Title
Studies on the Bursting Kinetics of even-skipped Transcription

Author
Berrocal Quezada, Nelson Augusto

Publication Date
2020

Peer reviewed|Thesis/dissertation

Studies on the Bursting Kinetics of *even-skipped* Transcription

By

Nelson Augusto Berrocal Quezada

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Molecular & Cell Biology

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Michael B. Eisen, Co-Chair
Professor Hernán G. García, Co-Chair
Professor Xavier Darzacq
Professor Murat Arcak

Spring 2020

Abstract

Studies on the Bursting Kinetics of *even-skipped* Transcription

by

Nelson Augusto Berrocal Quezada

Doctor of Philosophy in Molecular and Cell Biology

University of California, Berkeley

Professor Michael B. Eisen, Advisor

Professor Hernán G. García, Co-Advisor

Enhancers drive the transcription of genes in defined spatio-temporal patterns of gene expression that orchestrate the development of animal body plans. The spatial aspects of enhancer function and gene expression pattern formation have been elucidated with elegant experiments in *Drosophila melanogaster* and other organisms. However, our knowledge regarding the temporal dynamics of pattern formation is still incomplete.

In this work, I coupled the MS2 reporter system, which enables real-time, single-locus analysis of transcription, with the regulatory sequences of the gene *even-skipped (eve)*, a fine model of enhancer function active in seven antero-posterior stripes that ring the fruit fly embryo before gastrulation.

Our results show highly dynamic expression patterns that emerge from stochastic bursts of transcription in single loci. We found that the dynamic expression of *eve* stripes stems from shared transcriptional kinetics. Nuclei at the center of all stripes increase their frequency of transcriptional bursting. Nuclei in the inter-stripe boundaries remain quiescent and do not show any detectable bursting. All the while stripes shift

anteriorly, propelled by expression flow. The bursting kinetics in the ectopic patterns of *eve* expression that result from lifting the repression of *eve* enhancers are similar to the kinetics in the endogenous stripes. These observations show that the various *eve* enhancers generate similar bursting kinetics regardless of the myriad of distinct inputs that shape their activity.

This dissertation consists of four chapters. Chapter 1 describes the construction of our MS2-based *eve* reporter and the generation of the transgenic fruit flies carrying it. Chapter 2 goes through our imaging methods and their optimization. It ends with an account of our data analysis to turn movies of transcription in embryonic development into meaningful and quantifiable data. Chapter 3 summarizes our findings regarding the conserved kinetics of *even-skipped* transcriptional bursting. Finally, Chapter 4 explores the activity of *eve* enhancers in body regions where they are not normally active.

Dedication

Dedicated to my mother, Magdalena Quezada Baca; to my abuela, Leonides Baca Ribero; and her sister, Alfonsa Baca Ribero. They were the women who raised me and brought me here.

Acknowledgements

I want to acknowledge the support, advice, and encouragement from every person who made this dissertation possible.

First of all, my mother, my family, and my friends in Berkeley and abroad. They lifted my spirit when necessary and gave me the emotional support to complete my PhD.

I want to thank my advisors, Michael Eisen and Hernán García, for their support, their intellectual advice, and their scientific guidance. They enlightened me, inspired me, and taught me the craft of scientific research.

I express my gratitude to my thesis committee and the MCB faculty for their suggestions, and for teaching me by example the beauty of scientific pursuit.

Finally, I want to acknowledge my friends and colleagues in the Eisen, García, and Darzacq Labs, and my cohort in the MCB graduate program.
Their help, advice, and intellectually stimulating conversations left a very positive footprint on my work and my scientific thinking.

Table of Contents

Introduction

Chapter 1
Construction and Imaging of eveMS2-BAC

Chapter 2

Optimization and Analysis of MS2 Imaging Data

Chapter 3

eve Patterns Emerge from Stochastic Transcription in Single Nuclei

Chapter 4

Shared Bursting Kinetics in Endogenous and Ectopic Patterns of Gene Expression

Introduction

Transcriptional bursting

Transcription occurs in stochastic pulses, or bursts, of mRNA synthesis across domains in life (Chong et al. 2014; Chubb et al. 2006; Senecal et al. 2014; Dar et al. 2012; Skupsky et al. 2010; McKnight and Miller 1979). In 1979, (McKnight and Miller 1979) observed, perhaps, the first evidence of the discontinuous activation of transcription that leads to transcriptional bursting in *D. melanogaster.* With electron microscopy techniques, the authors described nascent mRNAs, which they referred to as 'ribonucleoprotein fibers', branching from the chromatid strands. They wrote: 'Active transcription units can have internal, fiber-free gaps which may result from interruptions in initiation of transcription'.

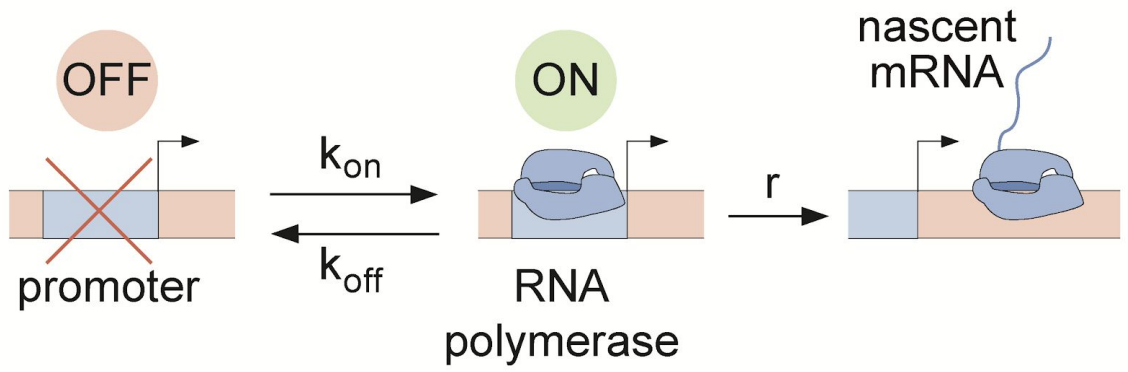
The stochastic occurrence of transcriptional bursts results from the stochastic nature of molecular events. *E.g.* (Chong et al. 2014) found that transcription in bacteria induces supercoiling in transcriptionally active DNA, which slows down and eventually stops transcription. mRNA synthesis can resume only after bacterial gyrases resolve supercoils. The time-lag between the stop of transcription and supercoil resolution results in transcriptional bursting. Other articles suggest that discontinuous transcriptional activity stems from the stochastic and dynamic interplay between chromatin, regulatory elements and transcription factors (Donovan et al. 2019; Raj et al. 2006; Boeger, Griesenbeck, and Kornberg 2008).

A two-step model of promoter activity captures the intermittent nature of mRNA synthesis (Bothma et al. 2014; Garcia et al. 2013; Lammers et al. 2020). A gene promoter switches between its inactive state, OFF, to an active state, ON, at a rate kon. When the promoter is in its ON state, it loads RNA Pol II molecules into its transcriptional unit at a rate r. Eventually, mRNA synthesis stops, and the promoter in ON state transitions to an OFF, inactive, state at a rate koff **(FIG i.1)**. The stochastic transitions between ON and OFF states generate transcriptional bursts. The kinetics of the two-step model of promoter activity reflect features of transcriptional bursting. A higher kon means more frequent transitions of the promoter from OFF to ON, that results in a higher bursting frequency. A lower koff means less frequent transitions from OFF to ON, that results in longer periods of promoter activity; thus, longer bursts. A

higher r means more mRNA Pol II molecules loaded into a promoter while it remains in the ON state; thus, resulting in a higher amplitude of transcriptional bursts.

FIGURE i.1

An inactive promoter (OFF) transitions to its active state (ON) at a kon rate. Frequent transitions from OFF to ON results in more frequent transcriptional bursts. When the promoter is in the ON state, it loads active RNA Pol II molecules at a rate r. Eventually, an ON promoter transitions to the OFF state at a koff rate. Less frequent transitions from ON to OFF (1/koff) result in longer bursts.



Several works have related transcriptional kinetics (kon, koff, r) with biological processes. *E.g. (Senecal et al. 2014)* found that a longer duration of transcription factor binding events leads to larger transcriptional bursts in mammal cells. (Fukaya, Lim, and Levine 2016) observed that stronger *Drosophila melanogaster* enhancers drive a higher frequency of transcriptional bursting in their target genes.

Advances in microscopy, brought by improved live-imaging capabilities such as MS2 and related techniques, enable us to observe transcriptional bursting in real-time at a single locus resolution (Ferguson and Larson 2013; Chubb et al. 2006; Garcia et al. 2013). In this work, we use the MS2 reporter system, which relies on the interaction between the MS2-Coat Protein-GFP fusion protein (MCP::GFP) and the MS2 transcripts to reveal nascent transcription in real time (See Details in Chapter 1 - Figure 1.5), to elucidate the temporal dynamics and behavior of transcriptional bursting en fruit fly embryos.

Transcriptional Bursting Leads to Gene Expression Pattern Formation

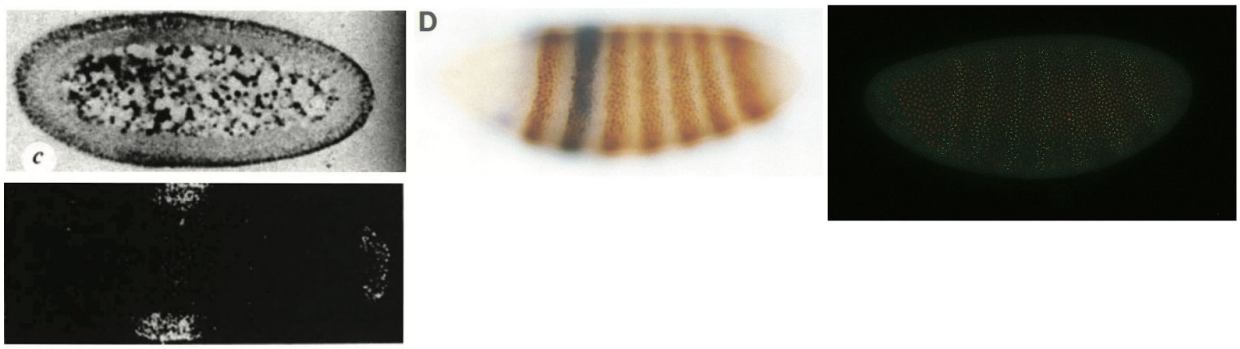
Despite the stochastic nature of transcriptional bursting, the interplay between spatio-temporal signalling, transcription factors and regulatory elements allows for the coordination of transcriptional bursts across many cells in a multicellular organism.

In many animal embryos, enhancers read spatial and temporal queues (Davidson 2010) and orchestrate the simultaneous transcription of a gene in many cells. The simultaneous synthesis of mRNA from a given gene over defined body areas of an embryo results in gene expression patterns.

In *D. melanogaster*, gene expression patterns have many shapes, and we can observe them as swaths, stripes or segments formed by coordinated transcriptional bursting in many nuclei (Garcia et al. 2013; Bothma et al. 2014; Berrocal et al. 2018). We can also see them as areas of higher concentration of a given mRNA or protein (Knipple et al. 1985; Goto, Macdonald, and Maniatis 1989; S. Small, Blair, and Levine 1992) **(FIG i.2)**.

FIGURE i.2

Some of the first observations of gene expression patterns in *D. melanogaster* embryos were made by autoradiography of tritium-labeled RNA probes complementary to mRNAs of developmental genes. On the left panel, we can see an autoradiography of Krüppel transcripts in fruit fly embryos. Adapted from (Knipple et al. 1985). On the middle panel, we can observe the seven-striped *eve* pattern labeled through immunostaining of anti-eve antibodies (brown). The second eve stripe (blue) results from the expression of a LacZ reporter labeled by *in situ* hybridization. Adapted from (S. Small, Blair, and Levine 1992). On the right panel, we can observe the seven stripes of eve as shown by the dynamic reporter MS2 (green spots). All embryos are oriented with their anterior tip to the left and their posterior end to the right.



Gene expression patterns seed the transcriptional programs that lead to the formation of animal body parts. Disruptions in expression patterns result in striking changes in body development, as exemplified by (Nüsslein-Volhard and Wieschaus 1980) in their Heidelberg screen of *D. melanogaster* larvae mutants.

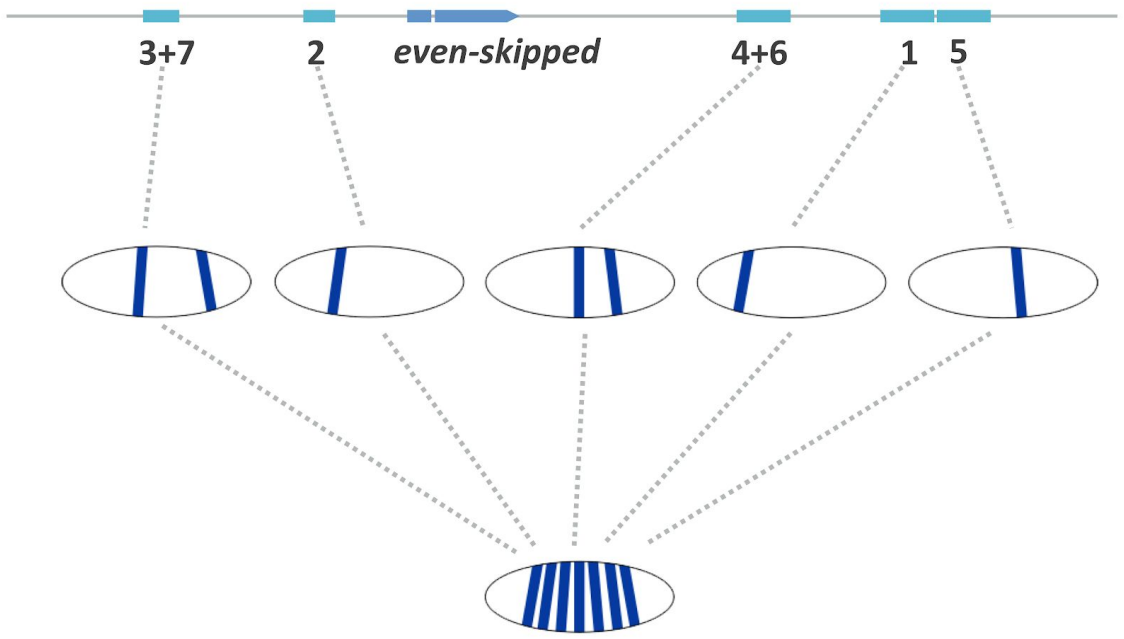
In this work, we aimed to explore and characterize how the stochasticity of transcriptional bursting in single nuclei yields dynamic patterns of gene expression, that change and refine in space and time.

We coupled the MS2 reporter with the *D. melanogaster even-skipped (eve)* locus as our model. The *eve* gene is embedded among five enhancers active in the early stages of embryonic development, and various other regulatory elements. The five early *eve* enhancers drive a regular pattern of stripes that ring *Drosophila* embryos after their fourteenth cycle of synchronous cell division and before gastrulation.

Each enhancer (3+7; 2; 4+6; 1; 5) is named after the one or two stripes it controls. *eve* stripe 3+7 enhancer drives the expression of the third and seventh stripe, *eve* stripe 2 drives the expression of the second stripe, and so forth **(FIG i.3)**. The study of *eve* expression patterns sheds light on the function of five distinct enhancers, each responding to various combinations of input transcription factors.

FIGURE i.3

The *D. melanogaster even-skipped* gene is embedded among multiple regulatory elements. Five of them are enhancers active in embryos after the fourteenth cell cycle after fertilization and before gastrulation. Each of the *eve* enhancers is named after the stripe it controls: *eve* stripe 3 + 7 drives the expression of the third and seventh stripes, *eve* stripe 2 drives the expression of the second stripe, and so forth.The first stripe is the most anterior and the seventh stripes appears at the posterior end. Using *eve* as our model allows the simultaneous observation of the function and transcriptional kinetics of five distinct enhancers at once. For DNA sizes, in kb, of *even-skipped* locus, see Chapter 1 - Figure 1A.



Classic experiments elegantly elucidated the spatial aspects of *eve* pattern formation (Harding et al. 1989; Sackerson, Fujioka, and Goto 1999; Frasch and Levine 1987; S. Small, Blair, and Levine 1992; Jiang, Hoey, and Levine 1991; Goto, Macdonald, and Maniatis 1989; Stephen Small, Blair, and Levine 1996). Nevertheless, the temporal components of *eve* expression patterns were more complicated to address, as the capabilities to image pattern formation in living embryos were limited. The MS2 reporter system used in this work shows with great detail the spatio-temporal transcriptional activity of single nuclei that leads to the emergence of complex *eve* expression patterns. Furthermore, the real-time, single-locus resolution of MS2 data enables the visualization of transcriptional bursting and the inference of the transcriptional kinetics (kon, koff, r) of promoter activity.

The chapters in this thesis go as follows. Chapter 1 describes the construction of the BAC-based reporter system that we used in this thesis, with the MS2 reporter coupled

to the *even-skipped* regulatory elements, and the struggles to establish homozygous fly lines for imaging. Chapter 2 tells the optimization steps of MS2 microscopy in developing embryos and the analysis of the imaging data to convert images into quantifiable and meaningful data. Chapter 3 tells the findings from the analysis of *even-skipped* expression and pattern formation as shown by MS2 transcription. Namely, the dynamic and stochastic nature of transcription that results in a complex pattern of gene expression, and the processes in single loci that shape this pattern. Chapter 4 could be considered the first draft of a paper analyzing MS2 expression in ectopic areas of gene expression to explore the kinetics of transcriptional bursting in body regions that are not normally active. The final goal is to learn about how enhancers function under different regulatory inputs and the origin of new expression patterns.

Bibliography

Berrocal, A., N. C. Lammers, H. G. Garcia, and M. B. Eisen. 2018. “Kinetic Sculpting of the Seven Stripes of the Drosophila Even-Skipped Gene.” *BioRxiv*. https://www.biorxiv.org/content/10.1101/335901v2.abstract.

Boeger, Hinrich, Joachim Griesenbeck, and Roger D. Kornberg. 2008. “Nucleosome Retention and the Stochastic Nature of Promoter Chromatin Remodeling for Transcription.” *Cell* 133 (4): 716–26. https://doi.org/10.1016/j.cell.2008.02.051.

Bothma, Jacques P., Hernan G. Garcia, Emilia Esposito, Gavin Schlissel, Thomas Gregor, and Michael Levine. 2014. “Dynamic Regulation of Eve Stripe 2 Expression Reveals Transcriptional Bursts in Living Drosophila Embryos.” *Proceedings of the National Academy of Sciences of the United States of America* 111 (29): 10598–603. https://doi.org/10.1073/pnas.1410022111.

Chong, Shasha, Chongyi Chen, Hao Ge, and X. Sunney Xie. 2014. “Mechanism of Transcriptional Bursting in Bacteria.” *Cell* 158 (2): 314–26. https://doi.org/10.1016/j.cell.2014.05.038.

Chubb, Jonathan R., Tatjana Trcek, Shailesh M. Shenoy, and Robert H. Singer. 2006. “Transcriptional Pulsing of a Developmental Gene.” *Current Biology: CB* 16 (10): 1018–25. https://doi.org/10.1016/j.cub.2006.03.092.

Dar, Roy D., Brandon S. Razooky, Abhyudai Singh, Thomas V. Trimeloni, James M. McCollum, Chris D. Cox, Michael L. Simpson, and Leor S. Weinberger. 2012. “Transcriptional Burst Frequency and Burst Size Are Equally Modulated across the Human Genome.” *Proceedings of the National Academy of Sciences of the United States of America* 109 (43): 17454–59. https://doi.org/10.1073/pnas.1213530109.

Davidson, Eric H. 2010. *The Regulatory Genome: Gene Regulatory Networks In Development And Evolution*. Elsevier. https://play.google.com/store/books/details?id=F2ibJj1LHGEC.

Donovan, Benjamin T., Anh Huynh, David A. Ball, Heta P. Patel, Michael G. Poirier, Daniel R. Larson, Matthew L. Ferguson, and Tineke L. Lenstra. 2019. “Live-Cell Imaging Reveals the Interplay between Transcription Factors, Nucleosomes, and Bursting.” *The EMBO Journal* 38 (12). https://www.embopress.org/doi/abs/10.15252/embj.2018100809.

Ferguson, Matthew L., and Daniel R. Larson. 2013. “Measuring Transcription Dynamics in Living Cells Using Fluctuation Analysis.” *Methods in Molecular Biology* 1042: 47–60. https://doi.org/10.1007/978-1-62703-526-2_4.

Frasch, M., and M. Levine. 1987. “Complementary Patterns of Even-Skipped and Fushi Tarazu Expression Involve Their Differential Regulation by a Common Set of

Segmentation Genes in Drosophila." *Genes & Development* 1 (9): 981–95. https://www.ncbi.nlm.nih.gov/pubmed/2892761.

Fukaya, Takashi, Bomyi Lim, and Michael Levine. 2016. "Enhancer Control of Transcriptional Bursting." *Cell* 166 (2): 358–68. https://doi.org/10.1016/j.cell.2016.05.025.

Garcia, Hernan G., Mikhail Tikhonov, Albert Lin, and Thomas Gregor. 2013. "Quantitative Imaging of Transcription in Living Drosophila Embryos Links Polymerase Activity to Patterning." *Current Biology: CB* 23 (21): 2140–45. https://doi.org/10.1016/j.cub.2013.08.054.

Goto, T., P. Macdonald, and T. Maniatis. 1989. "Early and Late Periodic Patterns of Even Skipped Expression Are Controlled by Distinct Regulatory Elements That Respond to Different Spatial Cues." *Cell* 57 (3): 413–22. https://doi.org/10.1016/0092-8674(89)90916-1.

Harding, K., T. Hoey, R. Warrior, and M. Levine. 1989. "Autoregulatory and Gap Gene Response Elements of the Even-Skipped Promoter of Drosophila." *The EMBO Journal* 8 (4): 1205–12. https://www.ncbi.nlm.nih.gov/pubmed/2743979.

Jiang, J., T. Hoey, and M. Levine. 1991. "Autoregulation of a Segmentation Gene in Drosophila: Combinatorial Interaction of the Even-Skipped Homeo Box Protein with a Distal Enhancer Element." *Genes & Development* 5 (2): 265–77. https://www.ncbi.nlm.nih.gov/pubmed/1671662.

Knipple, D. C., E. Seifert, U. B. Rosenberg, A. Preiss, and H. Jäckle. 1985. "Spatial and Temporal Patterns of Krüppel Gene Expression in Early Drosophila Embryos." *Nature* 317 (6032): 40–44. https://doi.org/10.1038/317040a0.

Lammers, Nicholas C., Vahe Galstyan, Armando Reimer, Sean A. Medin, Chris H. Wiggins, and Hernan G. Garcia. 2020. "Multimodal Transcriptional Control of Pattern Formation in Embryonic Development." *Proceedings of the National Academy of Sciences of the United States of America* 117 (2): 836–47. https://doi.org/10.1073/pnas.1912500117.

McKnight, S. L., and O. L. Miller Jr. 1979. "Post-Replicative Nonribosomal Transcription Units in D. Melanogaster Embryos." *Cell* 17 (3): 551–63. https://doi.org/10.1016/0092-8674(79)90263-0.

Nüsslein-Volhard, C., and E. Wieschaus. 1980. "Mutations Affecting Segment Number and Polarity in Drosophila." *Nature* 287 (5785): 795–801. https://doi.org/10.1038/287795a0.

Raj, Arjun, Charles S. Peskin, Daniel Tranchina, Diana Y. Vargas, and Sanjay Tyagi. 2006. "Stochastic mRNA Synthesis in Mammalian Cells." *PLoS Biology* 4 (10): e309. https://doi.org/10.1371/journal.pbio.0040309.

Sackerson, C., M. Fujioka, and T. Goto. 1999. "The Even-Skipped Locus Is Contained in a 16-Kb Chromatin Domain." *Developmental Biology* 211 (1): 39–52.

https://doi.org/10.1006/dbio.1999.9301.

Senecal, Adrien, Brian Munsky, Florence Proux, Nathalie Ly, Floriane E. Braye, Christophe Zimmer, Florian Mueller, and Xavier Darzacq. 2014. “Transcription Factors Modulate c-Fos Transcriptional Bursts.” *Cell Reports* 8 (1): 75–83. https://doi.org/10.1016/j.celrep.2014.05.053.

Skupsky, Ron, John C. Burnett, Jonathan E. Foley, David V. Schaffer, and Adam P. Arkin. 2010. “HIV Promoter Integration Site Primarily Modulates Transcriptional Burst Size Rather Than Frequency.” *PLoS Computational Biology*. https://doi.org/10.1371/journal.pcbi.1000952.

Small, S., A. Blair, and M. Levine. 1992. “Regulation of Even-Skipped Stripe 2 in the Drosophila Embryo.” *The EMBO Journal* 11 (11): 4047–57. https://www.ncbi.nlm.nih.gov/pubmed/1327756.

Small, Stephen, Adrienne Blair, and Michael Levine. 1996. “Regulation of Two Pair-Rule Stripes by a Single Enhancer in theDrosophilaEmbryo.” *Developmental Biology* 175 (2): 314–24. http://www.sciencedirect.com/science/article/pii/S0012160696901177.

