

## **Variable patterning in fruit fly embryos due to basins of attraction in underlying gene regulatory dynamics**

**Spirov, A.V.<sup>1,2</sup> and Holloway, D.M.<sup>3</sup>**

<sup>1</sup>The Sechenov Institute of Evolutionary Physiology and Biochemistry, Saint-Petersburg, Russia.

<sup>2</sup>Computer Science, State University of New York at Stony Brook, Stony Brook, NY, USA.

<sup>3</sup>Mathematics Department, British Columbia Institute of Technology, Burnaby, B.C., Canada.

E-mail: [alexander.spirov@gmail.com](mailto:alexander.spirov@gmail.com)

**Abstract:** Previous work on a reaction-diffusion model of a 4-gene regulatory network governing insect segmentation characterized the dynamical basis of robustness to perturbations in this system [1,2]. Here, we computationally study system behavior near bifurcation points corresponding to weak-allele mutant embryos (i.e. with altered gene regulation). Our computations suggest that the variable expressivity and incomplete penetrance observed in some gene mutations may stem from response of the dynamical system to variable input (regulatory genes) near such bifurcation points.

**Keywords:** spatially extended systems, reaction diffusion, pattern formation, gene network modeling, dynamical systems, bifurcation, phase diagrams.

### **1. Introduction**

Fruit flies (*Drosophila*) are model organisms for studying spatial pattern formation in animals. In the first few hours of development, a network of interacting genes forms expression patterns which determine the body plan. Data shows that wild-type (WT) development is remarkably robust, with various initial trajectories canalizing to an attracting state [3]. Dynamical systems analysis of a core nonlinear model of the anterior-posterior (AP) segmentation gene network has shown how this WT stability can arise as a trajectory through phase space [1,2]. The WT is stable only to a certain point, however. As CH Waddington and his colleagues showed, once an embryo's buffering capacity is overwhelmed by a sufficiently severe perturbation, altered phenotypes can arise in diverse organisms [4-6].

Strong gene mutations (deletions, insertions) can cause major (lethal) disruptions in the body plan. Our work focuses on using weaker perturbations of genes (weak alleles) to more continuously move the gene network from the WT trajectory. These gene variations can produce variable expressivity, in which the

---

Received: 14 April 2012 / Accepted: 2 October 2012

© 2012 CMSIM



ISSN 2241-0503

outcomes of a sample of embryos is not deterministic, but scatters between a selection of pathological outcomes (from nearly WT to strongly altered). Mutant patterns can be understood as bifurcations to pathological, non-WT, basins of attraction in a multi-stable phase space. Weak alleles bring the system to a bifurcation point, intrinsic variability in the developmental system can then produce the variable outcomes. In this way, weak alleles can provide a tool for mapping the fine structure of the underlying phase space.

A chief focus in recent years has been to use *Drosophila* to test the positional information hypothesis [7,8], that the local concentration of a spatially-distributed gradient can specify the differentiation of cell types in the correct positions. The Bicoid (Bcd) gradient is a classic case of such a ‘morphogen’, but it has been discovered that Bcd gradients, compared between embryos, show much greater variability than do the patterns of one of its primary downstream targets, the zygotically expressed *hunchback* (*hb*) gap gene [9-13]. This points to error correction at the level of the initial zygotic interpretation of the maternal signal. Using ‘coarse-grained’ reaction-diffusion modeling, in which gene-gene interactions are simplified to single signed connections, we can study the robustness in the zygotic segmentation network via dynamical systems analysis and computations. Understanding the model components and parameters which produce the experimental pattern perturbations allows us to map the ‘near-WT’ phase space, and by so doing, to create a detailed understanding of the biological regulatory dynamics used in body formation.

## 2. Methods and Approaches

*Modeling the segmentation gene network:* Four gap genes, (*Kr*), *giant* (*gt*), *knirps* (*kni*) and *hb*, are the core elements of our segmentation model. In *Drosophila*, these are transcriptionally activated by the maternal Bcd protein gradient in a concentration dependent manner. Three other gradients, Hb-maternal ( $Hb_{mat}$ ), Caudal (Cad), and Tailless (Tll), help determine the positions of the gap genes. The combination of this upstream specification and gap-gap cross-regulation results in sharp and precise gap patterns.

