Stochastic dynamics of gene expression in developing fly embryos

David M. Holloway
Mathematics Department
British Columbia Institute of Technology
Burnaby, B.C., Canada
David_Holloway@bcit.ca

Alexander V. Spirov
Sechenov Institute of
Evolutionary Physiology and Biochemistry
St. Petersburg, Russia

***Abstract*—Segmentation of the developing insect body is preceded by cell-specific gene expression. In fruit flies (*Drosophila*), pair-rule genes are expressed in spatial stripes specifying segment fates. Transcription of the *even-skipped* (*eve*) pair-rule gene was recently shown to proceed in noisy bursts. Here, we develop a stochastic model of *eve* transcription from DNA to mRNA. This indicates that *eve* transcription proceeds at two rates, with a slow rate providing basal production and a fast rate allowing for high mRNA output. This two-rate transcription may afford more reliability in mRNA output, and therefore the protein levels which specify cell type, than a simple on-off (one-rate) mechanism.**

Keywords—noise and fluctuations in biological systems; animal embryo development; gene transcription; stochastic simulations

I. INTRODUCTION

Cell-specific gene expression underlies the coordinated differentiation of cells during biological development. Cell types are defined by having particular types and amounts of protein building blocks. Embryonic development depends on regulated expression of genes, which make the mRNA to make these cell-specific protein profiles. Genes are activated (or repressed) by transcription factors (TFs) which are proteins which bind DNA upstream (generally) of the coding sequence for a gene. Multiple binding sites (BSs) for TFs generally exist for each gene; binding and unbinding of activating and repressing TFs in a gene's cis-regulatory sequence create complicated regulatory dynamics for initiation of transcription of DNA into RNA (the first stage of gene expression). Since genes generally have several to dozens of TF BSs, and TFs are in relatively low concentrations (on the order of 1000 molecules per cell), the bound state of the cis-regulatory region follows low-number statistics, and the kinetics of transcriptional initiation can be strongly stochastic.

While the stochastic nature of gene transcription has been studied for some time in populations of single cells, such as yeast or *E. coli* [1-3], the technology has recently been developed to image RNA production in living multicellular animal embryos, specifically in the fruit fly *Drosophila* [4-6]. This opens a new dimension in studying the coordinated expression of genes in spatial patterns across extended tissues, and studying how tissue borders can form reliably (critical, for example, in the correct formation of organs) from inherently stochastic transcriptional kinetics. Recently published work from Bothma et al. [7] is the first to show noisy 'bursting' of RNA production for the gene *even-skipped* (*eve*). Here, we develop a master equation model of *eve* transcriptional dynamics, and test the hypothesis that *eve* time series show evidence of two (slow and fast) transcriptional initiation rates.

II. MODELS AND METHODS

Table 1 summarizes the elementary reactions modelling *eve* transcription (specifically, the 2nd *eve* stripe, *eve2*): binding/unbinding of the activator TFs B (Bicoid) and H (Hunchback) to the regulatory region of the DNA (E[BH00], arguments denote the TF bound (1), unbound (0), or either (*x*); the 3rd and 4th positions are for repressors, as reported elsewhere [8]); initiation of transcription (making of nascent RNA), with a B bound LOW rate (k_{1000}) and a B+H bound HIGH rate (k_{1100}); and completion of transcription (release of a complete *eve* mRNA molecule). Rate constants are shown for each reaction. These are estimated from experimental data [7, 9, 10]. The binding, unbinding and initiation constants reproduce the observed 3 minute autocorrelation in nascent transcript number, with no autocorrelation in minute-to-minute changes; the fraction of minutes in which transcript number increases; the maximum observed per-minute increases; and the total mRNA output observed for B and B+H driven transcription.

TABLE I. KINETIC MODEL OF EVE GENE TRANSCRIPTION

cis-regulatory element binding	*Transcription*
$E[0xxx] + B \underset{k_{UNBIND}=3.3e-2}{\overset{k_{BIND-B}=2.6e6}{\rightleftharpoons}} E[1xxx]$	$E[1000] \xrightarrow{k_{1000}=9e-2} E[1000] + eve_{nascentRNA}$
$E[x0xx] + H \underset{k_{UNBIND}=3.3e-2}{\overset{k_{BIND-H}=1.8e5}{\rightleftharpoons}} E[x1xx]$	$E[1100] \xrightarrow{k_{1100}=5.6e-1} E[1100] + eve_{nascentRNA}$
	$eve_{nascentRNA} \xrightarrow{k_T=3.97e-3} eve_{mRNA}$

We thank the US NIH GM072022 for financial support. 978-1-5090-2760-6/17/$31.00 ©2017 IEEE.