Chapter 1
Construction and Imaging of eveMS2-BAC

eveMS2-BAC: a reporter system for *even-skipped*

Drosophila melanogaster even-skipped (eve) gene is embedded among various regulatory elements **(FIG 1.1A)**, including five enhancers active between the fourteenth cell division after fertilization (nc14) and before gastrulation. These five enhancers drive *eve* expression in a pattern of seven stripes that ring fruit fly embryos along their antero-posterior axis before the onset of gastrulation. *eve* enhancers are named after the one or two stripes that each controls: eve stripe 3+7 (eveS3+7) drives the expression of the third and seventh antero-posterior stripe of *eve* expression; eve stripe 2 (eveS2) drives the expression of the second antero-posterior stripe of *eve* expression, and so forth.

Thus, *eve* is a fine model to study enhancer function (P. Jiang et al. 2015; Ilsley et al. 2013; Goto, Macdonald, and Maniatis 1989; Fujioka et al. 1999; Arnosti et al. 1996; Hare et al. 2008; Ludwig, Patel, and Kreitman 1998; Crocker and Stern 2017), as analyzing the expression of *eve* stripes sheds light into the function of the underlying enhancers that drive their expression, *E.g.*, a closer analysis of expression at eveS2, combined with genetic disruption of the eveS2 enhancer, would elucidate the role of the sequence of the eveS2 enhancer in the shaping features of eveS2 expression.

We combined the advantages offered by spatial *eve* expression patterns to uncover enhancer biology, with the MS2 reporter system. MS2 relies on a DNA sequence derived from the phage MS2 that binds to MS2 Coat Protein fused with GFP (MCP::GFP). This interaction lights up MS2 mRNAs and enables us to follow the fate of nascent mRNAs (Bertrand et al. 1998; Garcia et al. 2013). Thus, we can visualize dynamic spatio-temporal transcriptional bursting, at a single locus level, in live *D. melanogaster* embryos. The details of MS2 function are explained below, in the section Visualization of eveMS2-BAC expression, in this chapter.

Genetic disruption of regulatory elements and further visualization of gene expression patterns has been very successful to uncover the biology of enhancers and other regulatory elements. Nevertheless, the mutation of genomic regulatory elements acting

upon developmental genes, as is the case for *eve* enhancers, might be hard to achieve and maintain, as their disruptions are lethal for fruit flies. Therefore, multiple works have used reporter genes coupled to a regulatory element of interest, including the classic and elegant experiments that dissected the spatial characteristics of *eve* expression patterns (J. Jiang, Hoey, and Levine 1991; Fujioka et al. 1999; Small, Blair, and Levine 1992; Harding et al. 1989).

Construction of eveMS2-BAC

We modified a bacterial artificial chromosome (BAC)-based reporter under the control of the entire *eve* regulatory sequence, and replaced the *eve* coding sequence with the reporter MS2 **(FIG 1.1B)**.

(Venken et al. 2009) generated the original *eve* BAC-reporter. They used a BAC backbone able to insert in the fruit fly genome through P-element and ΦC31 mediated recombination. Thus, this BAC backbone is named 'P/ΦC31 artificial chromosome for manipulation' (P[acman]) (Venken et al. 2006). They used the P[acman] BAC to carry a ~20kb DNA fragment around *D. melanogaster even-skipped* coding region. The ~20kb fragment was derived from a library of ~20kb-long fragments from the fruit fly genome, named CHORI-322 by (Venken et al. 2009). Therefore, the P[acman] BAC carrying the ~20kb fragment around *eve* is referred as P[acman] CHORI-322. Finally, (Venken et al. 2009) turned the *eve* P[acman]-CHORI-322 into a reporter construct, by replacing the coding sequence of *eve* with a GFP gene that recapitulates the seven-striped pattern of *eve* expression.

The Bellen Lab, where (Venken et al. 2009) developed their work, kindly provided us with the *eve* P[acman]-CHORI-322 version that had the *eve* coding sequence replaced with GFP. We used a two-step *galK* mediated recombineering (Warming et al. 2005) to substitute the GFP coding sequence with the MS2 reporter **(FIG 1.1B)**. I will refer to the final *eve* P[acman]-CHORI-322 carrying as eveMS2-BAC.

FIGURE 1.1A

The endogenous gene even-skipped (~1.5Kb) is embedded in a ~20Kb sequence that contains five enhancers and multiple regulatory elements.

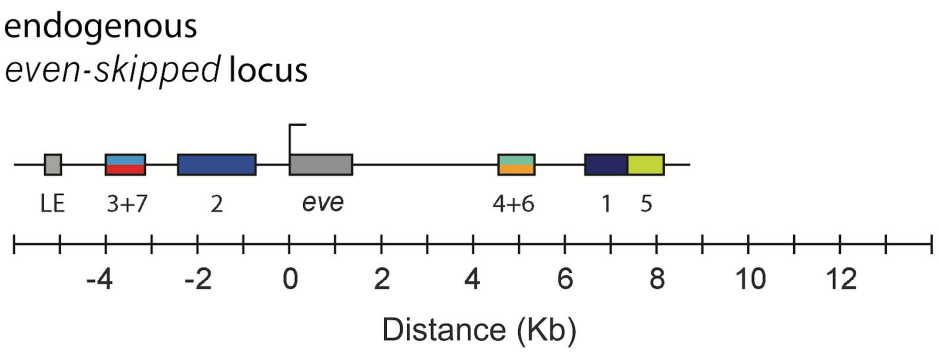
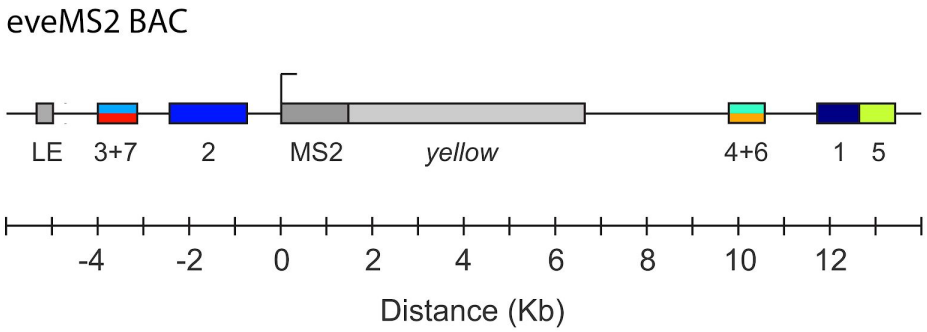


FIGURE 1.2B

We substituted *even-skipped* for the MS2 reporter (~6.5Kb) in the P[acman]-CHORI-322 BAC to observe the expression of MS2-labeled nascent mRNAs. We will refer to this construct as eveMS2-BAC.



Two-step *galK* mediated recombineering consists in replacing the GFP sequence in the *eve* P[acman]-CHORI-322 with a *galK* cassette through homology arm recombination, and then substitute *galK* cassette with MS2. **(FIG 1.2)**

The *galK* mediated recombineering method that we followed requires the bacterium *Escherichia coli* of strain SW102. It expresses the λ phage-derived recombinases *exo, bet*, and *gam* (Warming et al. 2005), under a temperature-sensitive repressor (*cI857*). The λ-derived recombinases are toxic, as they induce genomic rearrangements. Therefore, *E.coli* SW102 cells have to be grown at <30°C, and the recombinases have to be expressed by a sudden, 2 minutes long heat-shock (42°C). If *E.coli* SW102 cells grow at 37°C, they express the λ phage-derived recombinases constitutively. As the recombinases are toxic, this selects for mutations that inactivate the recombinases.

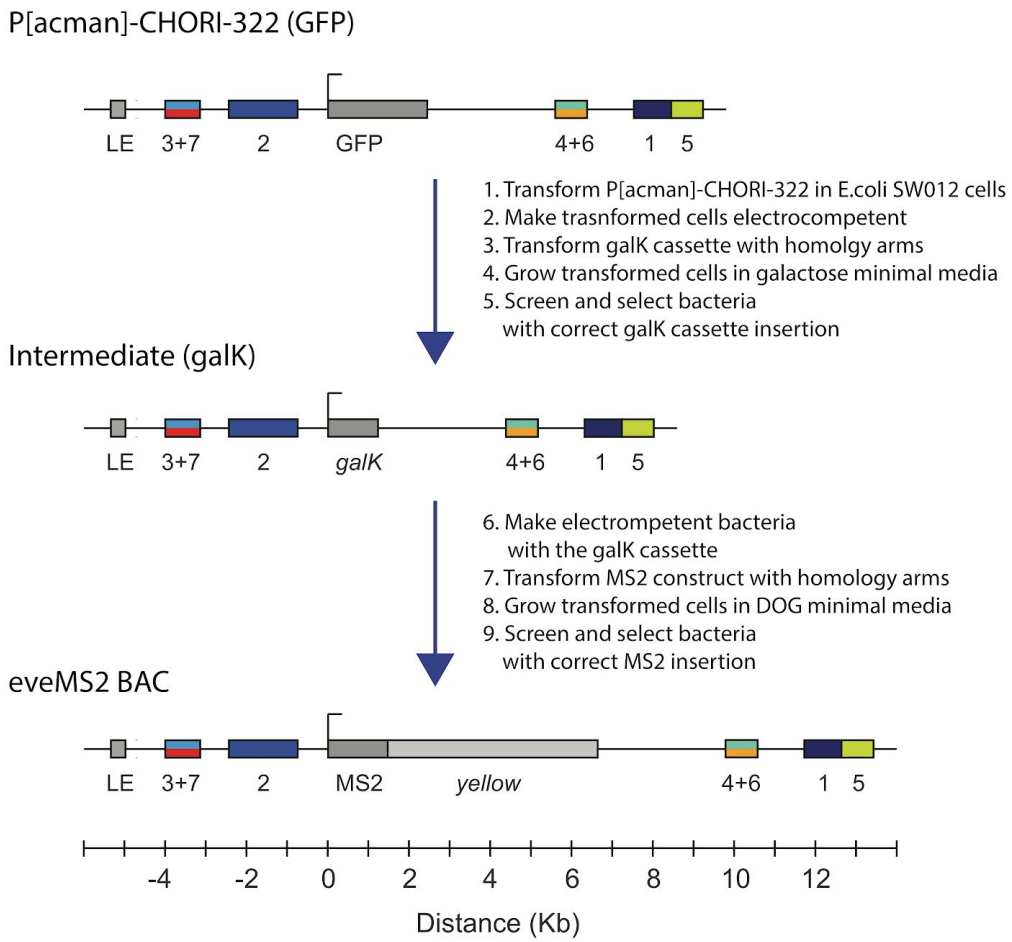
Thus, it is very important to grow *E.coli* SW102 cells at temperatures under 30°C to retain their functionality for recombineering.

The first step of two-step *galK* mediated recombineering **(FIG 1.2)** consists of transforming a *galK* cassette with homology arms into competent *E.coli* SW102 cells (Warming et al. 2005). In this case, the homology arms of the *galK* cassette were complementary to the sequence flanking the GFP gene in *eve* P[acman]-CHORI-322. As *E.coli* SW102 bacteria have their *galK* gene knocked out, minimal media with galactose as the only carbon source selects for bacteria with the *galK* cassette insertion (Warming et al. 2005). I screened colonies able to grow in galactose minimal media and made electrocompetent cells with those that underwent successful *galK* cassette insertions in the right position in P[acman]-CHORI-322. I tested and validated bacterial colonies for *galK* cassette insertions with restriction enzyme digestion (ApaLI), PCR-amplification of insertion sites and Sanger sequencing.

The second step of the two-step *galK* mediated recombineering **(FIG 1.2)** consisted of transforming the MS2 reporter with homology arms flanking the *galK* cassette into competent *E.coli* SW102 cells, in order to replace the *galK* cassette for MS2. Minimal media plates with 2-deoxy-galactose (DOG) select against bacteria with an active *galK* gene. The galactokinase encoded by the *galK* gene phosphorilates DOG and turns it into 2-deoxy-galactose-1-phosphate, a toxic molecule for bacteria. Thus, bacteria with an active *galK* gene are counterselected, and only those with a disrupted *galK* cassette are able to form colonies. (Warming et al. 2005)
I screened colonies able to grow in DOG minimal media through restriction enzyme digestion (ApaLI, XhoI, SmaI, and EcoRI), PCR-amplification of insertion sites and Sanger sequencing. The long size of the MS2 reporter (6.6Kb) made the insertion hard and low-yield, and I required to screen ~100 colonies to find one colony with the correct insertion of MS2.

FIGURE 1.2

We did two-step *galK*-cassette mediated recombineering to substitute the GFP reporter in P[acman]-CHORI-322 for MS2. The steps involved in the recombineering process are lengthy and complicated. They consist of a string of DNA transformations in various strains of *E.coli*. As described above.



I did the two-step *galK* mediated recombineering experiments during the first year after I joined the Eisen/Garcia labs. These experiments are not simple, they require many steps to be performed with utmost care. Due to my own inexperience, coupled with the teaching and qualifying exam commitments that I had in my first year in the Eisen/Garcia labs, it took me ~1 year to replace the GFP coding sequence for MS2 in the *eve* P[acman]-CHORI-322, to make the eveMS2-BAC reporter. I had to start the recombineering process from scratch, twice. Perhaps, I inadvertently grew the *E.coli* SW102 cells at 37°C, selecting for those with disrupted λ-derived recombinases. Many times, I found that the MS2 insert was present in colonies, but the PCR and Sanger sequencing validations showed that no recombination had happened to replace the

galK cassette with MS2. It was until I restarted the whole process, using fresh *E.coli* SW102 cells, that the experiment was successful.

It was hard and repetitive work to find a colony carrying the *eve* P[acman]-CHORI-322 with the right MS2 insertion.

The final result of the recombineering process is the eveMS2-BAC: a 39.2Kb construct that contains a ~20Kb DNA sequence with all the enhancer and regulatory elements surrounding the *eve* gene, embedded in a ~12.5Kb P[acman] backbone. The backbone contains a low-copy and an inducible high-copy origin of replication, a *white+* genetic marker, a chloramphenicol resistance gene and an *attB* site compatible with *attP* landing sites for ΦC31 mediated recombination in *Drosophila (Venken et al. 2006).*

Instead of the 1.5Kb *eve* coding sequence, the eveMS2-BAC has the MS2 reporter, a 6.5Kb-long reporter sequence that includes the 5' UTR from *eve*, a 24x repeats of the MS2 stable repeat ATCCC, and the *D. melanogaster yellow* reporter gene, to give stability and increase the size of the MS2 transcriptional unit. (Bothma et al. 2014)

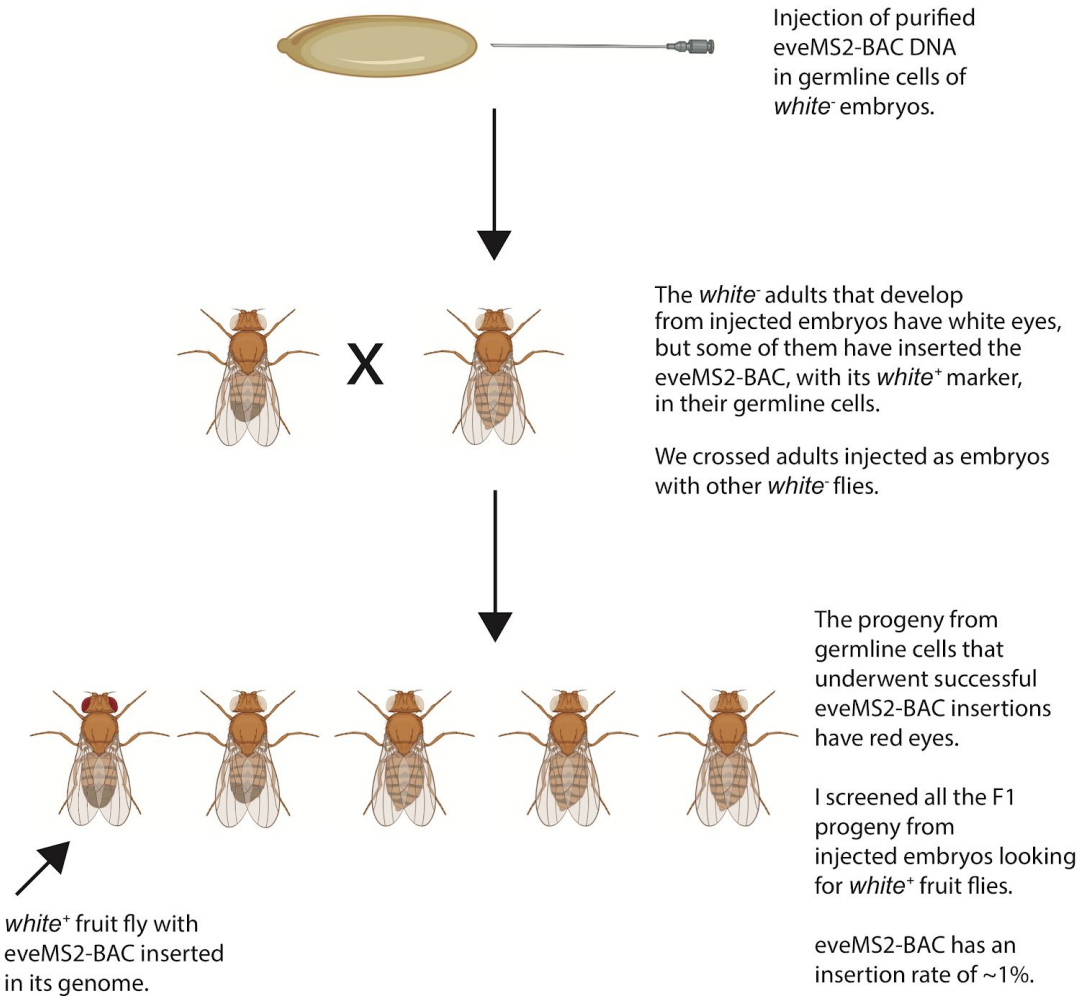
Injection of the eveMS2-BAC in the fruit fly genome

I validated that the eveMS2-BAC had the right sequence. I transformed the eveMS2-BAC in electrocompetent Epi300 cells to induce a high copy number. The λ-derived recombinases in SW102 cells can cause DNA rearrangements of the BAC if there is more than one copy of the eveMS2-BAC per bacterium. *E.coli* Epi300 bacteria do not express any recombinases, therefore, they are ideal to induce high copy number for purification of eveMS2-BAC. I purified the BAC with MIDI Prep (Qiagen plasmid Midiprep kit), and sent it to Rainbow Transgenic Flies Inc. for injection in *D. melanogaster* embryos bearing a ΦC31 *AttP* landing site on the chromosome 3L (Bloomington stock #24871; landing site VK00033; cytological location 65B2) **(FIG 1.3)**.

I screened the F1 progeny from the larvae shipped by Rainbow Transgenic Flies Inc. looking for red eyes, as a marker for the presence of the *white+* gene located in eveMS2-BAC **(FIG 1.3)**. Due to the large size of the eveMS2-BAC and the low rate of success of injections of this construct, I had to screen ~100 flies to find a single fly with red eyes. The 1% rate of eveMS2-BAC insertion in *D. melanogaster* was consistent across injections of many variations of the eveMS2-BAC.

FIGURE 1.3

We sent purified eveMS2-BAC DNA to Rainbow Transgenic Flies Inc. for injection in the germline cells of fruit fly embryos harboring the ΦC31 *AttP* landing site VK00033, located in the 3L chromosome.
Rainbow Transgenic Flies Inc. sent us back injected w^- embryos. We crossed the adult fruit flies from those embryos with w^- flies, and screened the F1 progeny for red eyes. Red eyes progeny from w- parents is indicative of eveMS2-BAC insertion, as it carries a w^+ marker. Then, we balanced the w^+ (red eyes) flies to generate a viable eveMS2-BAC homozygous line. (Flies, embryo and needle designs from https://biorender.com)



Visualization of eveMS2-BAC expression

The MS2 reporter system is very useful to track mRNA molecules. In particular, it enables the visualization of nascent mRNAs in real-time, in living organisms, including *Drosophila* embryos (Garcia et al. 2013).

MS2 is a DNA repetitive sequence derived from the phage MS2. When transcribed, it forms 24 mRNA stem-loops with high affinity for the MS2 Coat Protein (MCP), a structural protein derived from the phage MS2, fused with GFP (MCP::GFP) (Bertrand et al. 1998) **(FIG 1.4)**.

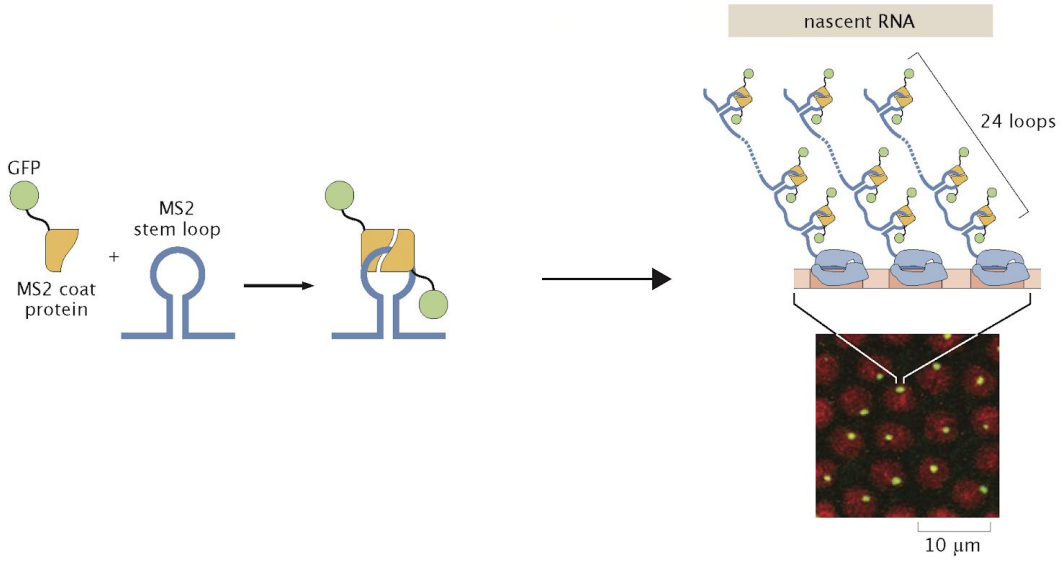
In our set up, the MCP::GFP fusion protein is maternally deposited and diffuses across the entire *D. melanogaster* embryo. It gives the embryo a uniform green fluorescence under the confocal microscope. As each of the 24 stem-loops in a nascent MS2 mRNA binds two MCP::GFP molecules, and many mRNAs are synthesized in a single bout of promoter activity, we observe a fluorescent spot of high concentration of GFP at sites of active MS2 transcription. By tracking the fluctuations in fluorescence of these spots, over time, we can observe, and analyze, the kinetics of transcriptional bursting.

The maternally deposited MCP::GFP fusion protein diffuses across the entire fruit fly embryo, and gives the embryo a uniform green (488nm wavelength) fluorescence under the confocal microscope. As each of the 24 stem-loops in a nascent MS2 mRNA binds two MCP::GFP molecules, and many mRNAs are synthesized in a single bout of promoter activity, we observe a fluorescent spot of high concentration of GFP at sites of active MS2 transcription. By tracking the fluctuations in fluorescence of these spots, over time, we can observe, and analyze, the kinetics of transcriptional bursting.

The maternally deposited His::RFP fusion protein labels chromatin in red (561nm wavelength). It enables a defined visualization of nuclei in *D. melanogaster* embryos at the blastocyst stage. Thus, we can track nuclei through early development and match them with their corresponding fluorescent spots of active MS2 transcription.

FIGURE 1.4

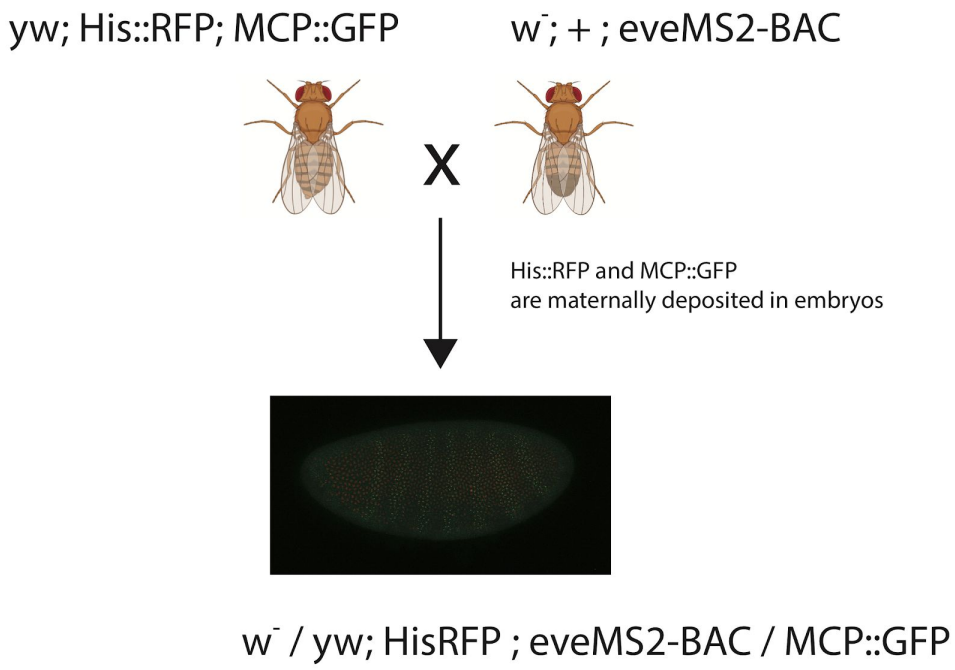
The MS2 reporter consists of a repetitive DNA sequence derived from the phage MS2, that when transcribed, forms 24 mRNA stem loops (MS2 stem loop). The MS2 coat protein (MCP) binds with high affinity and form dimers on these stem loops. MCP is fused to GFP (MCP::GFP) to fluoresce. Because the MCP binds to 24 stem loops per transcript, and many mRNA molecules are transcribed simultaneously, we can observe fluorescent spots at sites of MS2 transcription due to the high local concentration of GFP.



