THE UNIVERSITY OF CHICAGO

EXPERIMENTAL AND COMPUTATIONAL ANALYSIS OF CANALIZATION AND ITS RELATION TO EVOLUTION AND DEVELOPMENT

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE DIVISION OF THE BIOLOGICAL SCIENCES
AND THE PRITZKER SCHOOL OF MEDICINE
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

DEPARTMENT OF ECOLOGY AND EVOLUTION

BY
PENGYAO JIANG

CHICAGO, ILLINOIS
JUNE 2017

"What does not kill me, makes me stronger."

— Friedrich Nietzsche

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

First of all, I would like to express my sincere thanks to my advisor Dr. John Reinitz and co-advisor Dr. Marty Kreitman for their advice, mentorship, and support to me. Thanks John for being patient with me when I began to learn experiments and trouble-shooting for the right *in situ* hybridization conditions. He never rushed me and let me figure out things at my own pace. I also learned a lot from his rigorous attitude towards science. Marty is a near perfect human being in my opinion. Not only is his insightful ideas that inspire me almost every time I talk to him but also his sincere and kindness that helped me through some hard times. I cannot thank him enough for his help to me through my Ph.D.

I would also thank my committee members Dr. Chip Ferguson and Dr. Manyuan Long. I still remember when I was a first year graduate student with little background in biology, Chip encouraged me citing his own example of being a computer science undergraduate student and now mastering developmental biology. He is also kind to me by offering to read my manuscript on simulations of robustness. Manyuan is also kind to me always. If I need to talk to him for advice, he always finds time for me. He encourages me so much every time I talk to him.

Thank the past and current members that in Kreitnitz lab. Thanks Dr. Bin He for teaching me a lot of the basics when I was a first year student doing rotation in the lab. His positive attitude encourages me a lot. Thanks Kenneth Barr for being a pleasant and wonderful person to work with in the lab, who joined the lab the same time as me and is about to graduate the same time as well. Thanks Dr. Misha Ludwig for being a good friend and a teacher. I learned a lot from him for the fly crosses. He is also a nice a sincere person to talk to. If I forgot my keys on weekend or evening, he would come and rescue me. Thanks Dr. Zhihao Lou for helping me out with some computer-related problems. Thanks Dr. Xuan Zhuang and Dr. Soo Young Park for being around. I also learned more population genetics from Dr. Andrew Bergen. Thank Dr. Alexandre Ramos and Dr. Si Tang for helping me to get analytical results from my simulations, which is still ongoing. Thanks Dr. Loretta Au

for giving me a lot of useful suggestions during the SIAM meeting. Thanks Dr. Chun Wai Kwan for providing a lot of suggestion when I was trouble-shooting *in situ* hybridization experiments. Thank people working in E & E administration office who are always nice and helpful, Bonnie Brown, Connie Homan, Audrey Aronowsky, Mary Johnson, Jeff Wisniewski and Alison Anastasio.

I would also thank my dear friends who I share my happiness and sorrow with. Dr. Ziyue Gao, Dr. Rui Yang, Dr. Grace Yu, Tianyi Li, Anita Ng, Dr. Grace Lee, Dr. Qinwen Liu, Dr. Quan Gao, William Yee, Kanix Wang, Yi Zeng, Dr. Chen Shen, Dr. Ashish Jha, and also my cohort in E & E. Thanks people who has helped me but forgive me if I forgot to mention your name at this moment.

I would also thank my family for their support. Thanks my mom and dad, and also my grandfather, who cares for me all the time.

ABSTRACT

Canalization captures both cell differentiation processes during development as well as the reproducibility of traits despite environmental, developmental or genetic noise observed in wild-type organisms that have been under natural selection. On the one hand, canalization tends to keep phenotypes reproducible, but on the other hand, it is evident that phenotypic evolution has taken place in the past. We need to understand the role of canalization in an evolutionary context, specifically, how canalized traits could evolve. In this thesis, I address three questions that are related to different aspects of this question.

First, considering highly canalized traits at the molecular level, to what degree do variations exist that could pass through the sieve of development? What is the extent of phenotypic variation within a species? This question has been little studied because the conservation of the expression of these patterning genes across species gave rise to the idea that little variation could be expected within a species. In Chapter 2, I consider the the expression of an essential developmental patterning gene, *even-skipped*. I address this question by observing the expression of this gene in inbred lines derived from natural populations. The results demonstrate that appreciable expression variation exists within natural populations. In addition, I was able to characterize the molecular cause for one particular expression variant.

Second, given a trait that is highly conserved across species, are the same or different mechanisms being used to generate the same end phenotype? For a canalized trait which is conserved across large evolutionary distance, in this case *even-skipped* expression in *Drosophila melanogaster (D. mel)* and the sepsid fly *Themira putris (T. put)*, there is good evidence that expression is conserved in the face of significant divergence of regulatory sequence. We address this question using a computational model of transcription to help to understand the possible mechanism that drives different enhancer expression from two different species. Theoretical analysis indicated that Caudal binding sites had replaced those for Bicoid in the sepsid enhancer driving stripe 2. In Chapter 3, I test this conclusion by

constructing embryos that had no Caudal protein and performing a comparative assay of the expression of *even-skipped* stripe 2 driven by enhancers from two species. I found the *T. put* stripe 2 enhancer reduces its expression much more than the *D. mel* stripe 2 enhancer in the *cad* background, supporting the computational prediction.

Finally, in Chapter 4, I address the general question of how canalization could evolve using a computational model. I do this by simplifying canalization to a one-step process of genotype to phenotype map, but capturing its important property—robustness by a parameter in the model which can be explicitly controlled, either by the experimenter or by a genetic locus that can itself evolve. Fitness is determined by both individual's phenotype and the environment. I found high robustness has the selection advantage over low robustness populations under a stable environment to which the organism is optimally adapted, while low robustness has the selection advantage over high robustness ones when adapting to a new environment. The phenotypic space for high robustness populations is smaller than low robustness populations. When robustness itself is controlled by a genetic locus such that it evolves with the phenotype, low robustness is selectively advantageous in a transient manner immediately after an environmental shift. Shortly thereafter, however, and long before the population is well adapted, high robustness becomes selectively advantageous.

CHAPTER 1

INTRODUCTION

1.1 From genotype to phenotype

A central question in biology today is to figure out how genotype give rise to phenotype. With the huge advances in sequencing technologies in the last fifty years [73], genomic sequences from many species and populations are becoming available [29, 28, 83]. However, relatively little is known about how those genotype map onto phenotype. There are essentially two aspects concerning the genotype to phenotype relationship. One is the molecular mechanisms of how genotype gets translated into phenotype. The whole filed of developmental biology is devoted to tackle this question by figuring out how the genotype of a zygote gets unfold and develop into an adult [59]. On the other side of the problem, which is equally important, is how this process can occur reliably. There are many factors that could potentially disturb this process—environmental fluctuations, such as temperature; developmental noise, such as stochastic gene regulation and genetic mutations. Despite of that, most developmental processes occur reproducibly in every individual within a species.

1.2 Canalization

The idea of canalization was first proposed by C. H. Waddington, which captures both sides of the genotype to phenotype relationship. He proposed the idea by observing that there was much smaller phenotypic variation in wild-type organisms that have been under natural selection compared with mutants [185]. He made the famous analogy of an epigenetic landscape to describe the developmental process [187], which illustrates the potential to make different end cell types from a single static genotype. In this epigenetic landscape, a ball representing the undifferentiated state of a zygote is rolling downhill and ends up in one of several possible cell types, represented as valleys. The underlying landscape is determined by

the genotype states and gene-gene interactions. During the developmental process, different end cell types are made quite reliably in the face of environmental, developmental or genetic perturbations, manifested by the ball ending up in the same valley despite of these noise. The first mathematical formulation of such epigenetic landscape was proposed by Thom [180, 181]. He proposed a very general formulation of the dynamical system that evolves with time, in an attempt to capture the all behaviors of developmental processes. Recently, a mathematical model which is more coarse-grained but captured the essential biological structure (Gene regulatory network) has appeared [124, 76]. The network dynamics of gene expression states which changes over time determine the landscape. The final valleys in the landscape are the basins of attractors of the dynamical system.

There is considerable evidence indicating that canalization is a universal property of multi-cellular organisms. We see the reproducibility of many discrete traits, such that humans almost always have 2 legs, 2 hands, and *Drosophila melanogaster* always has head, thoracic, and abdominal segments from anterior to posterior in wild-type conditions. *Drosophila melanogaster* usually have 4 scutellar bristles [149]. *Caenorhabditis elegans* makes cell lineage fate decisions predictably from every cell division [174]. In the cases of recessive heterozygous mutants, where one functional gene is present, usually the phenotypic effect is buffered, which is an extreme case of the buffering effect. Many gene regulatory networks have been found to be robust as well, showing properties of canalization [114, 103].

In my thesis, I will focus on the second property of canalization, which is robustness, and explore its relation to evolution.

In this chapter, I will describe some background information for my experimental system under study. First I will describe the segmentation system of *Drosophila melanogaster*, in which the experiments in the thesis are based on. Then I will review the basis of eukaryotic gene regulation and a computational model of transcription which has been used in the lab, which will be used for prediction in Chapter 3. A short description of *Drosophila melanogaster even-skipped* stripe 2 enhancer is followed. Finally, I will point out the focus

of my thesis study and lay out the subjects of each of the following chapters.

1.3 *Drosophila* segmentation

Drosophila melanogster embryos, after fertilization, undergo 13 rapid and synchronous nuclear divisions over about 3 hours before the onset of gastrulation. At nuclear division 9, nuclei migrate to the outside, or cortex, of the embryo [43]. This stage is called the syncytial blastoderm: "syncytial" because the nuclei are not yet separated by cell membranes and "blastoderm" because of the morphology of a hollow shell. After the 13th nuclear division, celluarization occurs by membrane invagination. Future body segments are determined during the first three hours [163]. The genes that regulate this segmentation process occur in a hierarchy, and were first discovered by saturation mutagenesis [129, 130, 128, 195]. At the most upstream of this process are genes whose mRNA transcripts are deposited by the mother, called maternal genes. Those maternal mRNAs are translated in the scyncitial blastoderm, forming broad gradients, which define the anterioposterior or dorsal-ventral axis of the embryo [1]. Downstream of maternal genes, zygotic gap genes are then turned on, forming localized broad expression domain. Gap gene mutants usually develop large gap defects in their segments. Pair-rule genes, mutants of which lose structures with a periodicity of two segments, refine their expression over the blastoderm period to a characteristic pattern of seven stripes. Segment polarity genes are the most downstream genes in this hierarchy, in mutants of which segmental structures are deleted with one segment periodicity and replaced by mirrored image duplications. Another class of genes, the two Hox complex genes *Antennapedia-C* and *Bithorax-C*, confer identity on specific segments [1]. All the gap and pair-rule genes, together with their direct maternal regulators are transcription factors (TFs), which bind directly to DNA and regulate expression.

