Morphogenesis at criticality

Dmitry Krotov[a,1], Julien O. Dubuis[a,b], Thomas Gregor[a], and William Bialek[a,1]

[a]Joseph Henry Laboratories of Physics, Lewis–Sigler Institute for Integrative Genomics, and [b]Howard Hughes Medical Institute, Princeton University, Princeton, NJ 08544

Contributed by William Bialek, January 17, 2014 (sent for review September 28, 2013)

Spatial patterns in the early fruit fly embryo emerge from a network of interactions among transcription factors, the gap genes, driven by maternal inputs. Such networks can exhibit many qualitatively different behaviors, separated by critical surfaces. At criticality, we should observe strong correlations in the fluctuations of different genes around their mean expression levels, a slowing of the dynamics along some but not all directions in the space of possible expression levels, correlations of expression fluctuations over long distances in the embryo, and departures from a Gaussian distribution of these fluctuations. Analysis of recent experiments on the gap gene network shows that all these signatures are observed, and that the different signatures are related in ways predicted by theory. Although there might be other explanations for these individual phenomena, the confluence of evidence suggests that this genetic network is tuned to criticality.

genetic networks | *Drosophila* embryo

Genetic regulatory networks are described by many parameters: the rate constants for binding and unbinding of transcription factors to their target sites along the genome, the interactions between these binding events and the rate of transcription, the lifetimes of mRNA and protein molecules, and more. Even with just two genes, each encoding a transcription factor that represses the other, changing parameters allows for several qualitatively different behaviors (1). Strong mutual repression between two genes can lead to bistability, or switch-like behavior, such that the system can be stable in both "on–off" and "off–on" states (2). On the other hand, if interactions are weak, the two interacting genes have just one stable state, and the expression levels in this state are controlled primarily by the inputs. Importantly, if we imagine smooth changes in the strength of the repressive interactions, the transition from graded response to switch-like behavior is not smooth: the behavior is qualitatively different depending on whether the relevant interactions are stronger or weaker than a critical value. Here we explore the possibility that the gap gene network in the *Drosophila* embryo might be tuned to such a critical point.

Early events in the fruit fly embryo provide an experimentally accessible example of many questions about genetic networks (3–5). Along the anterior–posterior axis, for example, information about the position of nuclei flows from primary maternal morphogens to the gap genes, shown in Fig. 1 (6–8), to the pair rule and segment polarity genes. Although the structure of the gap gene network is not completely known, there is considerable evidence that the transcription factors encoded by these genes are mutually repressive (6, 9–11). If we focus on a small region near the midpoint of the embryo (near $x/L = 0.47$), then just two gap genes, hunchback (Hb) and Krüppel (Kr), are expressed at significant levels, and this is repeated at a succession of crossing points or expression boundaries: Hb–Kr, Kr–Kni (knirps; $x/L = 0.57$), Kni–Gt (giant; $x/L = 0.66$), and Gt–Hb ($x/L = 0.75$), as we move from anterior to posterior. In each crossing region, it is plausible that the dynamics of the network is dominated by the interactions among just the pair of genes whose expression levels are crossing.

We argue that criticality in a system of two mutually repressive genes generates several clear, experimentally observable signatures. First, there should be nearly perfect anticorrelations between the fluctuations in the two expression levels. As a result, there are two linear combinations of the expression levels, or "modes," that have very different variances. Second, fluctuations in the large variance mode should have a significantly non-Gaussian distribution, whereas the small variance mode is nearly Gaussian. Third, there should be a dramatic slowing down of the dynamics along one direction in the space of possible expression levels. Finally, there should be correlations among fluctuations at distant points in the embryo. These signatures are related: The small variance mode will be the direction of fast dynamics, and under some conditions the large variance mode will be the direction of slow dynamics; the fast fluctuations should be nearly Gaussian, whereas the slow modes are non-Gaussian; and only the slow mode should exhibit long-ranged spatial correlations. We will see that all of these effects are found in the gap gene network.

We begin by analyzing a two-gene network to make precise what we mean by criticality. Crucially, this analysis shows that there are quantitative, experimentally observable signatures of criticality that do not depend on the molecular details of the network. We then analyze experiments on the gap gene network, searching for these experimental signatures.

Criticality in a Two-Gene Network

One of the most important results in statistical mechanics and dynamical systems theory is that the behavior of systems near a critical point is classifiable, and independent of many details (12, 13). In this spirit, we want to identify behaviors of a genetic network that would emerge near criticality. We focus on a network of two genes, because this is the simplest case and because, as emphasized above, the gap gene network is dominated by two genes at each of the crossing regions identified in Fig. 1. The analysis that follows is rather standard, but we make it explicit to clarify how our predictions follow from the idea of criticality itself, rather than depending on the many (largely unknown) molecular parameters.