In order to visualize the expression of MS2 under control of the *eve* regulatory region, in real time, at single locus resolution, I crossed ~30 homozygous males carrying MS2 (w- ; + ; eveMS2-BAC) with ~100 homozygous virgin females of the genotype (yw; His::RFP; MCP::GFP), as the maternal source for the proteins necessary for visualization of MS2 transcription under a confocal microscope **(FIG 1.5)**.

FIGURE 1.5

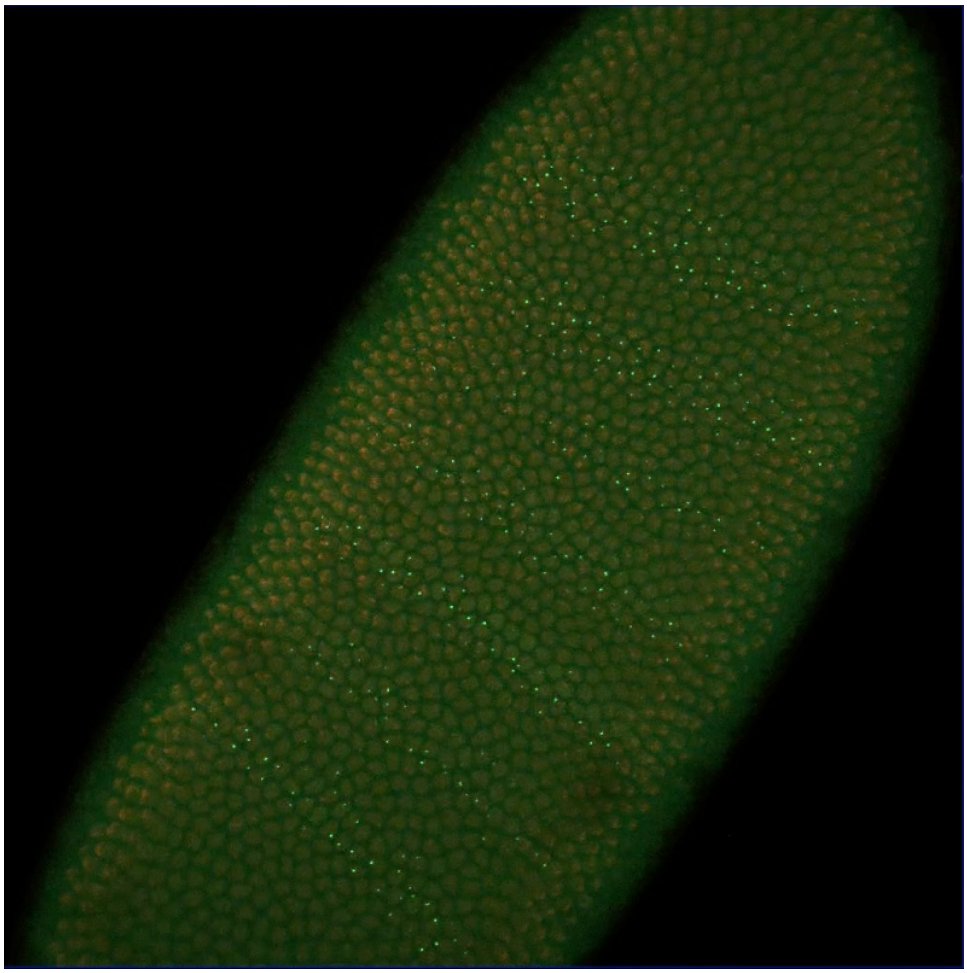
We crossed w⁻;+;eveMS2-BAC homozygous males with yw;His::RFP;MCP::GFP females to generate embryos for MS2 imaging under a confocal microscope. The w⁻;+;eveMS2-BAC homozygous males inherit one eveMS2-BAC allele. Therefore, we would see only one spot of MS2 fluorescence per nucleus.
MCP::GFP illuminates MS2 nascent transcripts in real-time, while His:RFP labels nuclei for reference and better visualization.
yw; HIS::RFP; MCP::GFP females deposit the fluorescent proteins His:RFP and MCP::GFP maternally in the embryos. This is very important, as these fluorescent proteins need ~30min to mature and fluoresce, and eveMS2 transcription starts ~1 hour after fertilization. Which gives enough time to GFP and RFP for maturations and allows them to capture the fast processes of cell division and transcription that occur in early embryonic development. (Fly designs from https://biorender.com)



After ~1 year of work, I finally achieved recording a movie of eveMS2-BAC expression in the early stages of *D. melanogaster* embryogenesis. There was a long way ahead before figuring out how to analyze and extract meaningful information of the eveMS2-BAC expression movies. However, since the very first moment that I observed the expression of my construct under the confocal microscope, I was able to see the canonical *eve* seven-stripe pattern drawn by the expression of MS2 **(FIG 1.6)**.
The eveMS2-BAC reporter is useful to visualize the dynamic expression of eve patterns in space and time, and the position of each of the seven stripes of *eve* expression corresponds to the position of the endogenous stripes as described in the literature.

FIGURE 1.6

This image comes from the first video of an embryo displaying the characteristic seven *eve* stripes with MS2 transcription. At this point, there was still a long way ahead to figure out how to analyze and extract meaningful information from this data.



Furthermore, eveMS2-BAC reporter enables the disruption and modification of its many enhancers, regulatory elements, and spacer sequences, without the deleterious effects in fruit fly viability associated with disrupting the endogenous sequences surrounding the gene *even-skipped*. Disrupting and modifying genes and regulatory sequences have been very powerful genetic tools to understand the genetics of gene regulation and embryonic development. The eveMS2-BAC reporter unleashes the possibility to disrupt and modify the regulatory DNA of *eve* to better understand the role of regulatory elements in shaping the kinetics of transcriptional bursting.

Bibliography

Arnosti, D. N., S. Barolo, M. Levine, and S. Small. 1996. “The Eve Stripe 2 Enhancer Employs Multiple Modes of Transcriptional Synergy.” *Development* 122 (1): 205–14. https://www.ncbi.nlm.nih.gov/pubmed/8565831.

Bertrand, E., P. Chartrand, M. Schaefer, S. M. Shenoy, R. H. Singer, and R. M. Long. 1998. “Localization of ASH1 mRNA Particles in Living Yeast.” *Molecular Cell* 2 (4): 437–45. https://doi.org/10.1016/s1097-2765(00)80143-4.

Bothma, Jacques P., Hernan G. Garcia, Emilia Esposito, Gavin Schlissel, Thomas Gregor, and Michael Levine. 2014. “Dynamic Regulation of Eve Stripe 2 Expression Reveals Transcriptional Bursts in Living Drosophila Embryos.” *Proceedings of the National Academy of Sciences of the United States of America* 111 (29): 10598–603. https://doi.org/10.1073/pnas.1410022111.

Crocker, Justin, and David L. Stern. 2017. “Functional Regulatory Evolution outside of the Minimal *even-Skipped* Stripe 2 Enhancer.” *Development* , July. https://doi.org/10.1242/dev.149427.

Fujioka, M., Y. Emi-Sarker, G. L. Yusibova, T. Goto, and J. B. Jaynes. 1999. “Analysis of an Even-Skipped Rescue Transgene Reveals Both Composite and Discrete Neuronal and Early Blastoderm Enhancers, and Multi-Stripe Positioning by Gap Gene Repressor Gradients.” *Development* 126 (11): 2527–38. https://www.ncbi.nlm.nih.gov/pubmed/10226011.

Garcia, Hernan G., Mikhail Tikhonov, Albert Lin, and Thomas Gregor. 2013. “Quantitative Imaging of Transcription in Living Drosophila Embryos Links Polymerase Activity to Patterning.” *Current Biology: CB* 23 (21): 2140–45. https://doi.org/10.1016/j.cub.2013.08.054.

Goto, T., P. Macdonald, and T. Maniatis. 1989. “Early and Late Periodic Patterns of Even Skipped Expression Are Controlled by Distinct Regulatory Elements That Respond to Different Spatial Cues.” *Cell* 57 (3): 413–22. https://doi.org/10.1016/0092-8674(89)90916-1.

Harding, K., T. Hoey, R. Warrior, and M. Levine. 1989. “Autoregulatory and Gap Gene Response Elements of the Even-Skipped Promoter of Drosophila.” *The EMBO Journal* 8 (4): 1205–12. https://www.ncbi.nlm.nih.gov/pubmed/2743979.

Hare, Emily E., Brant K. Peterson, Venky N. Iyer, Rudolf Meier, and Michael B. Eisen. 2008. “Sepsid Even-Skipped Enhancers Are Functionally Conserved in Drosophila despite Lack of Sequence Conservation.” *PLoS Genetics* 4 (6): e1000106. https://doi.org/10.1371/journal.pgen.1000106.

Ilsley, Garth R., Jasmin Fisher, Rolf Apweiler, Angela H. De Pace, and Nicholas M. Luscombe. 2013. “Cellular Resolution Models for Even Skipped Regulation in the

Entire Drosophila Embryo." *eLife* 2 (August): e00522. https://doi.org/10.7554/eLife.00522.

Jiang, J., T. Hoey, and M. Levine. 1991. "Autoregulation of a Segmentation Gene in Drosophila: Combinatorial Interaction of the Even-Skipped Homeo Box Protein with a Distal Enhancer Element." *Genes & Development* 5 (2): 265–77. https://www.ncbi.nlm.nih.gov/pubmed/1671662.

Jiang, Pengyao, Michael Z. Ludwig, Martin Kreitman, and John Reinitz. 2015. "Natural Variation of the Expression Pattern of the Segmentation Gene Even-Skipped in Melanogaster." *Developmental Biology* 405 (1): 173–81. https://doi.org/10.1016/j.ydbio.2015.06.019.

Ludwig, M. Z., N. H. Patel, and M. Kreitman. 1998. "Functional Analysis of Eve Stripe 2 Enhancer Evolution in Drosophila: Rules Governing Conservation and Change." *Development* 125 (5): 949–58. https://www.ncbi.nlm.nih.gov/pubmed/9449677.

Small, S., A. Blair, and M. Levine. 1992. "Regulation of Even-Skipped Stripe 2 in the Drosophila Embryo." *The EMBO Journal* 11 (11): 4047–57. https://www.ncbi.nlm.nih.gov/pubmed/1327756.

Venken, Koen J. T., Joseph W. Carlson, Karen L. Schulze, Hongling Pan, Yuchun He, Rebecca Spokony, Kenneth H. Wan, et al. 2009. "Versatile P[acman] BAC Libraries for Transgenesis Studies in Drosophila Melanogaster." *Nature Methods* 6 (6): 431–34. https://doi.org/10.1038/nmeth.1331.

Venken, Koen J. T., Yuchun He, Roger A. Hoskins, and Hugo J. Bellen. 2006. "P[acman]: A BAC Transgenic Platform for Targeted Insertion of Large DNA Fragments in D. Melanogaster." *Science* 314 (5806): 1747–51. https://doi.org/10.1126/science.1134426.

Warming, Søren, Nina Costantino, Donald L. Court, Nancy A. Jenkins, and Neal G. Copeland. 2005. "Simple and Highly Efficient BAC Recombineering Using galK Selection." *Nucleic Acids Research* 33 (4): e36. https://doi.org/10.1093/nar/gni035.

Chapter 2
Optimization and Analysis of MS2 Imaging Data

Optimization of MS2 confocal imaging

The MS2 system enables the analysis and visualization of transcription in real time and single locus resolution (Bertrand et al. 1998). Coupling an MS2 reporter to enhancers and other regulatory elements in embryos of *Drosophila melanogaster* adds a temporal component, often missed in previous analysis, to the studies of enhancer function and gene expression pattern formation.

Movies, recorded with high resolution microscopes, are the basic data for experiments addressing real-time molecular dynamics or transcription at single locus resolution, such as those that involve MS2.

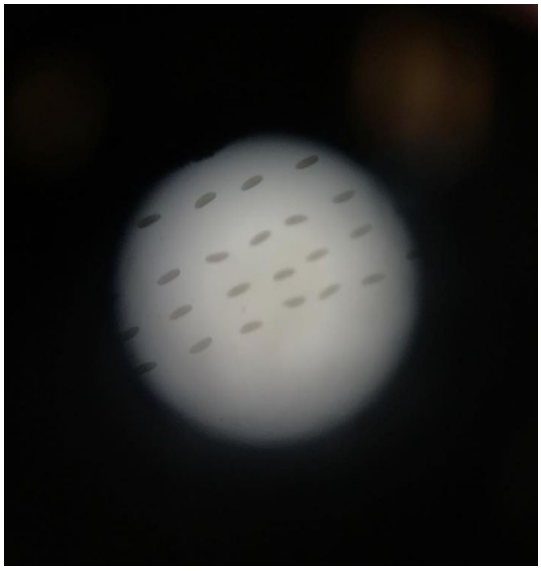
The MS2 reporter system relies on the interaction of MS2 stem loops and the fusion protein MCP::GFP. The latter forms clusters of high local concentration at sites of nascent MS2 transcription. We can observe MCP::GFP clusters under a fluorescence microscope as bright puncta of MS2 transcription. Following the intensity of these MS2 puncta over time will reveal transcriptional bursting.

Due to its commercial availability, flexibility for many kinds of experiments, and the robustness of its measurements, I used laser-scanning confocal microscopy for the real-time imaging experiments in my thesis regarding MS2 transcription.
The first step to record movies of MS2 expression in embryonic development is preparing *D. melanogaster* embryos for live imaging. These embryos result from crossing ~100 homozygous yw; His::RFP; MCP::GFP virgin females with ~30 homozygous w-; +; eveMS2-BAC males, carrying the eveMS2-BAC reporter, as explained in Chapter 1. These experiments require a steady pulse and manual skill, as embryos have to be handled with sensitivity, and kept alive, before putting them under the microscope.

The protocol of mounting fruit fly embryos on a microscope slide for imaging, done as specified by (Garcia et al. 2013; Bothma et al. 2014) consists on decoryonating embryos in chloride and embedding them in an oxygen-permeable Lumox film, that enables embryos to breath, soaked in halocarbon oil that keeps embryos hydrated through the lengthy (~3h long) imaging process. Finally, a glass coverslip protects the setting while the sample is under the microscope **(FIG 2.1).**

FIGURE 2.1

I set ~50 embryos per imaging session, to ensure that at least one of them could be recorded before the onset of nc14 and survive until the early signs of gastrulation. I lined the embryos on the microscope slide to record horizontal fields of view parallel to the antero-posterior axis of the embryo.



All experiments of MS2 expression in my thesis were conducted in a Zeiss 800 laser-scanning confocal microscope, following protocols to record movies of fruit flies developed by (Garcia et al. 2013) and improved by the Garcia Lab.

The most important aspects of our MS2 transcription imaging are capturing the puncta that result from a higher concentration of MCP::GFP at sites of nascent MS2 transcription, and the His:RFP labeled nuclei in the pre-cellularized embryo. Thus, we collected our movies in two channels. EGFP channel, that captures MCP::GFP, with a laser of 488nm wavelength and a laser power of 30μW. And a TagRFP channel that captures His::RFP, with a laser of 561nm wavelength, and laser power of ~7.5μW. We measured the laser powers with the 10X objective in our confocal microscope, before recording each embryo.

We optimized the settings in the confocal microscope to record high temporal resolution movies of the fourteenth nuclear cycle (nc14) of *D. melanogaster* embryogenesis, starting ~60 min before the onset of gastrulation. The final goal of this optimization was capturing as many of the seven *even-skipped* stripes, that ring the pre-cellularized embryo across its anterior-posterior axis, in a single field of view, without significantly sacrificing temporal or spatial resolution.

The microscope scanned 21 confocal Z-stack frames (633ms per frame), to record 3D imaging data across ~10µm, and capture the approximate volume of fruit fly embryo nuclei at the blastocyst stage. It took ~16.8 sec to record the whole set of 21 Z-stacks, before starting again. Hence, our analysis relies on time points spaced by 16.8 sec.

I optimized the imaging settings in the confocal microscope (laser power, microscopy gain, speed of confocal scanning, objectives, size of the confocal pinhole, size of imaging window, etc) to maximize the signal to noise ratio. We gauged the signal to noise ratio by directly looking at the frames of our movies. We optimized the 488nm-wavelength laser and set it at ~30µW power, to lower the levels of photobleaching and ensure the viability of embryos throughout the whole process. (See PreparingEmbryos and ZeissImaging protocol in https://github.com/meisenlab/BerrocalThesis)

Perhaps the most critical finding of this optimization process was choosing the size of field of view for recording movies of eveMS2 expression. Previous confocal movies of MS2 expression in fly development, as those performed by (Bothma et al. 2014), used a 63x magnification objective to cover a field of view in the shape of a rectangle of ~100µm x 50µm. This window size only captures ~2 stripes of eve expression, and I needed to record more embryos than with the optimized field of view to have ~5 independent datasets per stripe. Plus, the positioning of these windows varied along the dorsal-ventral axis of stripes to a greater degree than with the later optimal field of view. Thus, the ~100µm x 50µm movies required more effort, as more embryos were needed, and yielded noisy data hard to interpret.

I found that we could record movies in a field of view of 1024 x 256 pixel or 202.8µm x 50.7µm, at a time resolution of ~16.8 sec per full Z-stack, using the 63x magnification objective. These movies resulted in more clean data that enabled a better view of the patterns of gene expression and the transcriptional dynamics underlying stripe formation **(FIG 2.2A)**.

For each embryo that we analyzed, we recorded a movie in a field of view of 202.8µm x 50.7µm, encompassing 3-5 stripes. I recorded 11 embryos carrying the eveBAC-MS2 construct to capture each eve stripe 3-5 times.

The imaging protocol that I used, by (Garcia et al. 2013), also requires snapshots of the whole embryo to identify the position of the field of view along the embryo. The average length of an embryo is ~500µm. These snapshots of the whole embryo were recorded with the 40x magnification objective. (See ZeissImaging protocol in https://github.com/meisenlab/BerrocalThesis). We recorded movies from the surface of embryos **(FIG 2.2B)** and from the middle sagittal view **(FIG 2.2C)**, to better localize the position of our movies across the entire embryo.

FIGURE 2.2A

Z-maximum projection of the 21 frames in a full Z-stack. The field of view measures 1024 x 256 pixel, which equals 202.8µm x 50.7µm. We can observe the four most anterior *eve* stripes (*eve* stripe 1, 2, 3 and 4; anterior left, posterior right). Imaging windows of this size capture 3-5 stripes. The basis of our data are sequential sets of Z-stacks as this one, recorded at a resolution of 16.8 sec per full Z-stack.

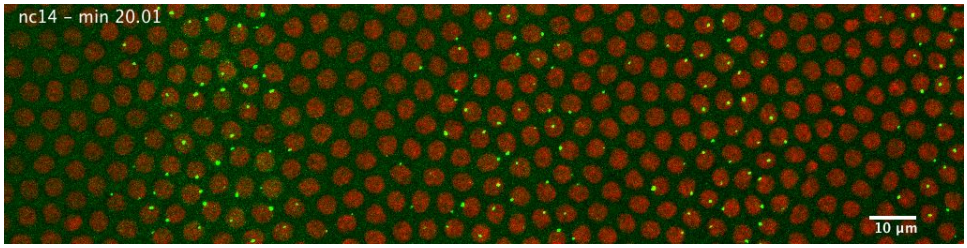


FIGURE 2.2B

Snapshot of whole embryo surface, ~5 min before the start of gastrulation. The nuclei in the full Z-stack shown in figure 2.2A match with their corresponding nuclei in the surface snapshot. Photobleaching of His::RFP causes the darker square on the left (anterior left, posterior right), and helps with the identification of the imaging area.

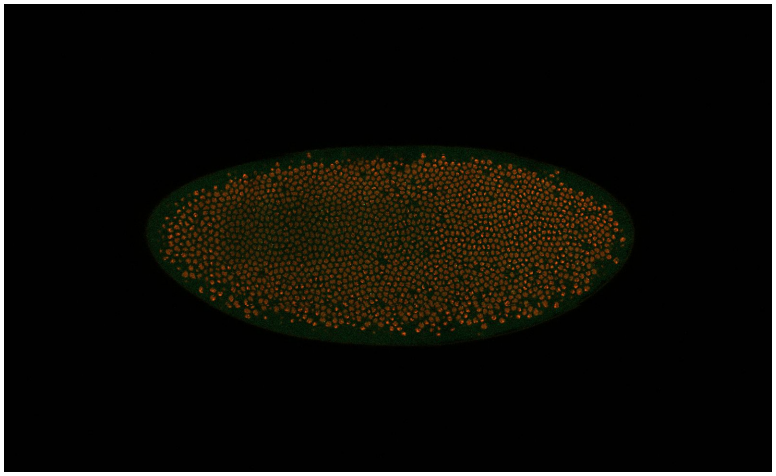
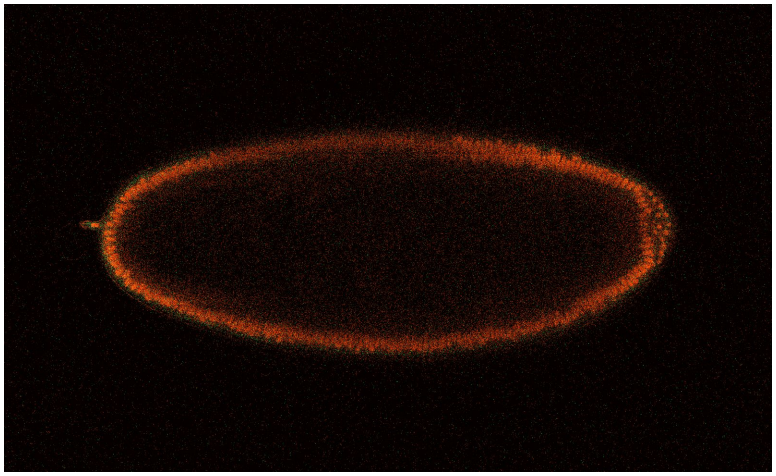


FIGURE 2.2C

Sagittal snapshot of the whole embryo ~5 min before the start of gastrulation. We take a sagittal view of the embryo to locate the position of the field of view in figure 2.2A, relative to the whole embryo length. As the surface snapshot might vary in size due to the pressure of the cover slip on the embryo. (anterior left, posterior right).



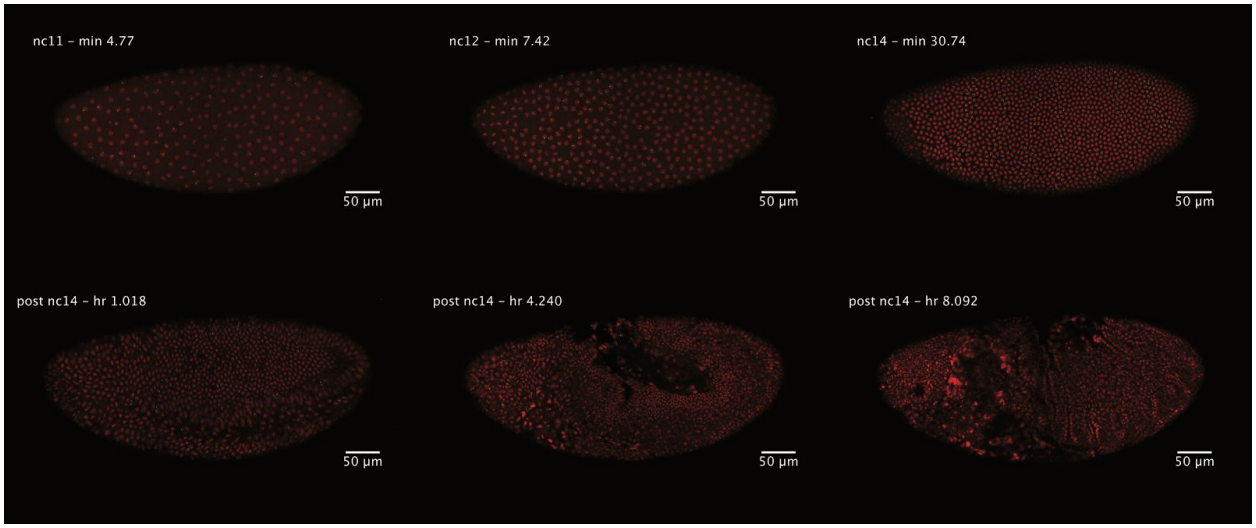
Collecting imaging data of eveMS2 transcription is time consuming and requires preparing fresh embryos every time that a new movie is about to be recorded. Oftentimes, embryos would die under the microscope. Therefore, many more embryos had to be recorded in order to achieve between 3 and 5 independent embryos for each

stripe. For the paper (Berrocal et al. 2018) we recorded 11 embryos that successfully survived throughout gastrulation.

We also recorded movies of development of the whole embryo to have a better idea of *eve* expression patterns across the embryo whole **(FIG 2.3)**. These movies were visually stunning, and show a highly dynamic expression of eve MS2, all of this while cells divide and migrate.

