Deriving a genetic regulatory network from an optimization principle

Thomas R Sokolowski, Thomas Gregor, William Bialek, Gašper Tkačik

Deriving a genetic regulatory network from an optimization principle

Thomas R Sokolowski[a,b], **Thomas Gregor**[c,d], **William Bialek**[c,e], **and Gašper Tkačik**[a]

[a]Institute of Science and Technology Austria, AT-3400 Klosterneuburg, Austria; [b]Frankfurt Institute for Advanced Studies, DE-60438 Frankfurt am Main, Germany; [c]Joseph Henry Laboratory of Physics & Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, NJ 08544, USA; [d]Department of Stem Cell and Developmental Biology, UMR3738, Institut Pasteur, 25 rue du Docteur Roux, FR-75015 Paris, France; [e]Center for Studies in Physics and Biology, Rockefeller University, New York NY 10065

This manuscript was compiled on February 14, 2023

Many biological systems approach physical limits to their performance, motivating the idea that their behavior and underlying mechanisms could be determined by such optimality. Nevertheless, optimization as a predictive principle has only been applied in very simplified setups. Here, in contrast, we explore a mechanistically-detailed class of models for the gap gene network of the *Drosophila* embryo, and determine its 50+ parameters by optimizing the information that gene expression levels convey about nuclear positions, subject to physical constraints on the number of available molecules. Optimal networks recapitulate the architecture and spatial gene expression profiles of the real organism. Our framework makes precise the many tradeoffs involved in maximizing functional performance, and allows us to explore alternative networks to address the questions of necessity vs contingency. Multiple solutions to the optimization problem may be realized in closely related organisms.

Gene regulatory networks | Optimization | Evolution | Drosophila

Optimization is the mathematical language of choice for a number of fundamental problems in physical and statistical sciences. Stochastic optimization likewise constitutes the foundation of evolutionary theory, where selection continually improves organismal fitness by favoring adaptive traits (1, 2). This evolutionary force pushes against quantifiable physical constraints and there are many examples where the organisms we see today operate very close to the physical limit: photon counting in vision (3), diffraction limited imaging in insect eyes (4), molecule counting in bacterial chemotaxis (5), and more. Experimental evidence for optimal performance can be promoted to an optimization principle from which one can derive non–trivial predictions about the functional behavior and underlying mechanisms, sometimes with no free parameters (6, 7). Attempts at such ambitious *ab initio* predictions include the optimization of coding efficiency in visual and auditory sensory processing (8–11); growth rates in metabolic networks (12); matter flux in transport networks (13); information transmission in regulatory networks (14); and the design of molecular machines and assemblies (15).

We are unaware of any successful optimization predictions for complex, multi-component biological systems whose interactions are described in molecular detail. Whether *any* first principles prediction is even possible at this level remains unclear. As a consequence, we cannot determine whether the existence of a particular gene, genetic interaction or regulatory logic is an evolutionary necessity or merely a historical contingency (16). This difficulty is not resolved by genetic tests for necessity, since these cannot rule out alternative evolutionary histories that would have unfolded without (or with modified) molecular components.

Here we address these issues during the early stages of development in the *Drosophila* embryo (17). About two hours post fertilization, the four major gap genes *hunchback*, *Krüppel*, *giant*, and *knirps* are expressed in an elaborate spatiotemporal pattern along the anterior–posterior (AP) axis of the embryo (18). The gap genes regulate one another, forming a network that responds to the anterior (Bicoid), posterior (Nanos), and terminal (Torso-like) maternal morphogen gradients (17, 19). The states of the gap gene network in turn drive the expression of pair rule genes in striped patterns that

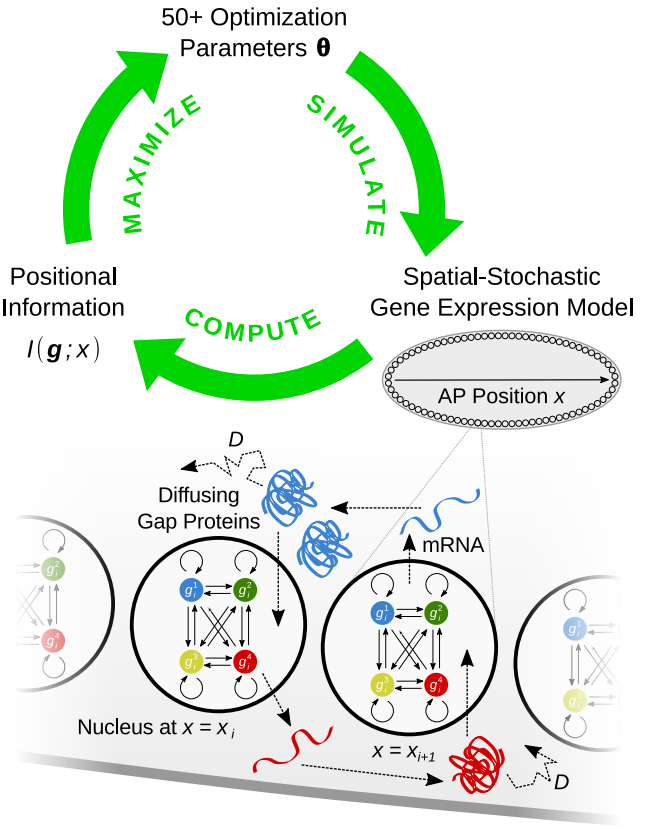


Fig. 1. Deriving a genetic regulatory network from an optimization principle. We simulate patterning during early fly development in a biophysically realistic, spatial–stochastic gap gene expression model (bottom; see Box 1) that accounts for the stochastic gene expression dynamics in individual nuclei along the anterior-posterior (AP) axis of the embryo. Regulatory interactions among four gap genes (arrows between colored circles in each nucleus), their response to three maternal morphogen gradients, and spatial coupling between neighboring nuclei are parameterized by a set of over 50 parameters $\boldsymbol{\theta}$. For each parameter set, we numerically simulate the resulting noisy gap gene expression patterns, compute the system's positional information $I(\mathbf{g};x)$, and adjust $\boldsymbol{\theta}$ using stochastic optimization to iteratively maximize the encoded I (top).

presage the segmented body plan of the fully developed organism (20). At readout time, about 40 minutes into nuclear cycle 14 (NC14), the local gap gene expression levels peak and encode 4.3 ± 0.1 bits of positional information (21–23). This information is necessary and sufficient for the specification of downstream pair–rule expression stripes and other positional markers with a positional error as small as $\sim 1\%$ of the embryo length (EL) (24), roughly corresponding to the nuclear spacing. Multiple lines of evidence further suggest that the flow of positional information through this system – comprising both its encoding into gap gene profiles and its readout by the pair–rule genes – is nearly optimal (22, 24, 25). These empirical observations lead us to the hypothesis that the gap gene network itself may be derivable from an optimization principle.

Quantitative experiments, genetic manipulations, and attempts to fit mathematical models of the gap gene network to data have uncovered a wealth of detail about this system (26–35). These facts are, in part, what an optimization theory for the gap gene network should explain. But there are also major conceptual questions: Is behavior of the network more constrained by its evolutionary history (36) or by the developmental constraints and physical limits that arise from the limited numbers of mRNA (37, 38) and protein (39) molecules? Are all three maternal morphogens and four gap genes necessary? Most importantly for our discussion, are the interactions among gap genes and the resulting expression patterns coincidental, or determined by some underlying theoretical principle (25)? In simpler terms, can we derive the behavior of the gap gene network, rather than fitting its parameters to data?