The detailed formulation of the segmentation is highly complex, but we outline here some key players that will of great importance to the work described herein. Prior to fertilization, maternal mRNA of *bicoid (bcd)* is deposited at the anterior tip of the embryo

by the mother. It is indispensable for forming all the anterior structures [35, 131]. Its protein gradient acts as an anterior determinant to activate the expression of downstream gap genes such as *hunchback (hb)* [173, 39]. *bcd* also translationally represses posterior maternal *caudal (cad)* mRNA [151]. *cad* is expressed both maternally and zygotically, and embryos without maternal and zygotic Cad protein exhibit severe shortening and the lost of abdominal segments [110]. *cad* was first identified by searching for genes that contain the homeobox domain [122]. *hb* also has both maternal and zygotic components. Maternal *hb* mRNA is uniformly distributed until nuclear cycle 8, and shortly after form an anterior to posterior gradient and eventually disappear before the blastoderm stage [177]. Zygotic *hb* transcripts are first observed in nuclear cycle 11 [178]. In early cell cycle 14A, *hb* is expressed strongly at the anterior half of the embryo, then with a cap of expression at the posterior pole. Later, expression at the anterior pole and posterior pole recedes. *Krüppel (Kr)* is expressed in the middle region of the embryo, and its expression overlaps with *hb* [55]. Giant (Gt) and Kr have mutually exclusive domains, with Kr expressed between the two Gt expression domains [79]. Similarly, Knirps (Kni) expresses between the two Hb domains. The regulation of the gap gene network has been explored by a dynamical model, which indicates that mutual repression from the gap genes is essential to set up their boundaries [80]. The terminal system is determined independently of the anterior-posterior domain, being set up by a complex of maternal genes whose final output to the segmentation gene system is through the zygotic genes *tailless* and *huckebein* [193].

Early work on gap gene expression and pair-rule gene expression was not quantitative. Later, fluorescent immunostaining made it possible to quantify relative mRNA and protein expression. These quantitative atlas data on gap genes and pair-rule genes includes Flyex (2D protein data) [145, 95] and VirtualEmbryo (3D RNA data) [46]. For some gap genes, the RNA and protein expression patterns differ a lot, such as *hb*. Its mRNA expression in late cycle 14A in the anterior of the embryo are two stripes [46]; however, its protein expression contains a broad high level of expression in the anterior half of the embryo with a stripe-like

pattern at the end of its anterior domain.

1.4 General understanding of transcription in eukaryotes

Transcription in eukaryotes is a complex process that requires specific interactions between proteins and DNAs [133]. As opposed to bacteria and archea, which only have one RNA polymerase [68], at least three RNA polymerases, I, II and III, have been identified in eukaryotes [30]. These three major RNA polymerases synthesize different categories of molecules: RNA pol I for ribosomal RNA, Pol II for messenger RNA, and Pol III for transfer RNA.

Pol II is the key polymerase in eukaryotic transcription, and only Pol II-catalyzed gene transcription will be further described below. Pol II contains a core with 10 subunits and peripheral heterodimer of Rpb4/7, which the later is only required for initiation but not elongation [68]. Pol II can use either DNA or RNA as a template for transcription. AT-rich DNA sites will hamper transcription, and Pol II can be trapped in transcriptional arrest. Transcription initiation starts at the core promoter. Pol II assembles TFIIS as well as TFIID, TFIIE, TFIIF, TFIIH into an initiation complex at the promoter to form a preinitiation complex (PIC). Pol II has a unique C-terminal domain (CTD), which contains 25-52bp heptad repeats YSPTSPS. CTD can recruit regulatory factors for transcription initiation, elongation and termination [68]. Transcription starts when 11-15bp DNA near the TSS are melted, and put in the complex in an open position [68]. After transcribing the first 30bp, Pol II leaves the core promoter and general factors, and enters transcription elongation. There are sequence elements found in the core promoters, such as TATA, BRE, Inr and DPE. In *D. melanogaster*, about 30% of the mRNA genes have TATA in their core promoters. For TATA promoters, TATA binding sites are essential for TATA-binding protein (TBP) to bind in order to initiate transcription. In TATA-less promoters, Inr and DPE motifs are usually found. TFIID is a big complex that consists of TBP and 14 TBP-associated factors (TAFs). Though functionally important, TFIID is not required in all promoters.

The genome is not a simple 1D sequence—DNA is wrapped onto histones to form nu-

cleosomes and nucleosomes are further folded into chromatin with higher structure. Early genome-wide studies found that nucleosomes are depleted near promoters [98]. Promoters of actively transcribed genes are enriched for histone acetylation and methylation. More recently, high resolution nucleosome mapping has revealed more detailed information [84]. +1 nucleosomes (first nucleosome that is downstream of TSS) pulled genomewide show the same relative distance to its TSS, and that signal decreases with +2, +3, ... nucleosomes.

It has been shown that many short pieces of DNA in non-coding regions can regulate mRNA expression [99, 10]. These sequences are called *cis*-regulatory elements (CRE), including enhancers and promoters [197, 133]. In many cases, a minimal DNA element that can drive gene expression pattern as (or a subset of) the endogenous gene expression independent of their position and orientation is called enhancer [166, 8]. Enhancers, which are bound by transcription factors, interact with the components of the mediator complex or TFIID to help recruit RNA Pol II to promoters [133]. Some well-studied enhancers include the *Drosophila even-skipped* stripe 2 enhancer [166], the sea urchin *endo16* enhancer [92], and the *Drosophila shaven-baby* enhancer [172]. There are two conceptual models on how the *cis*-regulatory information can be integrated—enhanceosome model and billboard model [169, 8]. The former emphasizes the cooperativity between enhancer-bound proteins. One example is the response of mammalian β-interferon gene to viral infection [99]. The latter one assumes a loose logic that each transcription factor contributes to the total gene transcription in a nearly additive way, and same output can be achieved by many configurations of transcription factors.

Shadow enhancers, which are enhancers showing overlapping expression patterns for the same gene have been identified in many cases, ensure robust expression pattern [99, 140]. They are also potential targets for evolution to generate novel expression domain [20]. In the post genomic era, genes are usually shown to be separated by insulator elements to ensure non-interference to nearby genes [99, 126], and insulator pairing is understood in more detail in a recent study [50].

Many genomics techniques have been developed to identify protein-DNA interactions, measure transcription or identify enhancers at the genomic level. ChIP-seq [86] is widely used to identify genomic binding profile of transcription factors. Open chromatin regions can be identified using DNase-seq [15], FAIRE-seq or ATAC-seq [16]. RNAseq is widely used to identify genomic RNA levels. More recently, STARR-seq has been developed to find putative enhancers genome-wide [6].

1.5 Short description of a transcription model

In order to understand the details of the transcription process, a quantitative model is needed to explicitly formulate our understanding and to test specific hypotheses. A quantitative model of transcription developed in the lab has been shown to have very good predictive power for enhancer expression that has not been trained using the model [81, 91]. The model has been utilized to model transcription of gene expression in *Drosophila* segmentation gene network, where the major transcription factor concentrations over time have been quantified. The framework of the model is quite general and can be in principal be used for other eukaryotic transcription systems. This model utilizes the CRE sequence, transcription factor binding preference and the transcription factor concentrations as input. Predicted mRNA levels are generated as output, based on a thermodynamic model capturing DNA-protein interactions and protein-protein interactions. The model first considers transcription factor binding equilibrium and calculates the fractional occupancy from each transcription factor. This is usually a reasonable approximation, when the transcription factor binding process occurs much faster than the transcription itself. Position weight matrices (PWM) are used in the model to calculate the binding affinity of each binding site. First the fractional occupancy for all activators are calculated, then short range repression from nearby repressors on the activators is taken into account. Some other known protein-protein interactions are considered in the model, such as Bcd-Bcd cooperation, Bcd-Hb co-activation. Finally, the rate of transcription is calculated to be proportional to the lowering of energy barrier in the

model. The model with free parameters is then fit with real experimental data to find a good fit of parameter set that will then be used for prediction.

1.6 An example of a well-characterized enhancer: *even-skipped* stripe 2 enhancer

Drosophila even-skipped (eve) stripe 2 enhancer is one of the most characterized enhancers. By examining the 5' sequence upstream of *eve*, a 1.7kb DNA that drives a stripe 2 and stripe 7 pattern was first discovered by Goto et al. [63] and Harding et al. [70]. Later, by a more refined deletion study in this region, Small et al. [166] found a minimal DNA sequence of 480bp to be necessary and sufficient to drive stripe 2 expression in *D. mel*, namely MSE2 (minimal stripe element). This short piece of DNA contain *in vitro* footprinted binding sites of the upstream gap genes [168, 166]. Specifically, there are 5 Bcd binding sites, 3 Kr sites, 4 Gt sites and 1 Hb site. Site-directed mutagenesis in the footprinted binding sites demonstrated that Bcd and Hb activate stripe 2 expression, while Gt and Kr are repressors for setting up anterior and posterior boarders respectively [166, 7]. Moreover, Bcd and Hb were found to act synergistically to activate stripe 2 expression [168]. A slightly longer piece of DNA (containing MSE2), which is defined by the sequence between two evolutionary conserved blocks (Block-A, Block-B), is called S2E (stripe 2 element) [105]. Quantification by fluorescent *in situ* hybridization (FISH) of the reporter expression by MSE2 and S2E show that the expression level of S2E is about 5 times higher than MSE2 (from Kenneth Barr, paper in review).

1.7 Canalization and evolution

The robustness property of canalization is to restrict the variation in some essential traits and maintain their reproducibility within a species. On the other hand, when looking at species that are separated by longer evolutionary history, huge diversity of those traits could

occur. This indicates that during evolutionary history, those extant canalized traits must have evolved and in certain lineages these traits are maintained afterwards. This poses a question of how canalization could evolve. Most studies concerning this question came from comparative Evo-devo studies, where the goal is to identify changes in the underlying genetic network in different species which give rise to different end phenotype [153, 179, 96]. However, this does not tell us how the evolutionary processes can lead to this phenomenon. According to Fisher's Fundamental Theorem of Natural Selection [41], selection acts on variance in fitness. If traits are fully canalized, there will be no phenotypic variation, and no evolution could occur. Therefore it is important to understand the detailed evolutionary process that leads to the evolution of canalization.

1.8 Outline of the following chapters

In the following three chapters of my thesis (Chapter 2-4), I will address three questions regarding different aspects of canalization and evolution.

In Chapter 2, I explore the phenotypic variation of a canalized trait at the molecular level within a species. It is useful to understand at the molecular level, to what degree do variations exist that could pass through the sieve of development. Phenotypic variations from natural genetic variations are usually explored for quantitative traits [21, 160, 184], such as body size, where a distribution of phenotype can be observed. Few studies explore the variation for canalized traits since usually very few variations have been seen within a species and usually a conserved pattern is observed across species. For example, there are 4 scutellar bristles in almost all species in the family of Drosophilidae [58, 194]. However, there are wild-type stocks that produce 5 bristles with not-so-rare probability in *Drosophila melanogaster* [149], but a detailed distribution of the scutellar bristle number from natural populations is lacking. If a deviation occurs in the phenotype, that individual needs to be at least viable for its phenotype to be seen and selected upon. In that case, whether there is a buffering mechanism that reduces the phenotypic effect of the initial change is not clear.

I consider the the expression of an essential developmental patterning gene, *even-skipped*. I address this question by observing the expression of this gene in inbred lines derived from natural populations. The results demonstrate that appreciable expression variation exists within natural populations. In addition, I was able to characterize the molecular cause for one particular expression variant.

In Chapter 3, I explore the mechanistic aspect of canalization, specifically, how a conserved pattern of gene expression could be regulated by different *cis*-regulatory configurations from different species. The seven-striped *even-skipped* expression is conserved in higher Diptera. The stripe 2 enhancer element from a distantly-related sepsid species *Themira putris (T. put)* is able to drive almost identical expression pattern as the *Drosophila melanogaster (D. mel)* stripe 2 enhancer when placed into *D. mel*. On the other hand, there is very little sequence conservation between the two enhancers, leading to the question of how *T. put* stripe 2 enhancer is able to form a stripe. In a collaboration with Dr. Ah-Ram Kim, we address this question using a computational model of transcription to help to understand the possible mechanisms. Theoretical analysis indicated in *T. put* stripe 2 enhancer, Caudal binding sites had replaced those Bicoid to activate stripe 2 expression. I test this prediction by constructing embryos that have no Caudal protein and performing a comparative assay of the expression of *even-skipped* stripe 2 driven by enhancers from two species.