Protein expression for the 4 gap genes is modeled using the gene circuit framework [1,2], producing AP concentration patterns (such as Fig. 1A). The model is computed for a one-dimensional row of nuclei, between 30 and 94% AP position (percent Egg Length, or %EL, where 0% is the anterior pole) during nuclear cleavage cycles 13 and 14A. Modeling each gene product *a* (*Kr*, *Hb*, *kni*, *Gt*) in each nucleus *i* defines a system of *number of proteins* times *number of nuclei* ODEs (Ordinary Differential Equations) given by

$$\frac{dv_i^a}{dt} = R_a g(u^a) + D^a [(v_{i-1}^a - v_i^a) + (v_{i+1}^a - v_i^a)] - \lambda_a v_i^a. \quad (1)$$

$R_a$  represents protein synthesis,  $D_a$  represents diffusion, and  $\lambda_a$  represents decay.  $g(u^a)$  is a sigmoidal regulation-expression function; for  $u^a$  below -1.5 and above

1.5  $g(u^a)$  rapidly approaches zero.  $u^a$  is given by  $u^a = \sum_b T^{ab} v_i^b + m^a v_i^{Bcd} + h^a$ .

Parameters  $T^{ab}$  constitute a gene interconnectivity matrix, representing activation of gene  $a$  by the product of gene  $b$  (with concentration  $v_i^b$ ) if positive, repression if negative, and no interaction if close to zero.  $v_i^{Bcd}$  represents the concentration of Bcd in nucleus  $i$ , which is constant in time.  $m^a$  describes the regulatory input of Bcd to each gene; Bcd is a general activator for all four gap genes considered.  $h^a$  represents regulatory input from ubiquitous factors. Our model includes  $Hb_{mat}$ , Cad, and Tll in a similar manner to Bcd, as time-independent parameters.

*Stability Analysis of the Gap Gene System:* The dynamics of an N variable gene circuit can be represented by behavior in an N-dimensional concentration or phase space. Time-varying solutions follow trajectories in the phase space; stable solutions are given by fixed points. In [1,2], the phase space was mapped numerically (Newton-Raphson method) and fixed points classified according to their stability ([2] Protocol S3). The positions of the fixed points and their stability properties determine the stability of a general time varying solution of the gene circuit, including bifurcation points between neighboring basins of attraction. In the present work, we computed a number of trajectories with different initial conditions to test stability and the reduction of variability. Eq. (1) was integrated for very long times in order to characterize the asymptotic behavior.

### 3. Results and Discussion

*Non-robust patterning under perturbed parameters:* Coarse-grained modeling allows us to investigate disturbances in gene-gene interaction strengths. Specifically, in [1,2] we found a set of gap network solutions (T interaction matrices, Eq. 1) which are robust to natural Bcd variability (shown in Fig. 1C). In the present work, we have used this solution set and systematically altered (in small steps) each of the 24  $T^{ij}$  values in the solution matrices. We find cases where small parameter changes cause abrupt changes in patterning, producing severe defects from the WT-like initial solutions (Fig. 1A vs. Fig. 1B). At the borders between WT and pathological cases, we found parametric points where the Bcd variability (Fig. 1C) produces a whole range of outcomes, from nearly normal to severely defective patterns (Fig. 1B). For example, changing  $T^{Kf-cad}$  (the Cad protein effect on the Kr gene) from 0.021 to 0.035 shifts the system to the mathematical bifurcation between nearly-WT and pathological solutions (see Fig. 2). Combined with natural Bcd variability, this produces both WT and pathological patterns, as in biological variable expressivity. Our computations allow us to quantify variable expressivity as the combination of maternal (Bcd) variability and particular alterations in regulator interactions.

Visual inspection and simple statistics show the splitting between WT-like and severely defective patterns. As an illustration, Fig. 1D shows the bi-modal distribution of the Hb gap protein concentration at AP coordinate 72.5 %EL

(from the data in Fig. 1B) compared to the uni-modal Bcd distribution (Fig. 1E; drawn from the data in Fig. 1C). That is, in such cases with altered  $T^{ij}$ , the natural variability in Bcd causes a subset of simulated embryos to be nearly WT, while another subset can be strongly disturbed. This behavior suggests a dynamical bifurcation.