Optimization in a realistic context

To answer the questions outlined above, we have formulated a detailed and realistic spatial–stochastic model of patterning that encompasses gap gene regulation by maternal morphogens; gap gene cross–regulation; discrete nuclei, including their divisions; transcription, translation, and degradation processes; and diffusion of gap gene products (Fig. 1 and Box 1). Within this class of models, we search for the networks that transmit the maximum positional information given limits on the number of molecules that can be synthesized. Here we give a preliminary account of this work, with subsequent analyses to be reported in a longer paper.

We considered the three maternal morphogens as well as maximal gap gene transcription, translation, and degradation rates to be physical constraints fixed to their measured or estimated values (Box 1). This leaves more than 50 parameters which govern how gap genes integrate transcriptional regulatory signals from other gap genes and from their morphogen inputs; we refer to all these parameters together as $\boldsymbol{\theta}$. As an example, for each gene regulated by another, there is a parameter that measures the concentration at which the regulator exerts half–maximal activating or repressive effect on its target, and another parameter that measures the strength of this regulatory interaction. Different points in this 50+ dimensional space describe a wide spectrum of regulatory networks and their diverse expression patterns, most of which are nothing like the real fly embryo but nonetheless are *possible* networks given the known component parts. For any set of parameters we simulate the time evolution of our model, evaluating the mean spatial pattern of expression for all four gap genes as well as the gap gene (co)variability at every nuclear location along the AP axis. These calculations, carried out in the Langevin formalism, are complex yet numerically tractable; they properly account for maternal morphogen gradient variability and intrinsic biochemical stochasticity.

Positional information $I(\mathbf{g}; x)$ can be formalized as the mutual information between the set of gap gene expression levels $\mathbf{g} \equiv \{g_1, g_2, g_3, g_4\}$ and the AP coordinate x (7, 22, 23, 25). This quantity can be computed from the means and covariances of gap gene expression, which are the results of our model simulation at fixed $\boldsymbol{\theta}$ (see Box 1). If the gap gene system indeed has been strongly selected to maximize positional information at some readout time T, then the real network should be near the optimal setting of parameters, $\boldsymbol{\theta}^* = \mathrm{argmax}_{\boldsymbol{\theta}} I\left(\mathbf{g}(T); x\right)$. This problem is well posed because there are physical limits: the maximal rates of molecular synthesis combine with degradation rates to limit the maximum number of molecules for each species, setting the scale of the noise which in turn limits information transmission. We have previously solved simplified versions of this optimization problem on small subnetworks (40, 41, 46–49), but understanding the whole network at the level where comparisons with data are possible required a new computational strategy (Box 1). This larger scale numerical approach, combining simulation and optimization (Fig. 1), provides a route to derive the first *ab initio* prediction for a gene network in a realistic context.

Comparing optimal networks with the real network

We used a custom simulated annealing code to optimize the gap gene system for positional information (Fig. 2A). We first biased the search towards solutions that might exist in the proximity of the wild–type (WT) *Drosophila* gap gene expression pattern using a recently developed statistical methodology (50), and then removed the bias to be sure that we have found a true optimum. Figure 2B compares the gap gene expression profiles generated by the optimized network to data. The match in mean expression profiles is very good (Fig. 2C), although not perfect. Mismatches—e.g., double–peaked anterior giant domain, posterior–most hunchback bump, linear anterior ramp of hunchback—likely trace their origin to the fact that the class of models we consider still is a bit too simple; even if we fit the parameters of the model to the data we cannot resolve these discrepancies. The predicted gap gene variability similarly recapitulates the measured behavior.

The mechanistic nature of our model allows us to rationalize how the optimal pattern emerges. For example, the precision of the system output, manifested in the low variability ($\sim 10\%$) of gap gene expression levels at fixed position, is achieved through a combination of temporal averaging and spatial averaging via diffusion, which substantially reduces noise components transmitted from upstream regulators and morphogens (40, 51, 52). The spatial patterns of expression in the optimal solution are shaped siginificantly by mutual repression and self–activation, closely mimicking what had been inferred about the structure of the network from genetic

All authors performed research.

The authors declare no competing interests.

[2]To whom correspondence should be addressed. E-mail: gtkacik@ist.ac.at

Box 1. Model description and assumptions.

We model the expression of gap genes up to the readout time, 40 min into nuclear cycle 14 ($T = 166$ min post fertilization). **First,** we assume that the dynamics of gap gene expression can be described by effective rates for mRNA and protein synthesis and degradation. mRNA dynamics is assumed to set the slowest time scale (lifetime $\tau_M = 20$ min), so that the corresponding protein concentrations (lifetime $\tau_P = 10$ min) track the mRNA levels. The maximal mRNA production rate at full activation, $\rho_{\max}$, reproduces the maximal mRNA counts per nuclear volume reported for gap gene *hb* in early nuclear cycle 14 ($M_{\max} \approx 5 \cdot 10^2$) (37). Proteins are produced from mRNA in bursts (burst size $\beta = 12$ per mRNA), leading to maximal average protein number per nucleus $N_{\max} \approx 6 \cdot 10^3$. These parameters are assumed to be the same for all gap genes (38). **Second,** while our model allows for a two-dimensional cylindrical embryo geometry, a one-dimensional approximation of pattern formation along the anterior-posterior axis, with $N = 70$ nuclei uniformly spaced at positions $x_i = i \cdot \Delta$ (with $\Delta = 8.5$ μm) along the length $L = N\Delta$ of the embryo, provides a tractable approximation (40). During the simulated time period the embryo is a syncytium, allowing expression levels in neighboring cells to be coupled via an effective diffusion constant D (baseline value $D = 0.5$ $\mu m^2/s$, varied in Fig 3E) that includes both cytoplasmic diffusion and transport across the nuclear membrane. **Third,** the spatial profiles of maternal inputs to the gap gene network (A = anterior, P = posterior, and T = terminal system; see Box image) are established early and are assumed to be constant throughout the relevant time period. **Fourth,** the rate of mRNA synthesis for each gene expressed at x_i is modulated between zero and $\rho_{\max}$ by local gap gene expression levels $\mathbf{g}_i = (g_i^1, g_i^2, g_i^3, g_i^4)$ and the local maternal input concentrations, $\mathbf{c}_i = (c_i^{\mathrm{A}}, c_i^{\mathrm{P}}, c_i^{\mathrm{T}})$, as described by regulation functions f_α that parametrically differ between gap genes α but are the same for all positions i. Regulatory functions are inspired by the Monod-Wyman-Changeux (MWC) model, where the expression of gene α switches between active (probability f_α) and inactive ($1 - f_\alpha$) states; these probabilities depend on regulatory effects as follows (41, 42):

$$f_\alpha(\mathbf{g}_i, \mathbf{c}_i, \theta) = \frac{\rho_\alpha}{1 + \exp\left(-F_\alpha(\mathbf{g}_i, \mathbf{c}_i, \theta)\right)}$$

$$F_\alpha(\mathbf{g}_i, \mathbf{c}_i, \theta) = \underbrace{\sum_{\kappa \in \{1..4\}} H_G^{\alpha\kappa} \log\left(1 + \frac{g_i^\kappa}{K_G^{\alpha\kappa}}\right)}_{\text{Self- and Cross-Reg.}} + \underbrace{\sum_{\zeta \in \{\mathrm{A,P,T}\}} H_M^{\alpha\zeta} \log\left(1 + \frac{c_i^\zeta}{K_M^{\alpha\zeta}}\right)}_{\text{Feed Forward (FF) Reg.}} + \underbrace{F_0}_{\text{Base Expr.}} .$$