This system is solved stochastically using the MesoRD software package [11], using the reaction-diffusion master equation approach with next-subvolume queuing [12, 13].

III. RESULTS AND DISCUSSION

Biological evidence indicates that *eve2* is activated at both a B-dependent basal LOW rate (k_{1000} in the model) and an enhanced HIGH B+H co-activated rate (k_{1100}) [9, 10]. Time series of this two-rate model (Fig. 1A) show the stochastic bursting observed in experimental time series [7]. The minute-to-minute changes in transcript number show a smooth distribution across rates (Fig. 1B), from common low initiation minutes to uncommon very high initiation minutes, similar to the experimental distribution.

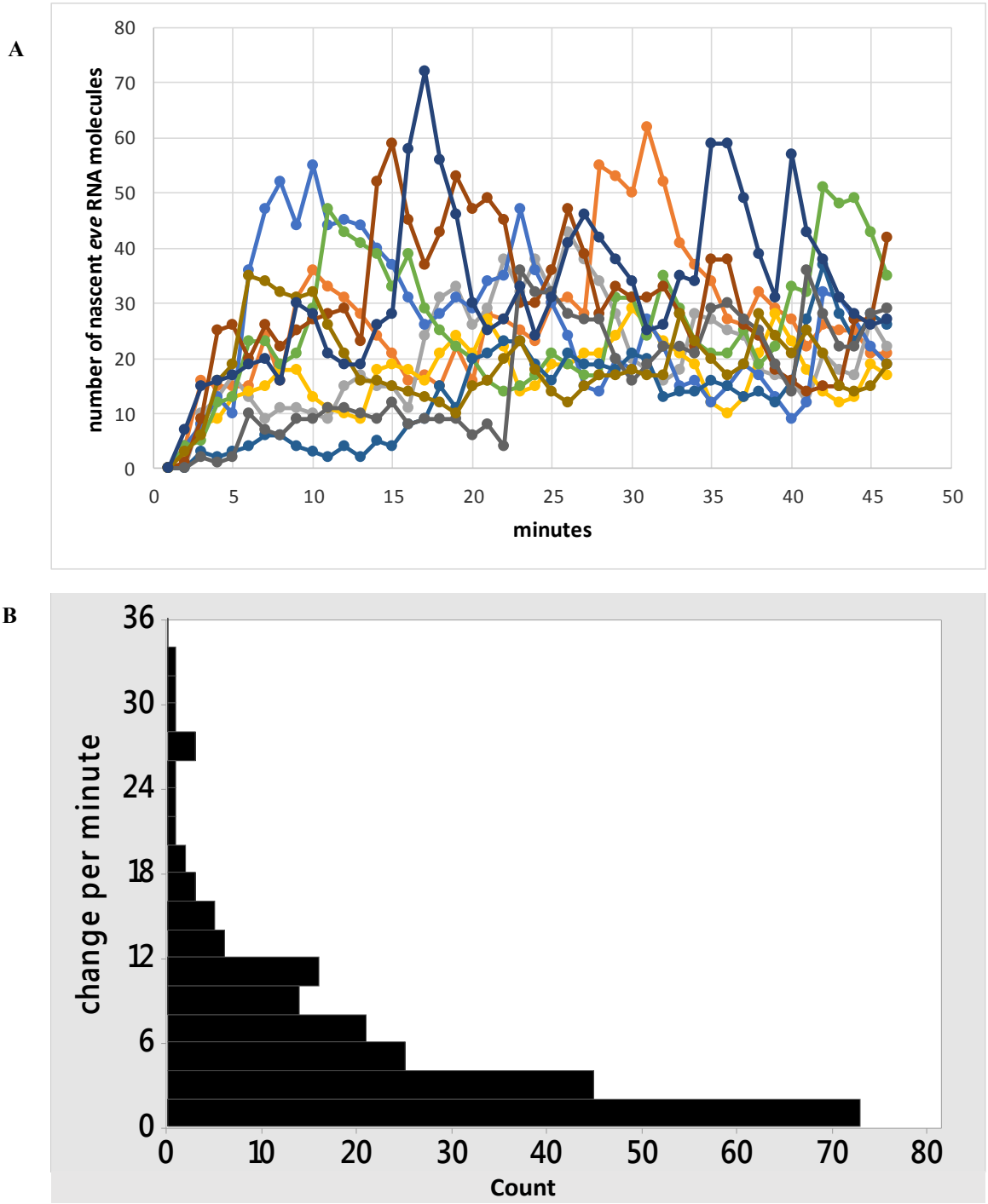

Fig. 1. Time series of *eve* gene transcription with two transcription rates (HIGH, LOW; k_{1000} and k_{1100}, Table 1). A) 45 minute time series for 10 simulations. B) Histogram of the minute-to-minute change in the time series.