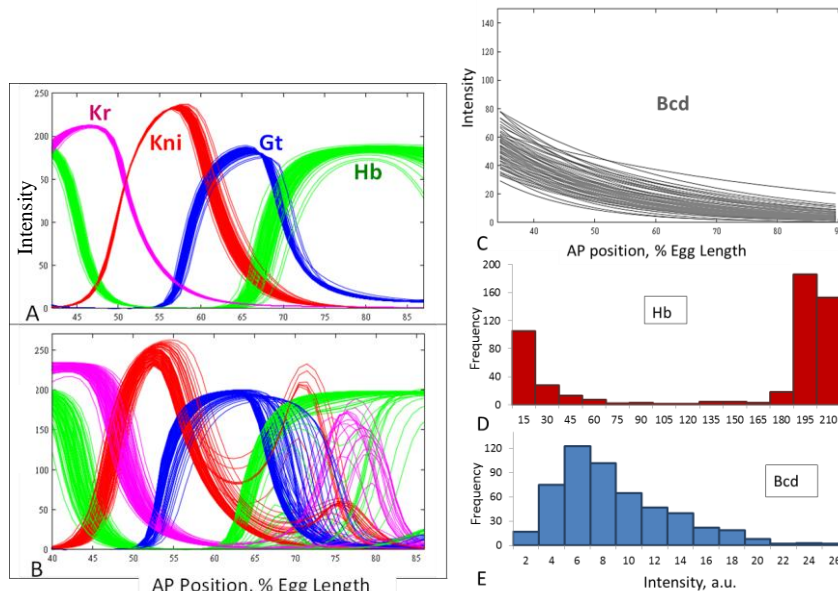


Fig. 1. Bcd-robust WT solutions vs non-robust patterning with perturbed parameters. A) From [1,2] we have a set of gap network solutions ( $T$  interaction matrices) which are robust to natural Bcd variability. B) Changing  $T^{\text{Kr} \leftarrow \text{cad}}$  (Cad protein effect on Kr gene) from 0.021 to 0.035 shifts the system to the mathematical bifurcation between WT and pathological solutions. C) The natural variability of the Bcd profiles used to test solution robustness. D,E) Comparison of the bi-modal distribution of the Hb gap protein with the uni-modal Bcd distribution, at AP coordinate 72.5 %EL.

*Bifurcation analysis of the incomplete penetrance solution* in comparison with the published results [1,2] shows that this new behavior does correspond to a bifurcation. For the example of changing (mutating)  $T^{\text{Kr} \leftarrow \text{cad}}$ , the system is shifted to the border between the old robust WT-like region of phase space and the new pathological one. Fig. 2 compares the two phase portraits at 72.5%EL in Gt, Hb, Kr coordinates, for the robust WT solution of [1,2] and for the mutant solution described here. The robust WT dynamics are characterized by a saddle-node combination (Fig. 2A; c.f. Fig. 4BE, [2]). For the mutant case, at the same AP position (72.5%EL), a second saddle-node combination appears by bifurcation (Fig. 2B). The particular attractor (node) the model reaches depends on the particular shape of the Bcd gradient. The two attractors correspond to the

bi-modal distributions for Hb & Kr (both, very low vs. very high; c.f. Fig. 1D and Fig. 2B), while Gt does not show such drastic differences (see Fig. 1B).

