Cellular automata for exploring gene regulation in Drosophila segmentation

Matthew J. Berryman[a], Andrew Allison[a], and Derek Abbott[a]

[a]Centre for Biomedical Engineering and
School of Electrical and Electronic Engineering,
The University of Adelaide, SA 5005, Australia

ABSTRACT

In this paper we present a 3D cellular automaton for exploring gene interactions in segmentation of *Drosophila* larvae. Beginning with the expression levels of maternally expressed genes such as *bicoid*, our simple model successfully produces the distinctive expression pattern of the *even-skipped* gene in the developing larvae. This work highlights how complex gene interactions in developing organism can nonetheless be accurately modeled using simple rules.

Keywords: Cellular automata, *Drosophila*, gene regulation, emergent behavior, redundancy

1. INTRODUCTION

In this paper we explore a cellular automaton for the segmentation of *Drosophila melanogaster*, commonly known as the fruit fly. The development of segments is controlled by a number of morphogens, proteins that act to control and regulate the development and shape of an organism.[1] Although partial differential equations have been used to explore morphogenesis,[2,3] we believe cellular automata offer a more powerful, flexible approach for capturing the key features of morphogenesis, in particular the segmentation of *Drosophila*. To quote John Holland, the inventor of genetic algorithms,

"Turing (1952) did manage to use PDE's to design a model that started from symmetric initial conditions, but produced an asymmetric variegated pattern, much like the color pattern of a Holstein cow. Even this simple formulation was mathematically intractable: Turing could observe specific examples of the dynamics, but he could derive no general consequences from the mathematical model. In fact, he depended on a computer-based version of the model to exhibit the dynamics of asymmetric pattern formation. Little has been done mathematically since then, and the problem remains much as it was."[4]

To be fair, Turing's work on morphogenesis has proven useful, including successes in describing *Drosophila*,[3] however we introduce a cellular automaton approach that is naturally suited to describing interactions within and between cells.

2. GENE EXPRESSION IN *DROSOPHILA*

2.1. Overview

The set of genes involved in *Drosophila* form a complex network with both positive and negative feedback and branching and converging pathways across and between levels in a multilevel network.[5] Although the network may appear simple (see Figure 1), such simplicity can give rise to highly non-linear behavior.[6,7] Here we consider a subset of genes from the maternal, gap, and pair-rule classes and detail the extensive research that has been undertaken into their interactions. In the following sections, we then build a simple model of these interactions and then show how this leads to the expression of even-skipped *stripe two* – where the term *stripe two* refers to the second of seven stripes that appear in the later stages of embryo development, in the formation of a segmented body.

Send correspondence to Derek Abbott
E-mail: dabbott@eleceng.adelaide.edu.au, Telephone: +61 8 8303 5748

BioMEMS and Nanotechnology, edited by Dan V. Nicolau, Uwe R. Muller,
John M. Dell, Proceedings of SPIE Vol. 5275 (SPIE, Bellingham, WA, 2004)
0277-786X/04/$15 · doi: 10.1117/12.522203

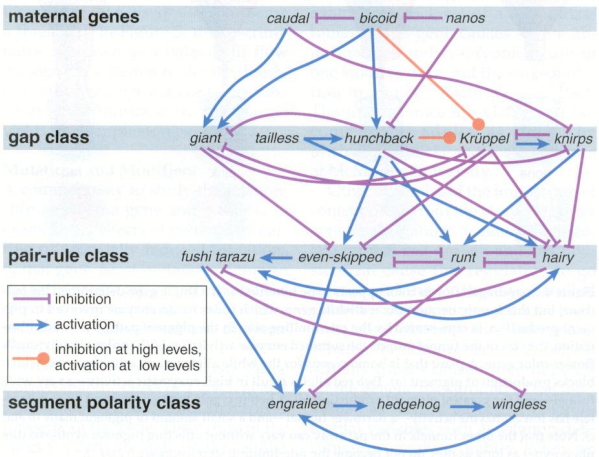

Figure 1. This figure (from Nijhout[5]) shows a network of some of the maternal, gap, pair-rule, and segment polarity class genes. Observe the branching both within and between layers, which gives rise to complex, non-linear behaviors.

2.2. Bicoid

Bicoid is a morphogen translated from maternally expressed mRNA (messenger ribonucleic acid) that is the first step in determining the anterior-posterior (AP) axis.[8] Bicoid expression is also affected by other maternal effect genes called *exuperantia*[*], *swallow*, and *staufen*.[9–11] Localization of *bicoid* mRNA begins during oogenesis, and is controlled by a number of genes including *homeless*.[12] As can be seen in Figure 2, bicoid expression follows an exponential decay curve.