In Chapter 4, I seek to address the general question of canalization and evolution by a computational model. Since canalization itself is a dynamical process and so as evolution, it would be difficult to consider both dynamical process at the same time. Therefore, canalization is simplified to a one-step process of genotype to phenotype map, but the most important property—robustness is captured by a parameter in the model which can be explicitly controlled, either by the experimenter or by a genetic locus that can itself evolve. Fitness is determined by both individual's phenotype and environment. I first explored the behaviors of populations with different fixed degrees of robustness under different environmental regimes. I further let robustness to be encoded by a genetic locus, which evolves with

the phenotype to see the interplay of the rate of evolution in robustness compared with the evolution in phenotype.

CHAPTER 2

AN OBSERVATION: NATURAL VARIATION IN EVEN-SKIPPED GENE EXPRESSION

2.1 Abstract

The evolution of canalized traits is a central question in evolutionary biology.[1] Natural variation in highly conserved traits can provide clues about their evolutionary potential. Here we investigate natural variation in a conserved trait—*even-skipped (eve)* expression at the cellular blastoderm stage of embryonic development in *Drosophila melanogaster*. Expression of the pair-rule gene *eve* was quantitatively measured in three inbred lines derived from a natural population of *D. melanogaster*. One line showed marked differences in the spacing, amplitude and timing of formation of the characteristic seven-striped pattern over a fifty-minute period prior to the onset of gastrulation. Stripe 5 amplitude and the width of the interstripe between stripes 4 and 5 were both reduced in this line, while the interstripe distance between stripes 3 and 4 was increased. Engrailed expression in stage 10 embryos revealed a statistically significant increase in the length of parasegment 6 and a decrease in the length of parasegments 8 and 9. These changes are larger than those previously reported between *D. melanogaster* and *D. pseudoobscura*, two species that are thought to have diverged from a common ancestor over 25 million years ago. This line harbors a rare 448bp deletion in the first intron of *knirps (kni)*. This finding suggested that reduced Kni levels caused the deviant *eve* expression, and indeed we observed lower levels of Kni protein at early cycle 14A in L2 compared to the other two lines. A second of the three lines displayed an approximately 20% greater level of expression for all seven *eve* stripes. The three lines are each viable and fertile, and none display a segmentation defect as adults, suggesting that early-acting variation in *eve* expression is ameliorated by developmental buffering mechanisms acting later in development. Canalization of the segmentation pathway

1. This chapter has been published in Developmental Biology, 2015 [85]

may reduce the fitness consequences of genetic variation, thus allowing the persistence of mutations with unexpectedly strong gene expression phenotypes.

In this work, I performed all the experiments as well as image analysis. Michael Ludwig taught me the techniques to stain Engrailed pattern and dissect *Drosophila melanogaster* stage 10 embryos.

2.2 Introduction

In 1942, Conrad Waddington introduced the idea of canalization, which involves the conservation of phenotype in the presence of extensive genetic and environmental variation [185]. The extent to which genetic variation can be buffered is currently unknown. Moreover, it is unclear if a canalized trait is maintained over evolution by phenotypically neutral mutations or a series of small compensatory phenotypic changes [116, 17]. The conservation of gene expression driven by enhancers from highly diverged species [71, 42, 154, 11, 108] shows that functional conservation does not require sequence conservation, but these observations shed little light on the detailed process of evolutionary change which conserved the phenotype. *Trans* changes in the above process complicate matters further, and hence have received little attention. Natural variation acts on extant individuals, and therefore can reveal limits on developmental constraints and provide clues about the evolutionary potential of a conserved trait. Here we investigate natural variation in a conserved trait, the formation of the seven-striped pattern of *even-skipped (eve)* RNA expression at the blastoderm stage of embryonic development in *Drosophila melanogaster*.

eve is one of the most well-characterized genes in *D. melanogaster*. It is essential for the formation of segments [129], and while classified as a pair-rule gene, it has the unique property that null mutations lead to a complete abolition of segments [109]. The segmentation function of *eve* is executed in the blastoderm stage of embryonic development. Transcripts can be reliably detected by cleavage cycle 12 and protein by cleavage cycle 13. After the 13th nuclear division, protein and RNA expression refine from a single broad domain to a characteristic pattern of seven transverse stripes [176, 45]. These dynamic changes in expression are a consequence of the activation of *eve* expression by broadly distributed maternal factors and its repression by more localized domains of zygotic gap gene expression [170, 148]. *eve* is necessary for the correct initiation of the expression of the segment polarity gene *engrailed (en)*, which stably demarcates the future parasegmental borders [59].

The seven-stripe pattern of *eve* before gastrulation is conserved in the suborder Brachyc-

era [32], albeit with different subcellular localizations in different species [18]. Within the genus *Drosophila*, the dynamic pattern of *eve* expression in *D. pseudoobscura* is very similar to that of *D. melanogaster* despite the fact that these species diverged 25-55 million years ago [45, 150].

The individual enhancers of *eve* are very well characterized in terms of function [70, 63, 166, 165, 167, 7, 48, 81, 91]. A 15kb segment of DNA (-6.4kb to +8.6kb of *eve*) can provide a normal segmentation phenotype and rescue an *eve* null mutant to hatching [48, 51]. If the stripe 2 enhancer is deleted from the construct, *eve* stripe 2 is greatly reduced in amplitude with a short parasegment 3 and vestigial En stripe 4, a lethal phenotype [107]. Other stripes have effects on survival that are marked but less severe. Transforming *eve* null flies with the *eve* locus bearing a deletion of the 4+6 enhancer results in viable flies missing certain abdominal segments [51]. The viability of these transformants may be a consequence of residual 4+6 expression driven from outside the classical enhancer, but it is equally possible that non-terminal abdominal segments are not absolutely required for viability.

Previous studies on the intraspecific variation of *eve* expression involved quantitative measurements of stripe placement but not amplitude in three lines of flies with differing egg size. *eve* expression among these lines scaled with egg size, demonstrating that intraspecific egg size variation can be compensated for by expression variation [104]. This point was reinforced by an experiment to artificially select for small or large embryos. In this case, the proportionality of *eve* stripe spacing to embryo length was not preserved, providing crucial evidence that *eve* stripe placement can be variable [121]. In both examples the phenotype under study was egg size, a complex genetic trait, with its consequences for *eve* expression a secondary effect. For this reason, the phenotypic alterations in these lines have not yet been fully mapped to sequence. In contrast, there exist two well characterized small deletions of the *D. melanogaster eve cis*-regulatory region in natural populations, but evidence linking them to phenotypic changes is ambiguous [136].

In this work we characterize intraspecific variation in *eve* expression in terms of quanti-

tative expression level, position, and timing. Our analysis of the three lines, while providing only a glimpse of the full range of *eve* phenotypic variation, demonstrates that significant quantitative variation exists, and can be provisionally assigned to specific changes in sequence.

2.3 Results

2.3.1 The dynamics of eve expression in three D. melanogaster lines

For reasons unrelated to the findings reported here, we examined *eve* expression in three lines from the Drosophila Genetic Reference Panel (DGRP) [111]. These were RAL-437 (denoted as L1 in this work), RAL-502 (L2), and RAL-365 (L5). We unexpectedly found that *eve* expression from L2 differed from that of the other two lines much more strongly than the previously reported *eve* expression differences between *D. melanogaster* and *D. pseudoobscura* [45], motivating the analysis presented here.

Image analysis of *eve* gene expression was carried out from 2D confocal scans of laterally-oriented embryos fluorescently stained for nuclei, Eve protein, and *eve* RNA. These scans were transformed into quantitative data at cellular resolution by image segmentation. The embryos were categorized into eight different time classes (T1-T8), each about 6.5 minutes long during cell cycle 14A [176]. Background staining was removed, and 1D data from the central 10% of dorso-ventral values was used for the detection of quantitative features of the expression pattern [82]. The analysis of *eve* expression in this paper was based on RNA expression. Protein expression resembles that of RNA with a lag of one time class (Figure 2.1; 2.2).

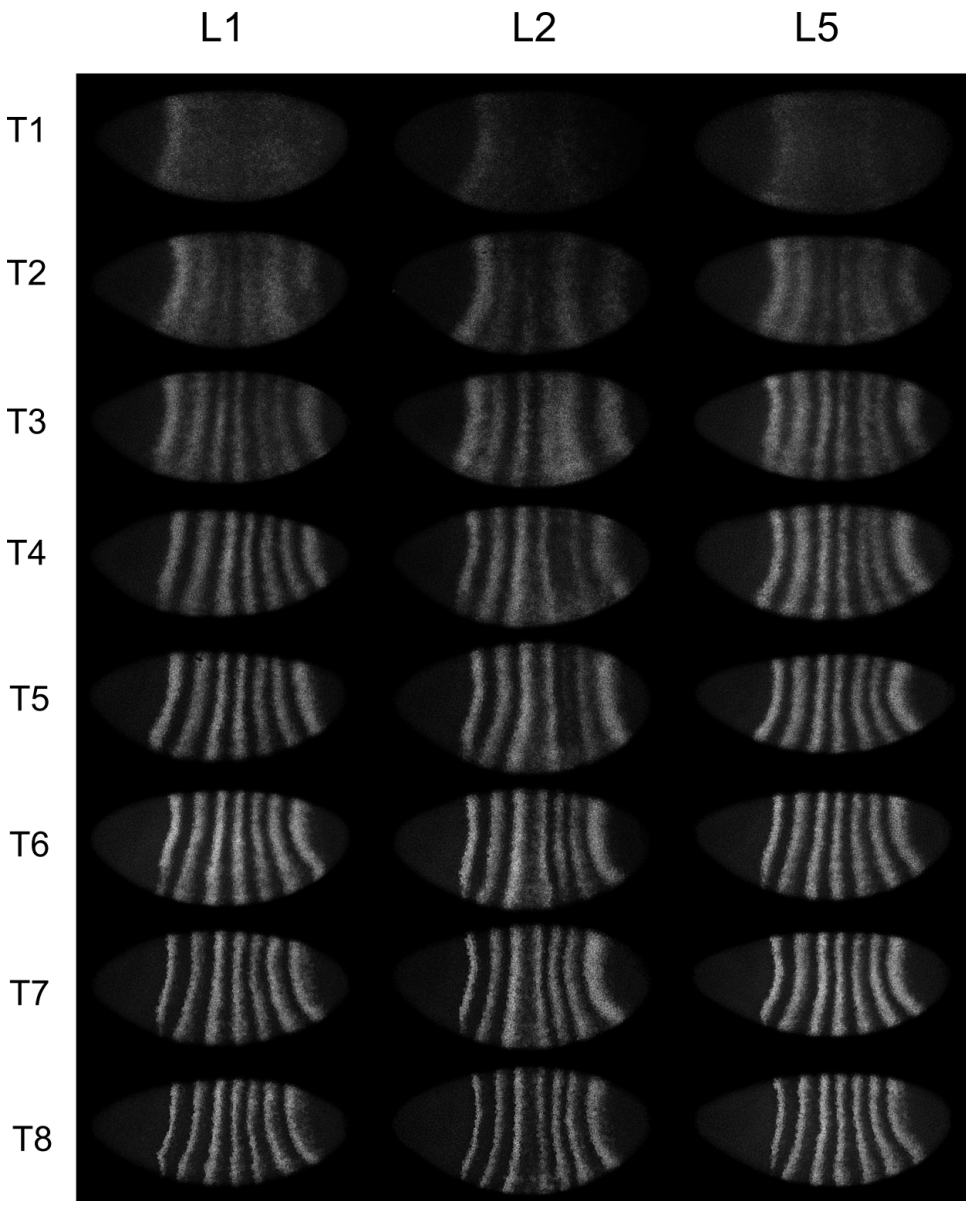


Figure 2.1: ***eve* RNA expression from 3 lines** Fluorescent *in situ* images of *eve* RNA expression of the same embryos chosen in Figure 2.3. Anterior is to the left and dorsal is up.

