Size Regulation in the Segmentation of *Drosophila*: Interacting Interfaces between Localized Domains of Gene Expression Ensure Robust Spatial Patterning

Sergei Vakulenko
Institute of Mechanical Engineering Problems, Institute of Print, St. Petersburg, Russia

Manu and John Reinitz
Department of Applied Mathematics and Statistics, and Center for Developmental Genetics, Stony Brook University, New York 11794-3600, USA

Ovidiu Radulescu
IRMAR, UMR CNRS 6025, University of Rennes 1, Campus de Beaulieu, 35042 Rennes, France
(Received 23 December 2008; published 15 October 2009)

We propose a new mechanism for robust biological patterning. The mechanism bears analogy to interface dynamics in condensed media. We apply this method to study how gene networks control segmentation of *Drosophila*. The proposed model is minimal involving only 4 genes and a morphogen gradient. We discuss experimental data for which developmental genes are expressed within domains spatially limited by kinks (interfaces) and the gene interaction scheme contains both weak and strong repulsion. We show how kink-kink interactions can be calculated from the gene interactions and how the gene interaction scheme ensures the control of proportions (size regulation).

DOI: 10.1103/PhysRevLett.103.168102 PACS numbers: 87.19.lx, 87.10.Ed, 87.18.Cf, 87.19.lp

In the last decades, a great deal of attention has been given to problems of pattern formation and control. Some general patterning principles emerged from theories using normal forms for dynamical systems [1]. In particular, it was understood that the interaction of kinks, vortices, and generally localized modes (LMs) is important in stabilizing complex equilibrium and nonequilibrium structures such as block-copolymer phases [2] and flow patterns of shear-banding fluids [3]. Typically patterning proceeds into two stages. The first stage is relatively fast growth of LMs, while the second one can be described as a slow motion of interacting modes [4]. Our goal is to apply these ideas to biological systems. There are two traditional approaches to patterning of biological systems. Diffusion-driven Turing instabilities control patterns in systems containing at least two substances with very different mobility [5,6]. Thresholding models are based on the existence of preestablished maternal morphogen gradients [7]. The maternal morphogen triggers zygotic gene expression in regions of the embryo where its concentration is larger than a threshold value. Our approach describes a new mechanism for patterning control. This mechanism is compatible with the thresholding hypothesis, which can be used to describe nucleation of the LMs at early patterning stages. However, the proposed interaction between mobile LMs represents a new patterning principle in biology.

To illustrate these new concepts, we consider the segmentation of insects such as *Drosophila* (fruit fly) which is in focus of many works, biological, mathematical and physical [8–12]. First, let us outline the segmentation process. During this process, developmental genes are expressed in localized domains distributed along the anterior-posterior (AP) axis of the embryo. The sizes and positions of the domains evolve in time. In *Drosophila*, the patterning is influenced by a maternal protein called bicoid whose concentration decreases exponentially from anterior to posterior. Bicoid directs the expression of gap genes. After rejecting the Turing mechanism, many biologists think now that segmentation of *Drosophila* is governed by a thresholding mechanism. This hypothesis implies that variation of position of zygotic gene expression domain borders should closely follow the variations of the maternal gradients. This simple explanation is disproved by recent quantitative studies [9]. Numerous theoretical works proposed hypothetical mobile determinants [10], additional maternal gradients [11], or size dependent concentrations [13] to explain the paradox.

We propose a new approach to these problems using LMs and their interaction. In segmentation processes, proteins produced by active genes are expressed in spatially localized domains. The local modes important for pattern stabilization are the kinks, representing transition regions between a domain expressing a gene and a neighboring domain where the same gene is not expressed. We start from the gene circuit model (GCM) proposed by one of us and now largely used to describe patterning by gene networks [14]. Recently we have used numerical simulation and the GCM to show that robust patterning in the embryo of *Drosophila* can be generated by the collective action of a genetic network [8]. Here we show that the observed robustness of the model can be explained by LMs interaction. We compute the kink-kink interactions for the GCM and relate them to genetic interactions. These interactions depend on the interkink distances and stabilize a

pattern that respects proportions. This mechanism also offers stability with respect to maternal gradient variations since this gradient serves mainly as an initial condition for attractor selection, stimulating kink nucleation and growth.