FIGURE 2.3

I recorded a movie of a whole embryo expressing the eveBAC MS2 construct, to have a better picture of *even-skipped* expression in development. We could observe *eve* expression since nc10, when the movie starts. We focused our analysis of *eve* expression in nc14, ~50 min before gastrulation, as this is when *eve* forms its characteristic seven stripe pattern, and the transcriptional programs set the identity of cells for gastrulation. Some *eve* nuclei active in stripes remain active after the cell migration that comes with gastrulation. Later in development, *eve* is expressed in the areas of the embryo that seem to form the gut and the brain. Sporadic *eve* expression lasts for 8 hours after gastrulation, at least.



In the incoming chapter, I will describe the analysis of the MS2 movies in detail.

Analysis of MS2 movies: converting images into numbers

The outcome of our eveMS2 imaging experiments are >50min movies recorded at a fixed rectangular field of view on the surface of fruit fly embryos at the pre-gastrulation stage.

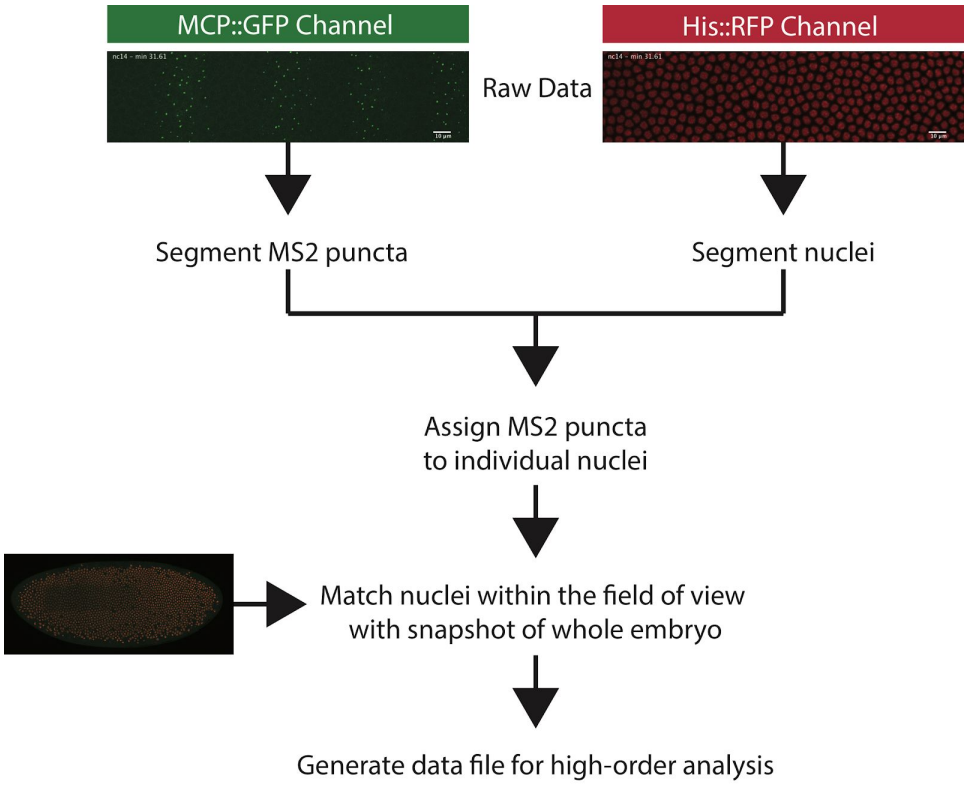
I used a MATLAB-based pipeline, developed and improved by the Garcia Lab on the basis of (Garcia et al. 2013), to extract numerical data from the movies of MS2 expression in eveMS2-BAC embryos.

As mentioned above, our movies of MS2 expression are recorded in two channels. The channel 1, EFGP, records the signal emitted by MCP::GFP fluorescence at sites of nascent MS2 transcription. The channel 2, TarRFP, records the signal emitted by His::RFP-labeled chromatin in the nuclei of fruit fly embryos.

The MATLAB-based pipeline that we used to extract numerical data is made of a series of scripts that analyze the two movie channels, separately, and then combine their spatial and temporal information to do a higher order analysis **(FIG 2.4)**.

FIGURE 2.4

We record movies of eveMS2 transcription in development in two channels. Ch1: MCP::GFP; ch2: His::RFP. We segment MS2 puncta in channel 1 and nuclei in channel 2, individually. Next, we assigned MS2 puncta to individual nuclei. We match the nuclei in the field of view to the whole embryo surface snapshot to locate the recorded region along the embryo. Finally, we generate a file that includes the dynamic position, fluorescence intensity, and corresponding nucleus of every MS2 punctum.



The analysis works as follows. The EGFP and TagRFP channels are analyzed individually. First, each frame in the EGFP channel is segmented to localize and track the position of MS2 fluorescent puncta at each time point. We used the Fiji-Weka Segmentation 3D software to segment MS2 puncta. This machine-learning based approach relies on the manual segmentation of a variety of MS2 puncta, of different intensity, size and position in the embryo, and across the Z-stack, to make a 'dictionary', or classifier, of MS2 puncta. This classifier serves as a model to segment, and track MS2 puncta of various sizes and shapes. Segmenting MS2 puncta in time and space enables us to assign sites of MS2 transcription to individual nuclei and extract their fluorescence. Thus, the information of transcriptional activity at single locus resolution, over time.

Then, we segment the TagRFP channel to localize and track His::RFP-labeled nuclei in fruit fly embryos before gastrulation. The eveMS2 BAC construct inserted in chromosome 3L is heterozygous in our experiments. Therefore, each nucleus has a single eveMS2 fluorescent punctum. It is important to know the position of nuclei to match them with their corresponding MS2 punctum.

Upon independent segmentation of EGFP and TagRFP channels, we assigned each MS2 punctum to a single nucleus, based on their overlapping positions.

After matching MS2 puncta with individual nuclei, we pair the nuclei in our field of view to an area along the embryo, based on the position of nuclei on the surface snapshot taken on the whole embryo at 50 min into nc14. We use the whole embryo sagittal snapshot to assign a number to the position of nuclei and MS2 punctum relative to the total length of the embryo. Thus, we could say that a given MS2 punctum is located at 0.4 AP along the embryo, where 0 is the anterior tip and 1 is the posterior end. Stripes are located at somewhat defined AP positions along the embryo. *E.g. eve* stripe 2 always appears at ~0.4 AP position (Bothma et al. 2014; Berrocal et al. 2018).

Finally, we generate a file that contains the dynamic position and intensity of MS2 puncta, and their corresponding nucleus, over time.

Extracting meaningful data from MS2 fluorescence

When we know the position and fluorescence of MS2 puncta over time, we can extract meaningful, quantitative information about the various enhancers of *even-skipped* that drive MS2 transcription.

A quick observation on the fluorescence of single puncta will reveal fluctuations in their intensity over time, *i.e.* transcriptional bursting. Pulses, or bursts, of transcription, are pervasive in life, and result from the stochastic activation and inactivation of gene promoters.

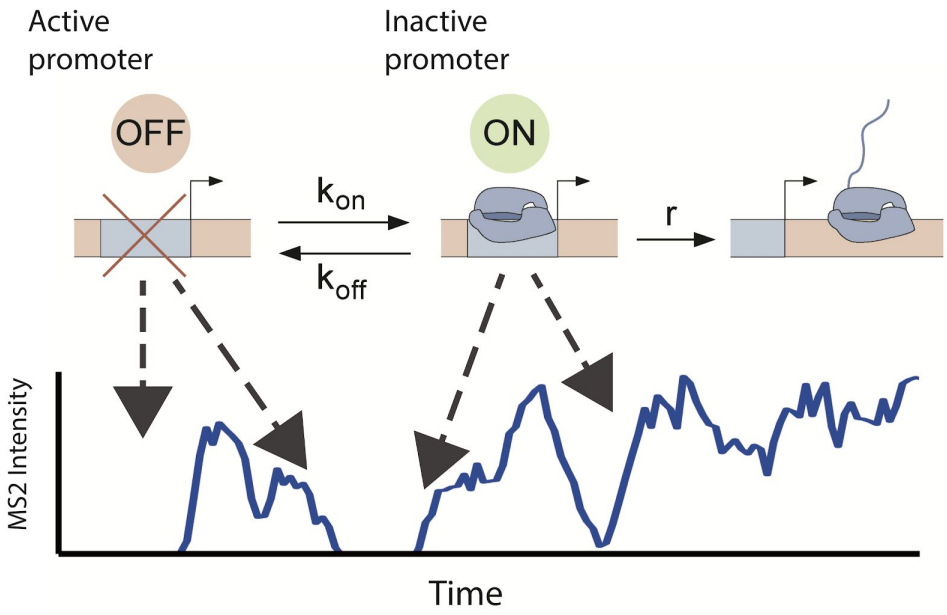
We can extract some information from the crude MS2 fluorescence fluctuations, as the position and movement of stripes of eveMS2 transcription, and the average levels of fluorescence intensity over time.

More importantly, we can infer the state of promoter activity in every nucleus with active transcription of the eveMS2-BAC construct. The two-state model of promoter activity states that gene promoters switch stochastically from an inactive state (OFF) to an active state (ON) at a rate kon, and from ON to OFF at a rate koff. When promoters are in the ON state, they load transcribing mRNA PolII molecules into the gene at a rate r. The stochastic transition of a promoter between ON and OFF states results in transcriptional bursting.

We can infer the state of a promoter from the fluctuations in fluorescence intensity of MS2 in a single active locus. When a promoter is ON and loads polymerases, MS2 intensity increases at a rate r. When the promoter is OFF, MS2 intensity decreases until no fluorescence signal is detected **(FIG 2.5)**.

FIGURE 2.5

Gene promoters switch stochastically between OFF state and ON state. When the promoter is ON, mRNA synthesis begins. Transcription occurs in bursts as promoters switch stochastically between ON and OFF states. These bursts can be observed as fluctuations in MS2 fluorescence. We can infer the state of a promoter from the fluctuations in MS2 fluorescence. When the promoter is ON, MS2 fluorescence increases. When the promoter is OFF, MS2 fluorescence decreases or there is no observable signal. Nevertheless, the MS2 signal that results from promoter activity is convoluted in time. A Hidden Markov Model developed by (Lammers et al. 2020) solves the problem of time convolution to infer the sequence states of a gene promoter over time.



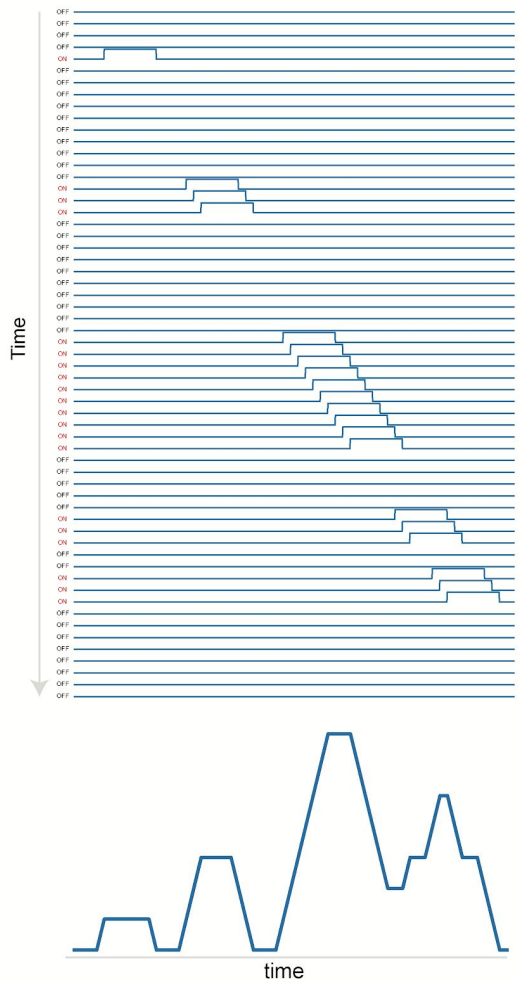
Inferring the sequence of promoter states, (*i.e.* 0 0 0 1 1 1 0 0 0 0 1 1 1, where 0 is OFF and 1 is ON) from transcriptional bursts would be a trivial problem if single bursts were clearly separated in time **(FIG 2.6)**. However, transcriptional bursts are convoluted in time, and it is often the case that the intensity of a single burst decreases momentarily, meaning that the promoter switches to OFF momentarily, and switches ON again, before the signal from the first burst disappears.

FIGURE 2.6

We can imagine promoter activity over time as a vector of states (OFF OFF OFF ON OFF OFF…) that produces a characteristic MS2 fluorescence signature. Hence, we can infer the vector of promoter states over time from its observable outcome: fluctuations in MS2 fluorescence.

A gene promoter loads mRNA Pol II molecules at every time point *t* that is on the ON state. By autocorrelation of the signal in adjacent time points of an MS2 fluorescent trace, we estimated that the fluorescent signal of a single polymerase bound to the MS2 gene lasts for 140 sec. The rationale for this analysis is that the signal of adjacent time points would have the lowest correlation when an active mRNA Pol II falls off the MS2 transcriptional unit (6563bp). This velocity is consistent with the rate of polymerase elongation measured by (Fukaya, Lim, and Levine 2016) of 2700bp/min.

The fluorescent signal of mRNA Pol II loaded at time point *t1* results in a fluorescent blip that lasts 140sec (Δt). If the promoter remains active 20sec later, at *t2* (*t1* + 20), a second fluorescent blip would last 140 sec, or t1 + 20 + Δt sec. From *t2* until *t1* + Δt, the fluorescent signal of both active blips will compound. Adding more and more discrete blips, each one of them not lasting more than 140sec, would result in the characteristic peaks in the fluorescent signal that we know as transcriptional bursts.



(Lammers et al. 2020) developed a Hidden Markov Model to infer the sequence of states of a promoter (*e.g.* 0 0 1 1 0 1 1 0 0 1 0 1) from the convoluted MS2 fluorescent signal. (Lammers et al. 2020) Hidden Markov Model enables us to infer the sequence of states of the *eve* promoter, over time, that better predicts its observable outcome: the fluorescence trace that results from MS2 expression. In a nutshell, a Hidden Markov Model (HMM) deduces the unobservable states of a system, or 'hidden' states. In this case, the model developed by Lammers *et al*. 2020, presumes two hidden states of the promoter. ON, when the promoter loads active polymerases into the gene; and OFF, when the promoter is inactive and does not load polymerases into the gene. These states are 'hidden' because we cannot directly observe the state of the actual promoter, we can only observe its outcome, the transcription of MS2. Furthermore, to explain the outcome, Hidden Markov Models must take into account the probability of transitioning from one state to another. The Hidden Markov Model by Lammers *et al*. 2020, estimates the probability of transitioning from the OFF state to the ON state (Kon) and the probability of transitioning from the ON state to the OFF state (koff). In the end, a sequence of promoter states that transition between ON and OFF, with Kon and koff probabilities, and load active mRNA polymerases to the promoter at a rate R when the promoter is ON, explains the observed MS2 traces.

The Hidden Markov Model enables us to take a glimpse into a biological process beyond our observation capabilities: the stochastic activation and inactivation of a gene promoter over time. Moreover, converting crude MS2 fluorescence intensity into a vector of sequential states of promoter activity enables a more quantitative and rigorous analysis of transcriptional bursting (Berrocal et al. 2018; Lammers et al. 2020).

The next chapter describes our findings regarding the dynamics of transcription and promoter activity from the quantitative analysis of MS2 under the control of *eve* regulatory sequences in *D. melanogaster* embryos.

Bibliography

Berrocal, A., N. C. Lammers, H. G. Garcia, and M. B. Eisen. 2018. “Kinetic Sculpting of the Seven Stripes of the Drosophila Even-Skipped Gene.” *BioRxiv*. https://www.biorxiv.org/content/10.1101/335901v2.abstract.

Bertrand, E., P. Chartrand, M. Schaefer, S. M. Shenoy, R. H. Singer, and R. M. Long. 1998. “Localization of ASH1 mRNA Particles in Living Yeast.” *Molecular Cell* 2 (4): 437–45. https://doi.org/10.1016/s1097-2765(00)80143-4.

Bothma, Jacques P., Hernan G. Garcia, Emilia Esposito, Gavin Schlissel, Thomas Gregor, and Michael Levine. 2014. “Dynamic Regulation of Eve Stripe 2 Expression Reveals Transcriptional Bursts in Living Drosophila Embryos.” *Proceedings of the National Academy of Sciences of the United States of America* 111 (29): 10598–603. https://doi.org/10.1073/pnas.1410022111.

Fukaya, Takashi, Bomyi Lim, and Michael Levine. 2016. “Enhancer Control of Transcriptional Bursting.” *Cell* 166 (2): 358–68. https://doi.org/10.1016/j.cell.2016.05.025.

Garcia, Hernan G., Mikhail Tikhonov, Albert Lin, and Thomas Gregor. 2013. “Quantitative Imaging of Transcription in Living Drosophila Embryos Links Polymerase Activity to Patterning.” *Current Biology: CB* 23 (21): 2140–45. https://doi.org/10.1016/j.cub.2013.08.054.

Lammers, Nicholas C., Vahe Galstyan, Armando Reimer, Sean A. Medin, Chris H. Wiggins, and Hernan G. Garcia. 2020. “Multimodal Transcriptional Control of Pattern Formation in Embryonic Development.” *Proceedings of the National Academy of Sciences of the United States of America* 117 (2): 836–47. https://doi.org/10.1073/pnas.1912500117.

Chapter 3
eve Patterns Emerge from Stochastic Transcription in Single Nuclei

Dynamic expression of eveMS2 patterns

The patterns of gene expression that shape embryonic development are dynamic at many levels, as they result from the interplay of different processes: transcription, mRNA decay and degradation, diffusion, active transport, and the migration of cells and tissues carrying the mRNAs that drive their transcriptional programs.
Here, we analyze the dynamic behavior of *even-skipped (eve)* transcription, the first step in the formation of the well-known seven stripes of *eve* expression.

Our analysis focused on the 14th cell cycle after fertilization (nc14) of *Drosophila melanogaster*, the last synchronous cell cycle in fruit fly embryos before the onset of gastrulation. Gastrulation occurs at ~60 min after the start of nc14. The transcriptional programs that drive cell migration and cell specification in gastrulation occur in nc14 (Davidson 2010). We characterized the kinetics of transcription of the seven *eve* stripes that ring the pre-cellularized embryo in nc14. The expression of the seven *eve* stripes is a fundamental step in the transcriptional programs that lead to the formation of antero-posterior segments in fruit flies. *even-skipped* mutant larvae lack the 2nd, 4th, 6th and 8th abdominal segments. (Nüsslein-Volhard and Wieschaus 1980)

In this chapter, I will describe the quantification and analysis of *eve* patterns, as shown by the MS2 reporter. First, we observed the raw fluorescent MS2 signal. Then, we inferred the kinetics of promoter activity that better described the MS2 signal at single locus resolution. Our results show dynamic stripe formation, driven by an increase in promoter kon (*i.e.* higher frequency of bursting) while koff (*i.e.* duration of bursts) remains constant. Stripes boundaries are defined by quiescent nuclei in the inter-stripe regions. All this, while stripes flow towards the anterior end of the embryo, propelled by a coordinated increase in kon at the anterior edge of stripes coupled with a decrease in kon at their posterior edge.

Quantification of raw fluorescence shows dynamic formation of *eve* stripes

We imaged 11 embryos carrying the eveMS2-BAC to have every stripe recorded 3-5 times, this is equivalent to ~3000 active nuclei. Previous works used a similar number of replicas (~5 embryos) per expression pattern (Bothma et al. 2014, 2015; Garcia et al. 2013). We converted movies into numerical values for quantification.

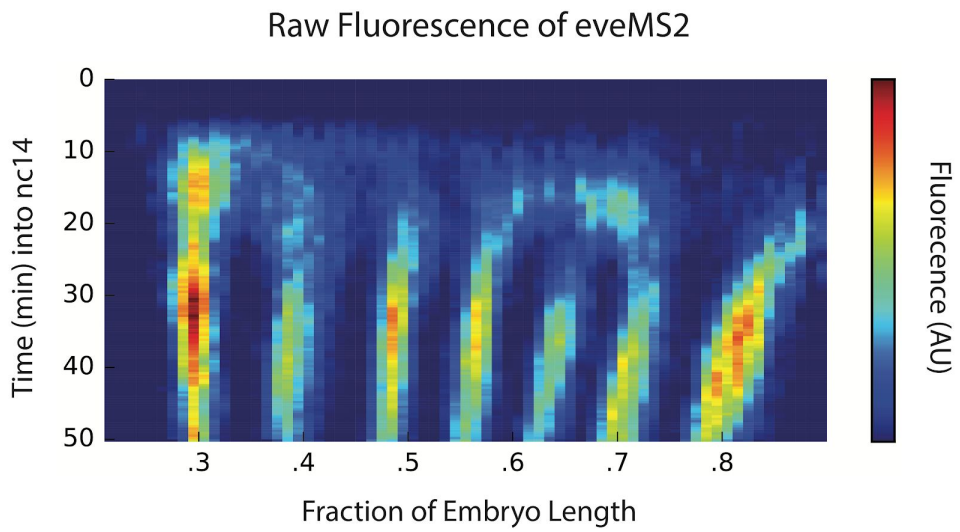
First, we normalized the antero-posterior position of MS2 puncta to be able to compare each of the seven stripes over many embryos. Then, we measured the average fluorescence of MS2 puncta across the 11 embryos in space and time **(FIG 3.1)**. The kymograph of eveMS2 fluorescence already gives a glimpse into the dynamic formation of *eve* expression patterns.

eve transcription starts 5 min after anaphase in a long region that goes from 0.25 to 0.75 AP egg-length. Empirically, I have observed that MS2 transcription always starts ~5 min after anaphase. This seems to be the time that it takes for chromatin to decondense after cell division in fruit fly early development.

By 10 min into nc14, we can observe two large swaths of MS2 expression in the embryo. An anterior swath, that goes from 0.25 to 0.5 AP egg-length, and a posterior swath that goes from 0.55 to 7.5 AP egg-length. The anterior swath of eveMS2 expression separates in three distinct stripes, *eve* stripe 1 (eveS1), eveS2, and eveS3, 20 min after the start of nc14. The posterior swath of eveMS2 expression refines into two distinct stripes, eveS4 and eveS6, with an space without transcription between them, 20 min after the start of nc14. By 30 min after the start of nc14, eveS5 appears in the area of low transcriptional activity between eveS4 and eveS6. eveS7 appears *de novo* at 0.85 AP egg-length, by 20 min into nc14, and shifts towards the anterior direction of the embryo. All the seven *eve* stripes are formed and have separate identities by 30 min into nc14. Stripes gradually increase in intensity, from the first time that they are distinguishable, at 20min into nc14, and reach a peak by 35 min into nc14.

FIGURE 3.1

We averaged the MS2 fluorescence values from 11 embryos to observe the average behavior of raw MS2 fluorescence along the antero-posterior axis (x-axis) over time (y-axis). The raw fluorescence from *eve* MS2 movies shows the seven stripes of *eve* expression. The positions of the stripes are consistent with the position of stripes in individual embryos. The MS2 fluorescence overtime shows the dynamic nature of eve expression.



The kymograph of average fluorescence from 11 embryos shows a dynamic picture of stripes, that move and change their patterns of expression and their MS2 fluorescence intensity over time. However, inferring the state of activity of promoters from the fluorescence of individual traces would let us glimpse into a molecular process beyond our capabilities of microscopy, the promoter behavior at single locus resolution. Moreover, it would enable us to quantify our MS2 fluorescence data in a more rigorous way.

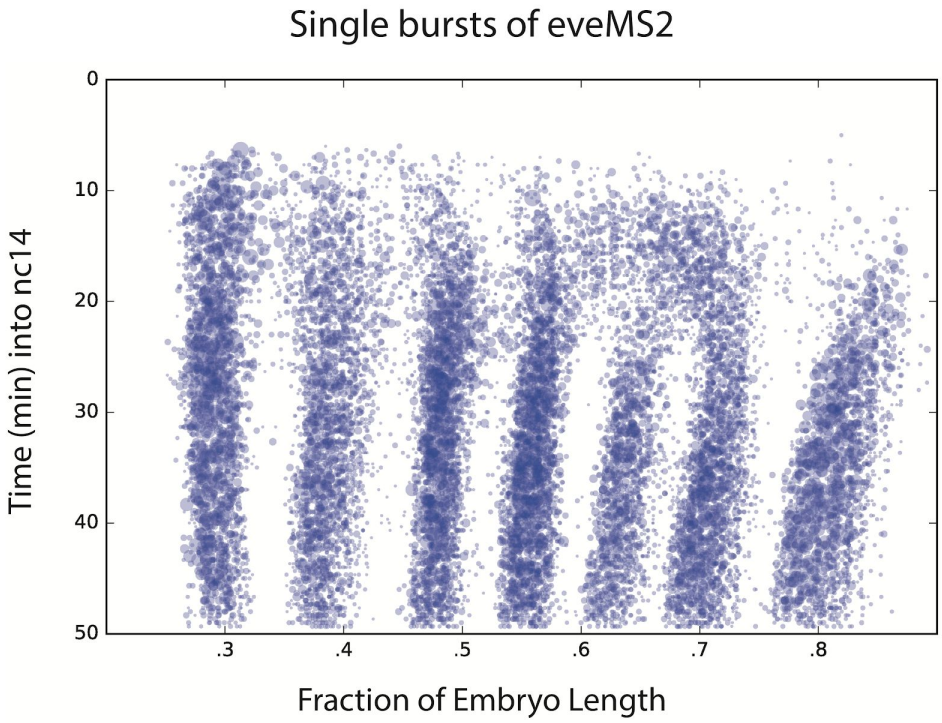
We can quantify the features of individual bursts by inferring the sequence of promoter states

We used the Hidden Markov Model developed by (Lammers et al. 2020) to infer the sequence of promoter states (OFF OFF ON ON OFF OFF) that better explains the fluctuations in fluorescence of MS2 traces.