Figure 2.3 and Figure 2.1 show the dynamics of *eve* RNA expression from typical individual embryos at each time class in the three lines during cell cycle 14A. The expression dynamics of *eve* in L2 are visibly different from the other two lines. For example, at T3 the presumptive stripes 4-6 are a single expression domain in L2 while L1 and L5 have already formed separate stripes. In T4 and T5, when L1 and L5 have formed 7 stripes, there is little

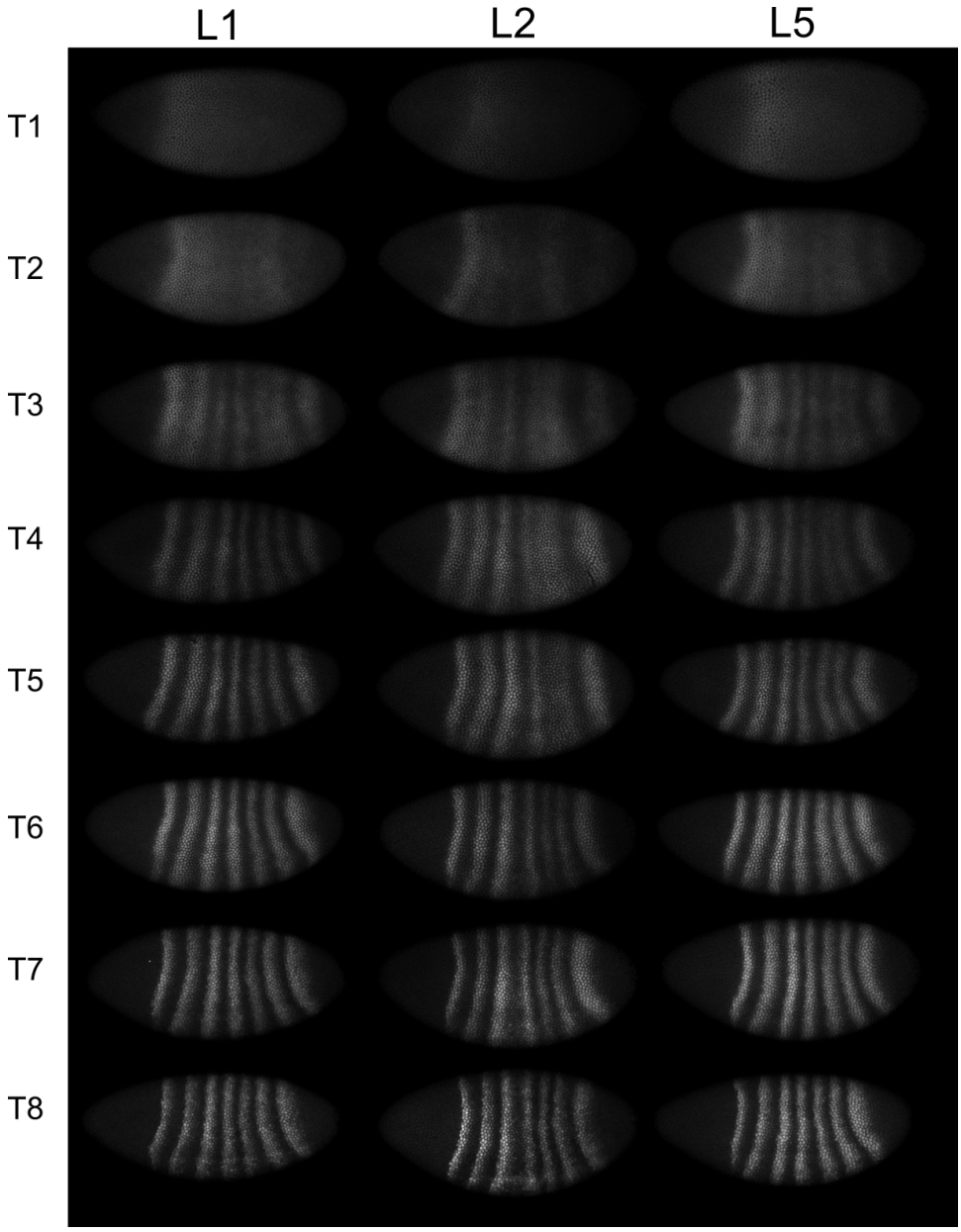


Figure 2.2: ***eve* protein expression from the 3 lines.** Antibody staining for Eve protein of the same embryos chosen in Figure 2.3.

or no stripe 5 *eve* expression in L2. By T6, stripe 5 *eve* expression in L2 rises to the levels seen in other stripes. Moreover, the 3/4 interstripe is wider in L2 than in the other two lines, with stripe 3 expressed at higher levels in L2.

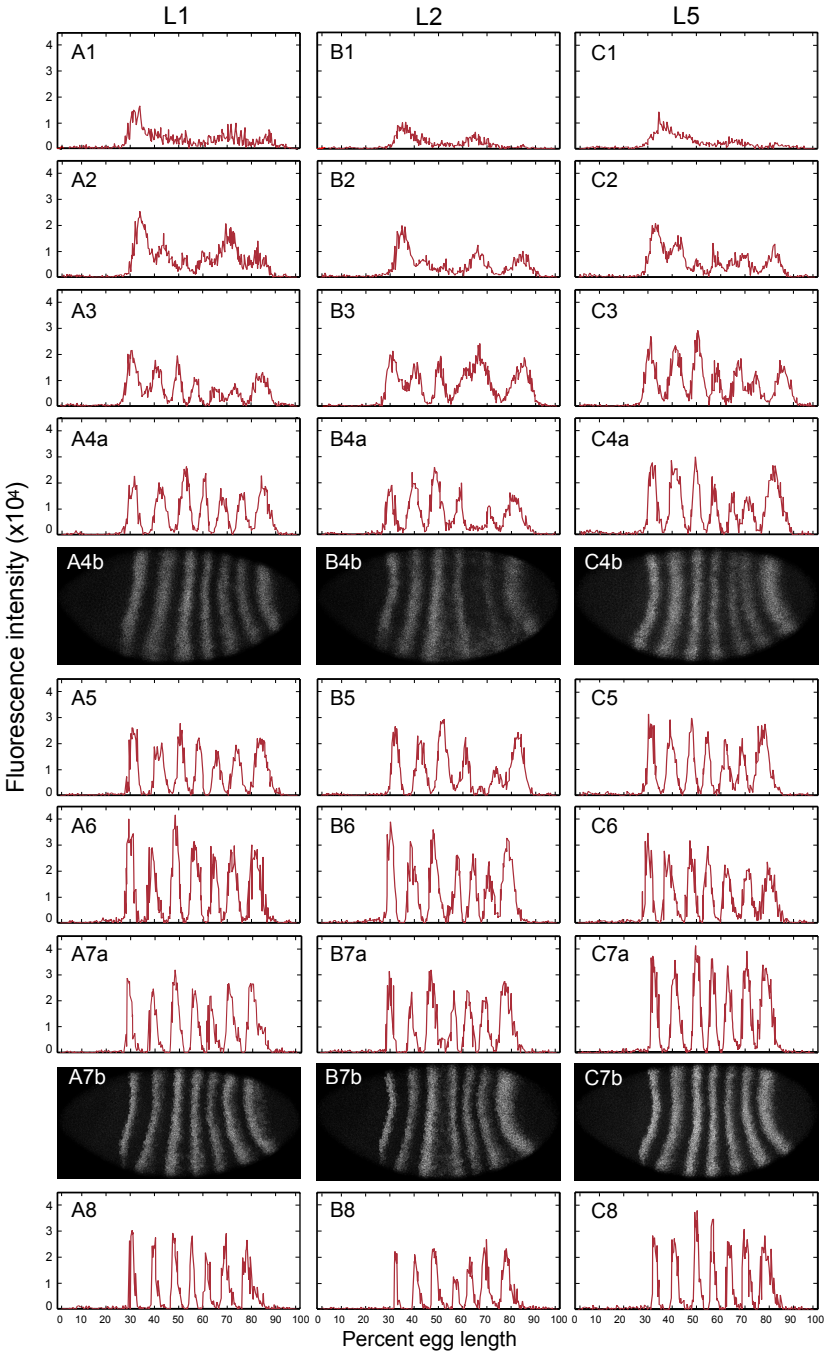


Figure 2.3: **1D *eve* RNA expression from single embryo in cycle 14 from 3 lines.** Each column shows *eve* RNA expression in one line (A: L1; B: L2; C: L5). The Arabic number following the letter represents the corresponding time class, T1 to T8. Each row shows the *eve* 1D RNA expression in its time class; for T4 and T7 embryos, both 1D expression (indicated by 'a') and the corresponding RNA fluorescent *in situ* embryo image (indicated by 'b') is shown. The expression is from 10% stripe from the middle of laterally-oriented embryo after segmentation and background removal.

We made a more precise analysis of differences in the *eve* expression patterns of the three lines by performing feature detection on the expression patterns. We performed the analysis on embryos from T4 to T8, because this is the period when reproducible features of the expression pattern can be characterized [176]. We made pairwise comparisons on stripe morphology features using the Wilcoxon rank sum test. We performed inference as to whether gene expression differed with respect to each class of feature as follows. There are $n = mk$ individual Wilcoxon tests, performed for each class of feature, where m is the number of such features (e.g., there are 6 interstripes) and k is the number of time classes they are measured in, which in this application is always 5 (T4-T8). After the tests, we did Bonferroni correction of n tests given the significance value of 0.05.

We considered three types of features: stripe height, interstripe width and stripe width (Figure 2.4). Stripe height is measured from the peak of the stripe to the minimum of the adjacent interstripe [106, 112]. Stripes 2 through 6 have two such height measurements, corresponding to the two adjacent interstripes; stripes 1 and 7 each has one. The A-P position at a point where expression was midway between stripe peak and interstripe minimum was taken to be location of the stripe border. We generated 6 interstripe distances and 7 stripe width from the 14 border positions.

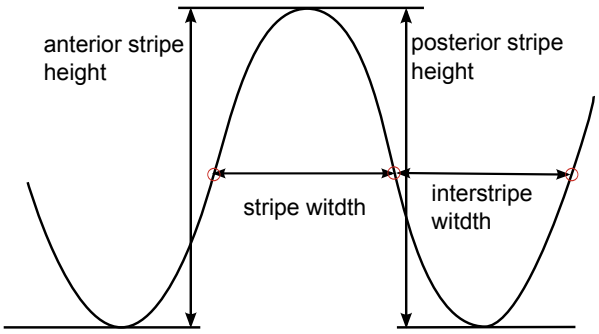


Figure 2.4: **Schematic graph of *eve* stripe features**

We first calculated the absolute stripe height in the three lines (Figure 2.5), and found that L5 had higher expression overall than the other two lines. The Wilcoxon rank sum test for pooled L5 stripe height was significantly different from L1 and L2 (in both cases,

$p < 2.2 \times 10^{-16}$), while L1 and L2 show more modest differences ($p = 0.045$), which supports our observation (boxplot shown in Figure 2.6). We also did pairwise Wilcoxon rank sum tests for the stripe height at each time class and each position (Table 2.1). As expected, most differences arose from the higher overall expression in L5. Specifically, for stripe 2 (positions 2 and 3 in the table), L5 expresses at a significantly higher level than L1 and L2. Notwithstanding this fact, the overall pattern of expression in L5 appeared very similar to L1 by visual inspection (Figure 2.5). In order to compare the features for individual stripes more closely, we normalized stripe height in the three lines by dividing by the mean fluorescence intensity in that line in a given temporal class to detect changes in specific features. After normalizing stripe height, as expected, the most significant differences arose from L2 with respect to L1 and L5 in more specific features (Figure 2.5, Table 2.2).

