum number of weight changes that were strictly necessary to achieve the observed fitness gain [Fig. 3(f)(ii)]. This evaluation reveals (1) when multiple

weight changes were strictly necessary for the original fitness gain and (2) when weight changes interact with one another to impact the phenotype. Many of the initial mutants ($n = 1$) carrying multiple weight changes achieve the associated fitness gain using only a single weight change [Fig. 3(f)(ii)]. This is puzzling as the postselection distribution of the number of weight changes per accepted mutant suggests a selective advantage for multiple weight changes. We reasoned that this selective advantage for multiple weight changes at $n = 1$ can be understood in the context of the minimality of the network at this stage of evolution. Because few interactions are initially present, new interactions are more likely to have no regulatory effect on the output gene and hence be phenotypically neutral. Introducing more weight changes merely increases the likelihood that an interaction between genes with a positive fitness effect is introduced (Fig. S3 in Ref. [36]).

Inversely, at $n = S_0$, the increasing number of weight changes carried by accepted mutants is necessary for the associated fitness advantage [Fig. 3(f)(ii)]. Given that multiple weight changes per mutant is a necessary feature for the evolution of new gene expression boundaries in our model, we sought to mechanistically understand how these interactions worked in combination to facilitate evolutionary trajectories. To this end, we investigated the quantitative effects of individual versus combined weight changes on fitness. Epistatic effects naturally arise within gene regulatory networks, since by definition multiple genes interact to activate or inhibit other nodes in the network [53,54]. Therefore, placing our analysis in the context of epistatic interactions is a promising framework for our analysis.

C. Epistatic effects explain the need for multiple weight-changes

A broad definition of epistasis implies nonadditive fitness effects between individual mutations [55]. Epistasis in our model emerges when the sum of the fitness changes associated with individual weight changes does not equal the fitness change induced by the combined weight change mutant. Epistatic effects can be classified into different types. In a 2-weight mutant which is adaptive (i.e., $\Delta\mathcal{F} > 0$), sign epistasis refers to a situation where at least one of the changes is deleterious in absence of the other. When both changes are deleterious individually this is known as reciprocal sign epistasis (RSE). Magnitude epistasis refers to the case where both individual changes are adaptive but together produce a fitness effect which differs from the sum of their individual contributions and can be positive or negative [55] (for quantitative definitions see Sec. 1.8.2 in Ref. [36]).

To understand if and how epistatic effects impact the selective bias towards more than one weight change per mutant, we focused on mutants carrying two weight changes and characterized their epistasis type across the four key evolutionary periods we have identified. First, we noted that the frequency at which two-weight mutants exhibited epistasis increased as evolution progressed Fig. 4(a). This is an intuitive result, as the addition of more interactions in a network creates dependencies between its connected components. Second, the point of stripe inception was strongly correlated with elevated sign and reciprocal sign epistasis, with the vast majority of two-weight mutants at this point in evolution showing some