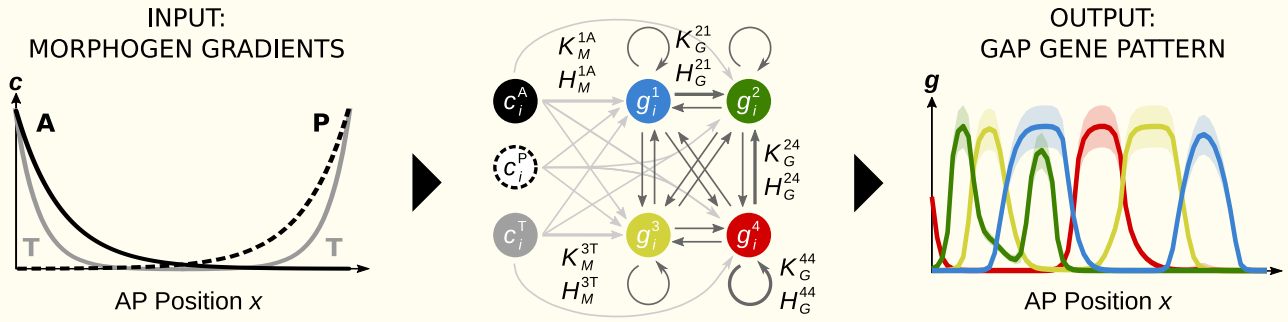
Here F_0 is the bias towards active state in absence of any regulatory signals; $H_G^{\alpha\kappa}$ and $H_M^{\alpha\zeta}$ are the strengths of regulatory action ($H > 0$ is activating and $H < 0$ repressive; $|H| < H_{\max}$ cf. Fig 4C), by gap proteins (for self– & cross–regulation) and maternal morphogen proteins (for feed forward regulation), respectively, while $K_G^{\alpha\kappa}$ and $K_M^{\alpha\zeta}$ are the associated concentration thresholds. Taken together, we obtain a system of coupled stochastic differential equations,

$$\partial_t g_i^\alpha = \underbrace{f_\alpha(\mathbf{g}_i, \mathbf{c}_i, \theta)}_{\text{Regulated Production}} - \underbrace{\frac{1}{\tau_\alpha} g_i^\alpha}_{\text{Degradation}} + \underbrace{\frac{D}{\Delta^2} \sum_n (g_n^\alpha - g_i^\alpha)}_{\text{Diffusion}} + \underbrace{\sum_k \Gamma_i^k(\theta)}_{\text{Noise}},$$

where n runs over all neighbors of nucleus i and Γ_i^k represent stochastic forces whose magnitude we derive in SI to account for the following noise processes: *(i)* "input noise" caused by the random arrival of TFs to gap gene CREs (43); *(ii)* "output noise" due to stochastic mRNA and protein synthesis and degradation; *(iii)* "diffusion noise" due to stochastic spatial exchange of gap gene proteins between neighbouring nuclei. Importantly, the variances of these noise terms typically scale (inversely) with the number of gap gene products. We phenomenologically add *(iv)* "extrinsic noise" due to maternal morphogen variability as well as other sources of embryo-to-embryo variation, where the contribution to the gap gene expression variance is assumed to be proportional to the squared mean expression (44), with the coefficient of variation ($\mathrm{CV}^2 = 0.02$) estimated directly from measurements.

The model is solved in two steps. In step one, we numerically integrate the deterministic part of the equation system defined above to obtain the mean expression levels at each position along the embryo axis at time T. Nuclear divisions are incorporated by doubling the maximal expression rate at the experimentally determined division times (i.e., $\rho_{\max}$ is only reached after the last division before T). In step two, the means are used to compute the full covariance matrix of the noise fluctuations in the gap protein levels, describing noise magnitude (on-diagonal matrix elements) and correlations (off-diagonal matrix elements) across space and gap gene species. These two quantities are used to compute the decoding map (24) and the corresponding positional information $I(\mathbf{g}; x)$ (23, 25).

Maximal copy number of gene products per nucleus and extrinsic noise imply that positional information must be upper bounded. To reach its absolute maximum value of $\log_2(N) \approx 6$ bits (error-free specification of N nuclei), these constraints would have to be lifted, implying a slower developmental process (due to lower protein and mRNA degradation rates) and/or a higher metabolic cost (due to higher transcription and translation rates). Within set rate constraints, various networks differ in the amount of actual gene expression, which we quantify by "resource utilization" (RU), the average expression across all gap genes and positions. RU = 1 means that gap gene expression is fully induced, proceeding at the maximal rate in every nucleus; for the *Drosophila* WT pattern, RU ≈ 0.2.



Spatial-stochastic model for gap gene expression. The gap gene regulatory network (center; colored circles = gap genes; grayscale circles = maternal morphogens) in each nucleus transforms maternal inputs with known spatial profiles (left) into a gap gene expression pattern at readout time T (right; solid lines = computed mean expression; shade = computed standard deviation). Each interaction in the network stands either for the feed forward (FF) regulation of a gap gene by a morphogen input (light gray arrows), or for the regulation of a gap gene by other gap genes or by itself (cross– and self–regulation, dark gray arrows), and is described by two parameters (concentration threshold K and regulatory strength H; several parameter pairs are shown, corresponding to nearby thicker arrows). All parameters denoted by regulatory arrows are jointly optimized to maximize positional information $I(\mathbf{g}; x)$.

interventions (Fig. 2D). Optimization correctly predicts strong mutual repression between *hunchback* and *knirps*, between *giant* and *Krüppel*, as well as most weak repressive interactions and self–activation of *hunchback* (52). Together, these factors combine to encode positional information nearly unambiguously, with a median positional error of $\sim 1.5\%$ (Fig. 2E); even the elevated positional uncertainty around the cephalic furrow and in the far posterior is consistent between the optimal