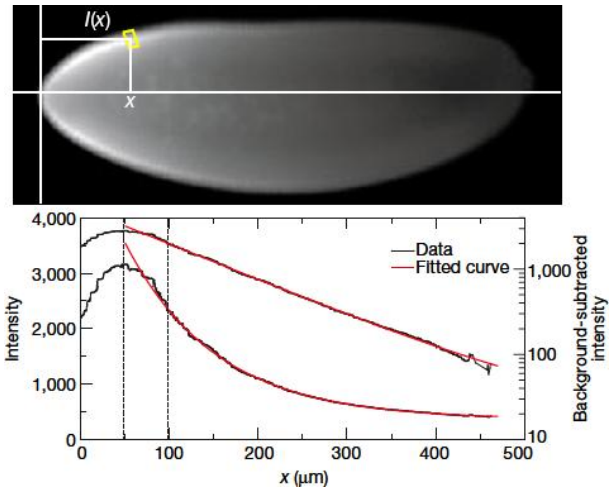

Figure 2. This figure (from Houchmandzadeh *et al.*[8]) shows the wild-type (wt) expression of the bicoid protein in a *Drosophila* larva. The top image shows the expression level using a grayscale intensity. The bottom image shows the numerical values of the intensity as a function of normalized length, determined from the image, and an exponential decay curve fitted to the data. The exponential curve takes the form $I = e^{-\lambda x}$ where I is intensity, x is position, and $\lambda = \sqrt{D/\omega}$ for D the diffusion coefficient and ω the protein degradation rate.

[*]We use *this style* of font to denote the gene, and normal style to denote the expressed protein.

2.3. Nanos

Nanos is another morphogen translated from maternally expressed mRNA that helps determine the posterior region of the *Drosophila* larva.[13] Although other maternally expressed genes are involved in setting up the posterior formation, such as *oskar* and *cappucino*[14, 15] , nanos plays a critical role in setting up the posterior region by repressing hunchback and bicoid.[13, 16] Figure 3 shows the expression of *nanos* in a *Drosophila* larva.

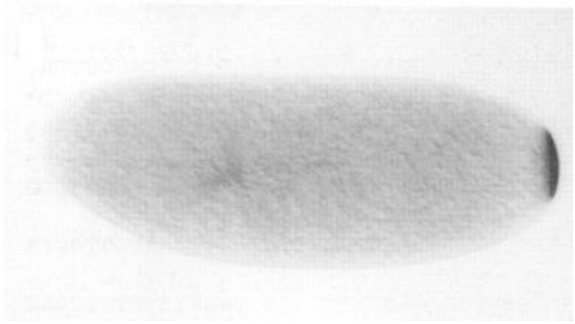

Figure 3. This figure (from Wang and Lehmann[13]) shows the maternally expressed *nanos* mRNA in *Drosophila*, which is highly localized to the posterior region.

2.4. Staufen

Another maternally expressed morphogen is staufen.[17] This is expressed in the pattern shown in Figure 4.

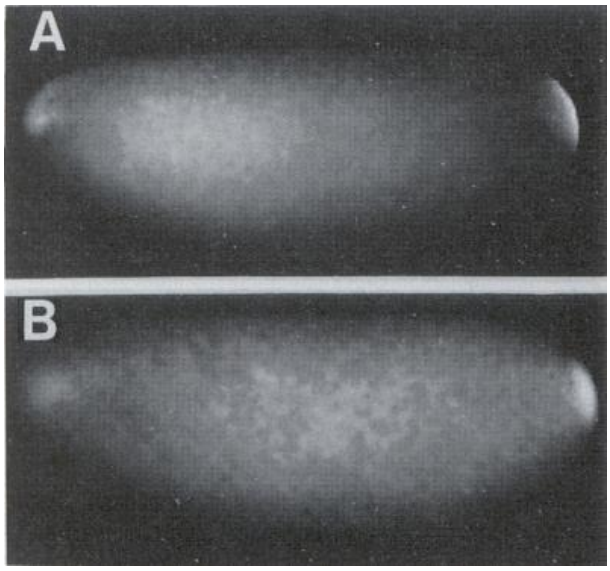

Figure 4. This figure (from St. Johnston *et al.*[17]) shows the wild-type expression of staufen protein in both a freshly-laid larva (A) and a mid-cleavage stage larva (B).