Many gene transcription processes have been modelled as a simple ON-OFF mechanism, with a single characteristic transcription rate when ON. To test the capacity for such a mechanism to model *eve* transcription, we set H-binding to zero in the model (Table 1), such that transcription was only through the B channel (Bicoid is necessary for transcriptional activation *in vivo*). k_{1000} was changed to the HIGH value of 0.56/s, corresponding to the highest observed per-minute change in the experimental data [7]. Fig. 2A shows time series from this simple ON-OFF mechanism. The histogram of per-minute changes (Fig. 2B), in contrast to the two-rate HIGH-LOW mechanism (Fig. 1) and the experimental data, shows a sharp change between the lowest two bars. That is, these time series have many minutes with only 0 to 1 new transcripts initiated; rates of 2 transcripts per-minute and higher are much less common. This trend can be seen visually in a number of the time series, with intervals at near zero initiation, interspersed with spikes of very high initiation. With a single

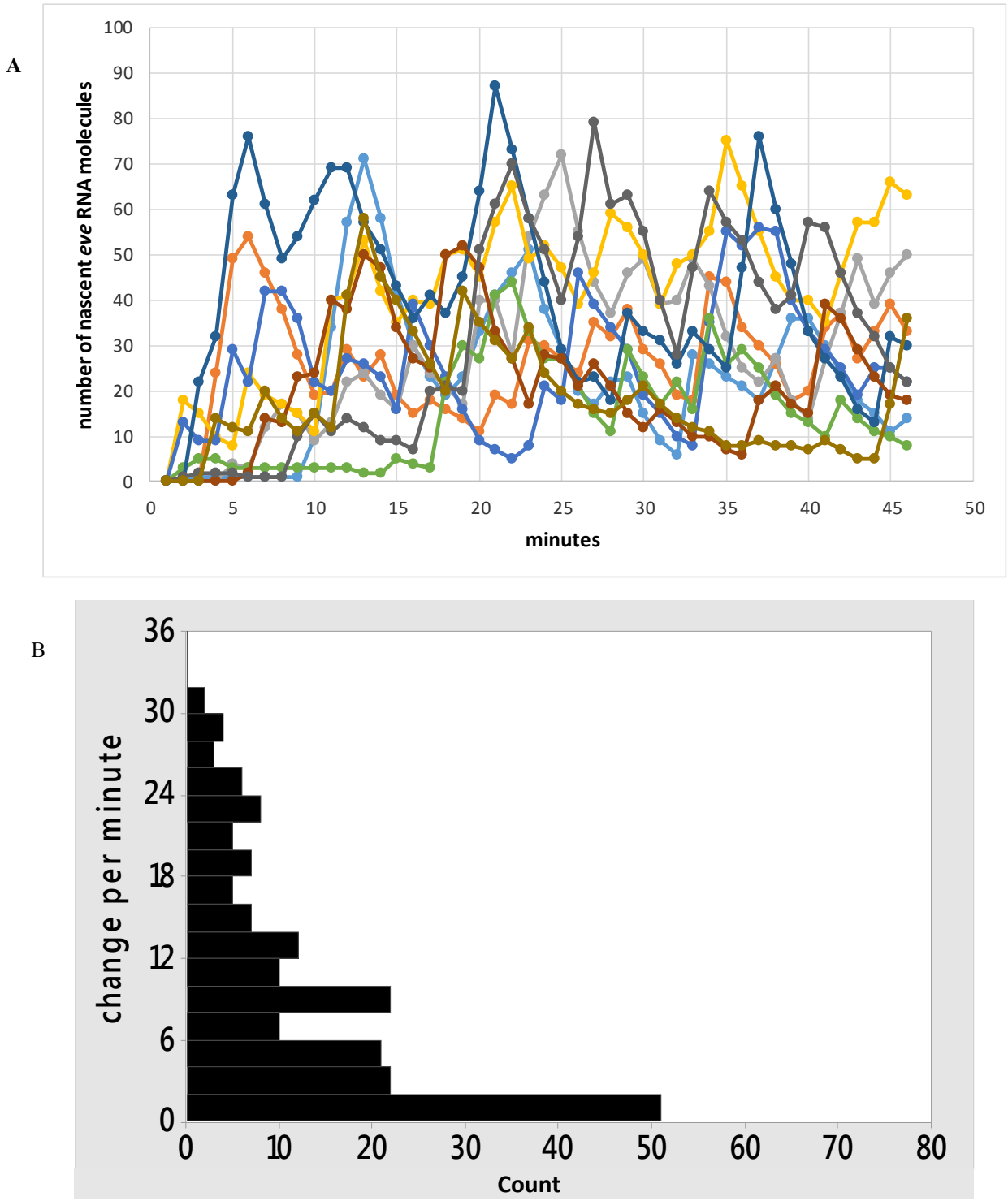

Fig. 2. Time series of *eve* gene transcription with one transcription rate (the HIGH k_{1000} rate), a simple ON-OFF model. A) 45 minute time series for 10 simulations. B) Histogram of the minute-to-minute change in the time series.

HIGH k_{1000} in these simulations, B-binding must be decreased (compared to Fig. 1) in order to generate the experimentally observed total *eve* mRNA. Therefore, the simple ON-OFF mechanism is characterized by strong spikes of transcription, in comparison to the two-rate mechanism and the data. This indicates that the combined LOW (B) and HIGH (B+H) activation contribute to a smoother distribution of initiation rates, and that this can be seen in the experimental time series.

The technology to image gene transcription in living embryos has recently been pioneered [4-6], leading to the first time series at per-minute resolution [7]. As this technology becomes increasingly established, and time series become available at this and higher resolution across developmental systems, stochastic modelling promises to play an important role in interpreting the experimental data and understanding how noise both arises and is controlled during gene expression. This is a critical piece in understanding the reliability of tissue differentiation during embryonic development.