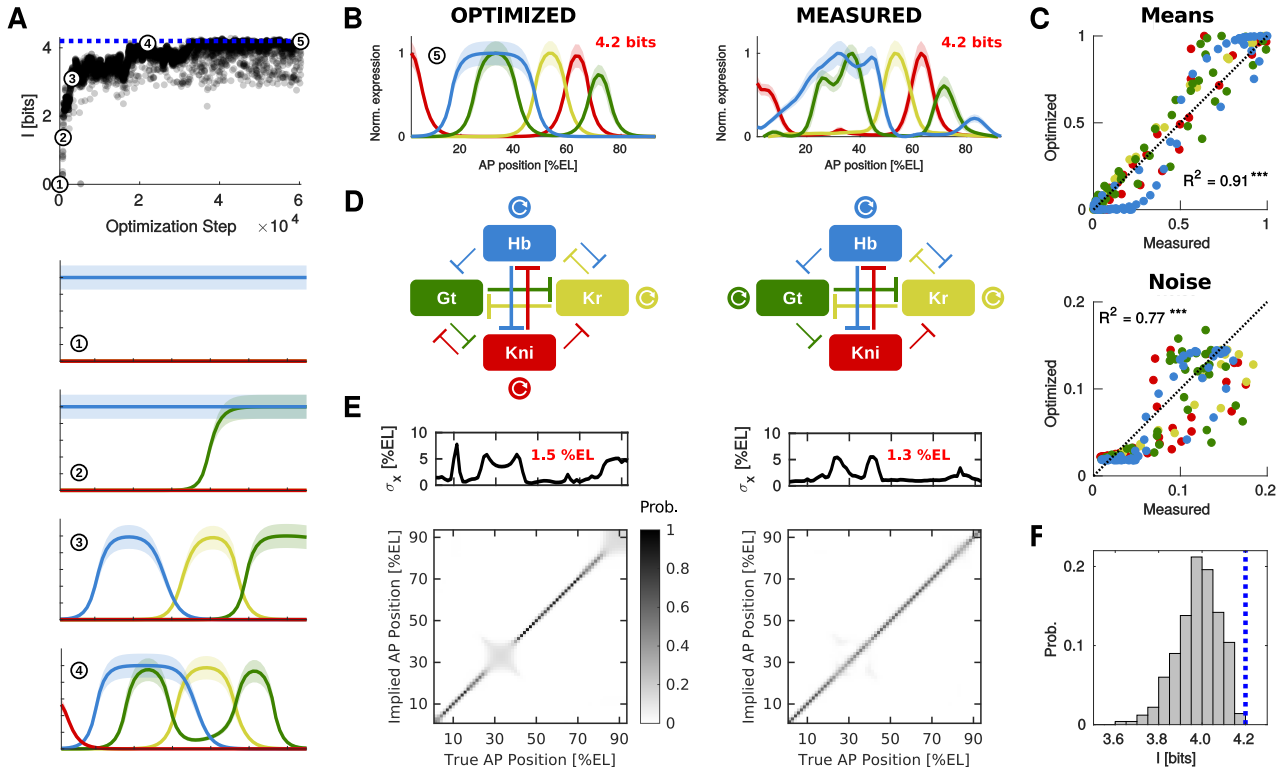


Fig. 2. Networks that maximize information transmission recapitulate the measured gap gene expression patterns and the regulatory network topology. (A) Positional information increases during a single optimization run, starting with the homogeneous profile at 0 bits (1), proceeding through more complex spatial patterns (2-4), to the final solution (5, pattern in panel B) that reaches ~ 4.2 bits (dashed blue line). **(B)** Predicted optimal (left) vs. measured gap gene expression pattern (right; (45)), 40 min into NC14 (blue = *hunchback*/Hb; green = *giant*/Gt; yellow = *Krüppel*/Kr; red = *knirps*/Kni; shade = standard deviation in gene expression). Positional information estimate from data is consistent with that reported in (22). **(C)** Measured vs. predicted mean expression (top) and variability (bottom) are highly correlated (color code as in B; Pearson $p < 10^{-3}$). **(D)** Predicted gap gene regulatory network (left; blunt arrows = repression; circular arrows = self–activation) vs. literature–based reconstruction (right; (18)). **(E)** Predicted (left) vs. measured (right) decoding map (bottom) shows a nearly unambiguous code (diagonal band) with $\sim 1.5\%$ median positional error and few outlier positions (top inset) (24). **(F)** Fitting the model to mean WT gap gene expression profiles yields a good fit but lower positional information values (black bars = distribution over replicate fits) compared to the optimized solution (blue dashed line).

prediction and the real embryo (24).

Ab initio optimization performed here makes minimal use of empirical data to derive a wide range of predictions, in stark contrast to traditional model fitting (50). This has three further important consequences. First, when fitting, objective functions are purely statistical (e.g., maximum likelihood, mean-square-error, etc.), lacking any biological interpretation. In contrast, positional information used in optimization constitutes a meaningful and independently measurable phenotype of the patterning system. For example, our optimal solution (Fig. 2A,B) reaches 4.2 bits, to be compared with $4.2-4.3$ bits estimated directly from data (22, 23). Second, if fitting is performed instead of optimization, e.g., by minimizing the mean-square-error of the predicted mean gap gene expression, the best fits underestimate the positional information (Fig. 2F). This is because fitting fails to take into account the functional consequences of noise and pattern variability. Third, optimization can identify locally optimal solutions that are qualitatively different from the gene expression patterns observed in the embryo but functionally near degenerate.

Multiple optimization runs indeed produce diverse solutions that locally maximize positional information while not exceeding the resource utilization of the wild type pattern. Together, these solutions constitute the *optimal ensemble*. A natural comparison is provided by the *random ensemble*, where parameters $\boldsymbol{\theta}$ are drawn independently and uniformly from broad but realistic intervals. Optimization for positional information automatically leads to significantly lower positional error (Fig. 3A), higher number of boundaries where gap gene expression switches from low to high or vice–versa ("slopes"), more uniform utilization of resources across gap genes, as well a slight but significant over–allocation of resources in the anterior, as can be seen in data as well. Within the optimal ensemble, solutions with higher information tend to be more dynamically stable at readout time, which we quantify by pattern rate–of–change (RoC), i.e., mean temporal derivative of gap gene expression (53). Low RoC is relevant since pair–rule genes appear to read out gap gene expression via fixed decoding rules (24, 27), implying that temporally varying solutions could cause larger spatial drifts in pair–rule stripes.

Networks in the random ensemble that transmit large amounts of information are exceedingly rare: the probability of drawing a network with positional information of 4 bits or more by chance is $\ll 10^{-6}$ (Fig. 3B). In contrast, optimization strongly and robustly enriches for solutions above 4 bits (Fig. 3B). In our optimization we have constrained the maximum numbers of molecules, and the real embryo uses $\sim 20\%$ (RU $= 0.2$) of this maximum, on average. This resource