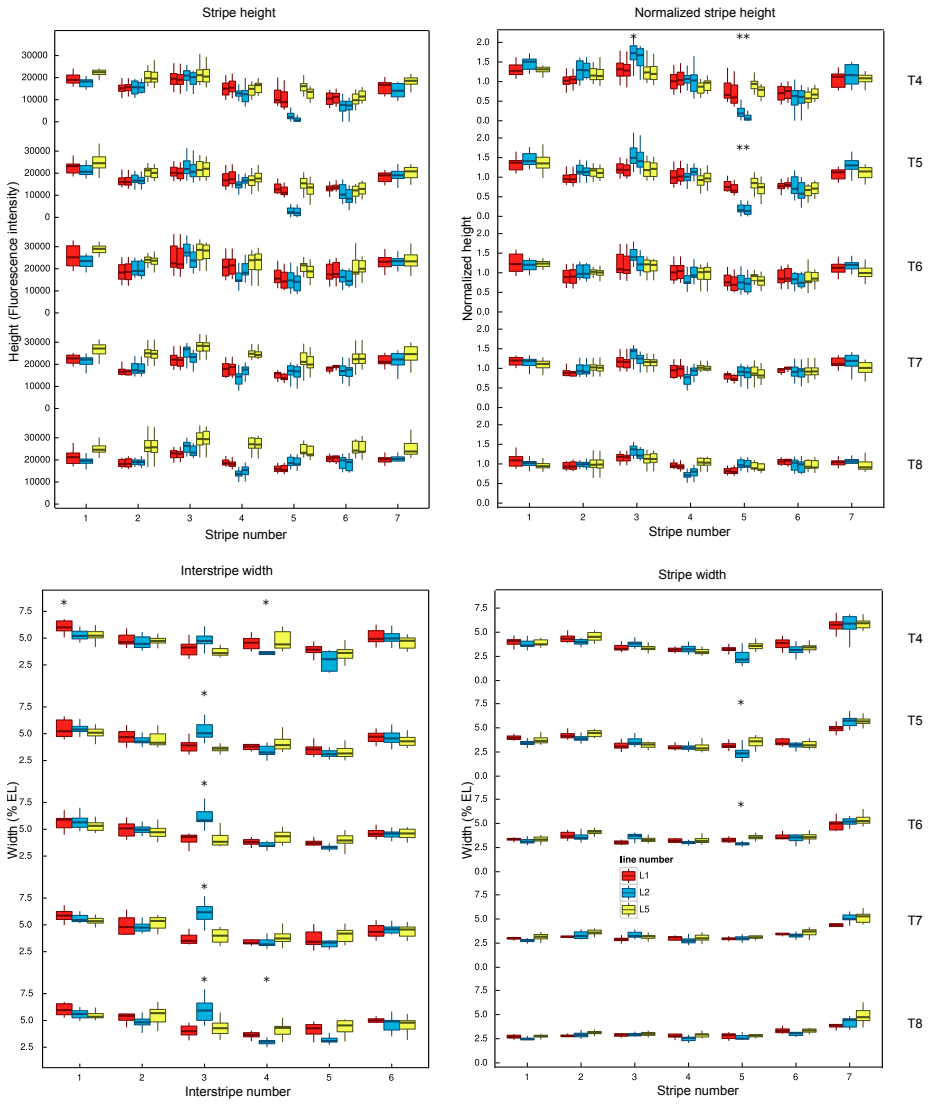


Figure 2.5: ***eve* features for 3 lines** The figure shows stripe height, normalized stripe height, interstripe width, and stripe width for the 3 lines. Red denotes L1, blue L2, and yellow L5. For stripe height, there are two boxes representing each stripe. The one on the left denotes the anterior half of the stripe and the right denotes the posterior half, except the first and last, which have only the posterior half of stripe 1 and anterior half of stripe 7. The asterisks mark the significant difference between one line to the other two lines after Bonferroni correction for each position at each time class (only mark in normalized stripe height, interstripe width and stripe width). Outliers in boxplot are not shown.

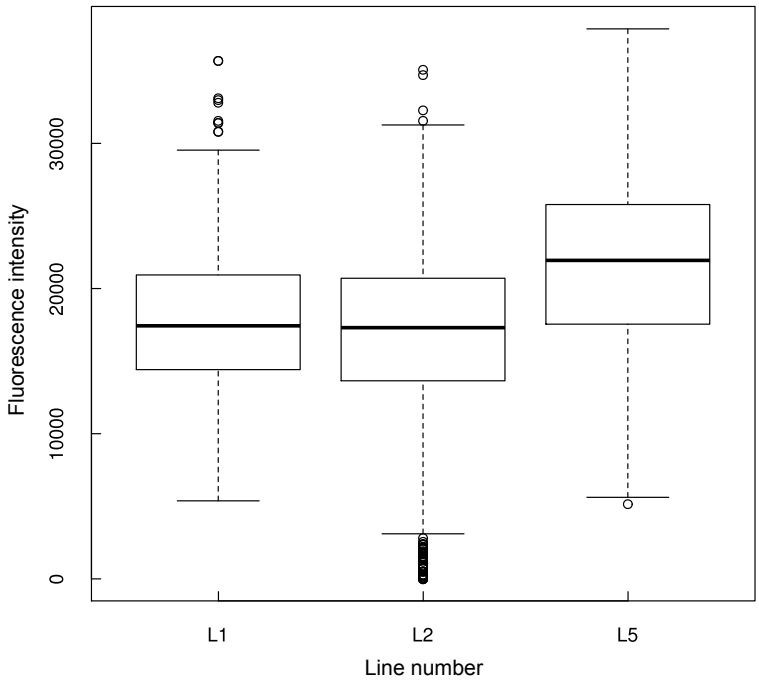


Figure 2.6: **Boxplot of overall stripe height among 3 lines.** Boxplot of absolute stripe height in fluorescence intensity for three lines from T4 to T8. L5 is significantly higher than L1 and L2 in overall stripe height. The boundary of box is first quartile, median, and third quartile. The whiskers extend to the most extreme data point which is no more than 1.5 times the interquartile range from the box. Outliers are marked in circles.

position	time class	p value	lines compared
8	4	2.578690×10^{-8}	L1 L2
8	5	2.312320×10^{-8}	L1 L2
9	4	1.309975×10^{-5}	L1 L2
9	5	2.312320×10^{-8}	L1 L2
2	4	6.549873×10^{-6}	L1 L5
2	5	5.278690×10^{-4}	L1 L5
2	7	7.342234×10^{-5}	L1 L5
3	4	2.654761×10^{-5}	L1 L5
3	7	1.101335×10^{-4}	L1 L5
6	7	4.699030×10^{-4}	L1 L5
6	8	5.292265×10^{-5}	L1 L5
7	8	5.292265×10^{-5}	L1 L5
8	8	2.646133×10^{-5}	L1 L5
9	8	2.646133×10^{-5}	L1 L5
12	8	7.938398×10^{-4}	L1 L5
1	6	2.862549×10^{-4}	L2 L5
1	7	2.010140×10^{-4}	L2 L5
2	7	2.712094×10^{-5}	L2 L5
2	8	7.654884×10^{-4}	L2 L5
3	4	2.756877×10^{-4}	L2 L5
3	5	8.017171×10^{-4}	L2 L5
3	7	2.712094×10^{-5}	L2 L5
5	7	6.611727×10^{-4}	L2 L5
6	6	4.483716×10^{-4}	L2 L5
6	7	3.988374×10^{-7}	L2 L5
6	8	2.835142×10^{-6}	L2 L5
7	7	2.712094×10^{-5}	L2 L5
7	8	2.835142×10^{-6}	L2 L5
8	4	2.578690×10^{-8}	L2 L5
8	5	4.242361×10^{-9}	L2 L5
8	6	5.565384×10^{-4}	L2 L5
8	8	8.505426×10^{-5}	L2 L5
9	4	7.736070×10^{-7}	L2 L5
9	5	8.484722×10^{-9}	L2 L5
10	8	3.402171×10^{-5}	L2 L5
11	4	5.278578×10^{-5}	L2 L5
11	7	2.586460×10^{-4}	L2 L5
11	8	1.984600×10^{-5}	L2 L5
12	8	2.750088×10^{-4}	L2 L5

Table 2.1: **Significance tests for stripe height** The table shows all the significant pairwise rank sum tests for stripe height after Bonferroni correction. The positions are numbered 1 to 12. '1' indicates the posterior half of stripe 1 height, '12' indicates the anterior half of stripe 7 height. '2-11' denote sequentially anterior and posterior of stripes 2-6 height. Time class is as indicated.

position	time class	p value	lines compared
4	4	2.214837×10^{-4}	L1 L2
8	4	3.094428×10^{-7}	L1 L2
8	5	2.312320×10^{-8}	L1 L2
9	4	1.112189×10^{-4}	L1 L2
9	5	2.312320×10^{-8}	L1 L2
4	4	8.731444×10^{-5}	L2 L5
5	4	6.293035×10^{-4}	L2 L5
6	7	5.276619×10^{-4}	L2 L5
6	8	5.386770×10^{-5}	L2 L5
7	8	7.654884×10^{-4}	L2 L5
8	4	2.578690×10^{-8}	L2 L5
8	5	4.242361×10^{-9}	L2 L5
9	4	5.278578×10^{-5}	L2 L5
9	5	2.969653×10^{-8}	L2 L5

Table 2.2: **Significance tests for normalized stripe height** The table shows all the significant pairwise rank sum tests for normalized stripe height after Bonferroni correction. See table 2.1 legend for details about position and time class.

The major result from individual tests of quantitative features (Figure 2.7; Tables 2.2-2.3) was that L2 has a delay of maturation in stripe 5, with an increase of the 3/4 interstripe width and a decrease of the 4/5 interstripe width as a consequence. Specifically, for normalized stripe height, stripe 3 was significantly higher in L2 than L1 or L5 in T4, while stripe 5 was lower in both T4 and T5 but reached similar levels as the other two by T6. The interstripe distance between stripes 3 and 4 in L2 was significantly wider in T5-T8, and that between 4 and 5 is narrower at T4 and T8 compared to the other two lines (Figure 2.5). Overall, the dynamic *eve* pattern in L1 and L5 resembled that seen in the lab stock Oregon R and Canton S [176, 45], with that in L2 deviating from this pattern. These results indicate that *eve* expression differs among the individual lines.

position	time class	p value	lines comed
1	4	2.214837×10^{-4}	L1 L2
3	5	4.448903×10^{-4}	L1 L2
3	6	1.748151×10^{-6}	L1 L2
3	7	4.824896×10^{-5}	L1 L2
3	8	1.058453×10^{-4}	L1 L2
4	4	1.309975×10^{-5}	L1 L2
4	8	5.027652×10^{-4}	L1 L2
1	4	1.186674×10^{-3}	L1 L5
3	4	9.618514×10^{-6}	L2 L5
3	5	4.242361×10^{-9}	L2 L5
3	6	2.136629×10^{-7}	L2 L5
3	7	9.671807×10^{-6}	L2 L5
3	8	5.528527×10^{-4}	L2 L5
4	4	1.761245×10^{-5}	L2 L5
4	5	1.293801×10^{-3}	L2 L5
4	6	4.483716×10^{-4}	L2 L5
4	8	1.275814×10^{-4}	L2 L5
5	6	4.844806×10^{-5}	L2 L5
5	8	3.940848×10^{-4}	L2 L5

Table 2.3: **Significance tests for interstripe width** The table shows all the significant pairwise rank sum tests for interstripe width after Bonferroni correction. The 'position' ranges from 1 to 6, where '1' represents the 1-2 interstripe and so on.