The sequence of promoter states for single MS2 traces shows an individual burst as a sequence of contiguous ONs interspersed by strings of OFFs. The spatio-temporal appearance of individual transcriptional bursts largely mimics the raw fluorescence from Figure 1. Individual bursts are more frequent at the center of stripes, while they are rarer at the edges of stripes **(FIG 3.2)**.

FIGURE 3.2

The position and timing of eveMS2 bursts largely mimic the average MS2 fluorescence values from Figure 1. 10 min into nc14, bursts occur interspersed in a wide area. As nc14 progresses, bursts become more frequent and concentrate at the center of stripes.



Interstripe regions remain in a quiescent state

A striking observation came upon closer examination of individual bursts over space and time. Stripe boundaries are sharp, separated by nuclei that lay in the inter-stripe regions. These nuclei never show detectable transcriptional bursting, despite being within the large swaths of eveMS2 expression with appreciable *eve* expression that occur between 0.25 and 0.75 AP egg-length in early nc14 **(FIG 3.3A)**.

Never-ON nuclei occur mostly in the inter stripe boundaries, save for the inter-stripe region between eveS5 and eveS6. The activity of never-ON nuclei is already different from nuclei whose fate is to be at the center of stripes at 15 min into nc14, before *eve* expression is refined into discernible stripes **(FIG 3.3B)**.

We hypothesize that nuclei in the inter-stripe boundaries enter into a quiescent state, where their kon is very low or 0, and their levels of eveMS2 transcription are undetectable. This 'quiescent state' might result from chromatin changes that inactivate the *eve* locus in the inter-stripe nuclei during nc14.

FIGURE 3.3A

We labeled in red the trajectories of never-ON nuclei over time. Individual bursts are labeled as black circles. Nuclei that lay in the inter-stripe boundaries are never active, despite the early appearance of large swaths of MS2 expression in a region that goes from 0.25 to 0.75 egg-length. Except nuclei in the inter-stripe boundary between eveS5 and eveS6.

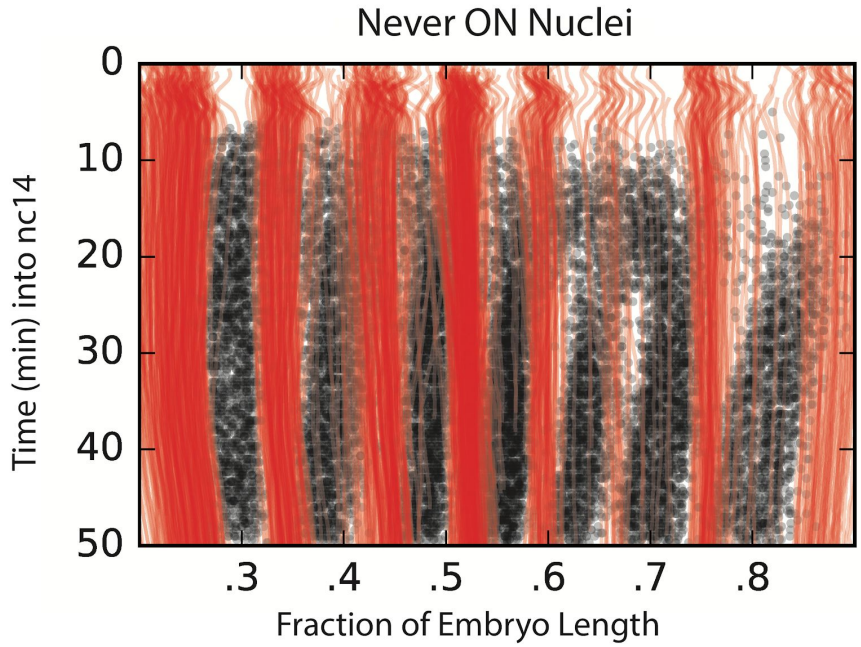
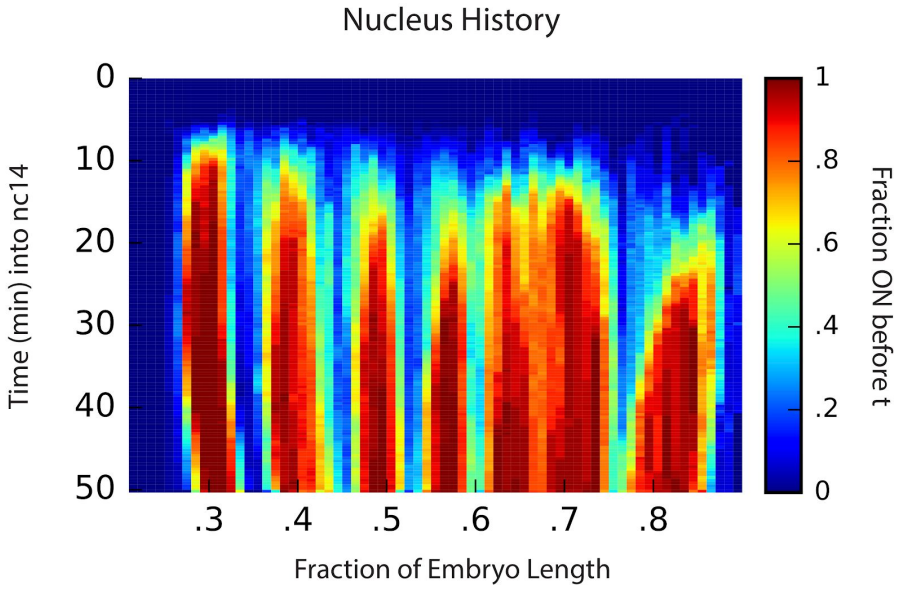


FIGURE 3.3B

We labeled the fraction of nuclei that have not turned ON along the embryo length (x-axis) before any given time into nc14 (y-axis). We can observe that nuclei in the inter-stripe boundaries are never-ON during all nc14. Moreover, 10 min into nc14, nuclei in the inter-stripe boundaries have not shown any detectable activity, while more than half nuclei in the center of stripes have had some activity. Despite the large swaths of eveMS2 expression that occur in a region that goes from 0.25 to 0.75 egg-length at ~10 min into nc14.



Stripes shift anteriorly due to expression flow

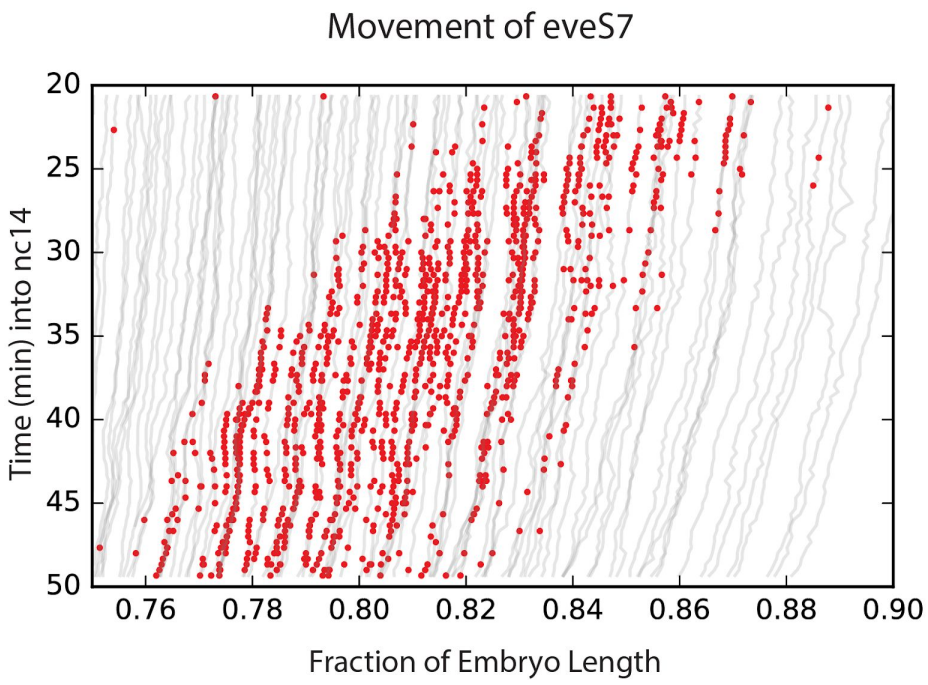
Previous studies had noticed that stripes shift towards the anterior tip of the embryo. Consistent with the shift in gap gene expression from the posterior towards the anterior side of the embryo (Keränen et al. 2006). However, it was not known whether this movement was due to nuclear flow, *i.e.* drift of nuclei towards the anterior side of the embryo, or expression flow, *i.e.* a treadmilling-like effect, where nuclei at the posterior edge of stripes cease transcription, while nuclei adjacent to the anterior edge of stripes initiate transcription. (Jaeger, Blagov, et al. 2004; Jaeger, Surkova, et al. 2004)

By tracking single bursts of transcription and the position of nuclei over time, we observed that nuclei remain mostly static, while a wave of transcription passes over them. Thus, we found that stripes shift anteriorly mostly due to expression flow **(FIG 3.4)**.

Expression flow is more evident for eveS7, as it moves anteriorly ~8% of embryo-length. All stripes shift anteriorly, although this movement is less evident for more anterior stripes.

FIGURE 3.4

We followed the position of individual nuclei, between 0.75 and 0.9 egg-length, over time (gray lines). These nuclei belong to eveS7. We tracked individual bursts in these nuclei. We observed that nuclei in the posterior edge of eveS7 cease bursting, while nuclei in the anterior edge of eveS7 start new bursts. The end result is a treadmilling-like effect that causes eveS7 to shift anteriorly 8% of the embryo length



Stripes result from an upregulation of kon

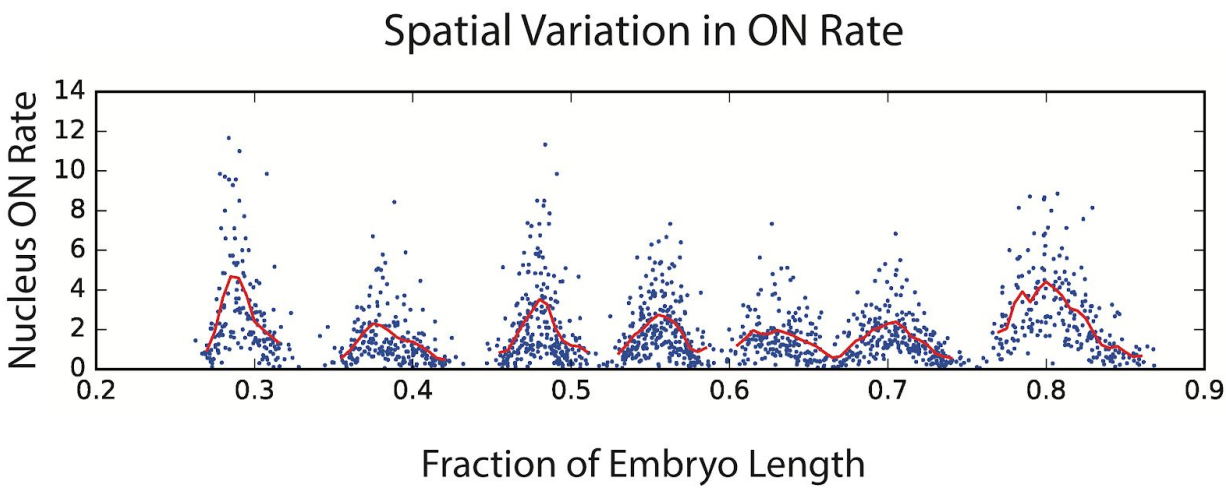
We can calculate the kon, the transition probability of an individual promoter from OFF to ON state, by dividing their total number of transitions from OFF to ON by the total number of time points between the first and last ON time point, between 25 min and 50 min into nc14. In the same manner, we can calculate the koff, the transition probability of an individual promoter from ON to OFF, by dividing their total number of transitions from ON to OFF by the total number of time points between the first and last ON time point, between 25 min and 50 min into nc14. The kon and koff rates describe *eve* bursting features in a single nucleus, and tell us the probability that the stochastic processes that drive transcription occur.

We measured the kon and koff kinetics that shape stripe formation, calculating the kon and koff of nuclei involved in the formation of individual stripes. We found that stripes are formed by an increase of kon in the nuclei at the center of stripes **(FIG 3.5)**. This means that nuclei in the center of *eve* stripes have more frequent bursts than nuclei at the edges of stripes. This is consistent with the work from (Fukaya, Lim, and Levine 2016), who showed that stronger enhancers increase the frequency of transcriptional bursting. The data regarding koff was harder to interpret, but in the end, we did not find that koff underwent spatial regulation. Instead, koff seems to be the same for all bursts across embryos.

FIGURE 3.5

We plotted the position of individual nuclei along antero-posterior axis of embryos (x-length) against the kon of individual nuclei, calculated as the total number of transitions from OFF to ON divided by the total number of timepoints between the first and the last time point between 25 min and 50 min into nc14.

Stipes are formed by nuclei with higher kon at the center of the stripes coupled with a lower kon at the edges of stripes.



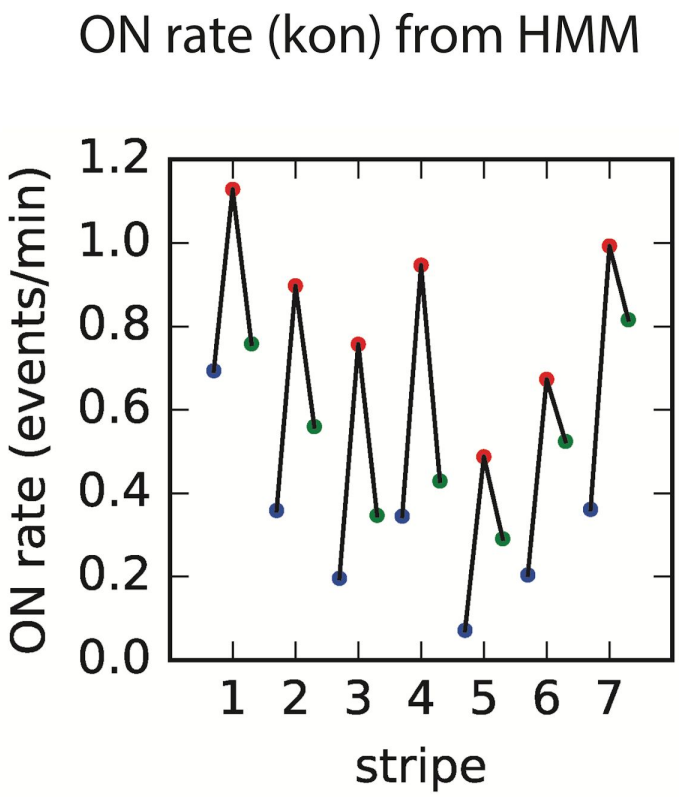
Lammer's *et al* Hidden Markov Model, which is based on the two-state model of promoter activity, yields the transition probability for a promoter to switch from OFF to ON state (kon) and from ON to OFF (koff). Kon and koff are proxies for burst frequency and duration, respectively. Kon and koff, as computed by the Hidden Markov Model, are inferred from the traces of many nuclei in particular body regions of the embryo. Therefore, they are more resilient to the variation of individual outliers. Kon and koff values enable us to understand better the kinetics of promoter activity that generate *eve* expression patterns.

We measured the kon and koff kinetics computed as transitions probabilities by Lammer's *et al* Hidden Markov Model. As we derived kon and koff from multiple nuclei in the center and edges of stripes, we gained robustness, but we lost the fine-grained, single-locus resolution that we observed for the analysis of the sequence of promoter states.

The ensemble kon kinetics calculated by the Hidden Markov Model supported the findings of kon kinetics calculated from the sequence of promoter states of individual nuclei. This is, the region at the center of stripes has a higher kon than the edges of stripes **(FIG 3.6)**.

Figure 3.6

We plotted kon, computed by Lammer's *et al.* Hidden Markov Model as the transition probability of going from OFF to ON for individual stripes and their anterior and posterior flanking regions. This kon measure is calculated from many nuclei in the center and edges of stripes. Thus, it is more robust to outlier variations. These results show that kon is higher in the center of stripes and lower at the edges of stripes.



We hypothesize that the increase in kon might result from the following process. Promoters might transition from OFF to ON as the molecular components necessary for transcription, such as TFs or mRNA Pol II molecules, are brought to promoters, perhaps by close contact with enhancers. Promoter transitions from OFF to ON occur at a higher and higher frequency, due to the readily availability of TFs and other elements of the transcriptional machinery at the center of stripes.

Taken together, these results highlight the dynamic nature of the seven stripes of *even-skipped*, and illustrate how the kinetics of promoter activity at single loci result in the emergence of complex patterns of gene expression. It is worth to notice that all seven *eve* stripes result from similar promoter kinetics. They are formed by an increase in kon coupled at the center of stripes; their stripe boundaries have many nuclei with the *eve* locus in a quiescent state that does not allow for detection of MS2 transcription; while *eve* stripes shift anteriorly due to expression flow. All, despite the myriad of

different inputs that the five enhancers, that control the expression of the seven *eve* stripes, receive. Perhaps, these strategies are not uncommon among enhancers.

It was known that gene expression is a dynamic process. However, now we possess the molecular, imaging, and conceptual tools to capture and quantify the dynamic features of transcription in real-time. This work opens the possibility for study a myriad of mutations and modifications of the *eve* locus: a fine model of enhancer function. These are exciting times in the fields of transcription and embryonic development.

Bibliography

Bothma, Jacques P., Hernan G. Garcia, Emilia Esposito, Gavin Schlissel, Thomas Gregor, and Michael Levine. 2014. “Dynamic Regulation of Eve Stripe 2 Expression Reveals Transcriptional Bursts in Living Drosophila Embryos.” *Proceedings of the National Academy of Sciences of the United States of America* 111 (29): 10598–603. https://doi.org/10.1073/pnas.1410022111.

Bothma, Jacques P., Hernan G. Garcia, Samuel Ng, Michael W. Perry, Thomas Gregor, and Michael Levine. 2015. “Enhancer Additivity and Non-Additivity Are Determined by Enhancer Strength in the Drosophila Embryo.” *eLife* 4 (August). https://doi.org/10.7554/eLife.07956.

Davidson, Eric H. 2010. *The Regulatory Genome: Gene Regulatory Networks In Development And Evolution*. Elsevier. https://play.google.com/store/books/details?id=F2ibJj1LHGEC.

Fukaya, Takashi, Bomyi Lim, and Michael Levine. 2016. “Enhancer Control of Transcriptional Bursting.” *Cell* 166 (2): 358–68. https://doi.org/10.1016/j.cell.2016.05.025.

Garcia, Hernan G., Mikhail Tikhonov, Albert Lin, and Thomas Gregor. 2013. “Quantitative Imaging of Transcription in Living Drosophila Embryos Links Polymerase Activity to Patterning.” *Current Biology: CB* 23 (21): 2140–45. https://doi.org/10.1016/j.cub.2013.08.054.

Jaeger, Johannes, Maxim Blagov, David Kosman, Konstantin N. Kozlov, Manu, Ekaterina Myasnikova, Svetlana Surkova, et al. 2004. “Dynamical Analysis of Regulatory Interactions in the Gap Gene System of Drosophila Melanogaster.” *Genetics* 167 (4): 1721–37. https://doi.org/10.1534/genetics.104.027334.

Jaeger, Johannes, Svetlana Surkova, Maxim Blagov, Hilde Janssens, David Kosman, Konstantin N. Kozlov, Manu, et al. 2004. “Dynamic Control of Positional Information in the Early Drosophila Embryo.” *Nature* 430 (6997): 368–71. https://doi.org/10.1038/nature02678.

Keränen, Soile V. E., Charless C. Fowlkes, Cris L. Luengo Hendriks, Damir Sudar, David W. Knowles, Jitendra Malik, and Mark D. Biggin. 2006. “Three-Dimensional Morphology and Gene Expression in the Drosophilablastoderm at Cellular Resolution II: Dynamics.” *Genome Biology* 7 (12): R124. https://doi.org/10.1186/gb-2006-7-12-r124.

Lammers, Nicholas C., Vahe Galstyan, Armando Reimer, Sean A. Medin, Chris H. Wiggins, and Hernan G. Garcia. 2020. “Multimodal Transcriptional Control of Pattern Formation in Embryonic Development.” *Proceedings of the National Academy of Sciences of the United States of America* 117 (2): 836–47.

https://doi.org/10.1073/pnas.1912500117.
Nüsslein-Volhard, C., and E. Wieschaus. 1980. “Mutations Affecting Segment Number and Polarity in Drosophila.” *Nature* 287 (5785): 795–801. https://doi.org/10.1038/287795a0.

Chapter 4
Shared Bursting Kinetics in Endogenous and Ectopic Patterns of Gene Expression

Introduction

The synthesis of mRNA is discontinuous. 40 years ago, electron micrographs of *Drosophila melanogaster* embryos (McKnight and Miller 1979) observed 'ribonucleoprotein fibers', or nascent mRNA transcripts, branching from transcriptionally active chromatids. The authors noted that 'Active transcription units can have internal, fiber-free gaps which may result from interruptions in initiation of transcription'. Further research confirmed the temporally discontinuous nature of mRNA synthesis, using techniques capable of assessing the static position of nascent mRNA molecules, such as RNA-FISH. Discontinuous mRNA synthesis results in transcription bursts. The transcriptional machinery initiates mRNA synthesis of a gene at stochastic time points, and mRNA synthesis lasts for a period of time until the transcriptional machinery falls off the gene, in the end we observe stochastic pulses, or bursts, of mRNA production. (Garcia et al. 2013; Bothma et al. 2014; Lammers et al. 2020) Nevertheless, the shape and frequency of transcriptional bursts had to be inferred from static data, until the advent of MS2 and related techniques that enable visualization of mRNA synthesis in living cells in real-time. (Ferguson and Larson 2013; Golding et al. 2005; Chubb et al. 2006; Garcia et al. 2013)

Transcriptional bursting is pervasive, and has been observed in many organisms from bacteria (Golding et al. 2005) to eukaryotes, such as yeast (Zenklusen, Larson, and Singer 2008), flies (Garcia et al. 2013; Fukaya, Lim, and Levine 2016; Bothma et al. 2014), humans (Senecal et al. 2014; Dar et al. 2012), and even eukaryotic viruses like HIV. (Skupsky et al. 2010) Transcriptional bursting results from discontinuous activation of gene promoters. According to this model, gene promoters switch stochastically between an inactive (OFF) and an active (ON) state, and when active, promoters load active RNA Pol II molecules at a rate r.

kon, defined as the rate at which promoters switch from OFF to ON, is a proxy for burst frequency. Higher frequency of OFF to ON transitions results in more bursts.
koff, defined as the rate at which promoters switch from ON to OFF, is a proxy for burst duration. Lower frequency of ON to OFF transitions results in active promoters loading polymerases for longer times, hence, it results in longer bursts.
r, defined as the rate at which promoters load RNA Pol II molecules when active, is a proxy for burst amplitude. A higher loading rate of RNA Pol II molecules into an active promoter results in more intense bursts of mRNA synthesis.

Albeit not well known, various studies have linked Kon, koff, and r to biological processes. For example, a lower koff, and thus, higher burst duration, seems to be related to higher transcription factor binding times on the promoter (Senecal et al. 2014). On the other hand, a higher bursting frequency, *i.e.* a higher kon, results from promoter activation by stronger enhancers. (Fukaya, Lim, and Levine 2016)

In principle, promoters and enhancers can regulate the shape and size of the bursts of mRNA synthesis by tweaking a myriad of combinations of kon, koff, and r parameters.
However, our previous work (Berrocal et al. 2018) observed that the five enhancers of the *Drosophila even-skipped (eve)* gene modulate the expression of *eve* stripes in a similar way. They increase promoter kon, *i.e.* they increase the frequency of bursting at the center of stripes.

We aimed to test two models of enhancer function. The first model posits that enhancers are under selection to generate particular bursting kinetics in their endogenous expression domains. While this selection would be absent from ectopic expansions of gene expression caused by derepression of enhancer activity. Thus, in the first model, endogenous expression domains would have distinct bursting kinetics, different from the bursting kinetics in ectopic expression domains.
On the contrary, the second model states that enhancers are under selection to shape the total amount of mRNA that enhancers express in time and space, but not the bursting kinetics that generate the enhancer-driven spatio-temporal patterns of gene expression. Thus, the second model predicts that bursting kinetics are largely the same in endogenous and ectopic gene expression domains. Similar bursting kinetics in ectopic and endogenous expression domains of the same gene suggest that bursting kinetics result from a common process in both areas. Perhaps their shared promoter or common constraints imposed by the transcriptional machinery.