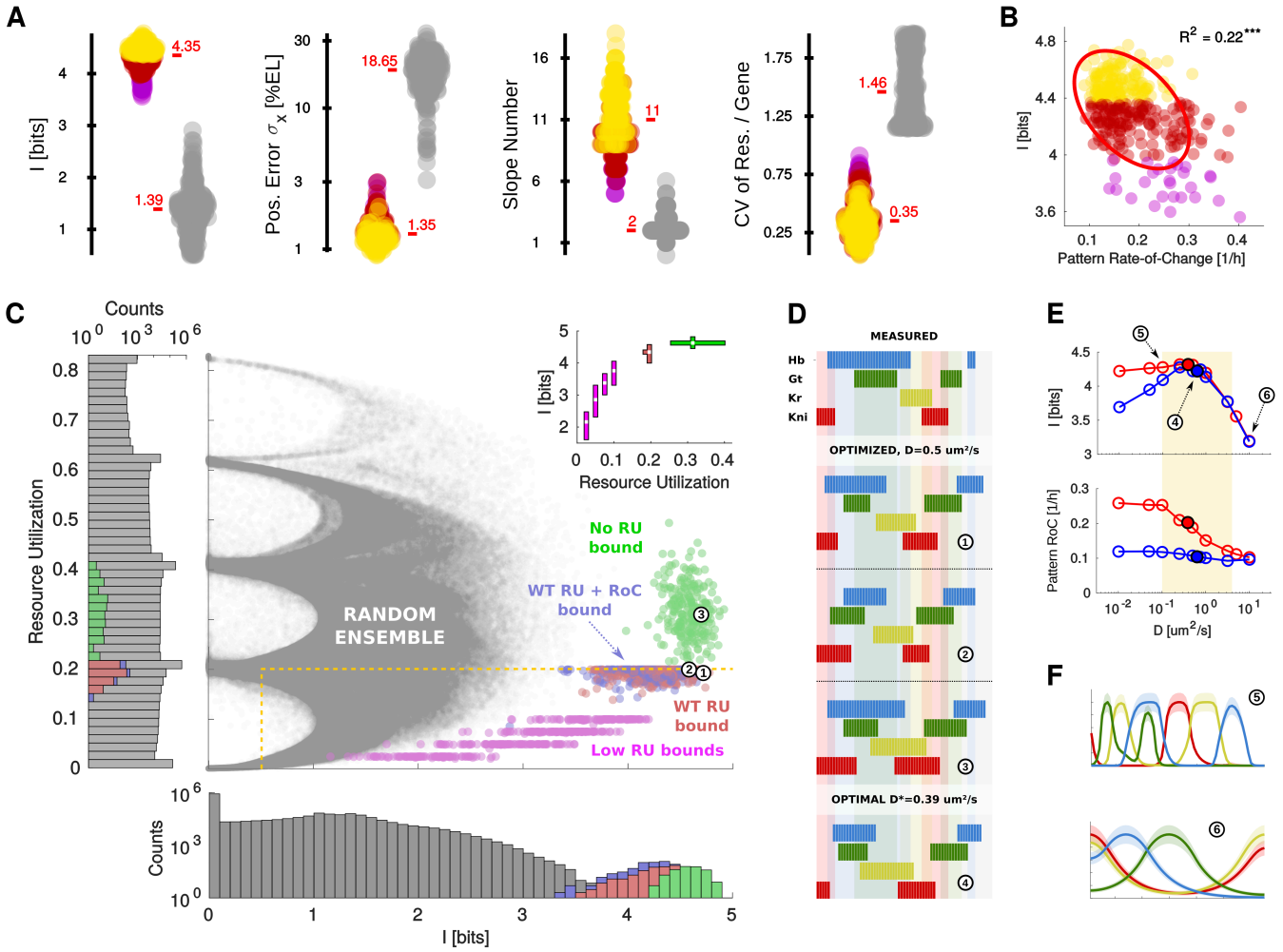


Fig. 3. Optimal and random gap gene network ensembles. (A) Patterning phenotypes for optimal ensemble (color, solutions from "WT RU" in panel C) vs. random ensemble (gray, including only solutions > 0.5 bit that are at or below WT resource utilization, delineated by dashed yellow lines in panel C) reveal that high positional information (leftmost; violet, red, yellow indicate lowest, middle, highest third of the information interval) correlates with low positional error, high number of gap gene "slopes", and a more uniform utilization of resources across the four gap genes (red numbers = ensemble medians). **(B)** Within the optimal ensemble, higher information correlates with higher dynamical stability, i.e., lower pattern rate–of–change (each dot = one optimal solution; red ellipse = 1 SD contour in the I vs. RoC plane; color code and optimal ensemble as in A). **(C)** Random (gray) and various optimal ensembles (red = resource utilization bounded by *Drosophila* WT denoted by dashed horizontal yellow line; magenta = progressively smaller resource utilization; blue = WT resource utilization plus a bound on pattern rate–of–change; green = no resource utilization or rate–of–change bounds) depicted in the information vs. resource utilization plane (each dot = unique parameter combination). Histograms in the margins show the raw counts of evaluated parameter combinations. Inset: Information vs. resource utilization (median, 0.1–0.9-quantile intervals over ensembles in the main panel shown as central white squares and ribbons, respectively). **(D)** Example optimal solutions (1–3) from panel C optimized at fixed gap product diffusion ($D = 0.5\ \mu\mathrm{m}^2/\mathrm{s}$), and an example solution (4) where D was also optimized from the ensemble in panel E, qualitatively match WT gap gene expression domains (top) and the regulatory network architecture. **(E)** Positional information (top) and pattern rate–of–change (bottom) as a function of gap gene diffusion constant D (empty circles = mean across optimal ensembles), capped at WT resource utilization (red) or with an additional rate–of–change constraint (blue). Solid circles = mean values for the case where D itself is also optimized; yellow shade = broad range of D consistent with literature reports. **(F)** Two example solutions optimized at lower-than-optimal (top) and higher-than-optimal (bottom) diffusion constant values.

utilization appears necessary for high-information solutions, whereas permitting more utilization within the same maximal rate limits does not noticeably increase positional information. In fact, among $> 10^3$ optimization runs we never found a solution exceeding 5 bits, indicating that such information values likely cannot be accessed within realistic constraints.

The random and the optimal ensemble are closely related to the evolutionary concepts of the *neutral* and the *selected* phenotype distributions (54). The random ensemble delineates what is physically possible in absence of selection for function, while the optimal ensemble delineates solutions that maximize function within fixed physical constraints. How closely natural selection *could* approach this optimality (as quantified by the selected phenotype distribution), or indeed *has* approached it (via the actual WT pattern), depends on selection strength and its history, genetic load, linkage disequilibrium, and other limitations that are of negligible concern to *in silico* optimization. Successful prediction of the pattern in Fig. 2B implies that selection was sufficiently strong to overcome such limitations and push the gap gene system beyond evolutionary tinkering (55) towards optimality (6, 50, 56). Even in strictly *ab initio* runs with zero bias towards the WT pattern we repeatedly find solutions that closely reproduce the overall size and placement of expression domains in *Drosophila* (Fig. 3D), the encoded positional information, as well as the regulatory interactions. Tantalizing early experimental work suggests that dipteran

species related to *Drosophila* may feature a broadly consistent gap gene domain arrangement whose expression domains are, however, shifted (57, 58) or swapped (59), as we find in our optimal ensembles.

Taken together, our results paint a nuanced picture of the "necessity vs. contingency" dichotomy. In the 50+ dimensional parameter space of possible networks, there is a highly non–random, locally optimal solution which produces expression patterns very similar to what we see in real fly embryos, but there are many other local optima that transmit about the same amount of positional information; all of these solutions are rare in the random ensemble. It is an open question whether alternative optima quantitatively recapitulate gap gene patterns seen in other dipterans or whether the degeneracy is removed by selection for additional phenotypes beyond positional information.

Alternatives to the real network

Our theory provides a framework within which we can explore tradeoffs beyond the structure of the gap gene network. As a first example, we have taken the effective diffusion constant of gap gene products to be a fixed physical property of the cytoplasm, $D = 0.5\ \mu\text{m}^2/\text{s}$, in line with existing measurements (39). But we can view D as one more parameter to be optimized, and remarkably we find that there is a broad optimum at the experimentally estimated value (Fig. 3E). Larger diffusion constants lead to a precipitous drop in information even when all other parameters $\boldsymbol{\theta}$ are re-optimized, because high diffusion smooths gap gene profiles to the extent that adjacent nuclei can no longer be distinguished reliably (Fig. 3F, bottom). On the other hand, slower diffusion does not serve as effectively to average over local super–Poisson noise sources; the optimization algorithm compensates by finding parameters that generate more and steeper transitions between high and low expression levels (Fig. 3F, top), but even these unrealistic patterns do not transmit quite as much information. Thus, a single parameter displaced away from its optimum causes significant decreases in positional information; to lessen the impact the optimization algorithm adjusts other network parameters, driving the predicted patterns of gene expression away from what we see in the real embryo.