Previous studies show that egg length varies within species [75, 66, 104, 114, 121]. Therefore, we measured embryo size for the three lines from confocal-scanned images. Our results show that embryos from L2 have reduced AP axis length and increased DV axis length compared to the other two lines (Figure 2.8). This change of embryo size in L2 could be related to the change of *eve* expression observed in L2, but we did not further address this possible relationship in this work.

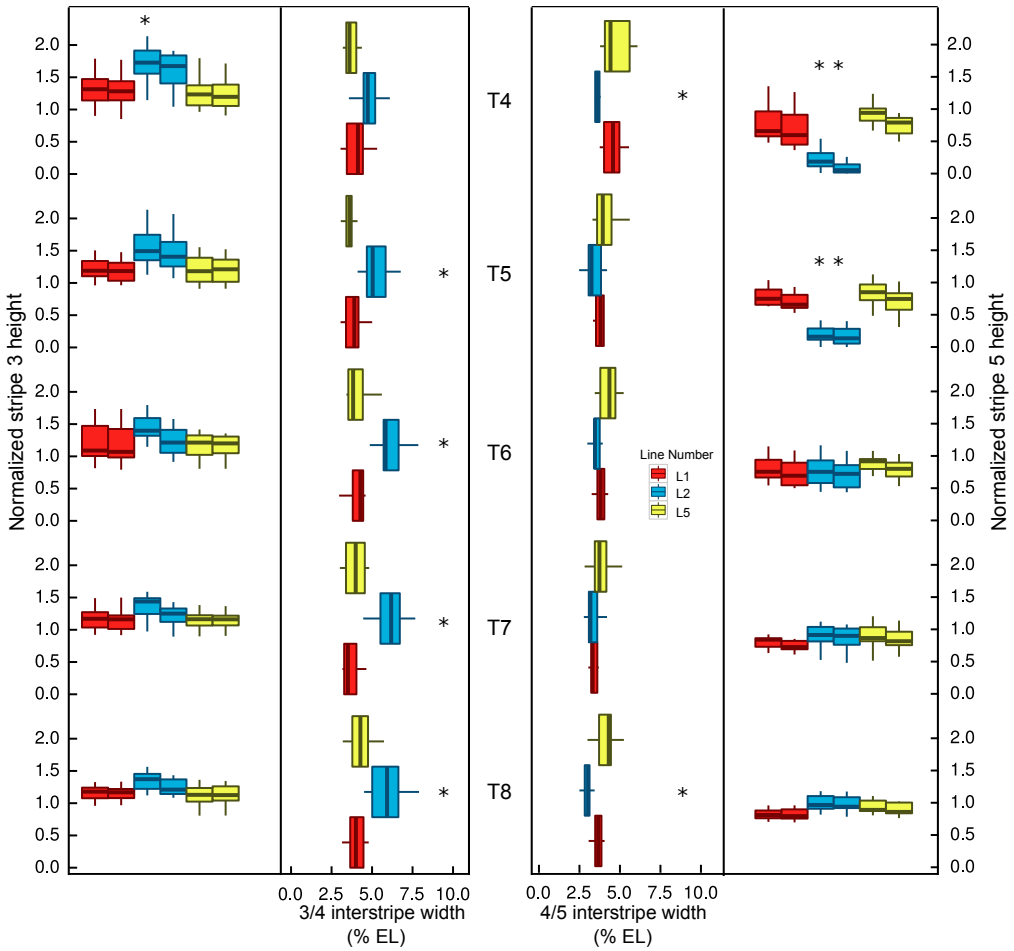


Figure 2.7: **Quantification of normalized stripe height and interstripe width.** This figure is a subset of Figure 2.5. Left to right: normalized stripe 3 height, 3/4 interstripe width, 4/5 interstripe width and normalized stripe 5 height from T4 to T8. If a feature is significantly different in two pairwise tests after Bonferroni correction, an asterisk is marked on top of the shared line for that feature. Two boxes represent each height, the one on the left shows the anterior height of the stripe, while the right shows the posterior height of the stripe. Outliers in boxplot are not shown.

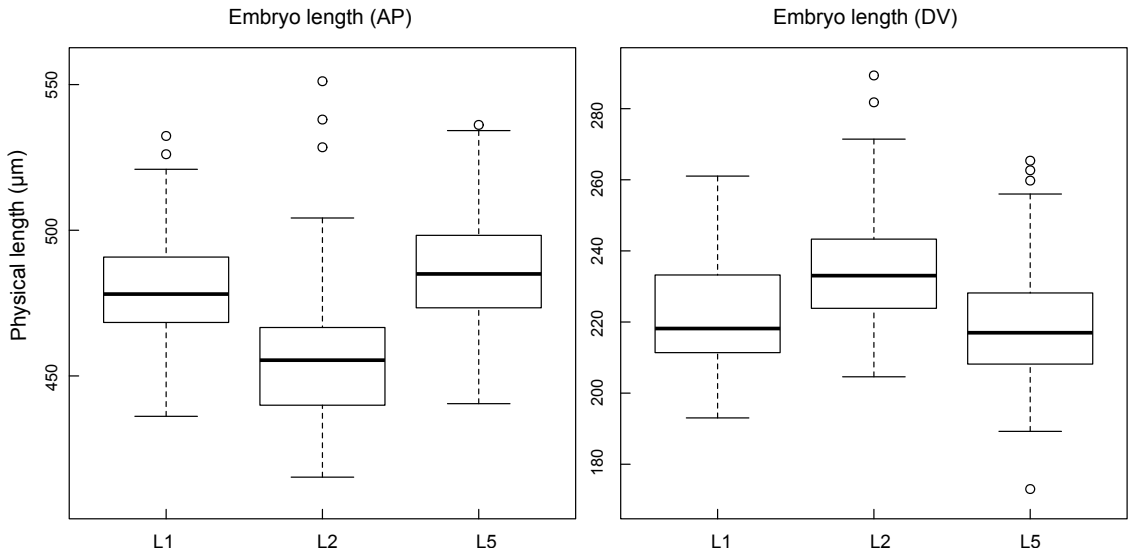


Figure 2.8: **Comparision of embryo size in 3 lines.**

2.3.2 Downstream effect of the eve pattern

We sought to measure the functional effects of these expression differences by assaying expression of *eve*'s functional target, *en*, which is a marker of parasegment boundaries. Stage 10 embryos were stained for En protein (Figure 2.9A), and parasegment length was measured as the distance between the anterior margins of two successive En stripes, normalized to the sum of all such lengths. The results of these measurements (Figure 2.9B) showed that parasegment 6 is significantly longer while parasegments 8 and 9 are shorter in L2 compared to the two other lines. The alternation of *eve* stripe widths in L2 thus has specific functional consequences.

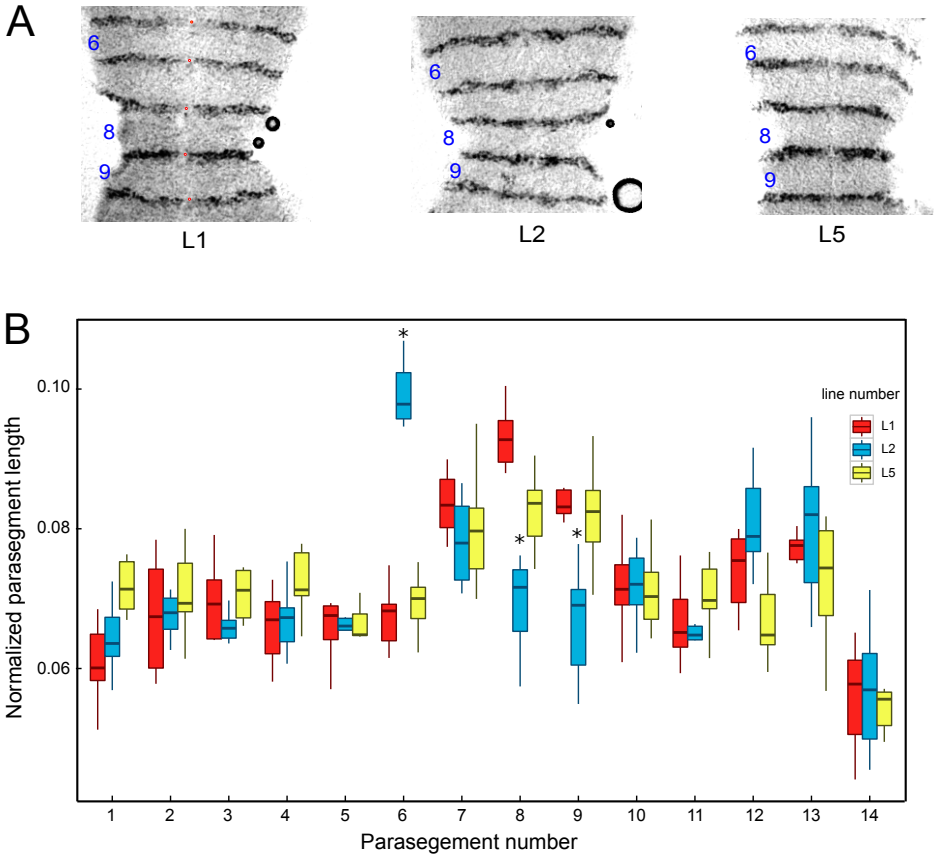


Figure 2.9: **Normalized parasegment lengths**

Figure 2.9: Continued. (A) Images of En expression in parasegments 6-9 of stage 10 dissected embryos from the three lines, as indicated. The red dots mark the anterior of the En stripes demarcating the parasegment, These fiducial marks were used to calculate the length of each parasegment. (B) shows the quantification of the normalized parasegment distance for each line. Significant differences of one line from the other two are marked with an asterisk on the boxplot. Outliers in boxplot are not shown.

2.3.3 Cause of altered expression in L2

We next explored the genetic basis of *eve* expression in L2. Given the marked alterations in stripe 5 expression, we first compared stripe 5 enhancer sequences [54]. Sequence alignment (Figure 2.10) revealed 5 SNPs in stripe 5 enhancer among the three lines, one of which is unique in L5. The remaining four SNPs are shared between L2 and L5 but differ with L1. Thus, there is no candidate SNP in the L2 stripe 5 enhancer to potentially account for the aberrant stripe 5 phenotype. For this reason, we consider it unlikely that the SNPs in the stripe 5 enhancer underlie the observed expression changes in L2.

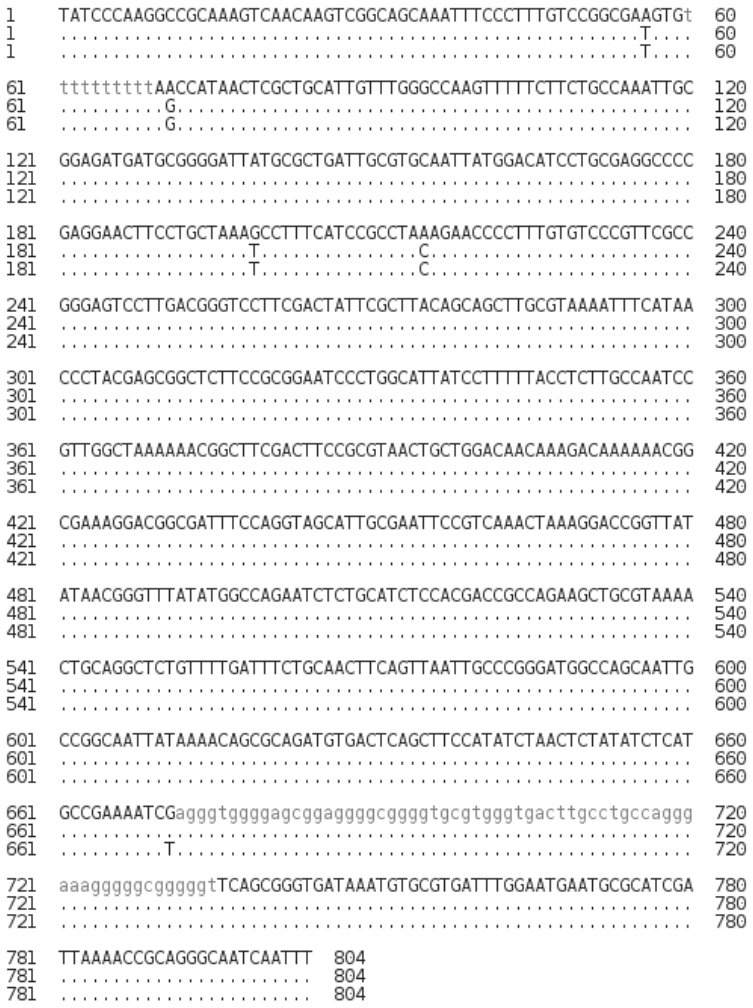


Figure 2.10: **Alignment of stripe 5 enhancer of the three lines** The first row is L1, second is L2 and third is L5.