Model.—To describe segmentation, we use a system of m reaction-diffusion equations, where m is the number of gene products (proteins). This model is a homogenized version of the space-discrete GCM [14]

$$\frac{\partial u_i}{\partial t} = d_i \frac{\partial^2 u_i}{\partial x^2} + R_i \sigma\left(\sum_{j=1}^{m} T_{ij} u_j + m_i \mu(x) + h_i\right) - \lambda_i u_i, \tag{1}$$

$u_i(x, t)$ are concentrations of proteins; d_i, λ_i, R_i are diffusion coefficients, degradation constants, and protein maximum production rates, respectively; σ is a fixed function of a typical sigmoidal form, smooth and monotonically increasing from 0 to 1; the matrix T describes pair interaction between induced (zygotic) genes; $\mu(x)$ is the maternal morphogen concentration, m_i defines the strength of the interactions between morphogen and induced genes, and the parameters h_i are thresholds.

Solutions of (1) contain domains of relatively constant expression, encompassed by kink-antikink pairs representing the domain borders. To simplify calculations we assume that σ is a step function. Then, kink positions are solutions of $\sum_{j=1}^{m} T_{ij} u_j(x, t) + m_i \mu(x) + h_i = 0$.

Equations for slow motion of kink positions can be obtained by using the Whitham principle [15].

Slow kink motion, one kink.—Let us first consider a single gene expressing in the external field $M(x)$. In this case we consider the one reaction-diffusion equation

$$u_t = d u_{xx} + R\sigma(Tu + M(x)) - \lambda u. \tag{2}$$

A kink solution of Eq. (2), corresponding to the right border of the expression domain, localized at q, is given by

$$U(x) = \frac{R}{\lambda}\begin{cases} 1 - \frac{1}{2}\exp(\gamma(x - q)), & x < q, \\ \frac{1}{2}\exp(-\gamma(x - q)), & x > q, \end{cases} \tag{3}$$

where $\gamma = \sqrt{\lambda/d}$ is the kink tail parameter. In *Drosophila*, whose egg size is about 500 μm, kink tails are about 0.1 μm^{-1}. Antikink solutions, i.e., left borders of expression domains are obtained by changing $x - q$ to $q - x$ in (3).

To obtain the slow kink motion note first that Eq. (2) can be rewritten in a variational form as

$$\frac{\delta D[u_t]}{\delta u_t} = -\frac{\delta F[u]}{\delta u}, \tag{4}$$

where D and F are the dissipation and energy functionals

$$D = \frac{1}{2}\int u_t^2 dx, \qquad F = \int\left[\frac{1}{2} d u_x^2 + \Phi(u, x)\right] dx.$$

The potential Φ is obtained from $\Phi_u = -R\sigma(Tu + M(x)) + \lambda u$.

To find a moving kink solution $U(x, q)$, depending on slow variables $q = q(t)$, we substitute the solution (3) into our functionals and we get the *Whitham averaged* functionals $\bar{D}$, $\bar{F}$ depending on q through U according to $\bar{D} = D[U]$, $\bar{F} = F[U]$. The equation for q time evolution follows from

$$\frac{d\bar{D}(q_t)}{dq_t} = -\frac{d\bar{F}(q)}{dq}. \tag{5}$$

We obtain the equation of motion of a kink under the influence of the field M

$$r\frac{dq}{dt} = s\left(\frac{M(q)}{RT} + \frac{1}{2\lambda}\right), \tag{6}$$

where $r = \gamma(2\lambda)^{-2}$ is a parameter. The equation of motion also contains the topological charge s. This takes two values $s = 1$ for kinks (right borders) and $s = -1$ for antikinks (left borders). Given the same field, kinks and antikinks move in opposite directions.

In order to pass from single kinks to multikink solutions, we apply a local field approximation. From Eqs. (1) and (2), the total local field acting on a gene i at x is

$$M_i(x) = \sum_{j \neq i}^{m} T_{ij} U_j(x) + m_i \mu(x) + h_i. \tag{7}$$

The field M_i gathers the influence of the maternal morphogen and of other zygotic genes on the gene i. In (7) $U_j(x)$ is the asymptotic expression for the field produced by gene j (typically the field of a kink-antikink pair).

Interacting kinks, alternating cushions.—A pattern is specified by a sequence of expression domains wherein gene 1 is expressed in $[0, q_1]$, gene 2 is expressed in $[q_2, q_3]$, etc. We assume that zygotic genes are all weakly self-activated $T_{ii}R_i/\lambda_i = T_A > 0$ and repress each other $T_{ij} < 0$, $i \neq j$. We also consider that there are two types of interactions: weakly repressive between genes in adjacent domains such that $T_{ij}R_j/\lambda_j = -\alpha_1 T_A < 0$ and strongly repressive between genes in next-adjacent domains $T_{ij}R_j/\lambda_j = -\alpha_2 T_A < 0$, $0 < \alpha_1 < \alpha_2$. This condition appears in the *Drosophila* gap gene system, where it is called "alternating cushions" (AC) [12]. In other insects that develop in a similar manner, there is limited information about genetic interactions. Nevertheless, it is known that in the midge *Clogmia* four gap genes that are paralogs of *hb*, *kni*, *Kr*, and *gt* are expressed in the strongly repressing pairs (*hb*, *Kr*) and (*kni*, *gt*) forming nonadjacent domains separated by "cushions," Fig. 1. Given that the minimal GCM satisfying AC contains at least four genes, this fact indicates that this mechanism may be evolutionarily conserved.