2.5. Hunchback

Hunchback expression is clearly regulated by bicoid, as can be seen in Figure 5. Hunchback expression is a positive feedback cycle, with both bicoid and hunchback itself driving further up-regulation of hunchback.[8, 18] Wu *et al.* suggest that positive feedback is the only mechanism for the second Hunchback stripe in the posterior region, however Houchmandzadeh *et al.* show that mutations in *staufen* affect the boundaries of hunchback by a mechanism other than by staufen changing regulation of bicoid expression. Further, staufen expression is localized to both the poles (see Figure 4). Hunchback expression is also repressed by nanos in the posterior region,[19] and possibly by knirps.[20]

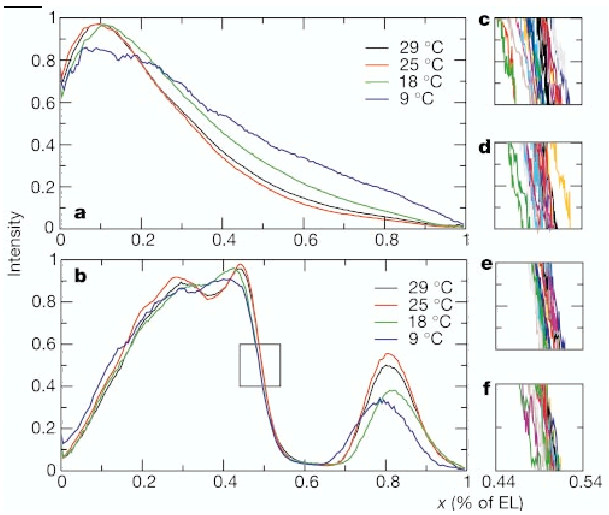

Figure 5. This figure (from Houchmandzadeh *et al.*[8]) shows the levels of bicoid (a) and hunchback (b) expression as a function of normalized length, averaged over 100 embryos for various different environmental temperatures at which they were growing. Note the small spread of hunchback levels for quite a large spread of bicoid levels, especially in the region highlighted in Subfigure (b) where hunchback falls sharply. Subfigures (c)-(f) show all the profiles for the boxed region in Subfigure (b) for temperatures of 9 °C, 18 °C, 25 °C, and 29 °C respectively.

2.6. Krüppel

Hoch *et al.* have done a detailed study of Krüppel activation and found that bicoid activates expression of Krüppel, while hunchback represses it.[21] In other work, they also found that the *Krüppel* promoter contains binding sites for the bicoid activator and the knirps repression.[22] The typical pattern of Krüppel expression is shown in Figure 6.

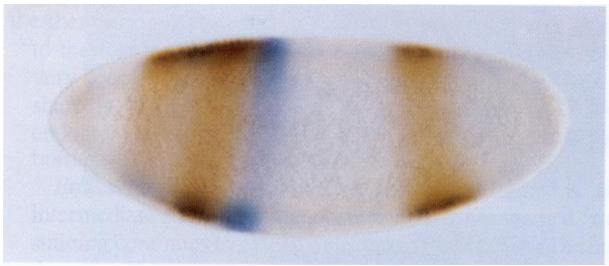

Figure 6. This figure (from Small *et al.*[23]) shows the expression of Krüppel in the darker regions.

2.7. Knirps

Knirps expression is activated by bicoid in the anterior end of the *Drosophila* larva.[24] Knirps expression in wild-type *Drosophila* is shown in Figure 7.

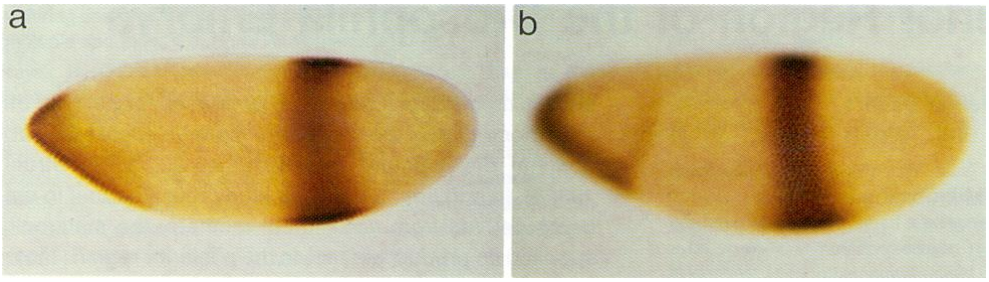

Figure 7. This figure (from Pankratz *et al.*[25]) illustrates the expression of knirps in early *Drosophila* development (a) and at a later stage (b) where the anterior knirps stripe has fully formed.