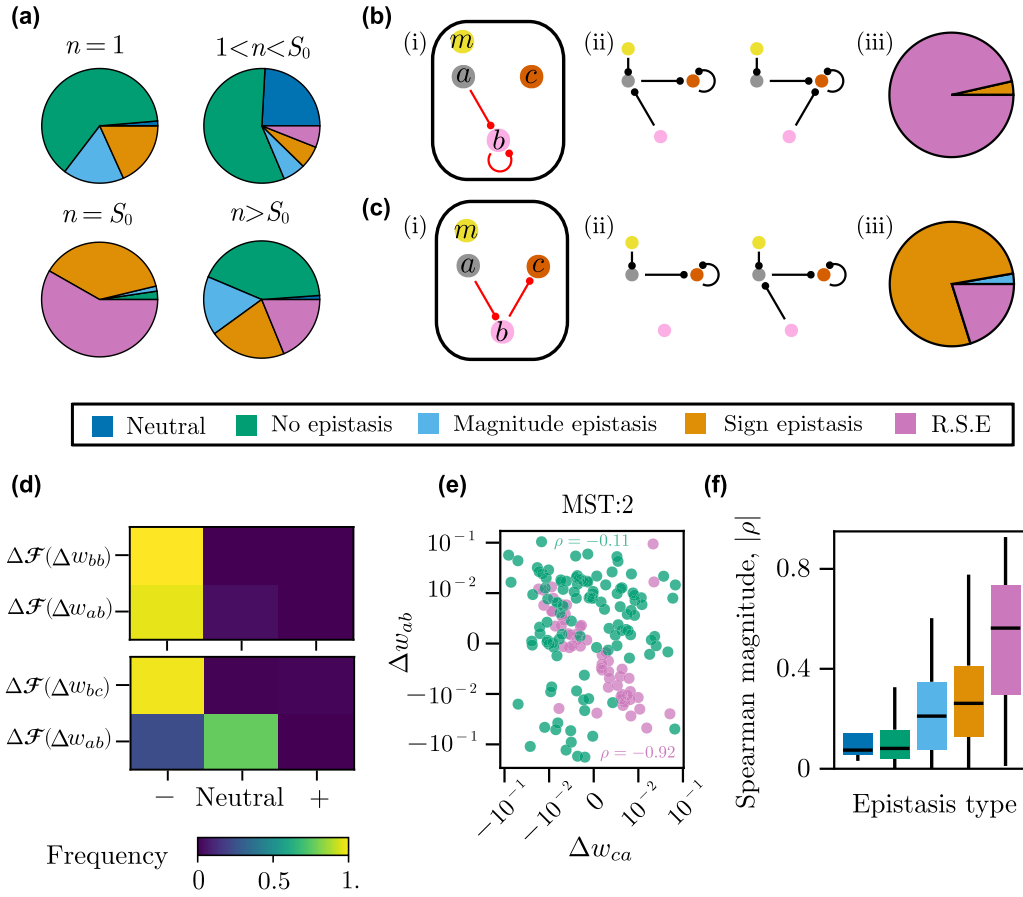


FIG. 4. Epistatic effects in 2-weight mutants. (a) The distribution of epistatic effect classifications, as defined in the main text and with color coding specified in legend, measured across 2-weight mutants in four different evolutionary stages. (b) An example of a mutant at the point of stripe inception, $n = S_0$, with high prevalence of RSE. (c) An example of an S_0 mutant with high prevalence of sign epistasis. In both panels (b) and (c): (i) The boxed GRN visualization shows the weight changes carried by an example 2-weight mutant and does not show the remainder of the existing network (ii) the top two most frequent network structures prior to mutation (iii) the distribution of measured epistatic effects for this S_0 mutant across our evolutionary simulations. (d) The heatmap shows the frequency of fitness effects associated with each individual accepted weight change for the two mutants in panels (b)(i) and (c)(i). Individual weight changes are classified as adaptive (+), neutral, or deleterious (−). For example, instances where $\mathcal{F}(w_{ij})$ and $\mathcal{F}(w_{kl})$ are deleterious would represent a 2-weight mutant with RSE. (e) Example calculation of weight change correlation. We calculated the weight changes ($\Delta w_{kl}^t, \Delta w_{ij}^t$) induced by accepted 2-weight mutants, where $\Delta w_{kl}^t = w_{kl}^t - w_{kl}^{t-1}$. We then measured the correlation ρ after grouping ($\Delta w_{kl}^t, \Delta w_{ij}^t$) by unique weight identity, MST, and epistasis type. (f) The distribution of spearman correlation magnitudes measured between pairwise weights, calculated from all 2-weight mutants conditioned on epistasis type.

form of sign epistasis Fig. 4(a). In the case of reciprocal sign epistasis, adaptive two-weight mutations are unlikely to fix through sequential mutations, as both of these mutants are deleterious individually. Similarly, sign epistasis means that only one of these two-weight mutations have positive fitness. Therefore, these forms of sign epistasis restrict single weight mutants, explaining the strong requirement for multiple weight changes observed at S_0 in our simulations.

To understand the mechanistic origin of these epistatic effects, first note that stripe formation in our evolved networks is achieved largely through incoherent feed-forward or negative feedback loops Fig. 2(b). Accordingly, at $n = S_0$ nearly all (99.6%) trajectories carried at least one weight change in the regulatory pathway that connects the gene that is activated by the morphogen (gray) to the stripe forming gene (orange) via interactions with the third gene in the network (pink) (Fig. S4 in Ref. [36]). These connections are necessary to form a

variant of feed-forward or feedback loop and hence to define the distal boundary.

To investigate why the evolution of these feedback loops comes hand-in-hand with epistatic effects, we examined the most frequent 2-weight mutants that led to stripe inceptions at $n = S_0$ (Fig. S5(a) in Ref. [36]). Our data indicate that reciprocal sign epistasis frequently arises at $n = S_0$ since gene b (pink) has an existing role in regulating the stripe forming gene c (orange) prior to stripe formation, either directly or through another gene (e.g., Fig. 4(b) and Fig. S6(a) in Ref. [36]). Therefore, any alteration of gene b activity must be appropriately compensated by the second mutation, else its individual effect is often deleterious. Sign epistasis at this point emerges when one of the two weight changes is neutral and the other is deleterious when introduced individually. To explain this observation, note that the construction of a feed-forward/feedback loop requires the inclusion of gene b

(pink). However, frequently just prior to stripe inception at $n = S_0 - 1$ gene b has no existing regulatory effect on the phenotype (e.g., Fig. 4(c) and Fig. S6(b) in Ref. [36]). Therefore, any new regulatory interaction which alters gene b without additionally introducing an interaction between gene b and the stripe forming gene c (orange) will be neutral.

It follows that an increased capacity for accepting neutral mutants may allow trajectories to bypass the constraints imposed by sign epistasis. To test this idea, we increased the likelihood of accepting neutral mutants and restricted the mutational process to generate only single weight changes per mutant. Under these conditions, new gene expression boundaries at $n = 1$ and $n = S_0$ were frequently established via the sequential fixation of a single weight change at a time (Figs. S7 and S8 in Ref. [36]). So while a selective advantage exists for mutants which induce multiple regulatory changes simultaneously, neutral mutations and drift contributes to the evolution of new gene expression boundaries. In this way our findings suggest a means through which neutral drift can contribute to the evolvability and diversification of GRN-driven patterning.

The overall increase of epistasis as evolution progressed in our model is not surprising, since stripe-forming mechanisms encompass connected, nonlinear regulatory interactions. Accordingly, the correct function of feed forward mechanisms requires interactions between multiple genes to be tuned with respect to each other [56,57]. We reasoned that this type of dependency between genes is likely to induce specific correlations in the underlying regulatory weight changes postselection. To test if this was true in our model, we calculated the weight changes ($\Delta w_{kl}^t, \Delta w_{ij}^t$) induced by accepted 2-weight mutants, where $\Delta w_{kl}^t = w_{kl}^t - w_{kl}^{t-1}$. We then measured the correlation ρ after grouping ($\Delta w_{kl}^t, \Delta w_{ij}^t$) by unique weight identity, MST, and epistasis type Fig. 4(e). As predicted, the correlations between weight changes postselection are higher in the presence of epistasis and maximal in mutants that exhibit RSE [Fig. 4(f)]. Furthermore, we found that RSE was particularly enriched when considering weight changes within the MST subnetwork (Fig. S9 in Ref. [36]). Therefore the minimal set of regulatory weights which support stripe formation induce the strongest pairwise correlations in their evolutionary trajectories, and these correlations manifest through strong epistatic effects. Together these results indicate that epistatic effects explain the selective bias towards multiple weight changes per mutant observed in the previous section.

D. Constraints and predictability during GRN evolution

Epistasis has been linked to the classic question of historical contingency [58], in which early chance mutations can restrict evolutionary paths [16,56,59,60], committing evolution to particular yet highly contingent outcomes. Given the prevalence of epistasis in our evolutionary trajectories, we wondered to what extent future evolutionary paths become contingent upon the weight changes accepted in earlier steps. Our framework generates a large number of evolutionary trajectories at high efficiency. Thus, we can obtain datasets that are sufficiently large to investigate the predictability of evolution using statistical methods. Using our large dataset of

trajectories, we aimed to identify mutations that hold a pivotal role in influencing evolutionary outcomes.