We consider a broad class of models for a genetic regulatory circuit in which the rate at which gene products are synthesized depends on the concentration of all of the relevant transcription

Significance

Biological networks are described by many parameters, and the behavior of a network is qualitatively different (monostable, bistable, oscillating, etc.) in different parts of parameter space. Critical points and surfaces are the borders between such qualitatively different regimes, as with phase transitions in equilibrium thermodynamics. We argue that, as expected from the thermodynamic case, genetic regulatory networks should exhibit behaviors near criticality that are independent of most molecular details. Analyzing recent experiments on the gap gene network in the early *Drosophila* embryo, we find that these signatures of criticality can be seen, quantitatively. This raises the question of why evolution has tuned this network to such a special point in its parameter space.

Author contributions: This is a collaboration between theorists (D.K. and W.B.) and experimentalists (J.O.D. and T.G.), involving the analysis of previously published data. All authors contributed to all aspects of the paper, to the best of their abilities.

The authors declare no conflict of interest.

Freely available online through the PNAS open access option.

[1]To whom correspondence may be addressed. E-mail: wbialek@princeton.edu or dkrotov@princeton.edu.

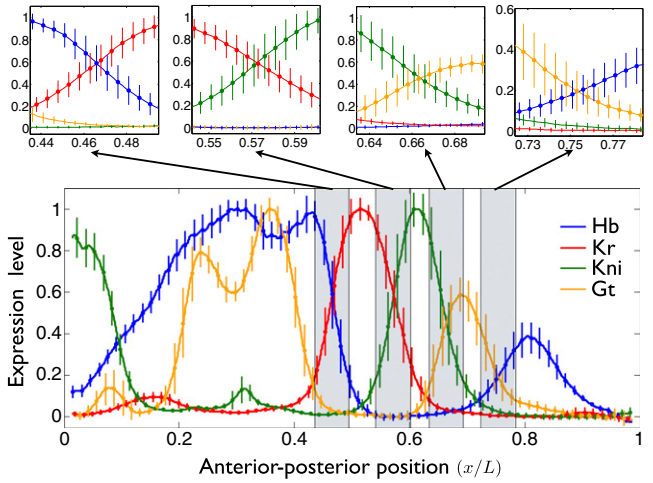

Fig. 1. Normalized gap gene expression levels in the early *Drosophila* embryo, from ref. 7. Measurements by simultaneous immunofluorescent staining of all four proteins, along the dorsal edge of the midsagittal plane of the embryo, 38–49 min into nuclear cycle 14; error bars are SDs across $N = 24$ embryos. (*Upper Left*) Expanded view of the shaded regions, near the crossings between Hb and Kr levels, where just these two genes have significant expression, and similarly for the Kr–Kni, Kni–Gt, and Gt–Hb crossings (*Upper*) from left to right.

factors, and the gene products are degraded. To simplify, we ignore delays, so that the rate at which the protein encoded by a gene is synthesized depends instantaneously on the other protein (transcription factor) concentrations, and we assume that degradation obeys first-order kinetics. We also focus on a single cell, leaving aside (for the moment) the role of diffusion. Then, by choosing our units correctly we can write the dynamics for the expression levels of two interacting genes as

$$\tau_1 \frac{dg_1}{dt} = f_1(c; g_1, g_2) - g_1 + \xi_1 \qquad [1]$$

$$\tau_2 \frac{dg_2}{dt} = f_2(c; g_1, g_2) - g_2 + \xi_2, \qquad [2]$$

where g_1 and g_2 are the normalized expression levels of the two genes, τ_1 and τ_2 are the lifetimes of the proteins, and c represents the external (maternal) inputs. The functions f_1 and f_2 are the "regulation functions" that express how the transcriptional activity of each gene depends on the expression level of all of the other genes; with our choice of units, the regulation function runs between zero (gene off) and 1 (full induction). All of the molecular details of transcriptional regulation are hidden in the form of these regulation functions (14), which we will not need to specify. Finally, the random functions ξ_1 and ξ_2 model the effects of noise in the system.

If the interactions are weak, then for any value of the external inputs c there is a single steady-state response, defined by expression levels $\bar{g}_1(c)$ and $\bar{g}_2(c)$. We can check whether this hypothesis is consistent by asking what happens to small changes in the expression levels around this steady state. We write $g_1 = \bar{g}_1 + \delta g_1$, and similarly for g_2, and then expand Eqs. **1** and **2** assuming that δg_1 and δg_2 are small. The result is

$$\frac{d}{dt}\begin{bmatrix}\delta g_1 \\ \delta g_2\end{bmatrix} = \begin{bmatrix}-\Gamma_1 & \gamma_{12} \\ \gamma_{21} & -\Gamma_2\end{bmatrix}\begin{bmatrix}\delta g_1 \\ \delta g_2\end{bmatrix} + \begin{bmatrix}\eta_1 \\ \eta_2\end{bmatrix}. \qquad [3]$$

Here Γ_1 and Γ_2 are effective decay rates for the two proteins, which must be positive if the steady state we have identified is stable. The parameter γ_{12} reflects the incremental effect of gene 2 on gene 1: $\gamma_{12} < 0$ means that the protein encoded by gene 2 is a repressor of gene 1, whereas $\gamma_{12} > 0$ means that the protein encoded by gene 2 is an activator of gene 1, and similarly for γ_{21}. The noise terms η_1 and η_2 play the same role as ξ_1 and ξ_2, but have different normalization.