2.8. Giant

Giant is activated by bicoid and repressed by hunchback,[26] and its expression is shown in Figure 8.

Figure 8. This figure (from Small *et al.*[23]) shows the expression of giant in the darker region, and the position of even-skipped stripe two in the narrow darkest region to the left of center.

2.9. Even-skipped

Even-skipped expression is controlled by activation of *even-skipped* gene translation by hunchback and bicoid.[23, 27] Knirps can act to repress bicoid-mediated activation by binding to promoter sites near *even-skipped* sites.[28] Pankratz *et al.* detail the importance of Krüppel and knirps in regulating stripe formation but acknowledge other gap genes may be involved.[25] Small *et al.* detail the involvement of Krüppel, giant, bicoid, and hunchback in the expression of even-skipped stripe two.[23] The pattern of even-skipped expression is shown in Figure 9.

3. CELLULAR AUTOMATON MODELING

3.1. Overview

For each of the expression level functions, we try to use the biological information as much as possible, where this is unclear or uncertain we make reasonable assumptions about the biology and/or leave such information out. We find that in determining the overall position of the stripes, the unused information makes little difference to the general trends when compared with the actual expression levels shown in Figures 2 to 9. Expression levels are generically functions of discrete cell position (x, y, z) and discrete time t. We use x to

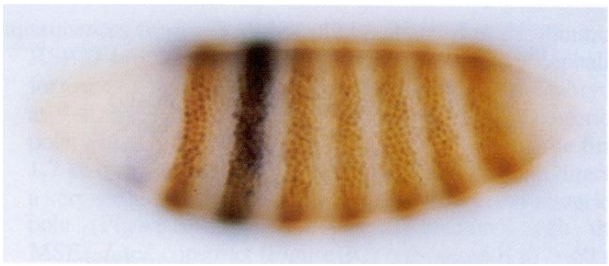

Figure 9. This figure (from Small *et al.*[23] shows even-skipped expression in *Drosophila*, with stripe two, regulated by Krüppel, giant, bicoid, and hunchback, shown in a darker color.

denote the normalized position along the anterior-posterior axis where $x = 0$ corresponds to the most anterior position and $x = 1$ corresponds to the posterior. Similarly, y is the normalized position in the ventral-dorsal axis, and z is the normalized position in the medial-lateral axis. Many of the formulae we use for computing the change in expression levels are based on simple thresholds, here we show that this can produce results concordant with biological observations, but cannot explain the observe robustness to variations in expression levels of the genes being thresholded, with much trial and error needed to find the settings of the thresholds.

3.2. Bicoid model

We treat the maternally expressed pattern of Bicoid expression as a fixed pattern, with an exponential decay function from the anterior end to the posterior end, of the form

$$E_b\left(x,y,z,t\right) = \exp\left(-2|0.05 - x|\right), \tag{1}$$

where $E_b\left(x,y,z,t\right)$ is the bicoid expression level at normalized position (x,y,z) at time t, $E_b \in (0,1)$. Note we consider two exponential decays from a normalized position of 0.1, to reflect better the true gradient as shown in Figures 2 and 5.

3.3. Nanos model

We localize the maternal nanos expression quite specifically and uniformly, in the pattern

$$E_n\left(x,y,z,t\right) = \begin{cases} 0.7, & x > 0.9 \\ 0, & \text{otherwise,} \end{cases} \tag{2}$$

where E_n is the expression level of nanos at position (x,y,z) at time t.

3.4. Staufen model

We treat the maternally expressed pattern of staufen as a pair of exponentially decaying functions, starting at both ends, roughly in line with the general trends observed St. Johnston *et al.*[17] (see Figure 4) but also making the assumption that this has an exponential trend in line with diffusion equations and similar to bicoid expression. Thus we use the following function for describing staufen expression,

$$E_s\left(x,y,z,t\right) = \frac{3}{4}\left(\exp\left(-3x\right) + \exp\left(3\left(-1+x\right)\right)\right), \tag{3}$$

where $E_s\left(x,y,z,t\right)$ is the expression level of staufen at position (x,y,z) at time t, $E_s \in (0,1)$.