In the absence of obvious differences in the stripe 5 enhancer, we checked the full *eve* sequence and the coding and flanking 2kb regions of plausible *trans*-regulators among the three lines (Table 2.4) for large deletions (30 bp minimum length) that are unique in L2 [88]. By visualizing the alignment in clustalx [97], we found large deletions or missing data in the DGRP sequences for *eve*, *giant (gt)*, *knirps (kni)* and *runt*. Direct experimental checks by PCR revealed that of the four putative deletions, only the one in *kni* is real (Table 2.5). We used BLAT [89] to map back the sequence and found that the deletion is 448bp in length and lies at the 5' end of the first intron of *kni* (Figure 2.11).

gene name	chromosome	begin	end
bcd	chr3R	2579564	2587188
cad	chr2L	20768731	20785135
ftz	chr3R	2688046	2693966
gt	chrX	2319146	2325000
hairy	chr3L	8666859	8674353
hb	chr3R	4514702	4525544
kni	chr3L	20683430	20689656
Kr	chr2R	21112134	21120917
odd	chr2L	3602224	3608756
paired	chr2L	12080995	12087827
runt	chrX	20559697	20570348
slp1	chr2L	3823675	3829099
slp2	chr2L	3834840	3841185
tll	chr3R	26676037	26682122
eve	chr2R	5860288	5876667

Table 2.4: **Gene list** The table shows the genes and their genomic locations used to detect large deletions or insertions that are present in L2 but absent in L1 and L5. The sequences are from 2kb upstream and downstream of each protein coding gene (RefSeq track from UCSC browser on 9.18.2014), except *eve*, which used 16kb whole-locus sequence.

TTTTTCTTTTAGTAACAAAAAATAAACAACGAGGGTTTTTGGGGCGACTCCTCCCACTTG
GTTTTTTCGCCGTGTACTTAGACACACACGAATATTCCCCTCATGGCACTAGCCGCATTG
TGGGCCACCGCGAAGCCACCGAAGTGGGTGGCCTCTGCACCACCTAGCTTCTGCCGCTTG
ACTTCCGCATGGGCGGTGGTGGTGTTCGGTGAGTGGGCGGCGGAGTGGGAGGAACTGCTG
CTGCTGGTGGTCAGGCACTCGGCCTCCAGTTGGTAGAACTTCCGGCGCGGAGCCACCTCC
ACTTCTTGATCCTCGGAGCCGCTGTCGTTGAAGCTGTGCACGGAGCTCCGCGAGGTCTTC
ATGCTCAGATCCATGGGTCCTTCTTGAGCGGAAACGGTGGGAGGATGCACTGGTACAACG
CTGGTGGTTTCGCCGCCAATGGAGCAAACCGAAACGGGCGACAAGCTCTGCATCTTGGCC
GCAAAACTCAGCGGGCTGCTGCTAGTCGGAGTGGTTGGCGTGGCTCGGATGGCAATGGGA
TGCAGGAGGCTTGCGGACGACTGATGGCTGTGCACGGAGTGAACATCCTCCTCCAGGCAG
ACATCGATGGGCGACTGGCGGGCCGAGGAGGTGGGCTGCACCACTGGTGGCTGGCTGGCG
GGACTAAAGCGGTTCTGCGACTCAACGGAATCCACGCTCTGGCGGTGCTTGTACATCTCC
TGCCGGTAAGCGGCATCGGCAGCAGAAGCCGCTGCAGCGGCAGCACTTGCATGGTAGCCT
GGGAAGAGGAGGTGTGGGGGCAGCTGGAAGGGCGACTGGTGCGGTACACCGCCCATCATG
CTGAAGAACGGCAGGGCGGCGGATGGGTCGGACAGATAGCTGGGATAGCCCAGCAGAGGC
ATATGTGGATGACGCGGCACCTGCTGCTGCTGCTGCTGCTGATGATGGTGGTGCAAATGG
GCGGCCATGTCCCCAAATCCGGGAGTGTGTGTCGAGCCCACCGGGGAAGAGGCCGACGGG
GCACCACCCACCGATACGCCACCCGCTAATGGAGGCGCCTTGCCCGCCGCTGCGGCGGCC
TGTTCGTGCTCCTGCAGCAGACAATGGATCTTGAACCAGTTGGAGCGACGTCCGTAGCGG
GATCCCCCCTTCGACATGCCCACGTTGTAGCACTTCCTCAAGCGGCACGCCTTGCAGGTG
GTGCGGTTCTTCTTGTCGATGATGCACTTGCCCTCGTTCTTGCACTCGCTGATGGTGCTG
ATGTTGTTGTAAGAGCGGCCAAAGAAGGACTGCAATGGCGAAGGGAAGGACGAATGTGGC
TTAGTTATGTGGTATGCAACTGATTAGCAAAAGAAGAGCAGAAATTCATTCAAGAACTTA
AGGCCAATGGTACAACGATTTGAATCGCCGATGGCCTGTTCCTCGAACATTTCTAAAACT
AACCCCTTGGGGGGTTACGTTCTGCTAATTTATTTGTCAATGGATTTCTTGAGACACAAC
TTACCTTGCAGCCCTCGCAGGTGAAGGCGCCAAAATGGAAGCCCGCCGCCGGCTCACCGC
ACACTTTGCATGTCTGGTTCATCTGGAAGATTGTAAAGAAATCCCGCGTTAGTAAGGGTT
TAATCCACTGGTCGAGAGGTATATGTGTAATCCACAAGTAGGCGAACGGCTCTGGACACT
AACCAAGTTGAACACCATTTTGGAATGATATTTGTTGGAATTTTGAATGCTTTTCTCACG
ATCGCTGTGAGGATCTGCTGTATGATCTGTTCGCTTGGAATGCTGCTGAGCAACTGATGC
TGAGCTCTAGGCGGGCCAGGATATATATAGGAAACTGGGAAAAACTAGACAGGTCTCGCT
CAGAACTTACCTTTTTGCTTTTTTTTGTTGGATGGGCACTAGGTTTATGACGACTTTTTA
ATTGCAGCGATTCGCGATCGTTGCTGCACG

Table 2.5: **Confirming L2 sequence expanding deletion in *kni* using PCR**

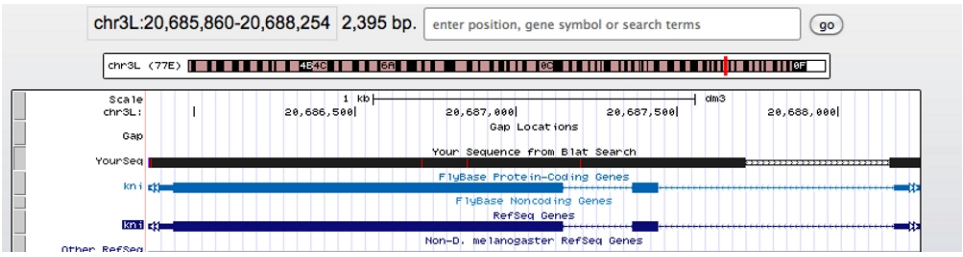


Figure 2.11: **BLAT of L2 showing a deletion in the intron of *kni***

Several lines of evidence suggest that the deletion in the *kni* intron in L2 could be the cause of the altered *eve* expression. First, in *kni* mutants, only *eve* stripes 4-6 are abolished, while stripes 1-3 and 7 are present in T5-T8 at the protein level, although their amplitude is reduced [175, Figure 1]. Second, in *kni* mutant embryos stripe 3 has a larger amplitude than other stripes. Finally, in *kni* heterozygotes, stripe 5 is completely absent until T5 but reaches a level close to wild-type level by the onset of gastrulation [47, Figure 5d], [175, Figure S7]. L2 has altered expression in a subset of the stripe 4-6 region, increased amplitude of stripe 3 expression early, and a reduction in stripe 5 expression before T6, similar to that seen in *kni/+*. The deletion in L2 comprises 30% of the *kni_+1* enhancer, which drives expression in both the head and putative abdomen [158]. The L2 deletion also removes sequences in the *kni* proximal shadow enhancer [140]. These observations led us to predict that Kni expression would be reduced in L2.

We tested our prediction by performing antibody staining of Kni in two independent experiments: L1 versus L2 and L2 versus L5. In each of these experiments, we compared embryos in early (T2-T3), middle (T4-T6), and late (T7-T8) cleavage cycle 14A. In early cleavage cycle 14A Kni expression is consistently lower in L2 than either L1 or L5 (Wilcoxon rank sum test of peak height between $45\% - 80\%$ AP position gives $p = 0.0001$ for L1 versus L2, and $p = 1.8 \times 10^{-6}$ for L2 versus L5; see Figure 2.12). In mid to late cycle 14A, L2 and L5 show indistinguishable levels of Kni expression, but L1 gives slightly higher expression than the other two lines. There exist multiple lines of evidence that the time of formation of *eve* stripes is determined by rising levels of repressive gap gene products

[170, 166, 148, 48, 81, 176], and hence it is extremely likely that the *kni* deletion in L2 is responsible for the *eve* phenotype we observe.

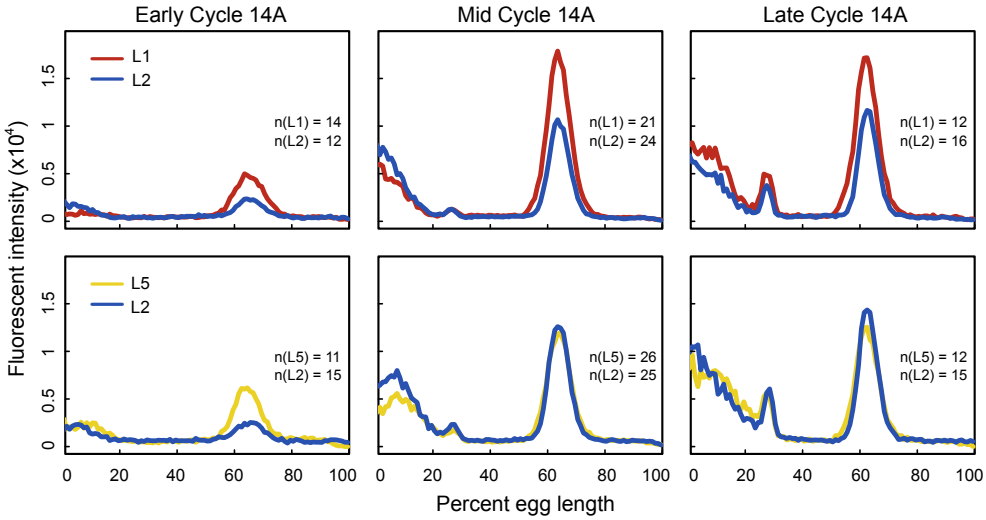


Figure 2.12: **Knirps expression in three lines.** Average 1D Kni protein expression in three lines determined by two antibody staining experiments: L1 versus L2 and L2 versus L5 (upper and lower panel respectively). Expression from the middle 10% of D-V coordinates of laterally oriented embryos is shown. The early cycle 14A temporal class comprises T2 and T3, the mid cycle 14A temporal class comprises T4, T5 and T6, and the late cycle 14A temporal class comprises T7 and T8. Axes are as labeled; the numbers of embryos imaged in each time class are shown.