We next address the question of evolutionary necessity and sufficiency. To this end, we make structural changes to the network and then re-optimize all of its parameters to explore "alternative evolutionary histories" that could have unfolded with changed molecular components or mechanisms. As an example, Figure 4A characterizes solutions obtained using $1, 2, \cdots, 5$ gap genes, subject to the same *total* resource utilization as the WT, plotting the positional information vs. the rate at which expression patterns are changing at readout time. Networks that transmit 4 bits or more—as in the real embryo—are completely inaccessible using only one or two gap genes, even though these networks are allowed to utilize the same total number of molecules as in the optimal four gene networks above. With three gap genes the optimized networks can transmit a total information comparable to what is seen experimentally, but detailed analysis reveals that three–gene networks all have local defects where the positional error spikes above $5-10\%$ of the embryo length, in contrast to the much more uniform distribution of precision along the length of the real embryo (22); we can quantify this by looking at the variations in the positional error along the AP axis (Fig. 4A, inset). This failure of the three gene networks arises because they cannot realize a sufficient number of slopes or switches between high and low expression levels. Four gap genes thus are necessary to ensure that high positional information translates into defect–free patterning not just on average, but across the entire AP axis of every embryo (22). The marginal benefit of the putative fifth gap gene appears small and may not be sufficient to establish the required additional regulatory mechanisms or to maintain them at mutation–selection balance (60).

We can explore, in the same spirit, the role of the multiple maternal morphogens. In the fly embryo, the anterior (A, Bicoid), posterior (P, Nanos), and terminal (T, Torso-like) systems jointly regulate gap gene expression (24). In our model, we can remove one or two of these inputs and re–optimize all the parameters of the gap gene network, and find that there are moderate yet statistically significant losses in both positional information and stability (Fig. 4B). The impact of primary morphogen deletions is limited because the optimization algorithm adjusts the gap gene cross–regulation parameters to restore informative spatial patterns. This ability, however, disappears entirely if gap gene cross–regulation is not permitted and the gap gene network is feed forward (FF) only (light gray arrows in Box figure, Fig. 4B); in the absence of feedback, removal of each primary morphogen system is associated with a large decrease in positional information.

Figure 4B suggests that stable, high information patterns could be generated by utilizing all three maternal morphogens even without the ability of gap genes to regulate one another. But in the absence of cross–regulation, the time scale for variations in the pattern is determined solely by the intrinsic lifetime of the most stable species (mRNA). In contrast, feedback in the gap gene network allows for the emergence of longer time scales which both slow the variations and can reduce noise by temporal averaging (47); possible evidence for these effects has been discussed previously (53). Evolutionarily, adding gap gene cross–regulation creates variability in the rate–of–change phenotype that could additionally be selected for. Indeed, the WT–like solution of Fig. 2B falls close to the accessibility frontier of Fig. 4B, suggesting such a preference.

Lastly, we varied the maximal allowed strength of regulatory interactions, H_{max} (see Box 1), in our model. This parameter determines how strongly each individual input, either a morphogen or a gap gene acting via self– or cross–regulation, can impact the expression of a target gap gene. In simple microscopic models, this parameter measures the number of transcription factor molecules that bind cooperatively to their target sites as they regulate a single gene, and correlates with the steepness (or sensitivity) of the resulting induction curve. Optimizations presented so far used $H_{\text{max}} = 50$, sufficiently high not to impose any functional constraint. As H_{max} is lowered and the constraint kicks in, the optimal feed forward solution of Fig. 4B (dark blue) suffers large drops in encoded positional information (Fig. 4C); optimal feed forward architectures are thus heavily reliant on levels of effective cooperativity that appear unrealistic. Further, one might have been tempted to interpret Fig. 4B by saying that cross–regulation and multiple input morphogens provide alternative or even redundant paths to high information transmission, but we see that this degeneracy is lifted when we limit the effective cooperativity to more realistic levels. From an evolutionary perspective, gap

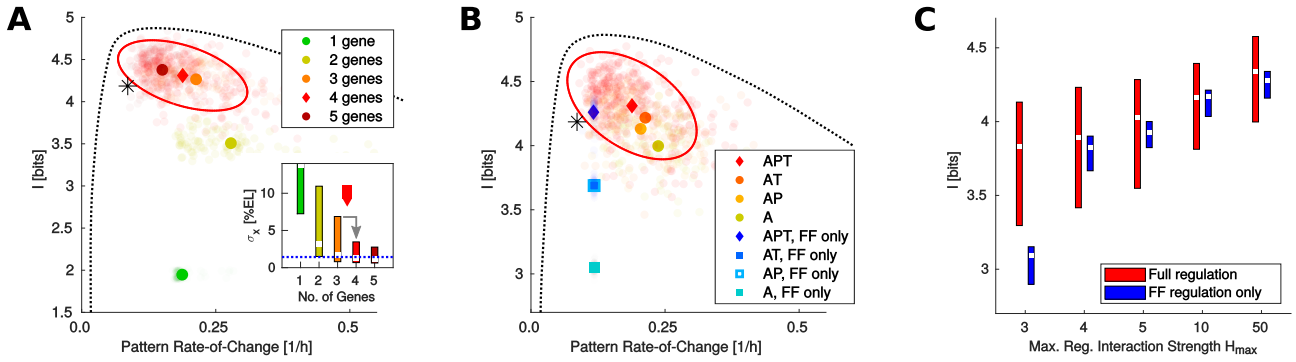


Fig. 4. Necessity and sufficiency of gap gene regulatory network mechanisms. **(A)** Optimal ensembles (transparent symbols = individual optimal solutions; solid symbols = ensemble medians) for networks with $1, 2, \cdots, 5$ gap genes (legend colors) optimized at the WT resource utilization (for reference, red diamond + red ellipse at 1 SD = WT–like optimal ensemble from Fig. 3B). Solutions delineate the accessibility frontier (dotted black line for visual guidance) in positional information (I) vs. pattern rate–of–change (RoC) plane. **(Inset)** While the median positional error (white squares) plateaus for optimal networks with three gap genes or more, the variability in positional error (ribbons denote 0.1- and 0.9-quantiles across AP positions in individual embryos) significantly shrinks only with 4 gap genes or more (red arrow). **(B)** Optimal ensembles for networks responding to different subsets of the three morphogens (A = anterior; P = posterior; T = terminal; red/yellow circle symbols = ensemble median; red dots, diamond, ellipse = WT–like ensemble as in A). Optimal ensembles for purely feed forward networks ("FF only", i.e., no gap gene self– or cross–regulation) denoted in bluish hues (legend). **(C)** Positional information in optimal ensembles with (red; white squares and ribbons denote median and 0.1–0.9-quantile intervals, respectively) or without (blue; FF networks) gap gene self– and cross–regulation (legend), for different maximal regulatory strength, H_{max}. Compared to feed forward networks, full regulation supports higher-information solutions, particularly at lower values for H_{max}.

gene cross–regulation therefore is favourable for two reasons: first, it generates temporally stable phenotypes at the accessibility frontier (as in Fig. 4B); and second, it permits high information solutions also in networks where the strength of individual regulatory interactions is limited (as in Fig. 4C).