We obtain the equations of motion for a system of interacting kinks by replacing in (6) the field (7) produced by the other domains. The various contributions are obtained from (3). For instance the field acting on the right border of the domain i has the three contributions

$$M(q_{2i-1}) = \underbrace{m_i\mu(q_{2i-1}) + h_i}_{\text{maternal}} - \underbrace{\alpha_1 T_A(1 - \tfrac{1}{2}e^{-\gamma_{i+1}\delta_{i+1}})}_{\text{domain } i+1} - \underbrace{\alpha_2 T_A e^{-\gamma_{i+2}\Delta_{i+1}}}_{\text{domain } i+2}.$$

A similar equation can be written for the left border of the domain $i + 1$. We get

$$\frac{dq_{2i-1}}{dt} = v_i[\alpha_1 e^{-\gamma_{i+1}\delta_{i+1}} - \alpha_2 e^{-\gamma_{i+2}\Delta_{i+1}} + \bar{h}_i],$$
$$\frac{dq_{2i}}{dt} = v_{i+1}[\alpha_2 e^{-\gamma_{i-1}\Delta_i} - \alpha_1 e^{-\gamma_i\delta_{i+1}} - \bar{h}_{i+1}]. \quad (8)$$

The distances $\delta_{i+1} = q_{2i-1} - q_{2i}$, $\Delta_{i+1} = q_{2i+2} - q_{2i-1}$ correspond to the overlap of neighboring domains and to the distance between next neighbor domains, respectively. The mobility $v_i = 4\sqrt{\lambda_i d_i}$ depends only on the gene in domain i. The terms $\bar{h}_i$ regroup nonexponential contributions to the field.

Various terms in the Eqs. (8) can be interpreted as forces acting on kinks. These forces are of two kinds. Terms $\bar{h}_i$ depend only weakly on the kink positions and can be interpreted by analogy as "pressure." Positive and negative pressures expand and shrink the domains, respectively. The second category are forces depending exponentially on the distances between kinks. These exponential terms, ensuring the main contribution to stability, will be called by analogy, "springs." AC architecture introduces two types of springs: between borders of next-adjacent domains that tend to prevent overlap, and between borders of adjacent domains that tend to increase overlap. Without springs, pattern dismantles. Domains with negative pressure shrink and disappear. Domains with positive pressure expand and occupy all available space.

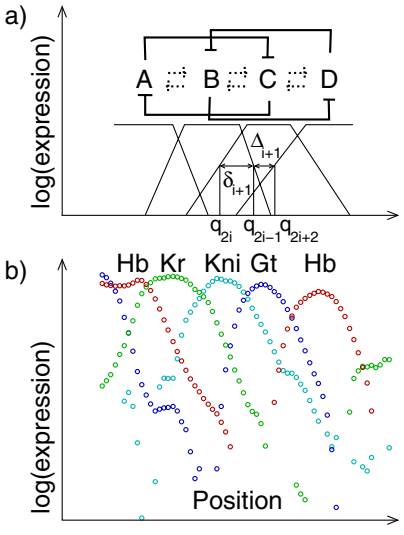

FIG. 1 (color online). (a) Minimal alternating cushions involve 4 genes. Weak and strong interactions are shown with dotted and continuous lines, respectively. (b) Experimental pattern in *Drosophila*: (*hb*, *kni*) and (*Kr*, *gt*) are pairs of strongly repulsive genes, repression is weak between genes belonging to distinct pairs.

An important property of developmental systems is size regulation (SR). For a patterning system this means that the pattern adapts proportionally when the size of the embryo is changed. Turing systems cannot reproduce SR without artificial assumptions [13]. An alternative explanation of SR [6], based on lateral induction and self-inhibition of the patterning substances, uses very mobile components. While such hypotheses may be sufficient in some cases, our model shows that they are not necessary in general. Our mechanism explains SR in a simplest way. When size varies, the kinks readjust their mutual distances by "spring" repulsion. Thus, if size increases then interkink distances increase as well. For periodic patterns with only one interkink distance, this mechanism ensures perfect SR. The cases with several interkink distances should be analyzed more carefully. For instance, the size increase can lead to uneven increases of domain widths and of domain overlaps. SR of patterns with two interkink distances is studied analytically. Then, we simulate numerically the dynamics of interacting kinks using interkink distances observed in *Drosophila*.