We used our *even-skipped* MS2 BAC (eveMS2 BAC) (Venken et al. 2009; Berrocal et al. 2018) as a model to investigate bursting dynamics in ectopic and endogenous

patterns of gene expression. eveMS2 BAC contains the MS2 reporter under the control of *eve* promoter and all *eve* enhancers and regulatory elements. The *eve* stripe 2 (eveS2) enhancer locus, in particular, is amenable for modification, as previous work (Small, Blair, and Levine 1992) mapped mutations in the eveS2 enhancer that cause an anterior ectopic expansion of eveS2. We disrupted three binding sites for the repressor protein Giant (Gt) in the eveS2 enhancer and we observed a transient expansion of eveS2, observable as ectopic *eve* active nuclei between the first and second *eve* stripes.
Then, we knocked-out the *eve* stripe 1 (eveS1) enhancer, coupled with the disruption of the Gt-binding sites on the eveS2 enhancer, to abrogate eveS1 expression and get a better view of the anterior expansion of eveS2. Notably, coupling both sets of mutations generated unexpected and complex ectopic patterns of *eve* expression. Namely, a wide swath of expression in the anterior part of the embryo encompassing the areas of eveS1 and eveS2 and the appearance of an extra anterior stripe that we called *eve* stripe 0 (eveS0).

The new ectopic patterns of *eve* expression are useful to compare their bursting kinetics against the endogenous stripes, and test our two models of enhancer function.

Ectopic patterns of *eve* expression show lower levels mRNA synthesis. Nevertheless, consistent with our second model of enhancer function, we observed that the bursting kinetics (kon, koff, r) of nuclei of equivalent transcriptional activity are consistently similar between ectopic and endogenous domains of *eve* expression.

More specifically, bursts of equivalent fluorescent intensity result from a similar mixture of kon, koff, and r parameters; regardless of whether they occur in endogenous or ectopic *eve*-active nuclei. This means that bursts of equivalent intensity occur at the same frequency, and have the same duration and amplitude, in endogenous and ectopic *eve*-active nuclei. Roughly speaking, bursts of similar intensity have a similar shape everywhere.

Preliminary data provides evidence that the lower fluorescence observed in ectopic expansions of *eve* expression results from a lower kon, as occurs in the regions of lower fluorescence at the edges of the endogenous *eve* stripes.

Taken together, these results show that *eve* enhancers interact in yet unknown ways to generate the characteristic *even-skipped* striped pattern, and enhancer disruptions have unpredictable effects that may result in unexpected patterns of ectopic transcription.

Furthermore, bursting kinetics do not vary among the well-known enhancer-driven endogenous stripes of *eve* expression, and neither between the endogenous and ectopic *eve* expression domains. The only evident differences among ectopic and endogenous *eve* domains of enhancer activity are lower levels mRNA synthesis at *eve* ectopic regions, perhaps due to the absence of optimal levels of activator and repressor transcription factors (TFs).

This data shows that enhancers are under selection to drive the expression of optimal mRNA levels in space and time. Whereas enhancers are not selected to shape the kinetics of transcriptional bursting.

Results

Enhancer mutations lead to the formation of new and unexpected ectopic patterns of gene expression

eveMS2 BAC contains the MS2-reporter gene under the control of the *eve* promoter and the five *eve* enhancers **(FIG 4.1A)** that generate the seven well-known *eve* stripes in early *D. melanogaster* development **(FIG 4.1B, FIG 4.1C)**. I used recombineering to generate three variations of the eveMS2 BAC reporter that yield ectopic patterns of gene expression (See Methods).

FIGURE 4.1A

Diagram of the eveMS2 BAC reporter (26.5kb). It contains the MS2 reporter, coupled with the gene *yellow* under the control of the entire *eve* regulatory sequence. The BAC reporter contains all five early enhancers of *eve* (3+7, 2, 4+6, 1, and 5), the late element (LE), and other regulatory elements.

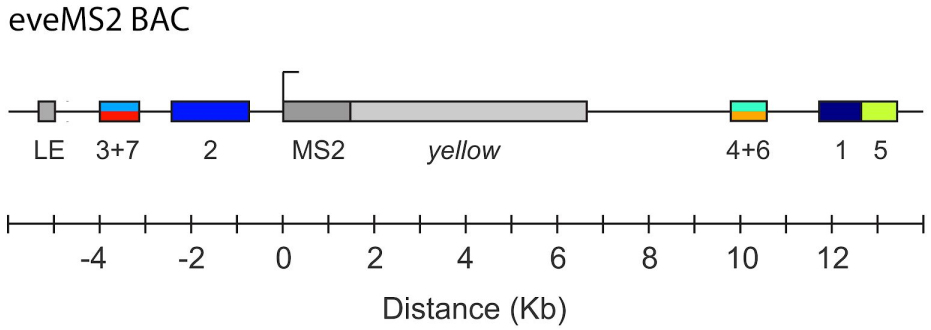


FIGURE 4.1B

This heatmap shows the combined fluorescence of 5-6 embryos carrying the wild-type eveMS2-BAC construct[-]. Average MS2 expression over many embryos shows a more general trend than single embryos. Here, we observe that MS2 expression in wt eveMS2 embryos starts at 5 min into nc14. A large and bright swath of anterior expression refines into eveS1, eveS2, and eveS3. By 20 min into nc14, eveS1 and eveS2 are different stripes, as opposed to the embryos carrying the eveS1wt - eveS2Gt[-] eveMS2BAC mutant.

Embryo Length refers to the position of nuclei relative to the anterior-posterior axis of *D. melanogaster* embryos. 0 is the anterior tip and 1 is the posterior tip of an embryo. Figure modified from (Berrocal et al. 2018)

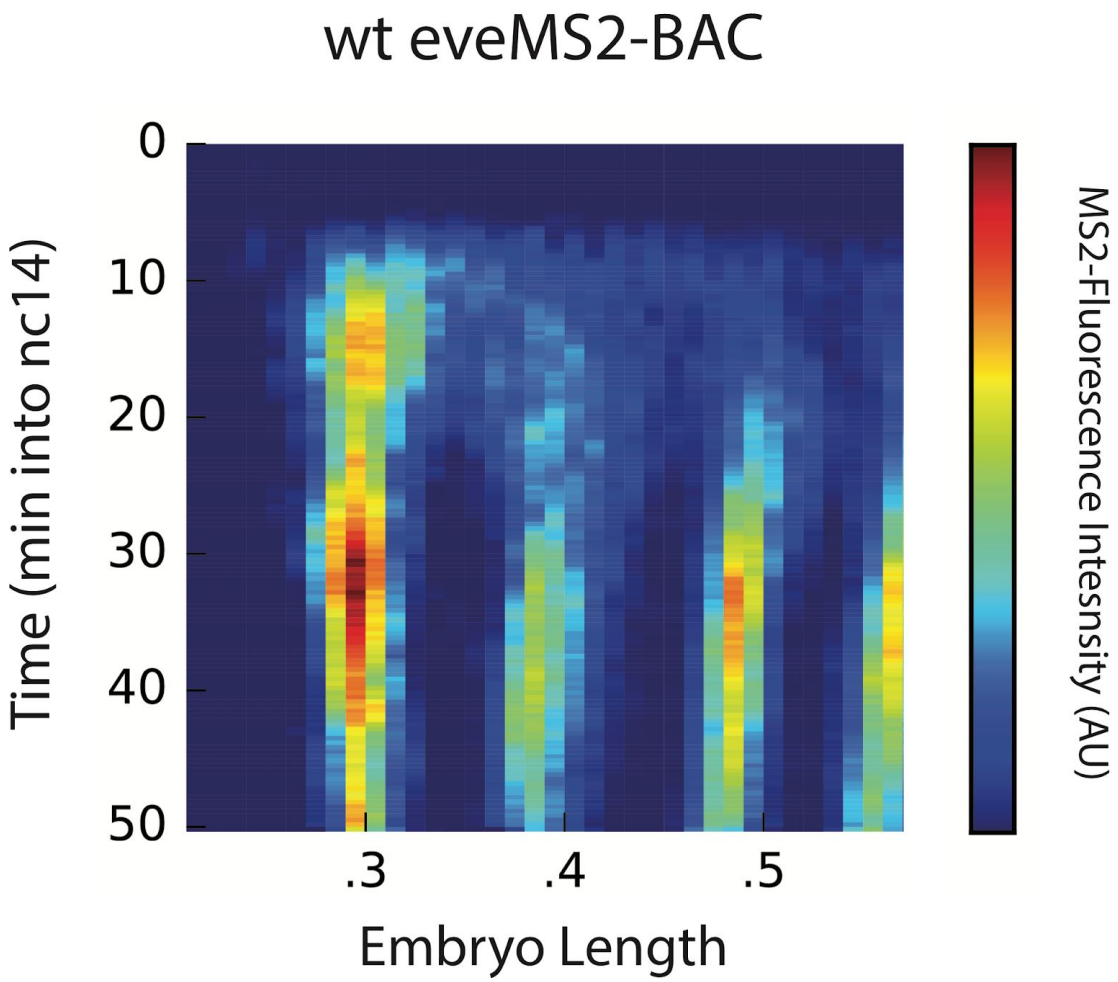
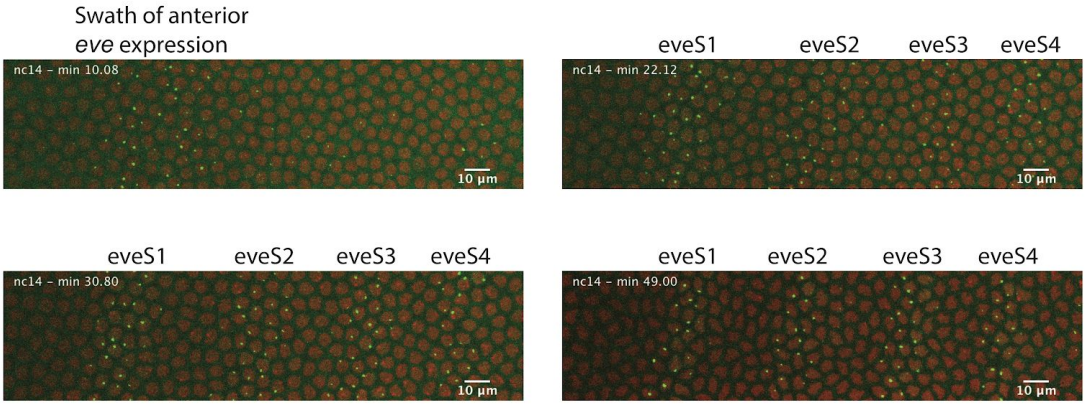


FIGURE 4.1C

MS2 expression snapshots of an embryo carrying the wild-type eveMS2 BAC at 10, 22, 30, and 49 min into nc14. Transcription starts at 5 min (not shown). At 10min, we observe a swath of anterior expression that refines into eveS1 and eveS2 by ~20min into nc14. After 20min into nc14, eveS1, eveS2, eveS3, eveS4, have appeared, and they last until gastrulation.



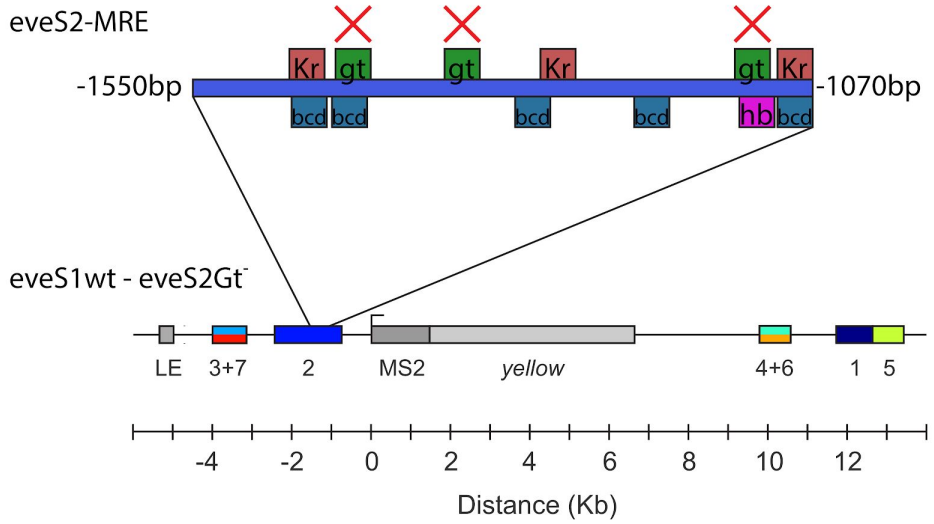
We disrupted three binding sites for the repressor protein Giant (Gt) in the eveS2 enhancer locus (Small, Blair, and Levine 1992) in our eveMS2 BAC reporter. Gt is a transcription factor that represses anterior expression of eveS2. Small, Levine *et al*. disrupted the aforementioned binding sites in the eveS2 Minimal Regulatory Element (eveS2-MRE), a 480bp regulatory sequence within eveS2 enhancer (~2kb total length) sufficient to drive the expression of the *eve* second stripe. These mutations lift the repression of Gt on eveS2 enhancer and enable the anterior expansion of eveS2.

The eveMS2 BAC reporter, with its eveS2 enhancer missing the three Gt binding sites, and an intact *eve* stripe 1 (eveS1) enhancer, will be referred to as eveS1wt - eveS2Gt$^-$ **(FIG 4.2)**.

FIGURE 4.2

Diagram of the eveS1wt - eveS2Gt⁻ construct, a variation of the eveMS2 BAC reporter. We deleted the three Gt-binding sites in the eveS2 Minimal Regulatory Element (eveS2-MRE*). Deleted sites are labeled with red crosses. eveS2-MRE contains binding sites for the TFs bicoid and hunchback, activators; and Giant and Kruppel, repressors. Experiments showed that eveS2-MRE is sufficient to drive the expression of eveS2.

*eveS2-MRE based on (Small, Blair, and Levine 1992).



Since eveS1wt - eveS2Gt⁻ is a 26.5kb eveMS2 BAC reporter that contains all *eve* enhancers and regulatory elements, and there are several interactions among them, we expected to see different ectopic expression patterns of gene expression than those previously observed by Small, Levine, *et al*., in their experiments on the eveS2-MRE; if any at all.

Nevertheless, eveS1wt - eveS2Gt⁻ embryos did show an anterior expansion of eveS2 expression, albeit transient **(FIG 4.3)**.

Since the eveMS2 BAC reporter contains an intact eveS1 enhancer, we observed the anterior expansion of eveS2 as inter-stripe active nuclei between eveS1 and eveS2.
The inter-stripe expression rescinds, and eveS1 and eveS2 eventually refine into two different stripes, perhaps due to the presence of additional Gt binding sites adjacent to the eveS2-MRE region within eveS2 enhancer. Nevertheless, inter-stripe MS2 expression lasts for longer than in embryos harboring the wild type eveMS2 BAC reporter. While wild type eveMS2 embryos repress expression between eveS1 and eveS2 after 20 min into the 14th nuclear cycle of cell division (nc14); eveS1wt -

eveS2Gt$^-$ repressed the inter-stripe expression between eveS1 and eveS2 at ~35 min into nc14 **(FIG 4.3A and FIG 4.3B)**. Eve stripe 3 (eveS3) seems to be normal, but we need further analysis to confirm this observation.

FIGURE 4.3A

MS2 expression snapshots of an embryo carrying the eveS1wt - eveS2Gt$^-$ construct at 10, 20, 28, and 40 min into nc14. Transcription starts at 5 min (not shown). At 10min, we observe a swath of anterior expression that refines into eveS1, eveS2, and eveS3, as in the wt eveMS2 BAC. Nevertheless, we observed MS2 active nuclei between eveS1 and eveS2 at ~20min into nc14. By this time, eveS1 and eveS2 have refined into separate stripes in the wt. Ectopic anterior expansion of eveS2, observable as inter-stripe expression between eveS1 and eveS2 lasts until ~30min into nc14. eveS1 and eveS2 become clear separate stripes after 30min into nc14.

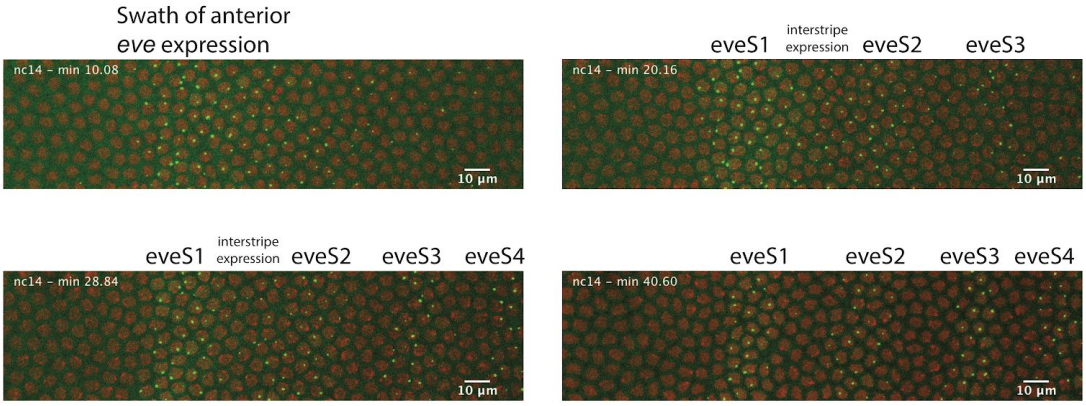
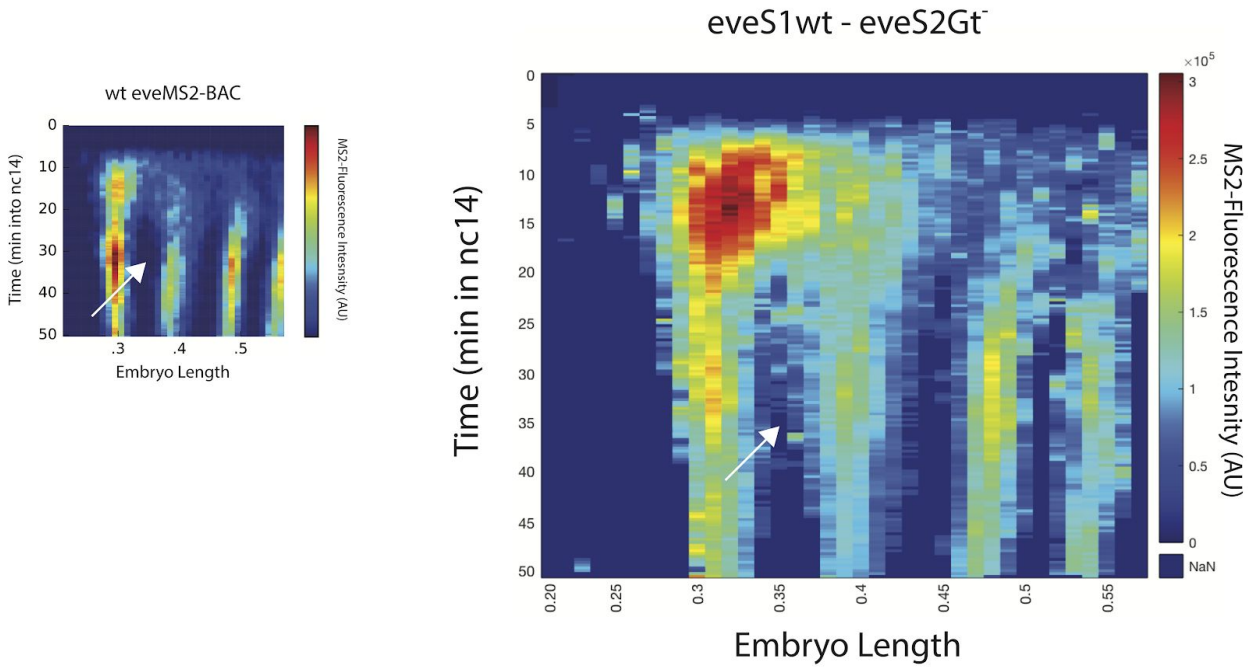


FIGURE 4.3B

This heatmap shows the combined fluorescence of 6 embryos carrying the MS2 construct eveS1wt - eveS2Gt$^-$. Average MS2 expression over many embryos shows a more general trend than single embryos. Here, we observe that MS2 expression in eveS1wt - eveS2Gt$^-$ embryos is similar to the expression in wt eveMS2 embryos. A large and bright swath of anterior expression refines into eveS1, eveS2, and eveS3. eveS1 is brighter, as it is the case for wt eveMS2 embryos. eveS3 and eveS4 appear as defined stripes at 15min into nc14.

However, eveS1wt - eveS2Gt$^-$ embryos show transient inter-stripe MS2 expression between eveS1 and eveS2 until ~30min into nc14 (white arrow). wt eveMS2 embryos, shown in the small heatmap on the left, cease inter-stripe MS2 expression between eveS1 and eveS2 at 20min into nc14.

Embryo Length refers to the position of nuclei relative to the anterior-posterior axis of *D. melanogaster* embryos. 0 is the anterior tip and 1 is the posterior tip of an embryo.



Since the eveMS2 BAC reporter eveS1wt - eveS2Gt$^-$ contains a functional eveS1 enhancer, we thought that eveS1 expression might be masking the anterior ectopic expansion of eveS2. We knocked out eveS1 and disrupted eveS2 Gt-binding sites in the same BAC reporter to abrogate expression of eveS1 and observe an unmasked anterior expansion of eveS2.

We knocked out eveS1 enhancer in our eveMS2 BAC reporter eveS1wt - eveS2Gt$^-$, as defined by the Zelda Chip-seq (Harrison et al. 2011) signal around the genomic region necessary for eveS1 expression (Fujioka et al. 1999; Sackerson, Fujioka, and Goto

1999; Haeussler et al. 2019) **(FIG 4.4A)**. *D. melanogaster* protein Zelda binds promoters and enhancers prior to their activation in the early stages of fruit fly embryogenesis. (Harrison et al. 2011; Liang et al. 2008) Therefore, the presence of Zelda bound to DNA is a predictor for active regulatory sequences. (Harrison et al. 2011)

The eveS2Gt^{-} eveMS2 BAC reporter with its eveS1 enhancer disrupted will be referred to as eveS1Δ - eveS2Gt^{-} **(FIG 4.4B)**.

FIGURE 4.4A

I disrupted the eveS1 regulatory element, as defined by the DNA sequence around eveS1 enhancer that shows evidence for Zelda binding in the Chip-seq experiments conducted by (Harrison et al. 2011)**.** The disrupted sequence shows strong binding for the TFs Kruppel (Kr), Bicoid (BCD) and Giant (Gt), likely regulators of eveS1 expression. I retrieved the DNA sequence for eveS1 enhancer from the UCSC Genome Browser (Haeussler et al. 2019) - *D. melanogaster* Genome Assembly Apr.2006 (BDGP R5/dm3).

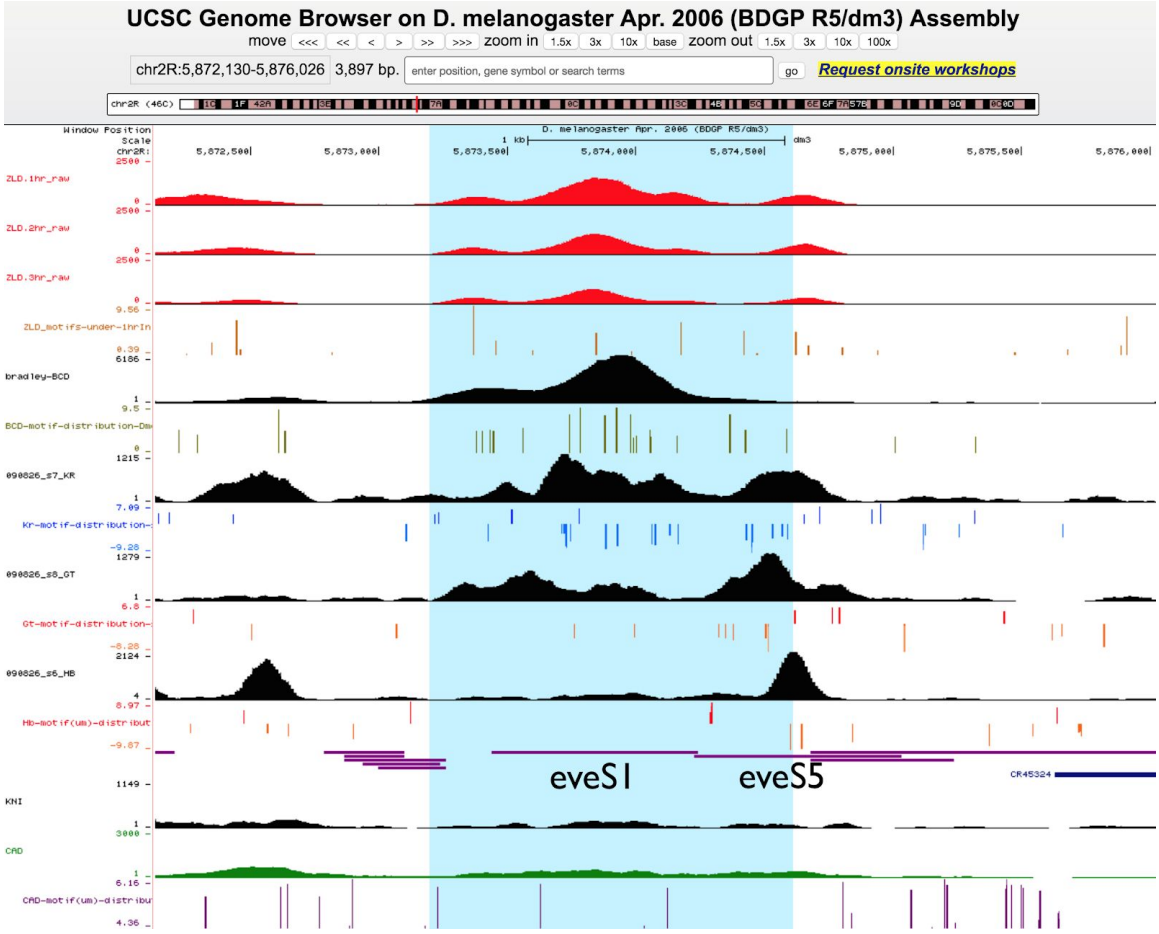
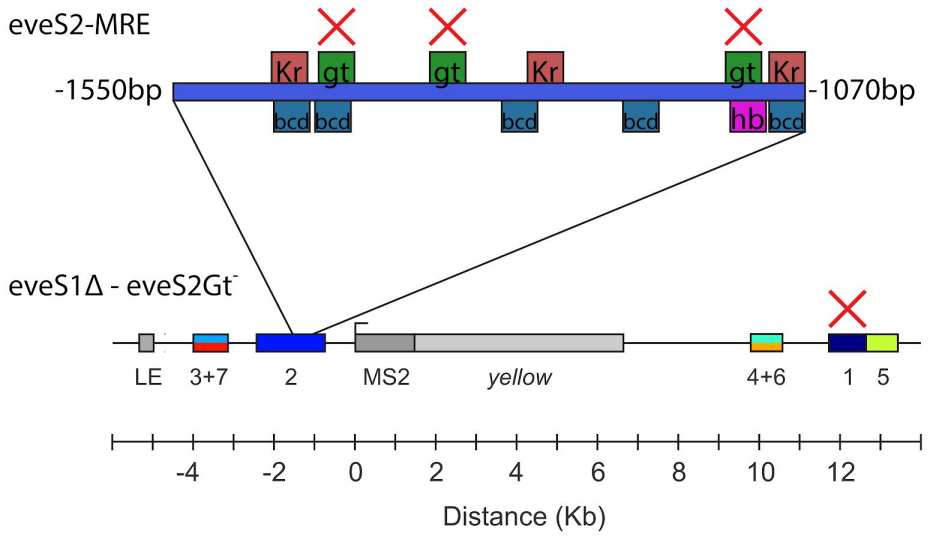


FIGURE 4.4B

Diagram of the eveS1Δ - eveS2Gt⁻ construct, a variation of the eveMS2 BAC reporter.
We deleted the three Gt-binding sites in the eveS2 Minimal Regulatory Element (eveS2-MRE), as in eveS1wt - eveS2Gt⁻. We coupled the Gt mutations with the knock-out of the eveS1 enhancer, as defined in FIG 4A. Deleted sites are labeled with red crosses.