If the steady state that we have identified is stable, then the matrix

$$\hat{M} \equiv \begin{bmatrix}-\Gamma_1 & \gamma_{12} \\ \gamma_{21} & -\Gamma_2\end{bmatrix} \qquad [4]$$

must have two eigenvalues with negative real parts. This is guaranteed if the interactions are weak ($\gamma_{12}, \gamma_{21} \to 0$), but as the interactions become stronger it is possible for one of the eigenvalues to vanish. This is the critical point. Notice that we can define the critical point without giving a microscopic description of all of the interactions that determine the form of the regulation functions.

The linearized Eq. **3** predicts that the relaxation of average expression levels to their steady states can be written as combinations of two exponential decays,

$$\begin{bmatrix}\langle g_1(t)\rangle \\ \langle g_2(t)\rangle\end{bmatrix} = \begin{bmatrix}\bar{g}_1(c) \\ \bar{g}_2(c)\end{bmatrix} + \begin{bmatrix}A_{1s} & A_{1f} \\ A_{2s} & A_{2f}\end{bmatrix}\begin{bmatrix}e^{\Lambda_s t} \\ e^{\Lambda_f t}\end{bmatrix}, \qquad [5]$$

where Λ_s and Λ_f are the slow and fast eigenvalues of $\hat{M}$. Thus, although we measure the two expression levels, there are combinations of these expression levels—directions $(\delta g_s, \delta g_f)$ in the (g_1, g_2) plane—that provide more natural coordinates for the dynamics, such that motion along each direction is a single exponential function of time. As we approach criticality, the dynamics along the slow direction becomes very slow, so that $\Lambda_s \to 0$.

The linearized Eq. **3** also predicts the fluctuations around the steady state. As we approach criticality, things simplify, and we find the covariance matrix

$$\begin{bmatrix}\langle(\delta g_1)^2\rangle & \langle\delta g_1 \delta g_2\rangle \\ \langle\delta g_1 \delta g_2\rangle & \langle(\delta g_2)^2\rangle\end{bmatrix} \to \sigma^2 \begin{bmatrix}1 & \Gamma_1/\gamma_{12} \\ \Gamma_1/\gamma_{12} & (\Gamma_1/\gamma_{12})^2\end{bmatrix}, \qquad [6]$$

where σ^2 is the variance in the expression level of the first gene. As with the dynamics, there are two "natural" directions in the (g_1, g_2) plane corresponding to eigenvectors of this covariance matrix (principal components). In this linear approximation, the critical point is the point where we "lose" one of the dimensions, and the fluctuations in the two expression levels become perfectly correlated or anticorrelated. In addition, the direction with small fluctuations is the direction of fast relaxation.

Eqs. **1** and **2** refer to a single cell, or, in the case of the syncitial *Drosophila* embryo, to a single nucleus. If we allow that the proteins encoded by genes 1 and 2 can move among neighboring nuclei, we should write coupled equations for the dynamics in these neighbors. A simple approximation is to imagine that nuclei are sufficiently dense that a plot of protein concentration vs. nuclear position becomes continuous, as with the profiles in Fig. 1. If the dynamics of molecular motion along the length of the embryo is diffusive, then the natural generalization of Eqs. **1** and **2** becomes

$$\tau_1 \frac{\partial g_1}{\partial t} = D\frac{\partial^2 g_1}{\partial x^2} + f_1(c; g_1, g_2) - g_1 + \xi_1 \qquad [7]$$

$$\tau_2 \frac{\partial g_2}{\partial t} = D\frac{\partial^2 g_2}{\partial x^2} + f_2(c; g_1, g_2) - g_2 + \xi_2, \qquad [8]$$

where for simplicity we assume that the diffusion constants of the two species are the same. These equations have the same qualitative form as those considered by Turing, but we consider

a system that does not cross the transition to spontaneous pattern formation considered in Turing's work (15).

The linearized version of Eqs. **7** and **8**, analogous to Eq. **3**, is

$$\frac{\partial}{\partial t}\begin{bmatrix}\delta g_1\\ \delta g_2\end{bmatrix} = D\frac{\partial^2}{\partial x^2}\begin{bmatrix}\delta g_1\\ \delta g_2\end{bmatrix} + \hat{M}\begin{bmatrix}\delta g_1\\ \delta g_2\end{bmatrix} + \begin{bmatrix}\eta_1\\ \eta_2\end{bmatrix}, \qquad [9]$$

where $\hat{M}$ is the same matrix as in Eq. **4**. In general this matrix depends on position x, and it is not easy to exhibit the solution of these equations in closed form. However, if we focus on small regions of the embryo, such as the crossings where we expect our two-gene model is most useful, we can make the approximation that $\hat{M}$ is constant. At the end of our discussion, we will try to stitch together these local approximations into a more global view.