We built a predictive model of MST outcome given the regulatory interactions present at any point in an evolutionary trajectory. Given an evolutionary trajectory which has not yet achieved a stripe, our model aims to predict a set of probabilities representing the likelihood of each possible mechanistic outcome becoming associated with the final evolved stripe. In other words, given some partial trajectory of genotype evolution $T_k^{(i)} = (W_0^{(i)}, W_1^{(i)}, W_2^{(i)}, \dots, W_k^{(i)})$ with $k < S_0^{(i)}$, we want to predict a set of probabilities $q^{(i)}(k) = [q_1^{(i)}(k), q_2^{(i)}(k), \dots, q_{N_T}^{(i)}(k)]$, where N_T is the number of MSTs identified. Here $q_j^{(i)}(k) = \mathbb{P}(M_{S_0}^{(i)} = j | T_k^{(i)})$ is the probability that the MST associated with the stripe forming mutant at $n = S_0$ is the j th MST, given that we have observed the trajectory up to the k th mutant. This is a probabilistic classification task, where the history of the evolutionary process up until the prediction point constitutes the explanatory features, and the initial MST at $n = S_0$ represents the “class outcome.” We can then use the trained classifier to yield probabilities for the different MST outcomes given the evolved network at any point in a trajectory. Note that the classifier predicts the MST driving stripe inception which can be several accepted mutants ahead of the present genotype. For the purpose of training a probabilistic classifier we limited ourselves to the top 4 MST’s [$\sim 65\%$ of trajectories; Fig. 2(b)], with an additional “Other” category representing the remaining MST examples. Further details of the training procedure can be found in Sec. 1.9 in Ref. [36].

An example trajectory and its associated predictions is shown in Fig. 5(a). This trajectory consisted of four accepted mutants before stripe inception at $n = S_0 = 4$. The initial prediction $q^{(i)}(0)$ simply reflects the proportion of each MST observed in the train-set—the null prediction learned by the model in absence of any information on the further evolution of the network. Following the first accepted mutant a new prediction, $q^{(i)}(1)$, indicates that the network changes have increased the likelihood that MST:1 will be adopted at the point of stripe formation. Following the second accepted mutant MST:4 becomes the most likely outcome and forms a correct prediction which persists until S_0 . We describe a mutant which changes the MST outcome prediction as branching (see Sec. 1.9 in Ref. [36]). Therefore in the above example, the second mutant was branching. In this way our model suggests that a single mutant can drastically change the likely evolved stripe-forming mechanism. These types of mutants represent the incorporation of new regulatory interactions which change the likely path of evolution. Therefore, this quantitative framework identifies network changes that drive specific mechanistic outcomes automatically and at a large scale.

In general, we were able to predict final MST outcome prior to S_0 , with a good out-of-sample predictive performance given relatively few network modifications (Fig. 5(b) and Fig. M1 in Ref. [36]). Interestingly, the predictive performance of the model deteriorated significantly when only topological information was considered by the classifier [Fig. 5(b)(i) versus 5(b)(ii)], indicating that network topology alone is not sufficient for determining the likely path of evolution and that the magnitude of the interactions between genes matter.