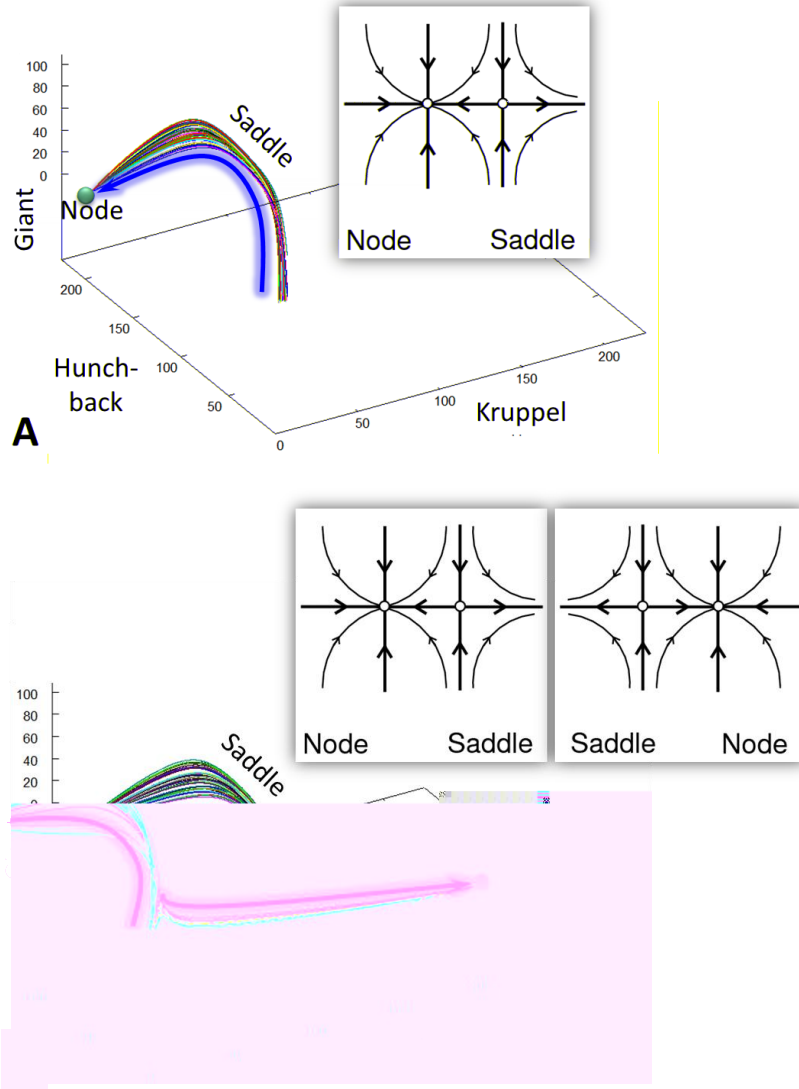


Fig. 2. Phase portraits at 72.5% EL (in Gt, Hb, Kr coordinates). A) Robust WT dynamics. The dynamics here are controlled by a saddle-node combination. B) Mutant behavior. A second saddle-node combination has appeared by bifurcation. The particular attractor (node) the model reaches depends on the particular shape of the Bcd gradient. Other nodes non-essential for this behavior have been omitted. The red bent arrows show the movements of all solutions tested, from initial points to final steady states (at the purple nodes).

#### 4. Conclusions

Building on our previous work, we have shown here how mutation of gene-gene interactions can lead the *Drosophila* segmentation gene network to a bifurcation point, at which natural maternal variability can push embryos into neighboring basins of attraction. Such variable expressivity or incomplete penetrance is observed in nature, but the causes have been elusive. Our work suggests a dynamical basis, in which a weak mutation takes the system to a bifurcation point, and the variable outcomes are a manifestation of natural variability in upstream control; i.e. the mutation removes the robustness of the gene network to maternal variability.

#### Acknowledgements

This work supported by NIH grant R01-GM072022.

#### References

1. Manu, Surkova S, Spirov AV, Gursky VV, Janssens H, Kim A-R, Radulescu O, Vanario-Alonso CE, Sharp DH, Samsonova M, and Reinitz J (2009a). Canalization of Gene Expression in the *Drosophila* Blastoderm by Gap Gene Cross Regulation, *PLoS Biology*, *PLoS Biol* 7(3): e1000049.
2. Manu, Surkova S, Spirov AV, Gursky VV, Janssens H, Kim A-R, Radulescu O, Vanario-Alonso CE, Sharp DH, Samsonova M, and Reinitz J (2009b). Canalization of Gene Expression and Domain Shifts in the *Drosophila* Blastoderm by Dynamical Attractors, *PLoS Computational Biology*, *PLoS Comput Biol* 5(3).
3. Surkova S, Kosman D, Kozlov K, Manu, Myasnikova E, Samsonova AA, Spirov A, Vanario-Alonso CE, Samsonova M, and Reinitz J (2008). Characterization of the *Drosophila* segment determination morphome. *Dev. Biol.* 313, 844-862.
4. Waddington CH (1953). Genetic assimilation of an acquired character. *Evolution* 7, 118-126.
5. Waddington CH (1956). Genetic assimilation of the bithorax phenotype. *Evolution* 10, 1-13.
6. McLaren A (1999). Signalling for germ cells. *Genes Dev.* 13, 373 – 376.
7. Wolpert L (1969). Positional information and the spatial pattern of cellular differentiation. *J Theor Biol* 25, 1-47.
8. Wolpert L (2002). *Principles of development*, 2nd Edition. Oxford University Press.
9. Houchmandzadeh B, Wieschaus E, and Leibler S (2002). Establishment of developmental precision and proportions in the early *Drosophila* embryo. *Nature* 415, 798-802.
10. Spirov A.V. and Holloway D.M., (2003), Making the body plan: precision in the genetic hierarchy of *Drosophila* embryo segmentation, *In Silico Biology*, 3, 89-100
11. Gregor T, Tank DW, Wieschaus EW, and Bialek W (2007a). Probing the limits of positional information. *Cell* 130, 153-164.

12. Gregor T, Wieschaus EW, McGregor AP, Bialek W, and Tank DW (2007b). Stability and nuclear dynamics of the bicoid morphogen gradient. *Cell* 130, 141-152.
13. He, F., Wen, Y., Deng, J., Lin, X., Lu, L.J., Jiao, R., and Ma, J. (2008). Probing intrinsic properties of a robust morphogen gradient in *Drosophila*. *Dev. Cell* 15(4): 558-567.