3.5. Hunchback model

For a cell at position (x, y, z, t), we set the expression level E_h of hunchback to be,

$$E_h(x,y,z,t) = \begin{cases} \text{sig}\left(E_h(t-1) + E_b(t-1) + E_s(t-1) - E_k(t-1)\right), \\ \text{if sig}\left(E_h(t-1) + E_b(t-1) + E_s(t-1) - E_k(t-1)\right) > 0.675 \text{ and } E_n(t-1) < 0.1 \\ 0, \text{ otherwise,} \end{cases} \tag{4}$$

where

$$\text{sig}(y) = \frac{1}{1+e^{-y}}, \tag{5}$$

and $E_h(x,y,z,t) \in [0,1]$ is the expression of hunchback at position (x,y,z) at time t, also E_k is defined below. Note we omit the position of the expression levels to save space, as these are all (x,y,z), and thus show that the expression of Hunchback in a cell depends only on the levels of the other proteins in the cell at the previous time step, $t-1$.

3.6. Krüppel model

For a cell at position (x, y, z, t), we set the expression level E_r of Krüppel to be,

$$E_r(x,y,z,t) = \begin{cases} 0.7, \\ \text{if } 0.5 < E_h(x,y,z,t-1) < 0.85 \text{ and } E_b(x,y,z,t-1) > 0.4 \text{ and } E_k(x,y,z,t-1) < 0.65 \\ 0.1, \text{ otherwise,} \end{cases} \tag{6}$$

3.7. Knirps model

For a cell at position (x, y, z, t), we set the expression level E_k of knirps to be,

$$E_k(x,y,z,t) = \begin{cases} \text{sig}\left(E_b(x,y,z,t-1)\right), & E_b(x,y,z,t-1) > 0.8 \\ 0.8, & E_b(x,y,z,t-1) > 0.4 \text{ and } E_h(x,y,z,t-1) < 0.55 \\ 0.1, & \text{otherwise,} \end{cases} \tag{7}$$

3.8. Giant model

For a cell at position (x, y, z, t), we set the expression level E_g of giant to be,

$$E_g(x,y,z,t) = \begin{cases} 0.7, & E_h(,x,y,z,t) > 0.75 \text{ and } E_k(x,y,z,t-1) < 0.6 \text{ and } E_r(x,y,z,t-1) < 0.6 \\ 0.1, & \text{otherwise} \end{cases} \tag{8}$$

3.9. Even-skipped model

Based on work by Small *et al.*[23] and Pankratz *et al.*,[25] we use the following equation for E_e, the expression level of even-skipped,

$$E_e(x,y,z,t) = \begin{cases} \text{sig}\left(\left(E_h(x,y,z,t-1) + 5E_b(x,y,z,t-1)/6\right)\right), & \frac{1}{|\mathcal{N}|}\sum_{\mathbf{p}\in\mathcal{N}} E_r(\mathbf{p},t) > 0.2 \\ & \text{and } \frac{1}{|\mathcal{N}|}\sum_{\mathbf{p}\in\mathcal{N}} E_g(\mathbf{p},t) > 0.2 \\ 0, & \text{otherwise,} \end{cases} \tag{9}$$

where $\mathcal{N}$ is the neighborhood of points
$\mathcal{N} = \{(x',y',z') : |x'-x| \le 1, |y'-y| \le 1, |z'-z| \le 1, (x',y',z') \neq (x,y,z)\}$ about the point (x,y,z).

4. RESULTS

4.1. Gene expression in the *Drosophila* model

Our results show the x-y plane in its usual cartesian arrangement, our choice of coordinates ensures the images produced by our software are in the standard orientation used for displaying *Drosophila* expression levels, with the anterior to the left and the dorsal to the top. Figure 10 shows the simulated bicoid expression levels. They appear quite similar to the actual levels of bicoid expression seen in real *Drosophila* as shown in Figure 2, although the exponential tail off towards the anterior end is more clear. Figure 11 shows the nanos expression,

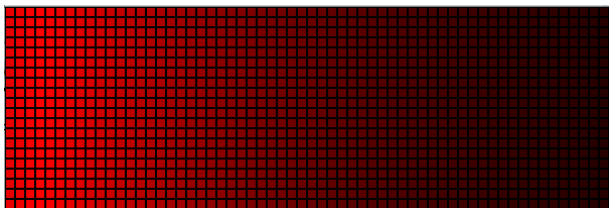

Figure 10. This figure shows the expression level of bicoid in a set of cells representing a cross-section (fixed z) through a *Drosophila* larva, with the anterior to the left and the dorsal to the top. The color intensity represents the expression level: darker for low levels of expression and lighter for high levels. Note the exponential decay of intensity as we move away from a normalized x position of 0.05, which is three cells from the left (anterior) side.