If the noise sources driving the randomness in gene expression arise from local processes, as seems plausible, then the "forces" $\eta_{1,2}(x,t)$ should be uncorrelated at different positions, although the output expression levels can become correlated via diffusion. These correlations again are expressed most naturally in terms of the slow and fast directions in the (g_1, g_2) plane. Concretely, we can solve Eq. **9** in the approximation of constant $\hat{M}$ to give

$$\langle \delta g_s(x,t)\delta g_s(x',t)\rangle \propto \exp\left(-\frac{1}{\xi_s}|x-x'|\right), \qquad [10]$$

where the "correlation length" $\xi_s = \sqrt{D/(-\Lambda_s)}$; similarly, for the fast mode we have $\xi_f = \sqrt{D/(-\Lambda_f)}$. Thus, connected to the slowing of the dynamics at criticality ($\Lambda_s \to 0$), we should see spatial correlations extending over long distances ($\xi_s \to \infty$).

To summarize, the approach to criticality generates slow dynamics ($\Lambda_s \to 0$ in Eq. **5**), strong local (anti-)correlations (Eq. **6**), and long-ranged correlations in space ($\xi_s \to \infty$ in Eq. **10**). These predictions are connected, so that the anticorrelated combination of expression levels is the mode that exhibits long-ranged correlations, and this mode is plausibly the same mode that has slow dynamics. As we will see below, there is one more somewhat subtle prediction, namely that the distribution of fluctuations along the slow mode will be non-Gaussian.

Local Correlations

Testing the predictions of criticality requires measuring the time dependence of gap gene expression levels, with an accuracy better than the intrinsic noise levels of the system. Absent live movies of the expression levels, the progress of cellularization provides a clock that can be used to mark the time during nuclear cycle 14 at which an embryo was fixed (16), accurate to within 1 min (7). Fixed embryos, with immunofluorescent staining of the relevant proteins, thus provide a sequence of snapshots that can be placed accurately along the time axis of development. Immunofluorescent staining itself provides a measurement of relative protein concentrations that is accurate to within $\sim 3\%$ of the maximum expression levels in the embryo (7).

In Fig. 2*A* we show the correlations between fluctuations in pairs of gap genes at each position. Gap gene expression levels plateau at ~ 40 min into nuclear cycle 14 (7), and the mean expression levels are shown as a function of anterior–posterior position in Fig. 1. At each position we can look across the many embryos in our sample, and analyze the fluctuations around the mean, as in ref. 8. We see that, precisely in the "crossing region" where Hb and Kr are the only genes with significant expression (marked "A" in Fig. 2*A*), the correlation coefficient approaches $C = -1$, perfect (anti-)correlation, as expected at criticality. This pattern repeats at the crossing between Kr and Kni (B), at the crossing between Kni and Gt (C), and, perhaps less perfectly, at the crossing between Gt and Hb (D). [Because we observe substantial anticorrelations both in the Gt–Hb pair and in the Hb–Kni pair, it is likely that the system is more nearly 3D in the neighborhood of crossing D, so that no single pair can achieve perfect correlation.]

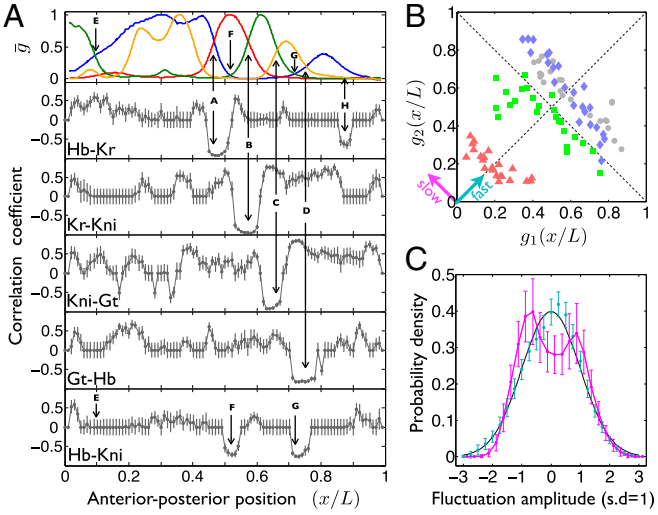

Fig. 2. Fluctuations in gap gene expression levels. (*A*) Pairwise correlation coefficients between fluctuations in the different gap genes vs. anterior–posterior position, from the same data as in Fig 1 (7). Mean expression levels at top to guide the eye, colors as in Fig 1; error bars are from bootstrap analysis. Correlations which are not significant at $p = 0.01$ are shown as zero. Major crossing points of the mean expression profiles are labeled A, B, C, and D; other points marked as described in the text. (*B*) Scatter plot of expression levels for pairs of genes in individual embryos: Kr vs. Hb at point A (gray circles), Kni vs. Kr at point B (blue diamonds), Gt vs. Kni at point C (green squares), and Hb vs. Gt at point D (red triangles). (*C*) Probability distribution of expression fluctuations. In each of the crossing regions from Fig 1, we form the combinations δg_f (fast modes, cyan) and δg_s (slow modes, magenta), and normalize the fluctuations across embryos to have unit variance at each position. Data from all four regions are pooled to estimate the distributions; error bars are from random divisions of the set of 24 embryos. Gaussian distribution (black) shown for comparison.