Uniform pressure terms, two interkink distances.—Size regulation is easy to study in the case of uniform pressure terms $\bar{h}_i = \bar{h}$. In this case stationary patterns are periodic with two interkink distances $\delta_i = \delta$, $\Delta_i = \Delta$, where $\Delta + \delta = L/N$, N is the number of domains. Let $\rho = \Delta/(\Delta + \delta)$ be the interkink distances ratio. Size regulation means that ρ is not sensitive to variations of L, i.e., $S_\rho = |d\log\rho/d\log L| < \epsilon$, where ϵ is a small number. If $\epsilon = 0.1$, then a relative variation of 10% of L leads to less than 1% variation of the ratio ρ. Simple calculations show that ϵ can be very small. From the stationary equations of Eq. (8) we obtain an implicit dependence of ρ on $\tilde{L}$ and by differentiation an expression for S_ρ,

$$\rho_\alpha \exp(-\tilde{L}(1-\rho)) + \tilde{h} = \exp(-\tilde{L}\rho), \quad (9)$$

$$S_\rho = \left| \frac{1 - (1-\rho)/\rho \exp(\tilde{L}(2\rho - 1))}{1 + \rho_\alpha \exp(\tilde{L}(2\rho - 1))} \right|, \quad (10)$$

with $\tilde{L} = \gamma(\Delta + \delta) = \gamma L/N$, $\tilde{h} = \bar{h}/\alpha_2$, $0 < \rho_\alpha = \alpha_1/\alpha_2 < 1$. $\tilde{L}$, ρ values can be obtained from experimental patterns. ρ_α, $\tilde{h}$ are model parameters. Most importantly, sensitivity (10) contains only one model parameter, ρ_α. Now, SR is easy to test. First we calculate the robustness domain (RD), defined as the set of $(\tilde{L}, \rho)$ for which $|S_\rho| < \epsilon$. For given ρ_α, SR is obtained when $(\tilde{L}, \rho)$ are inside RD (in blue in Fig. 2). For small ρ_α there are two disconnected RDs: one controlling bands with large overlaps (ρ small) and one controlling bands with narrow overlaps, or narrow interband spacings ($\rho = 1 \pm r$, $r \geq 0$ is small). These two domains are connected for larger ρ_α. Experimental $(\tilde{L}, \rho)$ were computed using quantitative gene expression data [16] for eight gap gene kinks and eight time points during cell cycles 13 and 14. Kink positions, calculated as

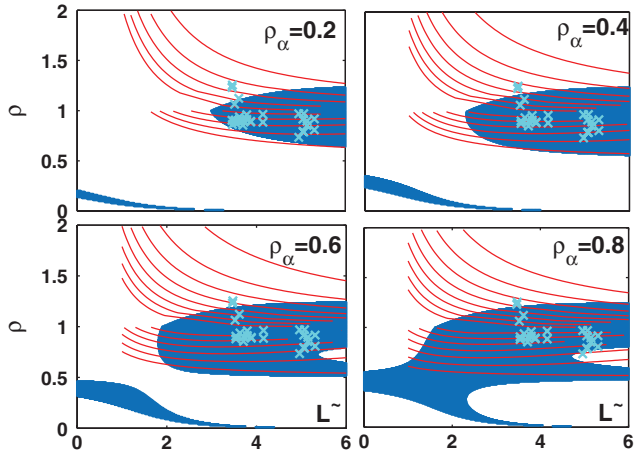

FIG. 2 (color). Size regulation. ρ, $\tilde{L}$ parameters should lie within the blue domains for size sensitivities <0.2. Experimental (ρ, $\tilde{L}$) (in cyan) for eight kinks in *D. melanogaster* embryos are represented for eight time classes in cycles 13 and 14. Nullclines (see text) are drawn in red for several negative $\tilde{h}$.

half maximum intensity positions, give direct access to ρ. For $\tilde{L}$ we need the value of γ. For each time point, γ was calculated as an average over the eight kinks. In Fig. 2 experimental values lie in the upper RD. This proves the size regulation by gene-gene interaction in *Drosophila*. ($\tilde{L}$, ρ) should also lie on the nullcline curves solutions of Eq. (9). Positions of experimental points with respect to the nullcline family in Fig. 2 suggest that pressure $\tilde{h}$ is negative for all kinks.