set to be expressed at the most posterior region, and not expressed elsewhere. With staufen, we also used a

Figure 11. The expression level of nanos in a set of cells representing a cross-section (fixed z) through a *Drosophila* larva is shown above. The light band on the right indicates high levels of expression and the darker region represents low expression levels.

fixed expression pattern, and this is shown Figure 12.

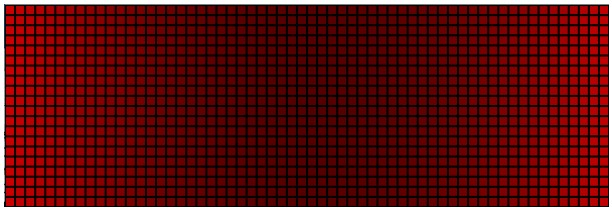

Figure 12. The expression level of staufen in a set of cells representing a cross-section (fixed z) through a *Drosophila* larva is shown in this figure. The color intensity represents the expression level: darker for low levels of expression and brighter for high levels.

In Figure 13, we show the hunchback expression after its expression has stabilized into a fixed pattern, and in Figure 14 we show both hunchback and bicoid simultaneously for the same point in time. Note that hunchback has a well defined boundary in the middle region, whereas bicoid has a continuous gradient.

Experimentation (not shown here) revealed that the position of this boundary varied little with changes in the bicoid gradient. This suggests that the proposals by Houchmandzadeh *et al.* regarding the effect of other genes including positive feedback from hunchback itself in Houchmandzadeh *et al.*[8] are correct.

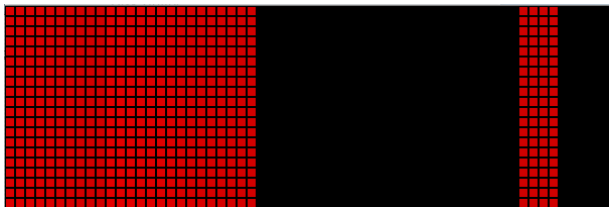

Figure 13. This figure shows the expression level of hunchback in a set of cells representing a cross-section (fixed z) through a *Drosophila* larva. The color intensity represents the expression level: darker for low levels of expression and lighter for high levels.

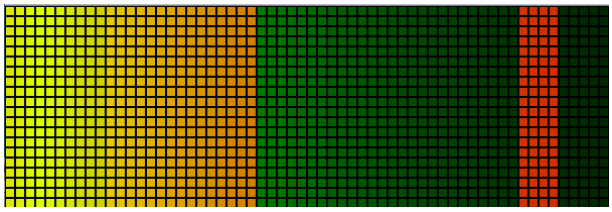

Figure 14. Here we show the expression level of both hunchback and bicoid in a set of cells representing a cross-section (fixed z) through a *Drosophila* larva. The pattern is simply an overlay of Figures 10 and 13, with the lightest regions corresponding to regions where hunchback is expressed.

In Figures 15, 16, and 17, we show the expression levels for the gap class genes *Krüppel*, *knirps*, and *giant*. These correspond well with the expression levels shown in Figures 6, 7, and 8, even though we have greatly simplified the set of interactions and the form these interactions take. The simplification has resulted in these stripes being very sensitive to minor changes in bicoid and hunchback expression, with variations in the normalized expression levels of 0.05 in hunchback and bicoid resulting in the absence of these strips in some cases. This highlights the fact that context of other genes as shown in the network in Figure 1 is important in adding robustness to the gape gene expression against variations in maternal gene expression. Robustness could also gained by mechanisms other than the simple thresholding we have used.[3,8]

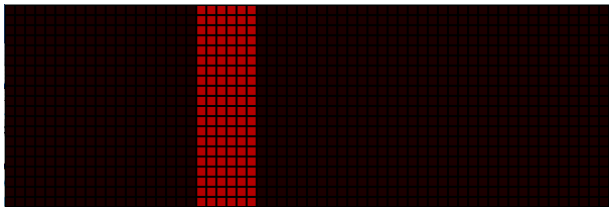

Figure 15. The expression level of Krüppel in a set of cells representing a cross-section (fixed z) through a *Drosophila* larva. The color intensity represents the expression level: darker for low levels of expression and lighter for high levels.