These strong anticorrelations are shown explicitly in Fig. 2*B*, where we plot the two relevant gene expression levels against one another at each crossing point. In all cases, the direction of small fluctuations is along the positive diagonal, whereas the large fluctuations are along the negative diagonal.

It is important that the strong anticorrelations tell us something about the underlying network, rather than being a necessary (perhaps even artifactual) corollary of the mean expression profiles. A notable feature of Fig. 2*A* thus is what happens away from the major crossing points. There is a Hb–Kni crossing at $x/L = 0.1$ (E), but this does not have any signature in the correlations, perhaps because spatial variations in expression levels at this point are dominated by maternal inputs rather than being intrinsic to the gap gene network (17, 18). This is evidence that we can have crossings without correlations, and we can also have correlations without crossings, as with Hb and Kr at point H; interestingly, H marks the point where an additional posterior Kr stripe appears during gastrulation (19, 20). We also note that strong correlations can appear when expression levels are very small, as with Hb and Kni at points F and G; there also are extended regions of positive Kr–Kni and Kni–Gt correlations in parts of the embryo where the expression levels of Kr and Kni both are very low. Taken together, these data indicate that the pattern of correlations is not simply a reflection of the mean spatial profiles, but an independent measure of network behavior.

It should be admitted that strong anticorrelations can be explained without criticality. As an example, if A represses B, and the dominant source of noise in the expression of B is the fluctuating concentration of A, then A and B will be perfectly anticorrelated. Alternatively, if both A and B respond to a common input (e.g., as

the gap genes respond to maternal morphogens), then if the dominant source of noise in the output is the response to variations in this input, again one can have strong (anti-)correlations. Although we could try to eliminate particular versions of these models as explanations for what we see in the gap genes, it seems more useful to note that none of these models has a natural way of generating the other signatures of criticality, to which we now turn.

Non-Gaussianity

If we transform Eq. **3** to a description in terms of the fast and slow modes g_f and g_s, then precisely at criticality there is no "restoring force" for fluctuations in g_s and formally the variance σ^2 in Eq. **6** should diverge. This divergence is stopped by higher-order terms in the expansion of the regulation functions around the steady state, and this leads to a non-Gaussian distribution of fluctuations in g_s. Although the data are limited, we do find, as shown in Fig. 2*C*, that fluctuations in the small variance (fast) direction are almost perfectly Gaussian, whereas the large variance (slow) direction shows significant departures from Gaussianity, in the expected direction.

Dynamics

The time dependence of Hb and Kr expression levels during nuclear cycle 14 is shown, at the crossing point $x/L = 0.47$, in Fig. 3. Criticality predicts that if we take a linear combination of these expression levels corresponding to the direction of small fluctuations in Fig. 2*B* (cyan), then we will see relatively fast dynamics, and this is what we observe. In contrast, if we project onto the direction of large fluctuations (magenta), we see only very slow variations over nearly 1 h. Indeed, the expression level along this slow direction seems almost to diffuse freely, with growing variance rather than systematic evolution. Thus, strong (anti-)correlations are accompanied by a dramatic slowing of the dynamics along one direction in the space of possible expression levels, and a similar pattern is found at each of the crossings, Kr–Kni, Kni–Gt, and Gt–Hb. Again, this is consistent with what we expect for two-gene systems at criticality.

If we move along the anterior–posterior axis in the vicinity of the crossing point, the sum of expression levels of two genes, which is proportional to the fast mode, remains approximately constant, whereas the difference, which is proportional to the slow mode, changes. Therefore, the dynamics of the slow mode, shown in Fig. 3, will generate motion of the pattern along the anterior–posterior axis. This slow shift is well known (7, 21).

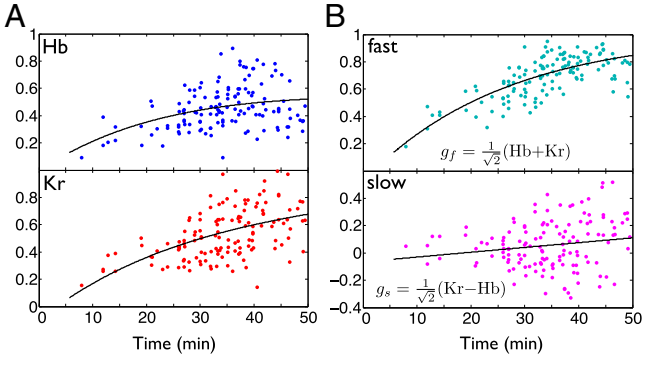

Fig. 3. Dynamics at the Hb–Kr crossing point. (*A*) Normalized expression levels of the individual genes, plotted vs. time during nuclear cycle 14; data from ref. 7. (*B*) Linear combinations of expression levels corresponding to the small variance (fast) and large variance (slow) directions in Fig 2*B*. Curves are best-fit single exponentials for each mode and are also shown projected back into the individual expression levels in *A*. Eigenvalues, as in Eq. **5**, are $\Lambda_f = -0.04 \pm 0.01$ min^{-1} and $\Lambda_s = -0.002 \pm 0.007$ min^{-1}.