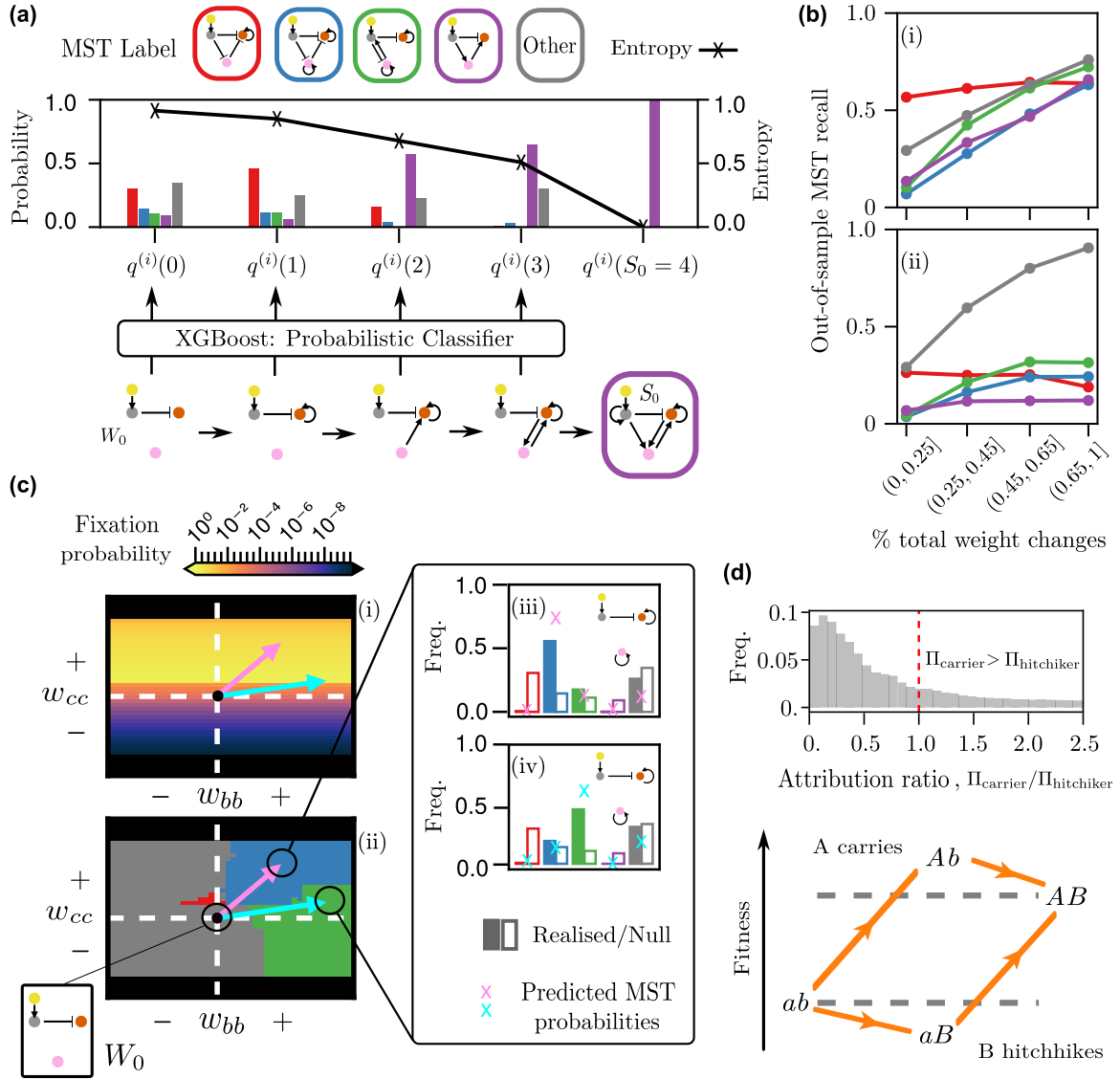


FIG. 5. Predictability of stripe evolution. (a) An example trajectory and the MST prediction probabilities for intermediate mutants. Colors (red, blue, green, and purple) represent MST labels and correspond to color coding used in Fig. 2, with gray representing the “other” MST category. A sequence of network configurations corresponding to fixed mutants in an example trajectory i is shown, with the MST realized at S_0 shown highlighted in purple, corresponding to MST:4. For each network $W_k^{(i)}$ in the i th trajectory we produce a probabilistic MST prediction $q^{(i)}(k)$ using a trained XGBoost classifier. Here $q^{(i)}(k) = [q_1^{(i)}(k), q_2^{(i)}(k), q_3^{(i)}(k), q_4^{(i)}(k), q_{\text{other}}^{(i)}(k)]$ where $q_j^{(i)}(k)$ is the probability that the j th MST is realized at S_0 given the k th mutant. Colored bars indicate the individual MST prediction probabilities $q_j^{(i)}(k)$. Black line indicates entropy associated with the prediction probabilities $q^{(i)}(k)$. (b) The recall for each MST outcome evaluated on out-of sample networks at different stages of evolution, evaluated using the complete predictive model (i) and the topology only predictive model (ii). Networks were first binned according to the number of weight edits they had accumulated as a percentage of the total accumulated at S_0 . The bins were chosen so that the number of networks in each bin was approximately equal. Next, the recall for each MST outcome was evaluated for the networks in each bin. Recall is the percentage of correct predictions for the outcome. A prediction is deemed correct if it matches the MST outcome and furthermore and persists from the point of prediction to S_0 ; for more details see Ref. [36]. (c) Two-dimensional fitness landscapes associated with pairwise weight changes carried by the ($n = 1$) mutant (i) with the associated MST predictions (ii). Pink and cyan arrows indicate weight changes associated with selected branching mutations. Bar charts (iii) and (iv) represent the MST outcome distributions associated with pink and cyan branching mutations. Filled bars indicate the distribution induced by the branching mutations, whereas nonfilled bars represent the null distribution ($\mathbb{P}(M_{S_0}^{(i)} = j | W_0)$). Scatter crosses indicate probability of each MST outcome given by predictive model for the branching mutants. (d) Histogram showing the distribution of prediction attribution ratios $\Pi_{\text{carrier}}/\Pi_{\text{hitchhiker}}$ associated with mutants carrying hitchhiking weight changes. Red dashed line indicates equal attribution between hitchhiking and carrier weight changes. The schematic shows that hitchhiking weight changes ($b \rightarrow B$) are those which are deleterious individually but adaptive when fix with a carrier mutation ($a \rightarrow A$).

We quantified how the statistical entropy, $H(\cdot)$, associated with the predicted probabilities $q^{(i)}(k)$ changes as mutants fix Fig. 5(a). Interpreting the entropy as a measure of uncertainty inherent to the MST outcome, this example highlights the intuitive notion that uncertainty decreases as evolution progresses. We found this to be a general principle, with mutants and their corresponding network changes most often decreasing uncertainty in eventual MST outcome (Figs. S10(a) and S10(b) in Ref. [36]). This is in agreement with previous work showing that the space of possible evolutionary paths that a trait can take tends to narrow during its evolutionary history [61,62]. While similar theoretical results have been established for evolution on abstract fitness landscapes [63], our analysis shows how the fitness landscapes associated with patterning GRNs constrain evolutionary paths.

We then wondered whether we could steer evolution towards a particular MST outcome by “engineering” the initial mutant. We used the predictive model to identify initial ($n = 1$) mutants that were branching towards a specific MST outcome and identified the branching mutant producing the largest negative entropy change, implying a highly certain mechanistic prediction. To aid tractability, we focused on mutants carrying 2-weight changes relative to the initial network. The mutants found through this procedure for MST:2 and MST:3 both induce the same change in fitness and hence have an equal probability of fixation [Fig. 5(c)(i)]. Although these mutants induce the same topological change, differences in the weight values result in alternative mechanistic predictions [Fig. 5(c)(ii)]. We then imposed the identified mutations at $n = 1$ and simulated 5000 independent trajectories from these new initial conditions. In both instances evolution was steered towards the predicted mechanism [Figs. 5(c)(iii) and 5(c)(iv)]. Similar results were observed for mutations predicted to steer evolution towards other MSTs (Fig. S11 in Ref. [36]). These findings suggest that regulatory interactions incorporated early in network evolution can significantly bias evolution towards a specific mechanistic outcome.

Recall that multiple weight changes were often selected for together at the first accepted mutant ($n = 1$), not all of which were strictly necessary [Fig. 3(f)]. Each adaptive weight change at $n = 1$ therefore likely fixed with a non-adaptive weight change, reminiscent of genetic hitchhiking whereby neutral or slightly deleterious mutations appear together with a beneficial mutation and increase in frequency as a result [64]. For each weight change we produce a prediction attribution, Π , which quantifies to what extent an individual weight change shifts the prediction for the MST outcome of evolution (see Sec. 1.9.5 in Ref. [36] for details). To then identify the relative importance of the hitchhiking versus adaptive weight change in each mutant, we calculate the ratio $\Pi_{\text{carrier}}/\Pi_{\text{hitchhike}}$. In the majority of cases, the hitchhiking weight change is far more instrumental in determining the mechanistic outcome than the weight change associated with the carrier Fig. 5(e). This suggests that regulatory networks can accumulate mutations which are selectively neutral but nonetheless significantly alter the course of subsequent evolution. A similar conclusion has been reached experimentally in a transcription network that controls the response to mating pheromone in yeast [65], where the gain of a neutral regulatory interaction permitted

changes in the overlapping circuit responsible for pheromone induction.

IV. DISCUSSION

In this work we have placed GRNs in an evolutionary context to investigate how gene network organization, mutation and natural selection together determine evolutionary outcomes. Changes at the regulatory level is a fundamental avenue through which developmental patterns diversify during evolution [2–8]. Although our analysis focuses on evolving stripes, forming and adjusting gene expression boundaries is fundamental to developmental patterning and its evolution. Therefore, our findings and conceptual framework offer general insights towards a mechanistic understanding of how developmental regulatory networks emerge and diversify.