Figure 18 shows the expression of even-skipped stripe two in our virtual *Drosophila* larva.

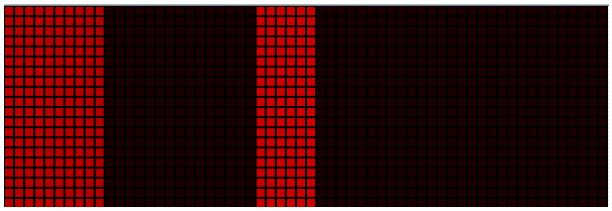

Figure 16. This figure shows the expression level of knirps in a set of cells representing a cross-section (fixed z) through a *Drosophila* larva. The color intensity represents the expression level: darker for low levels of expression and lighter for high levels.

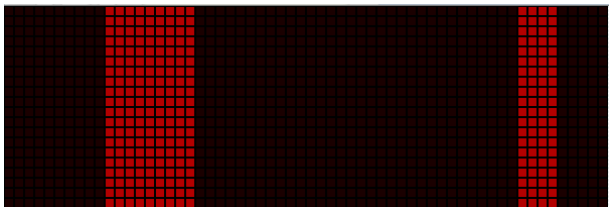

Figure 17. The expression level of giant in a set of cells representing a cross-section (fixed z) through a *Drosophila* larva. The color intensity represents the expression level: darker for low levels of expression and lighter for high levels.

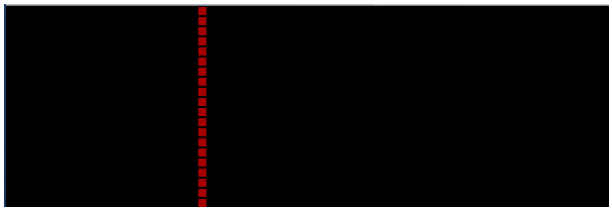

Figure 18. This figure shows the expression level of even-skipped in a set of cells representing a cross-section (fixed z) through a *Drosophila* larva. The color intensity represents the expression level: darker for low levels of expression and lighter for high levels.

5. CONCLUSIONS

This paper has presented a cellular automaton with a simplified set of genes and mostly simple rules governing interaction between those genes. Despite this simplicity, the cellular automaton is able to generate realistic patterns of stripes, up to the even-skipped stripe two. This suggests that *Drosophila* could be modeled quite accurately using a simple yet more powerful model taking into account the other gene interactions, and using interactions consisting of more than thresholding. This would give added robustness to fluctuations in expression in genes higher in the hierarchy. Our model does indicate that further work is needed to refine the mechanisms by which the gene promoters are acting, to give further clues as to how to best model the interactions.