Spatial Correlations

The slow dynamics associated with criticality implies that correlations should extend over long distance in space. In particular, as we approach criticality, the relaxation rate Λ_s vanishes and the associated correlation length ξ_s in Eq. **10** becomes infinitely long, in practice being limited by the size L of the embryo itself. Searching for these long-ranged correlations is complicated by the fact that the system is inhomogeneous, but we have a built-in control, because we should see the long-ranged correlations only in the slow, large variance mode δg_s, and not in δg_f. This also helps us discriminate against systematic errors that might have generated spurious correlations.

In Fig. 4*A* we show the normalized correlation function

$$C_{ss}(x,y) = \frac{\langle \delta g_s(x)\delta g_s(y)\rangle}{\left(\left\langle [\delta g_s(x)]^2\right\rangle \left\langle [\delta g_s(y)]^2\right\rangle\right)^{1/2}}, \quad [11]$$

with x held fixed at the Hb–Kr crossing and y allowed to vary. We see that this correlation function is essentially constant throughout the crossing region. In contrast, the same correlation computed for the fast mode decays rapidly, with a length constant $\xi/L \sim 0.02$, just a few nuclear spacings along the anterior–posterior axis.

The dominant slow mode corresponds to different combinations of expression levels in different regions of the embryo. Generally, we can write the slow mode as a weighted sum of the different expression levels,

$$g_s(x) = \sum_{i=1}^{4} W_i(x) g_i(x), \quad [12]$$

where we label the gap genes $i = 1$ for Hb, $i = 2$ for Kr, $i = 3$ for Kni, and $i = 4$ for Gt. Near the Hb–Kr crossing, labeled "A" in Fig. 2*A*, we have $W_i \approx W_i^A$, where W_i^A are the weights that give us the antisymmetric combination of Hb and Kr, as drawn in Fig. 2*B*: $W_1^A = -1/\sqrt{2}$, $W_2^A = 1/\sqrt{2}$, and $W_3^A = W_4^A = 0$. Similarly, near the Kr–Kni crossing, labeled "B" in Fig. 2*A*, we have $W_i \approx W_i^B$, with $W_2^B = -1/\sqrt{2}$, $W_3^B = 1/\sqrt{2}$, and $W_1^B = W_4^B = 0$, and this generalizes to crossing regions C and D. Using these weights, we obtain approximations to the slow mode,

$$\begin{bmatrix} g_s^A(y) \\ g_s^B(y) \\ g_s^C(y) \\ g_s^D(y) \end{bmatrix} = \sum_{i=1}^{4} \begin{bmatrix} W_i^A g_i(y) \\ W_i^B g_i(y) \\ W_i^C g_i(y) \\ W_i^D g_i(y) \end{bmatrix}, \quad [13]$$

and we expect that these approximations are accurate in their respective crossing regions. Now we can test for correlations over longer distances by computing, for example,

$$C_{ss}^{AB}(x,y) = \frac{\langle \delta g_s^A(x)\delta g_s^B(y)\rangle}{\left(\left\langle [\delta g_s^A(x)]^2\right\rangle \left\langle [\delta g_s^B(y)]^2\right\rangle\right)^{1/2}}, \quad [14]$$

holding $x/L = 0.47$ in the crossing region A while letting y vary through the crossing region B, and similarly for $C_{ss}^{AC}(x,y)$ and $C_{ss}^{AD}(x,y)$. The results of this analysis are shown in Fig. 4*B*; note that $C_{ss}^{AA}(x,y)$ is the correlation we have plotted in Fig. 4*A*.

Fig. 4*B* shows that the slow mode is correlated over very long distances. We can see, for example, in C_{ss}^{AC}, correlations between fluctuations in expression level at the Hb–Kr crossing region and at the Kni–Gt crossing region, despite the fact that these regions are separated by $\sim 20\%$ of the length of the embryo and have no significantly expressed genes in common. These peaks in the correlation functions appear also at points anterior to the crossing regions, presumably at places where our approximations in Eq. **13** come close to some underlying slow mode in the