Discussion

The idea that living systems can approach fundamental physical limits to their performance, and hence optimality, goes back at least to explorations of the diffraction limit in insect eyes and the ability of the human visual system to count single photons (6). The specific idea that biological systems optimize information transmission emerged shortly after Shannon's formulation of information theory, in the context of neural coding (7, 61). Despite this long history, most optimality analyses in biological systems have been carried out in very simplified contexts, using functional models with a small number of parameters. Here we have instantiated these ideas in a much more realistic setup, using mechanistic models for genetic regulatory networks that permit direct interpretation in terms of molecular mechanisms and interactions.

We focused on the *Drosophila* gap gene system, one of the paradigms for developmental biology and for physical precision measurements in living systems (62). Our work extends previous mathematical models of this system (26–34, 63–66), as well as attempts to predict it *ab initio* (67–69). In contrast to previously studied models, we systematically incorporate the unavoidable physical sources of noise, highlighting how patterning precision can emerge from noisy signals by a synergistic combination of multiple mechanisms. This novel contribution addresses a key question in developmental biology and provides a key mathematical ingredient for computing positional information. Crucially, we do not *fit* the parameters of the model to data, but rather *derive* them *ab initio* via optimization. In contrast to previous prediction attempts, our constraints and comparisons to data are not stylized, but fully quantitative and commensurate with the precision of the corresponding experiments.

We have found networks that maximize positional information with a limited number of molecules, and there is at least one local optimum quantitatively matching a large range of observations in the wild–type *Drosophila* system: its spatial patterns of expression and variability, the resulting decoding map, the molecular architecture of the network, as well as subtler biases in spatial resource utilization. Our optimization framework furthermore provides a platform for exploring the necessity and sufficiency of various network components that ensure maximal information transmission. Using this framework to deliver on our introductory questions, we have established that four gap genes appear necessary for defect–free patterning and that the apparent redundancy between the three maternal morphogens and gap gene cross–regulation is lifted under a developmental constraint on the strength of regulatory interactions.

Numerical optimization clearly is not evolutionary adaptation, yet its results nevertheless provide perspective on evolutionary questions. Discussions about the interplay of evolutionary optimization and developmental constraints, necessity versus contingency, and limits to selection have a venerable history (36, 70). Rather than discussing these questions in qualitative terms, here we explored the role of physical constraints and tradeoffs quantitatively, in the context of an expressive mechanistic model, using the powerful concepts of the random and the optimal ensembles. In the words of Jacob (55), the random ensemble delineates the space of the "possible." Within this space, our optimization principle acts as a proxy for strong selection for high positional information, thereby identifying a much more restricted optimal ensemble. It is surprising that this principle alone is sufficient to ensure that the optimal ensemble contains a solution very close to Jacob's "actual", the *Drosophila* gap gene network that we observe and measure.

ACKNOWLEDGMENTS. We thank Nicholas H Barton for his comments on the manuscript, Benjamin Zoller for inspiring discus-

sions, and Aleksandra Walczak and Curtis Callan for early discussions that shaped this work. This work was supported in part by the Human Frontiers Science Program, the Austrian Science Fund (FWF P28844), U.S. National Science Foundation, through the Center for the Physics of Biological Function (PHY-1734030); by National Institutes of Health Grants R01GM097275, U01DA047730, and U01DK127429; by the Simons Foundation; and by the John Simon Guggenheim Memorial Foundation.