Interestingly, eveS1Δ - eveS2Gt⁻ embryos showed more complex spatio-temporal transcription patterns than predicted. These patterns suggest that eveS1 enhancer influences eveS2 expression and represses anterior expression driven by other *eve* regulatory elements.

Three observations stand out **(FIG 4.5A and FIG 4.5B)**. First, 4/6 embryos show a broad anterior swath of expression encompassing the regions that correspond to eveS1 and eveS2, throughout nc14 and until gastrulation. 2/6 show discernible eveS1 and eveS2 stripes, albeit with observable inter-stripe expression, despite the absence of eveS1 enhancer. Although eveS1 expression might result from an incomplete deletion of eveS1, I consider this unlikely, since we disrupted a region that spans a sequence larger than the traditionally defined eveS1 enhancer. (Haeussler et al. 2019)
Our disruption of eveS1 enhancer affected half of the adjacent eveS5 enhancer, which abrogated eveS5 expression for the most part, resulting in only transient and weak expression of eveS5. eveS1 expression in the absence of eveS1 enhancer seems to result from the activity of the *eve* Late Element (LE). A ~500bp sequence located between -5.9 and -5.2 kb upstream of *eve* gene that positively feedbacks the expression of the *eve* stripes in the presence of the EVE protein (Jiang, Hoey, and Levine 1991).

The LE in the eveMS2 BAC is responsive to the EVE protein supplied by the endogenous *eve* locus.

Another noteworthy conclusion regarding eveS1Δ - eveS2Gt$^-$ embryos, is that they express MS2 in a broad swath that encompasses eveS1 and eveS2 areas and more anterior regions, through all nc14. While eveS1wt - eveS2Gt$^-$ embryos show well defined eveS1 and eveS2 stripes, and only transient inter-stripe MS2 expression between eveS1 and eveS2. This shows that knocking out eveS1 enhancer amplifies the anterior ectopic expression that results from lifting Gt-mediated repression on eveS2. This data suggests that eveS1 enhancer has a repressive role, repressing the expression of a regulatory element in the anterior part of early *Drosophila* embryos.

The second feature that we observed in eveS1Δ - eveS2Gt$^-$ embryos is the appearance of an additional region of expression more anterior than eveS1. The occurrence of this apparent 'eveS0' in eveS1Δ - eveS2Gt$^-$, but not in eveS1wt - eveS2Gt$^-$ embryos, is consistent with the idea that the eveS1 enhancer represses transcription driven by other regulatory elements in the anterior part of the embryo.

Finally, eveS2 shows stronger MS2 expression in eveS1Δ - eveS2Gt$^-$ than in wt eveMS2 BAC and eveS1wt - eveS2Gt$^-$ embryos. (Lim et al. 2018) noted that eveS2 has a broader domain that appears earlier and extends more anteriorly after disruption of eveS1 enhancer, than the eveS2 of embryos with an intact eveS1 enhancer. Our results confirm this observation and show additionally that eveS2 is brighter in eveS1Δ - eveS2Gt$^-$ embryos. A stronger expression of eveS2Gt$^-$ in the absence of eveS1 adds more evidence to the hypothesis that eveS1 has a repressive role in the anterior part of *Drosophila* embryos.

FIGURE 4.5A

MS2 expression snapshots of an embryo carrying the eveS1Δ - eveS2Gt^{-} construct at 10, 25, 30, and 50 min into nc14. Transcription starts at 5 min (not shown). At 10min, we observe a swath of anterior expression, that eventually refines into a wider eveS2. At 25min, the anteriorly expanded eveS2 appears as a wide swath that leaks into the region of eveS1. At 30 min into nc14, eveS2 is wider than wild-type eveS2, but it is a separate stripe from eveS1, which appears as a dim stripe, that seems to result from the expression of the LE. Some nuclei activate in the region that constitutes eveS0. At 50min into nc14, eveS1 is a defined stripe, perhaps due to the activity of the LE, and eveS0 appears more anterior than eveS1.

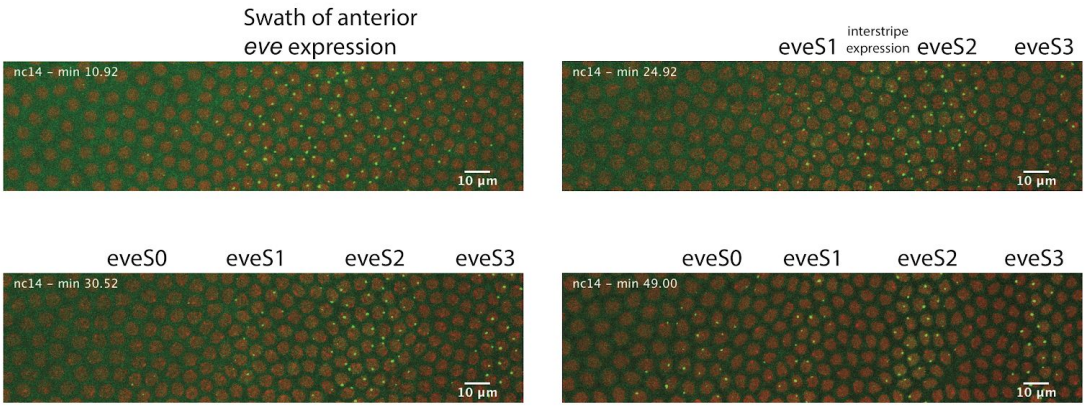
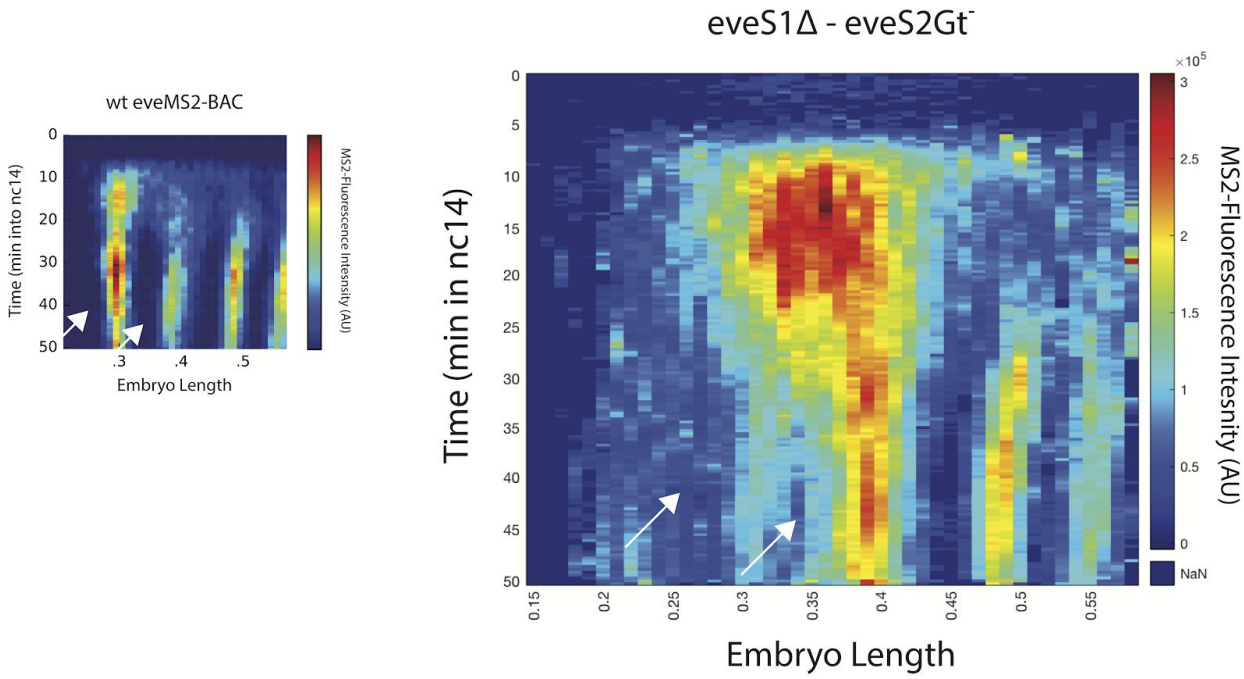


FIGURE 4.5B

This heatmap shows the combined fluorescence of 6 embryos carrying the MS2 construct eveS1Δ - eveS2Gt$^-$. MS2 expression in eveS1Δ - eveS2Gt$^-$ embryos differs from the anterior expression of wt eveMS2 embryos (wt eveMS2 fluorescence shown in the small heatmap on the left for comparison). eveS1Δ - eveS2Gt$^-$ embryos show a large swath of anterior expression (white arrows) that, in the absence of eveS1 enhancer, refines into a brighter and wider eveS2. Our data shows that eveS1 enhancer has repressive properties on the expression of eveS2. A wide area of ectopic expression covers most of the anterior region of eveS1Δ - eveS2Gt$^-$ embryos, up until 0.2 embryo length. eveS1 appears as a dim but defined stripe, most likely due to the activity of the LE. A defined eveS0 appears late in nc14. eveS3 seems to be stronger than in eveS1wt - eveS2Gt$^-$ embryos. It might be that eveS1 enhancer also plays a repressive role on eveS3, but more work is needed to confirm this observation.

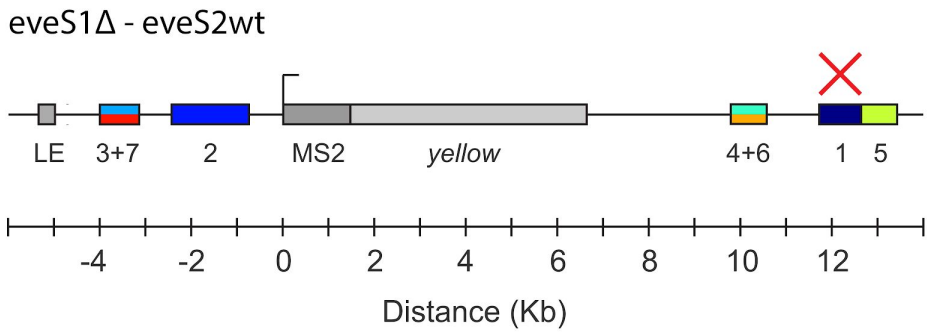
Embryo Length refers to the position of nuclei relative to the anterior-posterior axis of *D. melanogaster* embryos. 0 is the anterior tip and 1 is the posterior tip of an embryo.



We engineered a third modification of the eveMS2 BAC reporter. We disrupted eveS1 enhancer, as in the previously described eveS1Δ - eveS2Gt$^-$ BAC, while keeping an intact eveS2 enhancer sequence (eveS1Δ - eveS2wt) **(FIG 4.6)**. We devised this eveMS2 BAC mutant to test the role of eveS1 enhancer in the appearance of ectopic patterns of gene expression and explore its repressive properties.

FIGURE 4.6

Diagram of the eveS1Δ - eveS2wt construct, a variation of the eveMS2 BAC reporter.
We knocked-out the eveS1 enhancer, as defined in FIG 4A. Deleted sites are labeled with red crosses. The rest of the regulatory elements are wild-type.



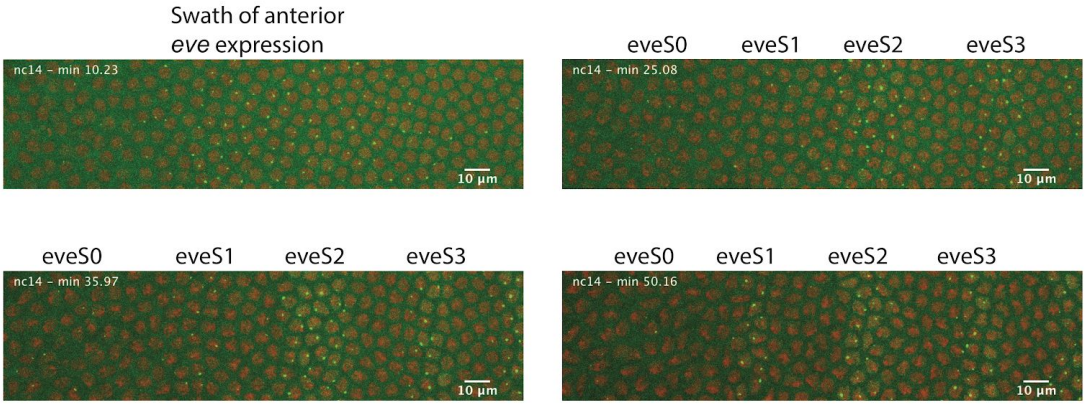
2 out of 5 eveS1Δ - eveS2wt embryos showed several transcriptionally active nuclei in a more anterior region than eveS1, which mimics the appearance of the eveS0 observed in
eveS1Δ - eveS2Gt^{-} embryos. 3 out 5 eveS1Δ - eveS2wt embryos only had between 2-4 active nuclei in the anterior region that corresponds to eveS0. This data indicates that deleting eveS1 enhancer lifts a repressive effect in the anterior part of the embryo **(FIG 4.7)**, that is further amplified by lifting the Gt repression on eveS2.

As (Lim et al. 2018) previously observed, eveS2 becomes broader and extends more anteriorly in the absence of eveS1 enhancer. At first glance, it seems that eves2 expresses brighter in eveS1Δ - eveS2wt embryos. However, we need further analysis to confirm this observation. This data adds further support to the idea that eveS1 enhancer exerts a repressive activity on the anterior part of the embryo, and when we lift the repressive activity of eveS1 enhancer eveS2 is active in a larger area.

Interestingly, we observed MS2 transcription in the region that corresponds to eveS1 in the absence of the eveS1 enhancer. As in wild type *eve* transcription, a broad swath of anterior expression appears at ~5 min after 14th nuclear cleavage. Nevertheless, in the absence of the eveS1 enhancer, this anterior swath refines only into eveS2, while eveS1 transiently disappears. eveS1 reappears at ~30 min into nc14, most likely as a result of the activity of the LE. It has been hypothesized that LE activity starts ~30 min into nc14 (Goto, Macdonald, and Maniatis 1989; Jiang, Hoey, and Levine 1991; Harding et al. 1989). The EVE endogenous protein present in the area may activate the LE in the eveMS2 BAC reporter.

FIGURE 4.7

MS2 expression snapshots of an embryo carrying the eveS1Δ - eveS2wt construct at 10, 25, 35, and 50 min into nc14. Transcription starts at 5 min (not shown). At 10min, we observe a swath of anterior expression that eventually refines into a wider eveS2. At 25min, eveS2 appears wider than eveS2 in the wt, perhaps due to the repressive activity of the eveS1 enhancer over eveS2. There are only a few nuclei active in the regions that correspond to eveS1 and eveS0. At 35 min, eveS2 is wider, and there is a marginal eveS1, perhaps as a result of the activity of the LE. A few nuclei are active in the region of eveS0. At 50min into nc14, eveS1 is a defined stripe, perhaps due to the activity of the LE, eveS2 is thinner than in previous time, and eveS0 appears more anterior than eveS1.



A tale of two body regions: transcription has the same kinetics in endogenous and ectopic *eve* active nuclei

Patterns of gene expression result from several regulatory elements acting in concert at a defined time and place in the body of a multicellular organism. Ectopic patterns of gene expression result from the unmasked activity of regulatory elements, often by lifting TF-mediated repression. Ectopic *eve* expression is normally repressed, while transcription in endogenous stripe patterns has been selected to yield optimal levels of mRNA at given locations in space and time.

The levels of ectopic MS2 mRNA synthesis that result from deleting eveS1 enhancer, coupled with the disruption of Gt-binding sites in eveS2 enhancer, are weaker than the levels of MS2 mRNA synthesis that we observe in the well-known, endogenous, seven *eve* stripes. Perhaps due to the absence of optimal levels of activator and repressor TFs in ectopic regions of *eve* expression.

Hence, we wondered whether the features of transcriptional bursting in ectopic patterns were different than those in the endogenous regions of gene expression, beyond the obvious variations in mRNA levels. Differences in transcriptional bursting between endogenous and ectopic, unmasked, *eve* patterns would point towards a selection for 'optimal' bursting kinetics
in endogenous stripes; a selection absent in ectopic regions. On the other hand, similar bursting kinetics between ectopic and endogenous regions of *eve* expression suggest similar underlying processes that are not under separate selective pressures.
As MS2 enables the analysis of transcriptional bursting driven by *eve* regulatory regions, we compared the MS2 traces from nuclei laying in ectopic and endogenous *eve* expression patterns.

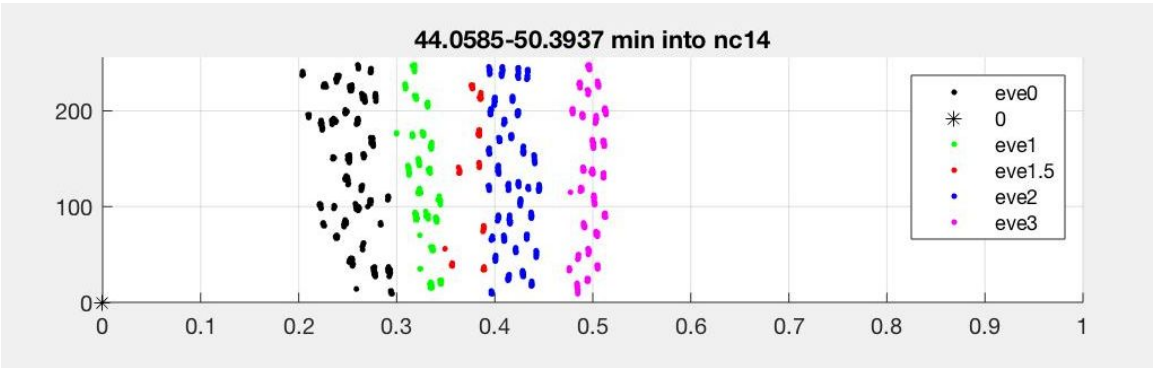
We manually segmented movies from eveS1wt - eveS2Gt$^-$ and eveS1Δ - eveS2Gt$^-$ embryos. I assigned an identity to regions of transcription in the movies (stripe 0, stripe 1, interstripe 1-2, stripe 2, etc) **(FIG 4.8).** As I mentioned above, eveS1wt - eveS2Gt$^-$ flies yielded very few ectopic active nuclei, insufficient to draw a meaningful profile of ectopic bursting kinetics.

I manually segmented the regions that correspond to different stripes, based on the position of stripes at 45 min into nc14. Then, we corrected the boundary between inter-stripe eveS1 and eveS2, based on the position of nuclei relative to the stripe observed in the heatmap in Figure 5.

eveS1Δ - eveS2Gt⁻ embryos showed large regions of ectopic MS2 expression, amenable for further analysis of transcriptional bursting.

FIGURE 4.8

Segmentation of the stripes of a single eveS1Δ - eveS2Gt⁻ embryo. I manually selected the regions that corresponded to each of the *eve* stripes (eveS0, eveS1, eveS2, eveS3) and the interstripe between eveS1 and eveS2 (eve1.5) at 45min into nc14. Active nuclei were assigned to their corresponding stripes). Unassigned nuclei appear as an asterisk and are labeled as 0 (not shown).



The analysis of bursts, developed by (Lammers et al. 2020), relies on a Hidden Markov Model that enables us to infer the sequence of states of the *eve* promoter (ON or OFF), over time, that better predicts its observable outcome: the fluorescence trace that results from MS2 expression, over time. In a nutshell, a Hidden Markov Model (HMM) deduces the unobservable states of a system, or 'hidden' states. In this case, the model developed by Lammers *et al*. 2020, presumes two hidden states of the promoter. ON, when the promoter loads active polymerases into the gene; and OFF, when the promoter is inactive and does not load polymerases into the gene. These states are 'hidden' because we cannot directly observe the state of the actual promoter, we can only observe its outcome, the transcription of MS2. Furthermore, to explain the outcome, Hidden Markov Models must take into account the probability of transitioning from one state to another. The Hidden Markov Model by Lammers *et al*. 2020, estimates the probability of transitioning from the OFF state to the ON state (kon) and the probability of transitioning from the ON state to the OFF state (koff). In the end, a sequence of promoter states that transition between ON and OFF, with kon and koff probabilities, and load active mRNA polymerases to the promoter at a rate r when the promoter is ON, explains the observed MS2 traces.

The formula below explains the relationship between promoter features and observed fluorescence outcome. The average fluorescence of MS2 bursts results from the time

that a promoter remains active multiplied by the number of polymerases loaded per unit of time while the promoter is active (kon * r), divided by the total amount of time that the promoter is active or inactive (kon + koff). **(FIG 4.9B)**

The kon, koff, and r parameters describe the activity of a promoter. Therefore, they shed light into the kinetics and features of transcriptional bursting.

kon is a proxy for bursting frequency, as every time that a promoter goes from OFF to ON, a new burst of mRNA synthesis starts. Increased kon results in more frequent initiation of new bursts of mRNA synthesis.

koff is a proxy for bursting duration: a lower probability that a promoter transitions from ON to OFF state means that a given burst would last for longer. Decreased koff results in longer bursts.

r is a proxy for burst amplitude, or the rate at which burst fluorescence increases per unit of time. A higher r means that burst fluorescence increases faster whenever the promoter is active. This implies that promoters with a higher r load active polymerases into the gene at a higher rate when the promoter is active.
We sought to compare the features of transcription in different regions of developing *D. melanogaster* embryos. Therefore, we aimed to compare the kon, koff, and r, *i.e*. the frequency, duration, and amplitude, of bursts in nuclei with equivalent fluorescence in different regions of developing *D. melanogaster* embryos. **(FIG 4.9A)**

FIGURE 4.9A

Two-step model of promoter activity. This model states that promoters switch between an inactive state (OFF) to an active state (ON) at a rate kon, and from the ON state to the OFF state at a rate koff. When promoters are in the ON state, they recruit active RNA Pol II molecules at a rate r. The details of these kinetics shape transcriptional bursting.

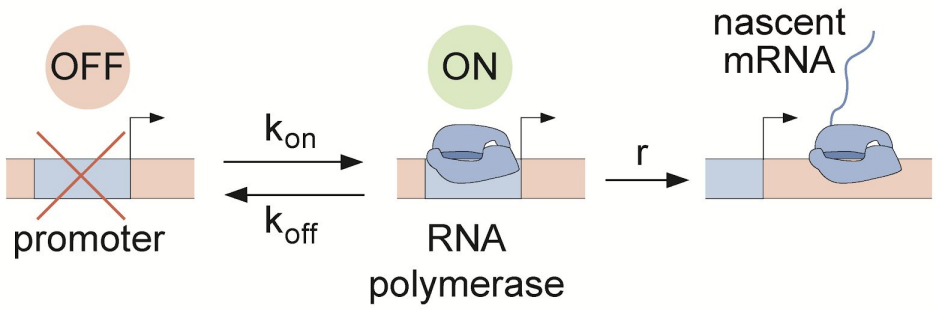


Figure 4.9B

This formula relates the kon, koff, and r values from the two-step model of promoter activity to the average fluorescence, *i.e.* the average levels of mRNA synthesis, that an active promoter yields.
The average fluorescence that results from an active promoter equals the amount of time that the promoter stays active (kon), multiplied by the mRNA Pol II molecules recruited to the promoter (r) while the promoter is active (r * kon). All, divided by the total amount of time that the promoter stays in its active or inactive states (kon + koff).

$$Fluorescence = r \frac{kon}{kon + koff}$$

In principle, promoters and enhancers can synthesize similar levels of mRNAs using different combinations of Kon, koff, and r parameters.