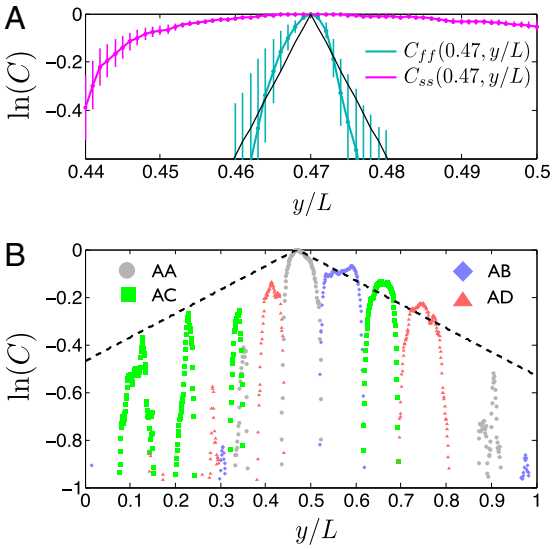

Fig. 4. Spatial correlations of the fluctuations in gene expression. (*A*) Autocorrelations of slow (magenta) and fast (cyan) modes near the Hb–Kr crossing, as defined in Eq. **11**. Line shows a fit to the exponential form in Eq. **10** with $\xi/L = 0.015 \pm 0.002$. (*B*) Correlations between the slow mode at the Hb–Kr crossing and slow modes at other crossings, as defined in Eq. **14**, indicated by the colors. All correlation functions are evaluated at $x/L = 0.47$ with y varying as shown. Middle peak is the same (auto-)correlation function C_{ss} as in *A*. Dashed line corresponds to Eq. **10** with $\xi = L$.

network. The pattern of correlations has an envelope corresponding to an exponential decay with correlation length $\xi = L$ (dashed line in Fig. 4*B*), and similar results are obtained for the correlation functions C_{ss}^{BC}, C_{ss}^{BD}, C_{ss}^{CD}, etc.. This means that fluctuations in expression level are correlated along essentially the entire length of the embryo, as expected at criticality.

Discussion

To summarize, the patterns of gap gene expression in the early *Drosophila* embryo exhibit several signatures of criticality: near-perfect anticorrelations of fluctuations in the expression levels of different genes at the same point, non-Gaussian distributions of the fluctuations in the large variance modes, slowing down of the dynamics of these modes, and spatial correlations of the slow modes that extend over a large fraction of the embryo. Although each of these observations could have other explanations, the confluence of results strikes us as highly suggestive. Note that we have focused on aspects of the data that are connected to the hypothesis of criticality in a very general way, independent of other assumptions, rather than trying to build a model for the entire network or fit parameters.

The possibility that biological systems might be poised near critical points, often discussed in the past (22), has been reinvigorated by new data and analyses on systems ranging from ensembles of amino acid sequences to networks of neurons to flocks of birds (23). Although we would like to avoid too much speculation, it does seem natural to ask how being tuned to a critical point could be functionally advantageous for the embryo.

Criticality may have a special meaning in the context of pattern formation. Usually, if we have a network of biochemical reactions and diffusion, there are length scales set by combinations of rate constants and diffusion constants, and these length scales set the size of features in the pattern. In contrast, the patterns formed in embryonic development are thought to scale in proportion to the size of the embryo (24, 25). One way of expressing the problem of scaling is that moving the boundary set by the posterior end of the egg must have an effect on the patterns of gene expression even near the anterior end, whereas conventional models predict that perturbations will penetrate only a short distance ξ from the point where they are applied, comparable to the distance over which fluctuations in expression are correlated. At criticality, however, this correlation length becomes the size of the embryo itself, as in Fig. 4. Whereas this does not provide an explicit mechanism for scaling, it may provide a different point of view on the problem.

At a critical point, fluctuations are large, and it might seem that this is a disadvantage. However, criticality also implies a very large gain in response to small changes in the input signals (e.g., the maternal morphogens), and the long time scales can serve to average out components of the noise in the system. Analysis of simple models shows that there is a regime in which transcriptional regulatory systems can maximize the flow of information from inputs to outputs by operating near a critical point (26). There is evidence that noise levels in the gap gene system are nearly as low as possible given the finite concentrations of the relevant molecules and the time available for decisions to be made (27), and that the representation of positional information by the gap genes is optimized by matching the distribution of input signals to the noise characteristics of the network (8, 28). Tuning to a critical point may provide another layer of optimization along the same functional axis. If correct, this would place criticality in the gap gene network alongside other examples of optimizing signal-to-noise or information flow in biological systems, ranging from molecule counting in bacterial chemotaxis (29) to photon counting in human vision (30), among many other examples (25).

Even leaving aside the possibility of criticality, the aspects of the data that we have described here are not at all what we would see if the gap gene network is described by generic parameter values. There must be something about the system that is finely tuned to generate such large differences in the time scales for variation along different dimensions in the space of expression levels, or to ensure that correlations are so nearly perfect and extend over such long distances.

ACKNOWLEDGMENTS. We thank S. Blythe, K. Doubrovinski, O. Grimm, J. J. Hopfield, S. Little, M. Osterfield, A. M. Polyakov, M. Tikhonov, G. Tkačik, and A. Walczak for helpful discussions. We are especially grateful to E. F. Wieschaus for discussions both of the ideas and their presentation. This work was supported in part by National Science Foundation Grants PHY-0957573 and CCF-0939370, by National Institutes of Health Grants P50 GM071508 and R01 GM097275, by the Howard Hughes Medical Institute, by the W. M. Keck Foundation, and by Searle Scholar Award 10-SSP-274 (to T.G.).