By framing the evolution of a stripe in gene expression levels as successively establishing and adjusting sharp gene expression boundaries, we identified network modifications commonly linked with specific phenotypic innovations. Our findings suggest that adjustments of existing boundaries typically occur through smaller changes in the strength of existing gene interactions. Inversely, novel boundaries are achieved through mutations that impact multiple gene interactions and have a larger effect on network organization. These findings are consistent with recent mutagenic studies in *Drosophila* embryos [9] where point mutations on native enhancers rarely resulted in expression outside of the native pattern, whereas inserting *de novo* enhancer elements often drove novel patterns. Our model offers a mechanistic theoretical framework to interpret these experimental observations. Point mutations in binding sites or TF coding regions typically induce smaller changes in transcription rates [41,66], whereas larger changes can arise through whole enhancer duplication [67]. Alternatively, a novel enhancer element may induce a new regulatory interaction into the network [68]. Therefore, although we do not explicitly model gene regulation at the sequence level, our model is able to capture key aspects of GRN response to mutation and corresponding phenotypic shifts.

The capacity to generate changes in multiple gene interactions within the same mutant was crucial to forming and adjusting gene expression boundaries in our model. In our analysis, we have interpreted this observation in the context of functional epistasis, that is known to characterize patterning GRNs [13,69]. Accordingly, epistatic effects were prominent throughout evolutionary simulations. Specifically, sign and reciprocal sign epistasis, whereby two changes in gene interactions are adaptive only when they appear in the same mutant, were key to the formation of new gene expression boundaries and boundary adjustment at later stages of evolution when the core network becomes established. The ubiquitous presence of these epistatic effects across more than 10^5 independent evolution trajectories indicates that they are likely a general property of GRN-driven pattern evolution. Given this prevalence of epistasis in our analysis, we speculate on its possible implications for the evolvability of GRN-driven patterning mechanisms.

A pervasiveness of sign and reciprocal sign epistasis during GRN evolution implies that the capacity to generate adaptive mutants depends on the likelihood of generating single

mutants where multiple gene interactions change. Therefore the evolvability of GRN-driven patterning, broadly defined as the ease with which new phenotypes are accessed, depends on the genetic constraints which determine how likely multiple gene interactions are to change together in single mutants. Indeed convergence rates in our simulations correlated with the likelihood of generating multiple weight changes at once. A key question then becomes the following: How does the molecular basis of transcription and regulation interact with mutation to rewire regulatory networks?

Experimental work has characterized the evolutionary potential of mutations at the TF level for network rewiring [70] and the role of enhancers in introducing novel genes into a regulatory network [68]. In our model, specific combinations of gene interactions were selected for together and in a highly correlated manner. Could genetic architecture itself facilitate the correlations that are necessary to resolve these constraints and hence facilitate the evolution of patterning GRNs? For example, theoretical work in a class of genetic algorithms suggests that correlations in system parameters may enable mutations to align optimally with the contours of a fitness landscape, resulting in faster convergence towards fitness optima [71]. In a regulatory network, these type of correlated changes could arise through mutations in enhancer binding sites with cooperative or competitive effects or when binding sites of independent TFs are in proximity of one another [72,73]. Our work provides grounds for studying these effects and their potential role in the evolution of patterning mechanisms. Advances in synthetic biology are beginning to shed light on how fragments on individual enhancers drive TF expression by engineering synthetic CREs and screening for their function [74]. This approach could be combined with our framework in the future to elucidate how the logical rules that underlie enhancer signaling impact GRN-driven spatial pattern evolution. Furthermore, insights from theoretical frameworks that study regulatory evolution by modeling mutation at the sequence level [75] could be incorporated in our framework to explore their consequences for patterning evolution.

We showed that in principle neutral drift could bypass certain epistatic constraints by allowing the accumulation of multiple weight changes in the same individual sequentially. Therefore in spite of the strong preference for multiple weight changes, we saw successful stripe evolution using only single weight changes given an increased propensity for neutral weight changes to fix. In this way, our work provides a mechanistic argument for the importance of neutral adaption in the evolution of regulatory networks [76]. The epistatic effects identified in the evolution of stripe forming mechanisms imply that the ability for neutral adaption to circumnavigate epistatic constraints could be a general property of evolution on patterning fitness landscapes. This may arise due to the inherent degeneracy in network topology for GRN-mediated patterning, whereby multiple network topologies can drive the

same spatial pattern [38] and the presence of so-called neutral networks, which have been shown using synthetic networks to enable new phenotypes to be accessed through mechanisms such as drift [12].

The dependence of phenotypic accessibility on GRN topology and the mutational process alludes to historical contingency as an important aspect of GRN evolution, so that network changes accumulated early may influence subsequent evolution [13,18]. We have built on this idea by investigating how the evolution of specific stripe forming mechanisms becomes contingent upon early network organization. By developing a predictive framework we identified “branching mutations” early in evolution, which we were able to show significantly altered the likely mechanistic outcome. This suggests that degeneracy in patterning mechanisms, in combination with the constraints imposed by the existing network on subsequent mutations drives historical contingency. By employing machine learning techniques in combination with a large dataset of evolutionary trajectories we present a novel method to investigate the importance of historical contingency, echoing existing probabilistic characterizations of historical contingency [77]. In particular we demonstrated how explainability techniques [78] offer a principled way to identify the specific motifs which introduce constraints in a network evolution setting. Furthermore, the identification of branching mutations using our framework could guide synthetic GRN evolution assays in the future.

In conclusion, we have developed a high-throughput and tractable computational approach for GRN-driven developmental patterning that we have used to probe general principles of the evolution of development. Our framework can be adapted in the future to integrate new information from synthetic engineering assays [74] and simulations at the sequence level [75] that probe the constraints that regulatory architecture introduces to the evolution of GRN-driven developmental patterning. In addition, the framework we have developed can be adapted to investigate other phenotypic transitions or more complex networks. In this way we hope that our findings will stimulate further research, driving advancements in our understanding of developmental pattern evolution.

ACKNOWLEDGMENTS

We thank Amy Bowen, James Briscoe, Jake Cornwall-Scoones, Léonard Dekens, James DiFrisco, Kabir Hussain, Taniya Mandal, and Lewis Mosby for discussions and feedback on this work. This work was supported by the Francis Crick Institute, which receives its core funding from Cancer Research UK, the UK Medical Research Council, and Wellcome Trust.

DATA AVAILABILITY

The code used to generate and analyse data can be found at [89].