REFERENCES

[1] M.B. Elowitz, A.J. Levine, E.D. Siggia, and P.S. Swain, "Stochastic gene expression in a single cell," Science, vol. 297, pp. 1183 – 1186, 2002.

[2] W.J. Blake, M. Kaern, C.R. Cantor, and J.J. Collins, "Noise in eukaryotic gene expression," Nature, vol. 422, pp. 633-637, 2003.

[3] I. Golding, J. Paulsson, S.M. Zawilski, and E.C. Cox, "Real-time kinetics of gene activity in individual bacteria," Cell, vol. 123, pp. 1025 – 1036, 2005.

[4] H.G. Garcia, M. Tikhonov, A. Lin, and T. Gregor, "Quantitative imaging of transcription in living *Drosophila* embryos links polymerase activity to patterning," Curr. Biol., vol. 23, pp. 2140 – 2145, 2013.

[5] T. Lucas, T. Ferraro, B. Roelens, J. de las Heras Chanes, A.M. Walczak, M. Coppey, and N. Dostatni, "Live imaging of Bicoid-dependent transcription in *Drosophila* embryos. Curr. Biol., vol. 23, pp. 2135 – 2139, 2013.

[6] T. Ferraro, T. Lucas, M. Clemot, J. de las Heras Chanas, J. Desponds, M. Coppey, A.M. Walczak, and N. Dostatni, "New methods to image transcription in living fly embryos: the insights so far, and the prospects," WIREs Dev. Biol., vol. 5, pp. 296 – 310, 2016.

[7] J.B. Bothma, H.G. Garcia, E. Esposito, G. Schlissel, T. Gregor, and M. Levine, "Dynamic regulation of *eve* stripe 2 expression reveals transcriptional bursts in living *Drosophila* embryos", Proc. Nat. Acad. Sci. USA, vol. 111, pp. 10598-10603, 2014.

[8] D.M. Holloway, and A.V. Spirov, "Transcriptional bursting in *Drosophila* development: stochastic dynamics of *eve* stripe 2 expression," PLoS ONE, in press.

[9] S. Small, A. Blair, and M. Levine, "Regulation of *even-skipped* stripe 2 in the *Drosophila* embryo," EMBO J., vol. 11, pp. 4047 – 4057, 1992.

[10] D.N. Arnosti, S. Barolo, M. Levine, and S. Small, "The *eve* stripe 2 enhancer employs multiple modes of transcriptional synergy," Development, vol. 122, pp. 205-214, 1996.

[11] J. Hattne, D. Fange, and J. Elf, "Stochastic reaction-diffusion simulation with MesoRD,: Bioinformatics, vol. 21, pp. 2923-2924, 2005.

[12] J. Elf, A. Dončić, and M. Ehrenberg, "Mesoscopic reaction-diffusion in intracellular signaling," Proc. SPIE, vol 5110, pp. 114-124, 2003.

[13] J. Elf, and M. Ehrenberg, "Spontaneous separation of bi-stable biochemical systems in spatial domains of opposite phases," Syst. Biol., vol. 2, pp. 230-236, 2004.