Nonuniform pressure terms, advantages of the alternating cushions.—The experimental data correspond to slightly dispersed interkink distances which means that pressure terms are not uniform. We investigated this situation numerically. We compared the cases with and without next nearest gene interaction. Stability was tested by random perturbations of the initial data. Starting with $t =$ 43 min during cell cycle 14, the point where gap gene kinks are fully formed, the kink dynamical equations were integrated up to $t = 68$ min, the onset of gastrulation. Additive noise uniformly distributed in the interval $\pm 5\%$ has been added to initial data. Size variations were simulated by multiplying γ by random factors uniformly distributed within $100 \pm 5\%$. For AC scheme the pattern is a stable attractor and can be reproduced identically in all samples (see Fig. 3 top left). The model parameters were chosen such that experimental and modeled kinks coincide at steady state of an average size embryo. Fitting the same positions without a next nearest genes interaction leads to unstable pattern: small perturbations in initial data are not corrected (Fig. 3 upper right). The lack of stability of this scheme can be understood analytically. Indeed, if $\alpha_2 = 0$ then the steady states of Eqs. (8) are degenerate: δ_i are fixed but Δ_i are free to change (only their sum is fixed).

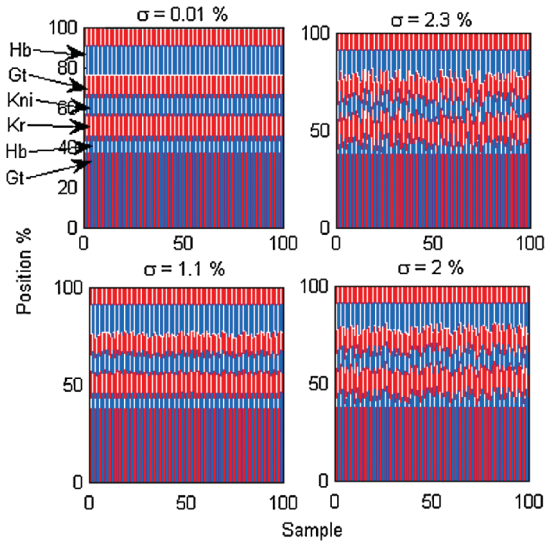

FIG. 3 (color). Pattern robustness in *D. melanogaster* with respect to variations in initial data (up), to variations in initial data and in size (down). Left: alternating cushions; Right: local repulsion scheme, in which nonadjacent genes do not interact. Expression domains were calculated for 100 random samples. σ is rms of the most variable position.

Conclusion.—We have presented an analytic study of the architecture of embryonic gap gene domains that indicates that the alternating cushions architecture has much better size-regulation properties, buffering against maternal gradient variations, and local stability of pattern than an arrangement of locally repulsive domains. Our calculations, while based on detailed experimental data for *Drosophila melanogaster*, apply to long germ band insects in general. The findings reported here are an evolutionary prediction that the AC architecture will be universal in this developmental mode.

[1] M. Cross and P. Hohenberg, Rev. Mod. Phys. **65**, 851 (1993).
[2] S. Vakulenko *et al.*, J. Chem. Phys. **106**, 3412 (1997).
[3] O. Radulescu *et al.*, Rheol. Acta **38**, 606 (1999).
[4] J. Carr and R. Pego, Commun. Pure Appl. Math. **42**, 523 (1989).
[5] A. Turing, R. Soc. Philos. Trans. London B **237**, 37 (1952).
[6] H. Meinhardt, *Models of Biological Pattern Formation* (AP, London, 1982).
[7] L. Wolpert, J. Theor. Biol. **25**, 1 (1969).
[8] Manu *et al.*, PLoS Biol. **7**, e1000049 (2009).
[9] B. Houchmandzadeh *et al.*, Nature (London) **415**, 798 (2002).
[10] T. Aegerter-Wilmsen, C. M. Aegerter, and T. Bisseling, J. Theor. Biol. **234**, 13 (2005).
[11] M. Howard and P. R. ten Wolde, Phys. Rev. Lett. **95**, 208103 (2005).
[12] R. Kraut and M. Levine, Development **111**, 611 (1991).
[13] S. Ishihara and K. Kaneko, J. Theor. Biol. **238**, 683 (2006).
[14] J. Reinitz and D. H. Sharp, Mech. Dev. **49**, 133 (1995).
[15] See EPAPS Document No. E-PRLTAO-103-041944 for Whitham's principle. For more information on EPAPS, see http://www.aip.org/pubservs/epaps.html.
[16] See http://flyex.ams.sunysb.edu/FlyEx/