[1] A. Kicheva and J. Briscoe, Control of tissue development by morphogens, *Annu. Rev. Cell Dev. Biol.* **39**, 91 (2023).

[2] I. D. Méndez-González, T. M. Williams, and M. Rebeiz, Changes in locus wide repression underlie the evolution

- of *Drosophila* abdominal pigmentation, *PLoS Genet.* **19**, e1010722 (2023).
- [3] A. Sadier, S. E. Santana, and K. E. Sears, The role of core and variable gene regulatory network modules in tooth development and evolution, *Integr. Comp. Biol.* **63**, 162 (2023).
 - [4] M. R. Johnson, S. Li, C. F. Guerrero-Juarez, P. Miller, B. J. Brack, S. A. Mereby, J. A. Moreno, C. Y. Feigin, J. Gaska, J. A. Rivera-Perez, Q. Nie, A. Ploss, S. Y. Shvartsman, and R. Mallarino, A multifunctional Wnt regulator underlies the evolution of rodent stripe patterns, *Nat. Ecol. Evol.* **7**, 2143 (2023).
 - [5] K. Onimaru, L. Marcon, M. Musy, M. Tanaka, and J. Sharpe, The fin-to-limb transition as the re-organization of a Turing pattern, *Nat. Commun.* **7**, 11582 (2016).
 - [6] B. Prud'homme, N. Gompel, A. Rokas, V. A. Kassner, T. M. Williams, S.-D. Yeh, J. R. True, and S. B. Carroll, Repeated morphological evolution through *cis*-regulatory changes in a pleiotropic gene, *Nature (Lond.)* **440**, 1050 (2006).
 - [7] V. F. Hinman, A. T. Nguyen, R. Andrew Cameron, and E. H. Davidson, Developmental gene regulatory network architecture across 500 million years of echinoderm evolution, *Proc. Natl. Acad. Sci. USA* **100**, 13356 (2003).
 - [8] I. S. Peter and E. H. Davidson, Evolution of gene regulatory networks controlling body plan development, *Cell* **144**, 970 (2011).
 - [9] R. Galupa, G. Alvarez-Canales, N. O. Borst, T. Fuqua, L. Gandara, N. Misunou, K. Richter, Mariana R P Alves, E. Karumbi, M. L. Perkins, T. Kocijan, C. A. Rushlow, and J. Crocker, Enhancer architecture and chromatin accessibility constrain phenotypic space during *Drosophila* development, *Dev. Cell* **58**, 51 (2023).
 - [10] X. C. Li, L. Gandara, M. Ekelöf, K. Richter, T. Alexandrov, and J. Crocker, Rapid response of fly populations to gene dosage across development and generations, *Nat. Commun.* **15**, 4551 (2024).
 - [11] T. Fuqua, J. Jordan, M. E. van Breugel, A. Halavatyi, C. Tischer, P. Polidoro, N. Abe, A. Tsai, R. S. Mann, D. L. Stern, and J. Crocker, Dense and pleiotropic regulatory information in a developmental enhancer, *Nature (Lond.)* **587**, 235 (2020).
 - [12] J. Santos-Moreno, E. Tasiudi, H. Kusumawardhani, J. Stelling, and Y. Schaerli, Robustness and innovation in synthetic genotype networks, *Nat. Commun.* **14**, 2454 (2023).
 - [13] Y. Schaerli, A. Jiménez, José M Duarte, L. Mihajlovic, J. Renggli, M. Isalan, J. Sharpe, and A. Wagner, Synthetic circuits reveal how mechanisms of gene regulatory networks constrain evolution, *Mol. Syst. Biol.* **14**, e8102 (2018).
 - [14] S Wright, The roles of mutation, inbreeding, crossbreeding and selection in evolution, in *Proceedings of the Sixth International Congress of Genetics, Ithaca, New York*, Vol. 1, edited by D. F. Jones (Brooklyn Botanic Garden, New York, 1932), pp. 356–366.
 - [15] J. Arjan G. M. de Visser and J. Krug, Empirical fitness landscapes and the predictability of evolution, *Nat. Rev. Genet.* **15**, 480 (2014).
 - [16] P. Shah, D. M. McCandlish, and J. B. Plotkin, Contingency and entrenchment in protein evolution under purifying selection, *Proc. Natl. Acad. Sci. USA* **112**, E3226 (2015).
 - [17] J. L. Payne and A. Wagner, Constraint and contingency in multifunctional gene regulatory circuits, *PLoS Comput. Biol.* **9**, e1003071 (2013).
 - [18] A. Jiménez, J. Cotterell, A. Munteanu, and J. Sharpe, Dynamics of gene circuits shapes evolvability, *Proc. Natl. Acad. Sci. USA* **112**, 2103 (2015).
 - [19] T. R. Sokolowski, T. Gregor, W. Bialek, and G. Tkačik, Deriving a genetic regulatory network from an optimization principle, *Proc. Natl. Acad. Sci. USA* **122**, e2402925121 (2025).
 - [20] B. Verd, A. Crombach, and J. Jaeger, Dynamic maternal gradients control timing and shift-rates for *Drosophila* gap gene expression, *PLoS Comput. Biol.* **13**, e1005285 (2017).
 - [21] J. Jaeger, S. Surkova, M. Blagov, H. Janssens, D. Kosman, K. N. Kozlov, Manu, E. Myasnikova, C. E. Vanario-Alonso, M. Samsonova, D. H. Sharp, and J. Reinitz, Dynamic control of positional information in the early *drosophila* embryo, *Nature (Lond.)* **430**, 368 (2004).
 - [22] M. Cohen, A. Kicheva, A. Ribeiro, R. Blassberg, K. M. Page, C. P. Barnes, and J. Briscoe, Ptch1 and Gli regulate Shh signalling dynamics via multiple mechanisms, *Nat. Commun.* **6**, 6709 (2015).
 - [23] N. Balaskas, A. Ribeiro, J. Panovska, E. Dessaud, N. Sasai, K. M. Page, J. Briscoe, and V. Ribes, Gene regulatory logic for reading the Sonic Hedgehog signaling gradient in the vertebrate neural tube, *Cell* **148**, 273 (2012).
 - [24] A. Pezzotta and J. Briscoe, Optimal control of gene regulatory networks for morphogen-driven tissue patterning, *Cell Syst.* **14**, 940 (2023).
 - [25] K. Exelby, E. Herrera-Delgado, L. G. Perez, R. Perez-Carrasco, A. Sagner, V. Metzis, P. Sollich, and J. Briscoe, Precision of tissue patterning is controlled by dynamical properties of gene regulatory networks, *Development* **148**, dev197566 (2021).
 - [26] B. Verd, N. A. M. Monk, and J. Jaeger, Modularity, criticality, and evolvability of a developmental gene regulatory network, *eLife* **8**, e42832 (2019).
 - [27] R. Mousavi and D. Lobo, Automatic design of gene regulatory mechanisms for spatial pattern formation, *nPJ. Syst. Biol. Appl.* **10**, 35 (2024).
 - [28] A. Henry, M. Hemery, and P. François, ϕ -evo: A program to evolve phenotypic models of biological networks, *PLoS Comput. Biol.* **14**, e1006244 (2018).
 - [29] P. François and E. D. Siggia, Predicting embryonic patterning using mutual entropy fitness and in silico evolution, *Development* **137**, 2385 (2010).
 - [30] P. François and V. Hakim, Design of genetic networks with specified functions by evolution in silico, *Proc. Natl. Acad. Sci. USA* **101**, 580 (2004).
 - [31] J. B. Rothschild, P. Tsimiklis, E. D. Siggia, and P. François, Predicting ancestral segmentation phenotypes from *Drosophila* to anopheles using *insilico* evolution, *PLoS Genet.* **12**, e1006052 (2016).
 - [32] P. François, V. Hakim, and E. D. Siggia, Deriving structure from evolution: Metazoan segmentation, *Mol. Syst. Biol.* **3**, 154 (2007).
 - [33] Renske M. A. Vroomans, P. Hogeweg, and Kirsten H. W. J. ten Tusscher, In silico evo-devo: reconstructing stages in the evolution of animal segmentation, *EvoDevo* **7**, 14 (2016).
 - [34] K. Fujimoto, S. Ishihara, and K. Kaneko, Network evolution of body plans, *PLoS ONE* **3**, e2772 (2008).
 - [35] L. S. Mosby, A. E. Bowen, and Z. Hadjivasilou, Morphogens in the evolution of size, shape and patterning, *Development* **151**, dev202412 (2024).