1. Hasty J, McMillen D, Isaacs F, Collins JJ (2001) Computational studies of gene regulatory networks: In numero molecular biology. *Nat Rev Genet* 2(4): 268–279.
2. Gardner TS, Cantor CR, Collins JJ (2000) Construction of a genetic toggle switch in Escherichia coli. *Nature* 403(6767):339–342.
3. Nüsslein-Volhard C, Wieschaus EF (1980) Mutations affecting segment number and polarity in Drosophila. *Nature* 287(5785):795–801.
4. Lawrence PA (1992) *The Making of a Fly: The Genetics of Animal Design* (Blackwell Scientific, Oxford).
5. Gerhardt J, Kirschner M (1997) *Cells, Embryos, and Evolution* (Blackwell Science, Malden, MA).
6. Jaeger J (2011) The gap gene network. *Cell Mol Life Sci* 68(2):243–274.
7. Dubuis JO, Samanta R, Gregor T (2013) Accurate measurements of dynamics and reproducibility in small genetic networks. *Mol Syst Biol* 9:639.
8. Dubuis JO, Tkačik G, Wieschaus EF, Gregor T, Bialek W (2013) Positional information, in bits. *Proc Natl Acad Sci USA* 110(41):16301–16308.
9. Jäckle H, Tautz D, Schuh R, Seifert E, Lehmann R (1986) Cross–regulatory interactions among the gap genes of Drosophila. *Nature* 324(6098):668–670.
10. Treisman J, Desplan C (1989) The products of the Drosophila gap genes hunchback and Krüppel bind to the hunchback promoters. *Nature* 341(6240): 335–337.
11. Hoch M, Seifert E, Jäckle H (1991) Gene expression mediated by cis-acting sequences of the Krüppel gene in response to the Drosophila morphogens bicoid and hunchback. *EMBO J* 10(8):2267–2278.

12. Guckenheimer J, Holmes P (1983) *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields* (Springer, Berlin).
13. Parisi G (1988) *Statistical Field Theory* (Addison–Wesley, Redwood City, CA).
14. Bintu L, et al. (2005) Transcriptional regulation by the numbers: Models. *Curr Opin Genet Dev* 15(2):116–124.
15. Turing AM (1952) The chemical basis of morphogenesis. *Philos Trans R Soc Lond B Biol Sci* 237(641):37–72.
16. Lecuit T, Samanta R, Wieschaus EF (2002) Slam encodes a developmental regulator of polarized membrane growth during cleavage of the Drosophila embryo. *Dev Cell* 2(4):425–436.
17. Rothe M, Wimmer EA, Pankratz MJ, González-Gaitán M, Jäckle H (1994) Identical transacting factor requirement for knirps and knirps-related Gene expression in the anterior but not in the posterior region of the Drosophila embryo. *Mech Dev* 46(3):169–181.
18. Kim Y, et al. (2013) Context-dependent transcriptional interpretation of mitogen activated protein kinase signaling in the Drosophila embryo. *Chaos* 23(2):025105.
19. Hoch M, Schröder C, Seifert E, Jäckle H (1990) cis-acting control elements for Krüppel expression in the Drosophila embryo. *EMBO J* 9(8):2587–2595.
20. Gaul U, Weigel D (1990) Regulation of Krüppel expression in the anlage of the Malpighian tubules in the Drosophila embryo. *Mech Dev* 33(1):57–67.
21. Jaeger J, et al. (2004) Dynamic control of positional information in the early Drosophila embryo. *Nature* 430(6997):368–371.
22. Bak P (1996) *How Nature Works: The Science of Self-Organised Criticality* (Copernicus, New York).
23. Mora T, Bialek W (2010) Are biological systems poised at criticality? *J Stat Phys* 144(2):268–302.
24. Wolpert L (1969) Positional information and the spatial pattern of cellular differentiation. *J Theor Biol* 25(1):1–47.
25. Bialek W (2012) *Biophysics: Searching for Principles* (Princeton Univ Press, Princeton).
26. Tkačik G, Walczak AM, Bialek W (2012) Optimizing information flow in small genetic networks. III. A self–interacting gene. *Phys Rev E* 85(4):041903.
27. Gregor T, Tank DW, Wieschaus EF, Bialek W (2007) Probing the limits to positional information. *Cell* 130(1):153–164.
28. Tkačik G, Callan CG, Jr., Bialek W (2008) Information flow and optimization in transcriptional regulation. *Proc Natl Acad Sci USA* 105(34):12265–12270.
29. Berg HC, Purcell EM (1977) Physics of chemoreception. *Biophys J* 20(2):193–219.
30. Rieke FM, Baylor DA (1998) Single-photon detection by rod cells of the retina. *Rev Mod Phys* 70(3):1027–1036.