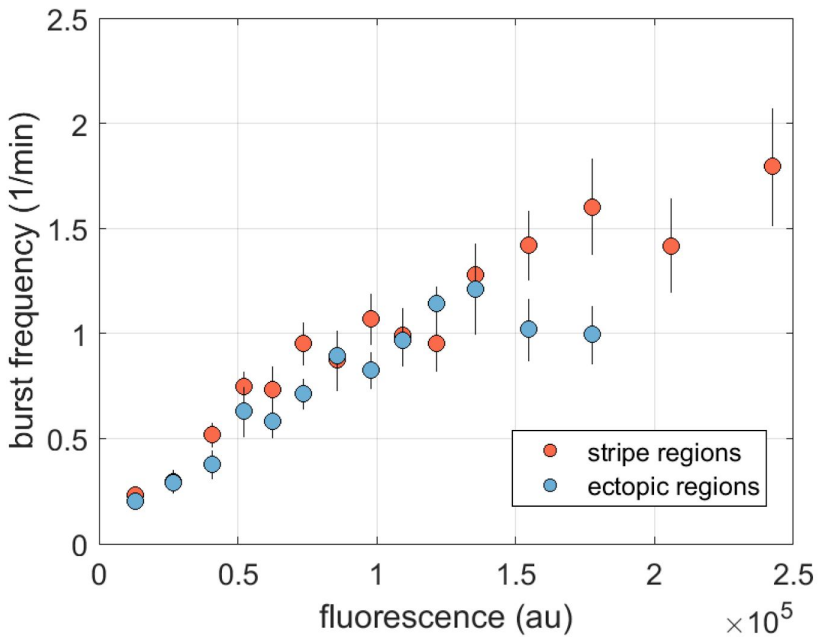
If we had observed that *eve* active nuclei yield equivalent levels of mRNA synthesis through varying mixtures of kon, koff, and r kinetics in different body regions, this would indicate that enhancers convey the information to tweak bursting kinetics in ways particular to each body region. Moreover, particular kon, koff, and r kinetics in endogenous *eve* stripes, different than the bursting kinetics observed in ectopic regions of gene expression, would suggest that enhancers are under selection to drive mRNA synthesis through 'optimal' bursting kinetics, and that such selection is not present at unmasked, ectopic areas of *eve* expression.

Interestingly, we observed the opposite. *eve* active nuclei with equivalent levels of mRNA synthesis achieve their transcriptional output through shared bursting kinetics, regardless of whether they are in the various endogenous *eve* stripe patterns or in the ectopic regions of *eve* expression.
This tells us that the many enhancers of *even-skipped* do not convey the information to customize bursting kinetics at particular regions of *eve* expression. Instead, the combination of kon, koff and r values common to *eve* active nuclei with similar levels of mRNA output stems from a shared molecular process, perhaps the shared *eve* promoter sequence. A second more interesting probability arises. Previous work by (Zoller, Little, and Gregor 2018) observed that transcription in the expression boundaries of all *D. melanogaster* Gap genes share the same bursting kinetics, despite the myriad of inputs that drive their expression. The authors propose that transcription may result from shared global strategies in all genes in *D. melanogaster*.
According to our data, enhancers are under selection to modulate mRNA output in space and time, while mRNA synthesis results from shared bursting kinetics in all *eve* active nuclei.

We sorted endogenous nuclei in 18 bins and ectopic nuclei in 8 bins, sorted according to their average fluorescence. Bins have the same number of *eve* nuclei.
Nuclei with similar average fluorescence show similar kon values, regardless of whether they are in the canonical endogenous *eve* stripes or in the ectopic regions of *eve* expression **(FIG 4.10)**. We also observe an increasing trend. More fluorescent nuclei have higher kon values. They have a higher frequency of bursting initiation.

FIGURE 4.10

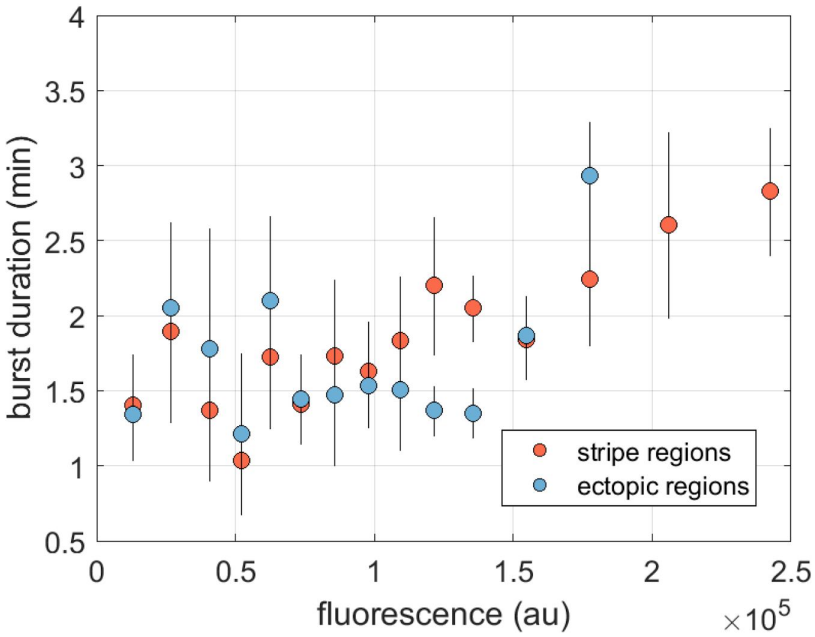
Nuclei with similar average fluorescence in ectopic and endogenous regions show similar kon values (labeled as burst frequency). *i.e.* nuclei with similar average fluorescence in ectopic and endogenous regions initiate a similar number of bursts per minute.



koff values, *i.e.* burst duration, do not show a considerable change across bins of nuclei with the same fluorescence. Consistent with our previous analyses of *eve* endogenous stripes, koff values of promoter activity seem to play only a marginal role in enhancer-mediated regulation of gene expression **(FIG 11).**

FIGURE 4.11

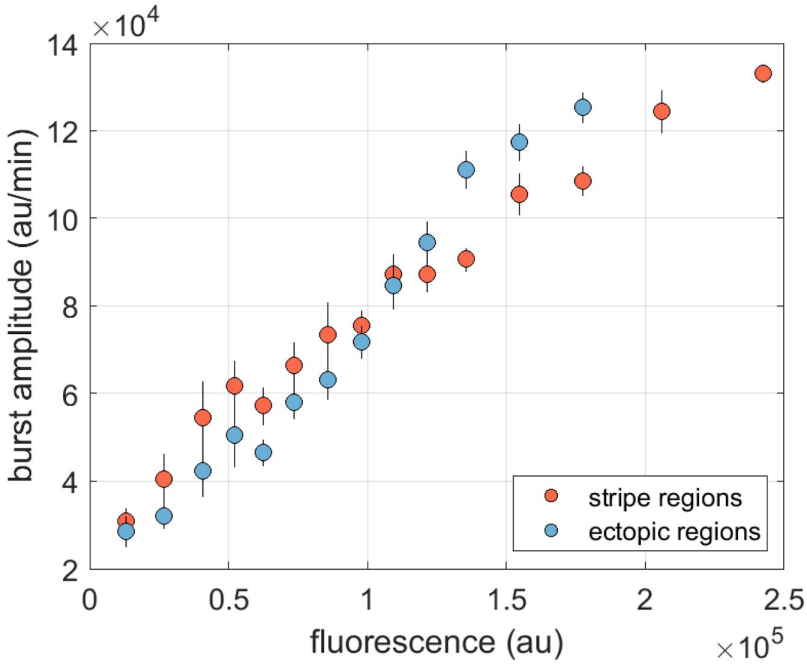
koff (burst duration) does not show a considerable change across *eve*-active nuclei sorted by bins of average fluorescence. *i.e.* bursts last for ~2 minutes in nuclei with similar average fluorescence, regardless on whether they are in ectopic or endogenous regions.



eve nuclei of similar average fluorescence have similar r values. *i.e.* burst amplitude. r values increase at higher values of mRNA synthesis. r can be interpreted as the increase of MS2 fluorescence intensity per minute, when *eve* promoter is in ON state **(FIG 4.12**).

FIGURE 4.12

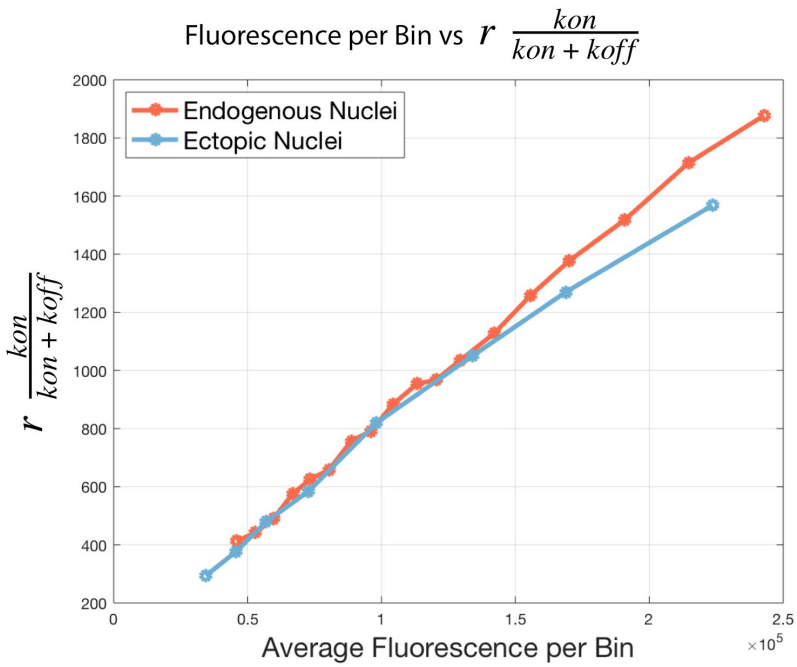
r (burst amplitude) increases as the average fluorescence intensity of *eve* active nuclei increases. However, r values for ectopic and endogenous regions of *eve* expression are almost identical.



According to our model, kon, koff, and r parameters must relate to the average fluorescence of nuclei through the formula $Fluorescence = r \frac{kon}{kon + koff}$. Therefore, the estimations of kon, koff, and r that resulted from the HMM by Lammers *et al.* 2020 must be consistent with the average fluorescence of nuclei in each of the fluorescent bins from figures 7, 8, and 9.

The estimations of kon, koff and r that followed from the HMM analysis were consistent with our model. Hence, we observe a linear relationship between average $Fluorescence$ and $r \frac{kon}{kon + koff}$ parameters from the nuclei within each fluorescence bin **(FIG 4.13)**. Deviations would indicate problems with our analysis. We observed a linear relationship between $Fluorescence$ and $\frac{kon}{kon + koff}$ for most fluorescent bins, except for the last 2 fluorescence bins from the ectopic combined data that deviate from the expected linear relationship. We defined bins of average fluorescence with the same number of nuclei per bin. Since there were few nuclei on the upper end of fluorescence in ectopic regions of expression, we believe that those two pesky data points result from HMM computations over outlier nuclei with very dissimilar mean fluorescence values.

FIGURE 4.13 - Relationship between average fluorescence per bin and the kon, koff, and r estimated for each bin in ectopic and endogenous regions. This graph was made with data from a previous round of analysis than that shown in figures 4.10, 4.11, and 4.12. For the final paper, we will replicate this graph using the results in the figures aforementioned. Nevertheless, I expect that the new version of this graph will be similar.



This data shows that the relationship between kon (proxy for burst frequency), koff (burst length), and r (burst amplitude), is the same in endogenous stripes, and between endogenous and ectopic regions of *eve* expression. Transcription has similar features, regardless of the enhancers that drive *eve* expression. Our results indicate that the kinetics of transcriptional bursts result from shared underlying processes in all *eve* active nuclei.

The mechanisms that generate the shared transcription features in *eve* expressing nuclei, regardless of whether they are ectopic or endogenous, may be a consequence of their shared gene promoter sequence, despite differences in enhancer function. On the other hand, the unified bursting features that we observe might result from the optimized activity of general transcription factors, mediator complex, RNA Pol II, and other elements of the transcriptional machinery. mRNA FISH experiments by (Zoller, Little, and Gregor 2018) observed that early *D. melanogaster* patterning genes are constrained to share similar transcriptional kinetics.

Preliminary data show that ectopic expansions of *eve* expression, which in general have lower levels of mRNA synthesis, have a lower kon and a higher koff than the more active nuclei at the center of the endogenous *eve* stripes. Expansions of *eve* expression may have lower levels of mRNA synthesis due to insufficient levels of activator TFs, higher levels of repressor TFs, or both.

The appearance of new and unexpected ectopic patterns of gene expression upon modification and disruption of eveS1 and eveS2 enhancers, suggests interesting possibilities regarding the evolution of expression patterns. Enhancers may have the potential to express in several nuclei. Nevertheless, their activity is masked at some places. Those nuclei with the right combination of activator and repressor queues express a given gene. The information to activate that gene is encoded in the DNA sequence of its genetic enhancers. Interestingly, disruption of binding sites in one enhancer drives expression of ectopic patterns in unexpected ways. Our data shows that disrupting two enhancers (eveS2 Gt binding sites and eveS1) has a different effect than disrupting each one of them alone. Mutations on existing enhancers can generate nouvelle patterns of gene expression, and thus new body plans in evolution.

Future analyses: The transcriptional kinetics of eveS1 enhancer-mediated repression

As mentioned above, we have generated sufficient data from embryos carrying a version of our eveMS2 BAC reporter with its eveS1 enhancer disrupted (eveS1Δ - eveS2wt). We are analyzing the *eve* MS2 bursting data from these embryos to understand how eveS1 enhancer mediated repression modifies transcriptional bursting, and what happens to transcriptional kinetics when we lift eS1 enhancer mediated repression. I predict that the bursting frequency (kon), and bursting duration (1/koff) of *eve* active nuclei in eveS2 will increase upon disrupting eveS1 enhancer. Thus lifting eveS1 enhancer-mediated repression. Nevertheless, the higher levels of transcription in eveS2 nuclei will show the same bursting kinetics (kon, koff, r) as any other *eve* active nucleus with equivalent levels of mRNA synthesis.

Further studies combining genetic disruption of *even-skipped* regulatory elements and live imaging of *eve* expression would shed light on the role of enhancers in orchestrating the stochastic process of transcription in time and space, the emergence

of gene expression patterns, and ultimately, the establishment of body parts in animal embryonic development.

Methods

VectorBuilder constructed and shipped the three modifications of the eveMS2 BAC through two-step ccdB-amp cassette mediated recombineering.

Two-step ccdB-amp mediated recombineering consists in replacing a DNA sequence in a BAC with a cassette through homology arm recombination, and then substitute that cassette for another DNA sequence. The cassette confers resistance to the antibiotic Ampicillin and expresses the bacterial toxin ccdB. Ampicillin resistance selects for bacteria with insertions of the ccdB-amp cassette. In the presence of the anti-toxin ccdA, ccdB toxin is innocuous for bacteria. The second step in the process consists in replacing the cassette with a customized DNA sequence. Since ccdB is toxic for bacteria in the absence of the anti-toxin ccdA, ccdB counter-selects in favor of bacteria that replaced the ccdB-amp cassette and successfully underwent the two recombineering steps.

VectorBuilder generated the eveS1wt - eveS2Gt$^-$ modification of eveMS2 BAC (Description of Vector Builder IDs and sequences is located in the file eveMS2BAC_Modifications in https://github.com/meisenlab/BerrocalThesis, section eveS1wt-eveS2Gt).
They knocked-out a 1861bp-long sequence corresponding to the wild type eveS2 enhancer, as defined by the Zelda Chip-seq signal around eveS2-MRE (Harrison et al. 2011; Haeussler et al. 2019). They inserted a ccdB-amp cassette in its place.
Then, they substituted the ccdB-cassette with a 1791bp-long eveS2 enhancer modified sequence, with three Gt-binding sites deleted. Integrated DNA Technologies (IDT) synthesized the modified eveS2 enhancer sequence.

VectorBuilder modified the eveS1wt - eveS2Gt$^-$ eveMS2 BAC to generate the variant eveS1Δ - eveS2Gt$^-$ (Description of Vector Builder IDs and sequences is located in the file eveMS2BAC_Modifications in https://github.com/meisenlab/BerrocalThesis, section eveS1Δ-eveS2Gt). They disrupted eveS1 in eveS1wt - eveS2Gt$^-$ by substituting the 1413bp eveS1 enhancer, as defined by the Zelda Chip-seq signal around eveS1 (Harrison et al. 2011; Haeussler et al. 2019), with a 1422bp neutral ccdB-amp bacterial DNA cassette.

VectorBuilder modified the wt eveMS2 BAC to generate the variant eveS1Δ - eveS2wt (Description of Vector Builder IDs and sequences is located in the file eveMS2BAC_Modifications in https://github.com/meisenlab/BerrocalThesis, section eveS1Δ-eveS2wt). They disrupted eveS1 in wt eveMS2 BAC by substituting the eveS1 enhancer with a neutral ccdB-amp bacterial DNA cassette, as described for above for eveS1Δ - eveS2Gt^{-}.

Bibliography

Berrocal, A., N. C. Lammers, H. G. Garcia, and M. B. Eisen. 2018. “Kinetic Sculpting of the Seven Stripes of the Drosophila Even-Skipped Gene.” *BioRxiv*. https://www.biorxiv.org/content/10.1101/335901v2.abstract.

Bothma, Jacques P., Hernan G. Garcia, Emilia Esposito, Gavin Schlissel, Thomas Gregor, and Michael Levine. 2014. “Dynamic Regulation of Eve Stripe 2 Expression Reveals Transcriptional Bursts in Living Drosophila Embryos.” *Proceedings of the National Academy of Sciences of the United States of America* 111 (29): 10598–603. https://doi.org/10.1073/pnas.1410022111.

Chubb, Jonathan R., Tatjana Trcek, Shailesh M. Shenoy, and Robert H. Singer. 2006. “Transcriptional Pulsing of a Developmental Gene.” *Current Biology: CB* 16 (10): 1018–25. https://doi.org/10.1016/j.cub.2006.03.092.

Dar, Roy D., Brandon S. Razooky, Abhyudai Singh, Thomas V. Trimeloni, James M. McCollum, Chris D. Cox, Michael L. Simpson, and Leor S. Weinberger. 2012. “Transcriptional Burst Frequency and Burst Size Are Equally Modulated across the Human Genome.” *Proceedings of the National Academy of Sciences of the United States of America* 109 (43): 17454–59. https://doi.org/10.1073/pnas.1213530109.

Ferguson, Matthew L., and Daniel R. Larson. 2013. “Measuring Transcription Dynamics in Living Cells Using Fluctuation Analysis.” *Methods in Molecular Biology* 1042: 47–60. https://doi.org/10.1007/978-1-62703-526-2_4.

Fujioka, M., Y. Emi-Sarker, G. L. Yusibova, T. Goto, and J. B. Jaynes. 1999. “Analysis of an Even-Skipped Rescue Transgene Reveals Both Composite and Discrete Neuronal and Early Blastoderm Enhancers, and Multi-Stripe Positioning by Gap Gene Repressor Gradients.” *Development* 126 (11): 2527–38. https://www.ncbi.nlm.nih.gov/pubmed/10226011.

Fukaya, Takashi, Bomyi Lim, and Michael Levine. 2016. “Enhancer Control of Transcriptional Bursting.” *Cell* 166 (2): 358–68. https://doi.org/10.1016/j.cell.2016.05.025.

Garcia, Hernan G., Mikhail Tikhonov, Albert Lin, and Thomas Gregor. 2013. “Quantitative Imaging of Transcription in Living Drosophila Embryos Links Polymerase Activity to Patterning.” *Current Biology: CB* 23 (21): 2140–45. https://doi.org/10.1016/j.cub.2013.08.054.

Golding, Ido, Johan Paulsson, Scott M. Zawilski, and Edward C. Cox. 2005. “Real-Time Kinetics of Gene Activity in Individual Bacteria.” *Cell* 123 (6): 1025–36. https://doi.org/10.1016/j.cell.2005.09.031.

Goto, T., P. Macdonald, and T. Maniatis. 1989. “Early and Late Periodic Patterns of Even Skipped Expression Are Controlled by Distinct Regulatory Elements That

Respond to Different Spatial Cues." *Cell* 57 (3): 413–22. https://doi.org/10.1016/0092-8674(89)90916-1.

Haeussler, Maximilian, Ann S. Zweig, Cath Tyner, Matthew L. Speir, Kate R. Rosenbloom, Brian J. Raney, Christopher M. Lee, et al. 2019. "The UCSC Genome Browser Database: 2019 Update." *Nucleic Acids Research* 47 (D1): D853–58. https://doi.org/10.1093/nar/gky1095.

Harding, K., T. Hoey, R. Warrior, and M. Levine. 1989. "Autoregulatory and Gap Gene Response Elements of the Even-Skipped Promoter of Drosophila." *The EMBO Journal* 8 (4): 1205–12. https://www.ncbi.nlm.nih.gov/pubmed/2743979.

Harrison, Melissa M., Xiao-Yong Li, Tommy Kaplan, Michael R. Botchan, and Michael B. Eisen. 2011. "Zelda Binding in the Early Drosophila Melanogaster Embryo Marks Regions Subsequently Activated at the Maternal-to-Zygotic Transition." *PLoS Genetics* 7 (10): e1002266. https://doi.org/10.1371/journal.pgen.1002266.

Jiang, J., T. Hoey, and M. Levine. 1991. "Autoregulation of a Segmentation Gene in Drosophila: Combinatorial Interaction of the Even-Skipped Homeo Box Protein with a Distal Enhancer Element." *Genes & Development* 5 (2): 265–77. https://www.ncbi.nlm.nih.gov/pubmed/1671662.

Lammers, Nicholas C., Vahe Galstyan, Armando Reimer, Sean A. Medin, Chris H. Wiggins, and Hernan G. Garcia. 2020. "Multimodal Transcriptional Control of Pattern Formation in Embryonic Development." *Proceedings of the National Academy of Sciences of the United States of America* 117 (2): 836–47. https://doi.org/10.1073/pnas.1912500117.

Liang, Hsiao-Lan, Chung-Yi Nien, Hsiao-Yun Liu, Mark M. Metzstein, Nikolai Kirov, and Christine Rushlow. 2008. "The Zinc-Finger Protein Zelda Is a Key Activator of the Early Zygotic Genome in Drosophila." *Nature* 456 (7220): 400–403. https://doi.org/10.1038/nature07388.

Lim, Bomyi, Takashi Fukaya, Tyler Heist, and Michael Levine. 2018. "Temporal Dynamics of Pair-Rule Stripes in Living Drosophila Embryos." *Proceedings of the National Academy of Sciences of the United States of America* 115 (33): 8376–81. https://doi.org/10.1073/pnas.1810430115.

McKnight, S. L., and O. L. Miller Jr. 1979. "Post-Replicative Nonribosomal Transcription Units in D. Melanogaster Embryos." *Cell* 17 (3): 551–63. https://doi.org/10.1016/0092-8674(79)90263-0.

Sackerson, C., M. Fujioka, and T. Goto. 1999. "The Even-Skipped Locus Is Contained in a 16-Kb Chromatin Domain." *Developmental Biology* 211 (1): 39–52. https://doi.org/10.1006/dbio.1999.9301.

Senecal, Adrien, Brian Munsky, Florence Proux, Nathalie Ly, Floriane E. Braye, Christophe Zimmer, Florian Mueller, and Xavier Darzacq. 2014. "Transcription Factors Modulate c-Fos Transcriptional Bursts." *Cell Reports* 8 (1): 75–83.

https://doi.org/10.1016/j.celrep.2014.05.053.

Skupsky, Ron, John C. Burnett, Jonathan E. Foley, David V. Schaffer, and Adam P. Arkin. 2010. “HIV Promoter Integration Site Primarily Modulates Transcriptional Burst Size Rather Than Frequency.” *PLoS Computational Biology*. https://doi.org/10.1371/journal.pcbi.1000952.

Small, S., A. Blair, and M. Levine. 1992. “Regulation of Even-Skipped Stripe 2 in the Drosophila Embryo.” *The EMBO Journal* 11 (11): 4047–57. https://www.ncbi.nlm.nih.gov/pubmed/1327756.

Venken, Koen J. T., Joseph W. Carlson, Karen L. Schulze, Hongling Pan, Yuchun He, Rebecca Spokony, Kenneth H. Wan, et al. 2009. “Versatile P[acman] BAC Libraries for Transgenesis Studies in Drosophila Melanogaster.” *Nature Methods* 6 (6): 431–34. https://doi.org/10.1038/nmeth.1331.

Zenklusen, Daniel, Daniel R. Larson, and Robert H. Singer. 2008. “Single-RNA Counting Reveals Alternative Modes of Gene Expression in Yeast.” *Nature Structural & Molecular Biology*. https://doi.org/10.1038/nsmb.1514.

Zoller, Benjamin, Shawn C. Little, and Thomas Gregor. 2018. “Diverse Spatial Expression Patterns Emerge from Unified Kinetics of Transcriptional Bursting.” *Cell* 175 (3): 835–47.e25. https://doi.org/10.1016/j.cell.2018.09.056.