REFERENCES

1. W. Wolpert, R. Beddington, J. Brockes, T. Jessell, P. Lawrence, and E. Meyerowitz, *Principles of Development*, Oxford University Press, 1998.
2. A. Turing, "The chemical basis of morphogenesis," *Philosophical Transactions of the Royal Society of London* **B 237**, pp. 37–72, 1952.
3. D. Holloway, L. Harrison, and A. Spirov, "Noise in the segmentation gene network of *Drosophila*, with implications for mechanisms of body axis specification," in *Proc. SPIE: Fluctuations and Noise in Biological, Biophysical, and Biomedical Systems*, S. Bezrukov, H. Fraunfelder, and F. Moss, eds., **5110**, pp. 180–191, 2003.
4. J. Holland, *Hidden Order: How Adaptation Builds Complexity*, Perseus Books, 1995.
5. H. Nijhout, "The importance of context in genetics," *American Scientist* **91**(5), pp. 416–423, 2003.
6. H. Nijhout, "The nature of robustness in development," *BioEssays* **24**(6), pp. 553–563, 2002.
7. J. Holland, *Emergence: From Chaos to Order*, Oxford University Press, 1998.
8. B. Houchmandzadeh, E. Wieschaus, and S. Leibler, "Establishment of developmental precision and proportions in the early *Drosophila* embryo," *Nature* **415**, pp. 798–802, Feb. 2002.
9. H. Fronhöfer and C. Nüsslein-Volhard, "Maternal genes required for the anterior localization of *bicoid* activity in the embryo of *Drosophila* embryo," *Genes and Development* **1**, pp. 880–890, 1987.
10. E. Stephenson, Y. Chao, and J. Fackenthal, "Molecular analysis of the *swallow* gene of *Drosophila melanogaster*," *Genes and Development* **2**, pp. 1655–1665, 1988.
11. D. St. Johnston, W. Driever, T. Berleth, S. Richstein, and C. Nüsslein-Volhard, "Multiple steps in the localization of *bicoid* RNA to the anterior pole of the *Drosophila* oocyte," *Development* **107**, pp. 13–19, 1989. Supplement.
12. R. Ray and T. Schüpbach, "Intercellular signaling and the polarization of body axes during *Drosophila* oogenesis," *Genes and Development* **10**, pp. 1711–1723, 1996.
13. C. Wang and R. Lehmann, "Nanos is the localized posterior determinant in *Drosophila*," *Cell* **66**, pp. 637–647, 1991.
14. R. Lehmann and C. Nüsslein-Volhard, "Abdominal segmentation, pole cell formation, and embryonic polarity require the localized activity of *oskar*, a maternal gene in *Drosophila*.," *Cell* **47**, pp. 141–152, 1986.
15. L. Manseau and T. Schüpbach, "*cappucino* and *spire*: two unique maternal-effect loci required for both the anteroposterior and dorsoventral patterns of the *Drosophila embryo*," *Nature* **329**, pp. 167–170, 1989.
16. R. Lehmann and C. Nüsslein-Volhard, "The maternal gene *nanos* has a central role in posterior pattern formation of the *Drosophila* embryo," *Development* **112**, pp. 679–693, 1991.
17. D. St Johnston, D. Beuchle, and C. Nüsslein-Volhard, "*staufen*, a gene required to localize maternal RNAs in the *Drosophila* egg," *Cell* **66**, July 1991.
18. X. Wu, V. Vasisht, D. Kosman, J. Reinitz, and S. Small, "Thoracic patterning by the *Drosophila* gap gene *hunchback*," *Developmental Biology* **237**, pp. 79–92, 2001.

19. V. Irish, R. Lehmann, and M. Akam, "The *Drosophila* posterior-group gene *nanos* functions by repressing *hunchback* activity," *Nature* **338**, pp. 646–648, 1985.
20. F. Sauer and H. Jäckle, "Heterodimeric *Drosophila* gap gene protein complexes acting as transcriptional repressors," *The EMBO Journal* **14**, pp. 4773–4780, 1995.
21. M. Hoch, E. Seifert, and H. Jäckle, "Gene expression mediated by *cis*-acting sequences of the Krüppel gene in response to the *Drosophila* morphogens *bicoid* and *hunchback*," *The EMBO Journal* **10**(8), pp. 2267–2278, 1991.
22. M. Hoch, N. Gerwin, H. Taubert, and H. Jäckle, "Competition for overlapping sites in the regulatory region of the *Drosophila* gene *Krüppel*," *Science* **256**, pp. 94–97, 1992.
23. S. Small, A. Blair, and M. Levine, "Regulation of even-skipped stripe 2 in the *Drosophila* embryo," *The EMBO Journal* **11**(11), pp. 4047–4057, 1992.
24. M. Rothe, E. Wimmer, M. Pankratz, M. Gonzalez-Gaitan, and H. Jäckle, "Identical transacting factor requirement for knirps and knirps-related gene expression in the anterior but not the posterior region of the *Drosophila* embryo," *Mechanisms of Development* **46**(3), pp. 169–181, 1994.
25. M. Pankratz, E. Seifert, N. Gerwin, B. Billi, U. Nauber, and H. Jäckle, "Gradients of *Krüppel* and *knirps* gene products direct pair-rule gene stripe patterning in the posterior region of the *Drosophila* embryo," *Cell* **61**, pp. 309–317, 1990.
26. E. Eldon and V. Pirrotta, "Interactions of the *Drosophila* gap gene *giant* with maternal and zygotic pattern-forming genes," *Development* **111**, pp. 367–378, 1991.
27. M. Frasch and M. Levine, "Complementary patterns of even-skipped and fushi tarazu expression involve their differential regulation by a common set of segmentation genes in *Drosophila*," *Genes and development* **1**, pp. 981–995, 1987.
28. D. Arnosti, S. Gray, S. Barolo, J. Zhou, and M. Levine, "The gap protein knirps mediates both quenching and direct repression in the *Drosophila* embro," **15**(14), pp. 3659–3666, 1996.