- [36] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/q79n-k93j> for further details on methodologies and numerical results, which include Refs. [8,21,37,38,42,45–49,55,70,73,78–88].
- [37] J. Reinitz and D. H. Sharp, Mechanism of eve stripe formation, *Mech. Dev.* **49**, 133 (1995).
- [38] J. Cotterell and J. Sharpe, An atlas of gene regulatory networks reveals multiple three-gene mechanisms for interpreting morphogen gradients, *Mol. Syst. Biol.* **6**, 425 (2010).
- [39] T. Friedlander, A. E. Mayo, T. Tlustý, and U. Alon, Mutation rules and the evolution of sparseness and modularity in biological systems, *PLoS ONE* **8**, e70444 (2013).
- [40] A. J. Stewart, S. Hannenhalli, and J. B. Plotkin, Why transcription factor binding sites are ten nucleotides long, *Genetics* **192**, 973 (2012).
- [41] T. Duque and S. Sinha, What does it take to evolve an enhancer? A simulation-based study of factors influencing the emergence of combinatorial regulation, *Genome Biol. Evol.* **7**, 1415 (2015).
- [42] S. Jiménez-Delgado, J. Pascual-Anaya, and J. García-Fernández, Implications of duplicated *cis*-regulatory elements in the evolution of metazoans: the DDI model or how simplicity begets novelty, *Brief. Funct. Genom.* **8**, 266 (2009).
- [43] S. A. Teichmann and M. Madan Babu, Gene regulatory network growth by duplication, *Nat. Genet.* **36**, 492 (2004).
- [44] D. Arendt, J. M. Musser, C. V. H. Baker, A. Bergman, C. Cepko, D. H. Erwin, M. Pavlicev, G. Schlosser, S. Widder, M. D. Laubichler *et al.*, The origin and evolution of cell types, *Nat. Rev. Gen.* **17**, 744 (2016).
- [45] J. H. Gillespie, Molecular evolution over the mutational landscape, *Evolution* **38**, 1116 (1984).
- [46] B. S. Khatri, Tom C. B. McLeish, and R. P. Sear, Statistical mechanics of convergent evolution in spatial patterning, *Proc. Natl. Acad. Sci. USA* **106**, 9564 (2009).
- [47] S. F. Greenbury, A. A. Louis, and S. E. Ahnert, The structure of genotype-phenotype maps makes fitness landscapes navigable, *Nat. Ecol. Evol.* **6**, 1742 (2022).
- [48] M. Kimura, On the probability of fixation of mutant genes in a population, *Genetics* **47**, 713 (1962).
- [49] L. McInnes, J. Healy, N. Saul, and L. Großberger, UMAP: Uniform manifold approximation and projection, *J. Open Source Softw.* **3**, 861 (2018).
- [50] D. E. Clyde, Maria S. G. Corado, X. Wu, A. Paré, D. Papatsenko, and S. Small, A self-organizing system of repressor gradients establishes segmental complexity in *Drosophila*, *Nature (Lond.)* **426**, 849 (2003).
- [51] S. Mangan and U. Alon, Structure and function of the feed-forward loop network motif, *Proc. Natl. Acad. Sci. USA* **100**, 11980 (2003).
- [52] W. Shi, W. Ma, L. Xiong, M. Zhang, and C. Tang, Adaptation with transcriptional regulation, *Sci. Rep.* **7**, 42648 (2017).
- [53] A. L. Tyler, F. W. Asselbergs, S. M. Williams, and J. H. Moore, Shadows of complexity: What biological networks reveal about epistasis and pleiotropy, *BioEssays* **31**, 220 (2009).
- [54] P. Nghe, Marjon G.J. de Vos, E. Kingma, M. Kogenaru, F. J. Poelwijk, L. Laan, and S. J. Tans, Predicting evolution using regulatory architecture, *Annu. Rev. Biophys.* **49**, 181 (2020).
- [55] K. Crona, D. Greene, and M. Barlow, The peaks and geometry of fitness landscapes, *J. Theor. Biol.* **317**, 1 (2013).
- [56] J. Cotterell and J. Sharpe, Mechanistic explanations for restricted evolutionary paths that emerge from gene regulatory networks, *PLoS ONE* **8**, e61178 (2013).
- [57] G. T. Reeves, The engineering principles of combining a transcriptional incoherent feedforward loop with negative feedback, *J. Biol. Eng.* **13**, 62 (2019).
- [58] S. J. Gould, *Wonderful Life: The Burgess Shale and the Nature of History* (W. W. Norton & Company, New York, 1989).
- [59] Merijn L. M. Salverda, E. Dellus, F. A. Gorter, Alfons J. M. Debets, John van der Oost, R. F. Hoekstra, D. S. Tawfik, and J. Arjan G. M. de Visser, Initial mutations direct alternative pathways of protein evolution, *PLoS Genet.* **7**, e1001321 (2011).
- [60] S. Kryazhimskiy, D. P. Rice, E. R. Jerison, and M. M. Desai, Global epistasis makes adaptation predictable despite sequence-level stochasticity, *Science* **344**, 1519 (2014).
- [61] T. Ijäs and R. Koskinen, Exploring biological possibility through synthetic biology, *Euro. Jnl. Phil. Sci.* **11**, 39 (2021).
- [62] J. DiFrisco and G. P. Wagner, Body plan identity: A mechanistic model, *Evol. Biol.* **49**, 123 (2022).
- [63] S. Kauffman and S. Levin, Towards a general theory of adaptive walks on rugged landscapes, *J. Theor. Biol.* **128**, 11 (1987).
- [64] N. H. Barton, Genetic hitchhiking, *Phil. Trans. R. Soc. Lond. B* **355**, 1553 (2000).
- [65] T. R. Sorrells, L. N. Booth, B. B. Tuch, and A. D. Johnson, Intersecting transcription networks constrain gene regulatory evolution, *Nature (Lond.)* **523**, 361 (2015).
- [66] M. Lässig, From biophysics to evolutionary genetics: Statistical aspects of gene regulation, *BMC Bioinf.* **8**, S7 (2007).
- [67] M. Martínez-Ara, F. Comoglio, and B. van Steensel, Large-scale analysis of the integration of enhancer-enhancer signals by promoters, *eLife* **12**, RP91994 (2024).
- [68] P. Majic and J. L. Payne, Enhancers facilitate the birth of *denovo* genes and gene integration into regulatory networks, *Mol. Biol. Evol.* **37**, 1165 (2020).
- [69] F. Baier, F. Gauye, R. Perez-Carrasco, J. L. Payne, and Y. Schaerli, Environment-dependent epistasis increases phenotypic diversity in gene regulatory networks, *Sci. Adv.* **9**, eadf1773 (2023).
- [70] C. Iglér, M. Lagator, G. Tkačik, J. P. Bollback, and C. C. Guet, Evolutionary potential of transcription factors for gene regulatory rewiring, *Nat. Ecol. Evol.* **2**, 1633 (2018).
- [71] O. M. Shir and A. Yehudayoff, On the covariance-hessian relation in evolution strategies, *Theor. Comput. Sci.* **801**, 157 (2020).
- [72] S. Kim and J. Wysocka, Deciphering the multi-scale, quantitative *cis*-regulatory code, *Mol. Cell* **83**, 373 (2023).
- [73] S. Weingarten-Gabbay and E. Segal, The grammar of transcriptional regulation, *Hum. Genet.* **133**, 701 (2014).
- [74] J. Cornwall-Scoones, D. Benzinger, T. Yu, A. Pezzotta, A. Sagner, L. Gerontogianni, S. Bernadet, E. Finnie, Giulia L. M. Boezio, H. T. Stuart, M. Melchionda, Oliver C. K. Inge, B. Dumitrascu, J. Briscoe, and M. Joaquina Delás, Predictable engineering of signal-dependent *cis*-regulatory elements (2025), <https://www.biorxiv.org/content/10.1101/2025.03.07.642002v1>.
- [75] T. Friedlander, R. Prizak, N. H. Barton, and G. Tkačik, Evolution of new regulatory functions on biophysically realistic fitness landscapes, *Nat. Commun.* **8**, 216 (2017).