1. GA Parker, JM Smith, Optimality theory in evolutionary biology. *Nature* **348**, 27–33 (1990).
2. B Walsh, M Lynch, *Evolution and selection of quantitative traits*. (Oxford University Press), (2018).
3. FM Rieke, DA Baylor, Single photon detection by rod cells of the retina. *Rev Mod Phys* **70**, 1027–1036 (1998).
4. HB Barlow, The size of ommatidia in apposition eyes. *J Exp Biol* **29**, 667–674 (1952).
5. HC Berg, EM Purcell, Physics of chemoreception. *Biophys J* **20**, 193–219 (1977).
6. W Bialek, *Biophysics: searching for principles*. (Princeton University Press, Princeton, NJ), (2012).
7. G Tkačik, W Bialek, Information processing in living systems. *Annu. Rev. Condens. Matter Phys.* **7**, 89–117 (2016).
8. Y Karklin, EP Simoncelli, Efficient coding of natural images with a population of noisy linear-nonlinear neurons. *Adv. neural information processing systems* **24**, 999 (2011).
9. BA Olshausen, DJ Field, Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature* **381**, 607–609 (1996).
10. EC Smith, MS Lewicki, Efficient auditory coding. *Nature* **439**, 978–982 (2006).
11. W Młynarski, JH McDermott, Ecological origins of perceptual grouping principles in the auditory system. *Proc. Natl. Acad. Sci.* **116**, 25355–25364 (2019).
12. RU Ibarra, JS Edwards, BO Palsson, Escherichia coli k-12 undergoes adaptive evolution to achieve in silico predicted optimal growth. *Nature* **420**, 186–189 (2002).
13. A Tero, et al., Rules for biologically inspired adaptive network design. *Science* **327**, 439–442 (2010).
14. G Tkačik, CG Callan, W Bialek, Information flow and optimization in transcriptional regulation. *Proc. Natl. Acad. Sci.* **105**, 12265–12270 (2008).
15. Y Savir, E Noor, R Milo, T Tlusty, Cross-species analysis traces adaptation of rubisco toward optimality in a low-dimensional landscape. *Proc. Natl. Acad. Sci.* **107**, 3475–3480 (2010).
16. ZD Blount, RE Lenski, JB Losos, Contingency and determinism in evolution: Replaying life's tape. *Science* **362**, eaam5979 (2018).
17. C Nüsslein-Volhard, E Wieschaus, Mutations affecting segment number and polarity in drosophila. *Nature* **287**, 795–801 (1980).
18. J Jaeger, The gap gene network. *Cell. Mol. Life Sci.* **68**, 243–274 (2011).
19. J Briscoe, S Small, Morphogen rules: design principles of gradient-mediated embryo patterning. *Development* **142**, 3996–4009 (2015).
20. PA Lawrence, , et al., *The making of a fly: the genetics of animal design*. (Blackwell Scientific Publications Ltd), (1992).
21. L Wolpert, Positional information and the spatial pattern of cellular differentiation. *J. Theor. Biol.* **25**, 1–47 (1969).
22. JO Dubuis, G Tkačik, EF Wieschaus, T Gregor, W Bialek, Positional information, in bits. *Proc. Natl. Acad. Sci.* **110**, 16301–16308 (2013).
23. G Tkačik, JO Dubuis, MD Petkova, T Gregor, Positional information, positional error, and readout precision in morphogenesis: a mathematical framework. *Genetics* **199**, 39–59 (2015).
24. MD Petkova, G Tkačik, W Bialek, EF Wieschaus, T Gregor, Optimal decoding of cellular identities in a genetic network. *Cell* **176**, 844–855 (2019).
25. G Tkačik, T Gregor, The many bits of positional information. *Development* **148**, dev176065 (2021).
26. L Sánchez, D Thieffry, A logical analysis of the drosophila gap-gene system. *J. theoretical Biol.* **211**, 115–141 (2001).
27. J Jaeger, et al., Dynamic control of positional information in the early drosophila embryo. *Nature* **430**, 368–371 (2004).
28. J Jaeger, et al., Dynamical analysis of regulatory interactions in the gap gene system of drosophila melanogaster. *Genetics* **167**, 1721–1737 (2004).
29. TJ Perkins, J Jaeger, J Reinitz, L Glass, Reverse engineering the gap gene network of *drosophila melanogaster*. *PLoS Comput. Biol.* **2**, e51 (2006).
30. Manu, et al., Canalization of gene expression and domain shifts in the drosophila blastoderm by dynamical attractors. *PLoS Comput. Biol.* **5**, e1000303 (2009).
31. Manu, et al., Canalization of gene expression in the drosophila blastoderm by gap gene cross regulation. *PLoS Biol.* **7**, e1000049 (2009).
32. M Ashyraliyev, et al., Gene circuit analysis of the terminal gap gene huckebein. *PLoS Comput. Biol.* **5**, e1000548 (2009).
33. B Verd, A Crombach, J Jaeger, Dynamic maternal gradients control timing and shift-rates for drosophila gap gene expression. *PLOS Comput. Biol.* **13**, e1005285 (2017).
34. B Verd, et al., A damped oscillator imposes temporal order on posterior gap gene expression in drosophila. *PLOS Biol.* **16**, e2003174 (2018).
35. R Seyboldt, et al., Latent space of a small genetic network: Geometry of dynamics and information. *Proc. Natl. Acad. Sci.* **119**, e2113651119 (2022).
36. JM Smith, et al., Developmental constraints and evolution: a perspective from the mountain lake conference on development and evolution. *The Q. Rev. Biol.* **60**, 265–287 (1985).
37. SC Little, M Tikhonov, T Gregor, Precise developmental gene expression arises from globally stochastic transcriptional activity. *Cell* **154**, 789–800 (2013).
38. B Zoller, SC Little, T Gregor, Diverse spatial expression patterns emerge from unified kinetics of transcriptional bursting. *Cell* **175**, 835–847 (2018).
39. T Gregor, DW Tank, EF Wieschaus, W Bialek, Probing the limits to positional information. *Cell* **130**, 153–164 (2007).
40. TR Sokolowski, G Tkačik, Optimizing information flow in small genetic networks. iv. spatial coupling. *Phys. Rev. E* **91**, 062710 (2015).
41. AM Walczak, G Tkačik, W Bialek, Optimizing information flow in small genetic networks. ii. feed-forward interactions. *Phys. Rev. E* **81**, 041905 (2010).
42. R Grah, B Zoller, G Tkačik, Nonequilibrium models of optimal enhancer function. *Proc. Natl. Acad. Sci.* **117**, 31614–31622 (2020).
43. G Tkačik, T Gregor, W Bialek, The role of input noise in transcriptional regulation. *PLoS ONE* **3**, e2774 (2008).
44. PS Swain, MB Elowitz, ED Siggia, Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proc. Natl. Acad. Sci.* **99**, 12795–12800 (2002).
45. JO Dubuis, R Samanta, T Gregor, Accurate measurements of dynamics and reproducibility in small genetic networks. *Mol. systems biology* **9**, 639 (2013).
46. G Tkačik, AM Walczak, W Bialek, Optimizing information flow in small genetic networks. *Phys. Rev. E* **80**, 031920 (2009).
47. G Tkačik, AM Walczak, W Bialek, Optimizing information flow in small genetic networks. iii. a self-interacting gene. *Phys. Rev. E* **85**, 041903 (2012).
48. P Hillenbrand, U Gerland, G Tkačik, Beyond the french flag model: exploiting spatial and gene regulatory interactions for positional information. *PLoS One* **11**, e0163628 (2016).
49. TR Sokolowski, AM Walczak, W Bialek, G Tkačik, Extending the dynamic range of transcription factor action by translational regulation. *Phys. Rev. E* **93**, 022404 (2016).
50. W Młynarski, M Hledík, TR Sokolowski, G Tkačik, Statistical analysis and optimality of neural systems. *Neuron* **109**, 1227–1241 (2021).
51. T Erdmann, M Howard, PR Ten Wolde, Role of spatial averaging in the precision of gene expression patterns. *Phys. Rev. Lett.* **103**, 258101 (2009).
52. TR Sokolowski, T Erdmann, PR Wolde, Mutual repression enhances the steepness and precision of gene expression boundaries. *PLoS Comput. Biol.* **8** (2012).
53. D Krotov, JO Dubuis, T Gregor, W Bialek, Morphogenesis at criticality. *Proc. Natl. Acad. Sci.* **111**, 3683–3688 (2014).
54. M Hledík, N Barton, G Tkačik, Accumulation and maintenance of information in evolution. *Proc. Natl. Acad. Sci.* **119**, e2123152119 (2022).
55. F Jacob, *The Possible and the Actual*. (University of Washington Press), (1982).
56. G Sella, AE Hirsh, The application of statistical physics to evolutionary biology. *Proc. Natl. Acad. Sci.* **102**, 9541–9546 (2005).
57. A Crombach, MA García-Solache, J Jaeger, Evolution of early development in dipterans: reverse-engineering the gap gene network in the moth midge clogmia albipunctata (psychodidae). *Biosystems* **123**, 74–85 (2014).
58. KR Wotton, et al., Quantitative system drift compensates for altered maternal inputs to the gap gene network of the scuttle fly megaselia abdita. *eLife* **4** (2015).
59. Y Goltsev, W Hsiong, G Lanzaro, M Levine, Different combinations of gap repressors for common stripes in anopheles and drosophila embryos. *Dev. Biol.* **275**, 435–446 (2004).
60. U Gerland, T Hwa, On the selection and evolution of regulatory dna motifs. *J. molecular evolution* **55**, 386–400 (2002).
61. HB Barlow, *Sensory mechanisms, the reduction of redundancy, and intelligence*. (HM Stationery Office, London), pp. 537–574 (1959).
62. T Gregor, HG Garcia, SC Little, The embryo as a laboratory: quantifying transcription in drosophila. *Trends Genet.* **30**, 364–375 (2014).
63. MD Schroeder, et al., Transcriptional control in the segmentation gene network of drosophila. *PLoS biology* **2**, e271 (2004).
64. E Segal, T Raveh-Sadka, M Schroeder, U Unnerstall, U Gaul, Predicting expression patterns from regulatory sequence in drosophila segmentation. *Nature* **451**, 535–540 (2008).
65. J Bieler, C Pozzorini, F Naef, Whole-embryo modeling of early segmentation in drosophila identifies robust and fragile expression domains. *Biophys. journal* **101**, 287–296 (2011).
66. T Duque, et al., Simulations of enhancer evolution provide mechanistic insights into gene regulation. *Mol. biology evolution* **31**, 184–200 (2014).
67. P François, V Hakim, ED Siggia, Deriving structure from evolution: metazoan segmentation. *Mol. systems biology* **3**, 154 (2007).
68. P François, ED Siggia, Predicting embryonic patterning using mutual entropy fitness and in silico evolution. *Development* **137**, 2385–2395 (2010).
69. JB Rothschild, P Tsimiklis, ED Siggia, P François, Predicting ancestral segmentation phenotypes from drosophila to anopheles using in silico evolution. *PLoS Genet.* **12**, e1006052 (2016).
70. AS Wilkins, *The evolution of developmental pathways*. (Sinauer Associates), (2002).