- [76] M. Lynch, The frailty of adaptive hypotheses for the origins of organismal complexity, *Proc. Natl. Acad. Sci. USA* **104**, 8597 (2007).
- [77] T. Y. William Wong, Sources of evolutionary contingency: chance variation and genetic drift, *Biol. Philos.* **35**, 36 (2020).
- [78] S. M. Lundberg, G. G. Erion, and S.-I. Lee, Consistent individualized feature attribution for tree ensembles, [arXiv:1802.03888](https://arxiv.org/abs/1802.03888).
- [79] J. Bezanson, A. Edelman, S. Karpinski, and V. B. Shah, Julia: A fresh approach to numerical computing, *SIAM Rev.* **59**, 65 (2017).
- [80] C. Rackauckas and Q. Nie, DifferentialEquations.jl—A performant and feature-rich ecosystem for solving differential equations in Julia, *J. Open Res. Softw.* **5**, 15 (2017).
- [81] Ch. Tsitouras, Runge–Kutta pairs of order 5(4) satisfying only the first column simplifying assumption, *Comput. Math. Appl.* **62**, 770 (2011).
- [82] G. Datseris, J. Isensee, S. Pech, and T. Gál, Drwatson: The perfect sidekick for your scientific inquiries, *J. Open Source Softw.* **5**, 2673 (2020).
- [83] V. Sundaram, Y. Cheng, Z. Ma, D. Li, X. Xing, P. Edge, M. P. Snyder, and T. Wang, Widespread contribution of transposable elements to the innovation of gene regulatory networks, *Genome Res.* **24**, 1963 (2014).
- [84] Z. Burda, A. Krzywicki, O. C. Martin, and M. Zagorski, Motifs emerge from function in model gene regulatory networks, *Proc. Natl. Acad. Sci. USA* **108**, 17263 (2011).
- [85] F. Proulx-Giraldeau, T. J. Rademaker, and P. François, Untangling the hairball: Fitness-based asymptotic reduction of biological networks, *Biophys. J.* **113**, 1893 (2017).
- [86] T. Chen and C. Guestrin, XGBoost: A scalable tree boosting system, in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (ACM, New York, 2016), pp. 785.
- [87] A. Tharwat, Classification assessment methods, *Appl. Comput. Inform.* **17**, 168 (2021).
- [88] S. M. Lundberg and S.-I. Lee, A unified approach to interpreting model predictions, *Adv. Neural Inf. Proc. Syst.* **30**, 4765 (2017).
- [89] <https://github.com/harrybooth/Gene-Network-Organization>.

Correction: Contact author information has been added.