Only L2 bears this deletion among the DGRP collection of approximately 200 lines. In the core DPGP2 lines [143], which consist of lines originating from African and a few European populations, the full deletion is absent, although two small deletions (7-8 bp each) segregate in these populations. The 448bp deletion in L2 may, therefore, be a low frequency variant with limited geographic range.

2.4 Materials and Methods

2.4.1 Fly culture, embryo collection and fixation

Flies were grown and embryos were collected at 25^o C. For *eve in situ* hybridization and antibody staining, embryos were collected after 1.5 hours, and aged for another 2 hours. Fixation was performed as described [94] with a fixation time of 25 minutes. En antibody staining was performed on embryos aged for 8 hours and fixed for 10 minutes.

2.4.2 Fluorescent in situ hybridization and antibody staining

Fluorescent *in situ* hybridization followed the published protocol [94]. FITC-labeled *eve* antisense RNA probe was generated from p48-X1.4 [109] using SP6 polymerase. Following hybridization, embryos were incubated with rabbit anti-FITC (1:1000) and Guinea Pig (1:1000) anti-Eve antibody [9]. After washing, embryos were incubated in Alexa Fluor 647 Goat Anti-Rabbit IgG (Life Technologies, 1:1000) and Alexa Fluor 555 Goat Anti-Guinea Pig IgG (Life Technologies, 1:1000). Nuclei were stained with DAPI (Life Technologies Cat. No. P36935). En protein is stained with mouse MAb 4D9 antibody (1:3) followed by Goat Anti-Mouse IgG-HRP (1:300) [138]. For Kni staining, embryos were incubated with anti-Kni Guinea Pig antibody (1:1000), followed by Alexa Fluor 555 Goat Anti-Guinea Pig IgG (1:1000) and anti-Eve Rabbit (1:2000) followed by Alexa Fluor 647 Goat Anti-Rabbit IgG (1:1000).

2.4.3 Imaging

Fluorescence data was acquired on a Leica SP5 confocal microscope. Gain was set to produce several saturated pixels after averaging, and offset was set so that approximately half of the background pixels outside the embryo displayed zero intensity and half non-zero intensity. This procedure was carried out with the 5 brightest embryos for each line, and the setting for the brightest line was used to standardize all data collection. Images were taken from

the surface of lateral embryos, using a 20X apo objective (HC PL APO 20x/0.70NA lens [dry]). The images for *eve* are 12 bits per pixel and have 8 line averages. Note that 12 bit images are converted to 16 bit during processing, so that the fluorescent intensity is of 16 bit range. Five z-sections of 0.5 μm each were acquired for laterally-oriented embryos. From these 5 z-sections, the three that best traverse the layer of blastoderm nuclei that is close to cover slide were selected for further processing. A DIC image of the middle of the embryo was acquired to visualize membrane invagination to aid in time classification. Images for Kni staining were acquired under the same microscope, with 8 bit depth (converted to 16 bit showing in graph), two z-sections which are 1μm apart.

Images of En staining were acquired using Zeiss Axioskop light microscope.

2.4.4 Time classification

Time classification of younger (T1-T3) embryos was performed by inspection of protein expression patterns as described [176]. Older (T4-T8) embryos were categorized by the degree of cell membrane invagination. We noticed the membrane at the ventral side is more mature than the dorsal side, and we based on our membrane invagination on the ventral side membrane. The time classification by membrane invagination is quite robust, while the early patterns are not as reliable. For T1-T3 embryos, we noticed that characteristic stages of the RNA pattern occur about one time class earlier than the same stage of the protein pattern. This fact was useful in the classification of the altered patterns seen in some lines. T1 embryos could always be unambiguously classified by protein pattern. The RNA patterns of these embryos were indicative of the expected T2 protein pattern of a particular line. Continuation of this procedure provided consistent and reliable temporal classification for all lines.

2.4.5 Feature detection

For observations of *eve* and Kni expression, segmentation, background removal, and the extraction of data from the central 10% of dorso-ventral positional values were performed as described [176]. Embryos stained for Kni were registered using Eve stripes as described [176]. For investigations of *eve* expression, cubic splines were used to detect extrema and borders of the stripes for time class T4 to T8, when all seven *eve* stripes are detectable [106]. A border is defined to be the position with expression midway between that of a stripe peak and interstripe. Each stripe, except 1 and 7, has an anterior and posterior height defined by the difference between expression at the stripe peak and the adjacent interstripe. The width of a stripe was taken to be the distance between its borders in percent egg length (% EL), with 0% at the anterior pole. A small number of embryos that were detected to have fewer or more than the full set of 13 extrema were included in the analysis after manual addition or removal of extraneous features. The number of embryos collected for each time class is summarized in Table 2.6.

	T4	T5	T6	T7	T8
L1_original	24	12	14	10	13
L1_OK_orientation	15	12	10	9	12
L1_final	15	12	10	9	8
L2_original	19	19	14	13	15
L2_OK_orientation	18	19	13	13	13
L2_final	14	18	13	13	11
L5_original	16	14	15	18	12
L5_OK_orientation	15	14	15	14	11
L5_final	15	14	15	14	11

Table 2.6: **Number of embryos collected for each line** There are three numbers indicated for each line. The label ending with "original" denotes the total number of embryos imaged at the confocal microscope. The label ending with "OK_orientation" denotes those embryos which have approximately lateral orientation and were analyzed further. The label ending with "final" denotes the number of embryos used in feature detection for extrema and borders. Only T4 to T8 embryos were used. This number includes manually recovered embryos that have more than or less than the normal number of extrema.

2.4.6 Measurement of En parasegments

Points were marked at the midline of the anterior of each En stripe using imageJ [157] (See the red dots in Figure 2.9A), and parasegment lengths were measured between these points. The normalized parasegment length was taken to be the length of each parasegment divided by the sum of length from all the parasegments in that embryo.

2.4.7 Measurement of embryo size

In the data processing step, we generate an outline of the entire embryo with the length in pixels of the major and minor axes. These are converted to microns using metadata from confocal imaging.

2.4.8 PCR primers

For validation of deletions for

eve: forward:`GGTCGCTTGGAGAAGGAGTT` reverse: `CACACCCAGTCCGGTATAGC`

gt: forward: `TGGCACAAGAGCTCGATGTT` reverse: `TAAATGCAGGGGGTTCCGAC`

kni: forward: `CCTAAGTGTGAGCGAGCACA` reverse: `TGAGAAAACGTGCAGCAACG`

runt: forward:`ACATGACCTACGGCTATGCG` reverse: `TAATTTTTGCCCGCTTGCCG`

2.5 Discussion

The three lines displayed three distinct *eve* stripe expression patterns. This variability was large enough to depart from the evolutionarily conserved reference pattern. The overall higher expression of *eve* in L5 could be caused by genetic variants controlling overall *eve* expression. In *cis*, these might involve variants in the *eve* autoregularoty elements or some unknown chromatin regulation mechanism at the whole-locus level. In *trans*, the cause could be the consequence of expression level changes of other genes in the network. In the absence of a clear genomic signature, further experimentation will be required to elucidate the cause of increased expression. In addition to the higher overall expression of *eve* in L5, we observed in L2 the absence of *eve* stripe 5 formation at T4, the time at which stripe 5 typically becomes visible [176]. Although the seven-stripe pattern is restored in the blastoderm embryo by T6, alternations in interstripe distances persist well into gastrulation. These differences are consistent from embryo to embryo, and are supported by relatively conservative statistical methods that do not assume normality and take into account multiple pairwise tests of significance. The unusual early *eve* expression pattern has measurable consequences later in the pattern formation cascade, as evidenced by a significant change in the spacing of En stripes. The altered pattern of *eve* expression in L2 strongly resembles that seen in a *kni* heterozygote [175], and we found that this line bears a 448 bp deletion in its *kni* intron, which removes part of the *kni_+1* enhancer. This mutation appears to be rare in *D. melanogaster*, as it is not present in any other DGRP line, nor in any DPGP2 line, the majority of which are African lines representing the ancestral range of that species.

We predicted this deletion would lead to reduced *kni* expression, and verified this prediction by quantitative measurement of Kni expression. These observations strongly suggest the contribution of the *trans* background to *eve* stripe variation.

Odd numbered *en* stripes are expressed at the anterior margin of *eve* stripes, while the even numbered stripes are expressed at the anterior margins of *ftz* stripes [77]. The expansion of the 3/4 *eve* interstripe leads to an increase of the distance between *en* stripes 6 and 7.

The position of *en* stripe 6 relative to 5 and 7 depends on *ftz* expression [77, 49] and may indicate that *ftz* stripes are also altered in L2. Similarly, the reduction in the length of the *eve* 4/5 early interstripe causes a decrease in the total distance between the *eve*-dependent *en* stripes 7 and 9. The effect of the reduction of stripe 5 expression on parasegment 9 (the distance between *en* stripes 9 and 10) is reminiscent of that seen in enhancer deletions, but much milder [51]. This effect is also consistent with observations of *ftz* expression in *kni* heterozygotes [22, Figure 3F]. In these embryos, *ftz* stripes 4 and 5 are markedly closer, consistent with reduction in distance between *en* stripes 8 and 10, while the distance between *ftz* stripes 3 and 4 increases, implying an increase of distance between *en* stripes 6 and 8.

In *kni* and *Kr* heterozygotes, altered gap and pair-rule expression are largely corrected by gastrulation [175]. In L2 we did observe an En expression phenotype attributable to *eve* misexpression, but adults do not display any obvious phenotypic defect, and the line is viable and fertile. In *bicoid* copy number variants, the expansion of the head region in flies with extra *bicoid* copies is corrected by differential cell apoptosis [19, 125]. Multiple layers of canalization may buffer misexpression of early genes in the segmentation pathway, making variations more common than previously suspected based on the strong evolutionarily conservation of the pathway. Canalization theory predicts such variation in strongly buffered pathways [119].

2.5.1 Evolutionary Implications

The most surprising result reported here is that phenotypic variation in *eve* expression within *D. melanogaster* is of larger magnitude than previously-reported inter-species variation between *D. melanogaster* and *D. pseudoobsecura*. We believe *eve* expression is under relatively strong stabilizing selection, and therefore it is generally conserved across relatively large evolutionary distance. However, the large variation within species indicates that conservation is not complete, and that there is a large potential for expression to change. There must be limits to natural variation, however, and the fact that the *kni* deletion is a rare variant

suggests that it is deleterious in nature. The similarity of the L2 *eve* phenotype to that of *kni* heterozygote mutant embryos, which are viable, suggests that the permissive threshold may be around half of wild-type expression. In natural populations, the *kni* deletion variant will be present almost entirely in heterozygotes because it is rare, and we expect that the phenotypic effect in heterozygotes will be subtler than the homozygous effects measured in L2.

Examples of *cis* and *trans* coevolution are pervasive [12, 61, 107], and binding sites turnover, rearrangement, and change of spacing within enhancers is rampant for phenotypically conserved traits. Both may be manifestations of compensatory evolution. There are also examples of developmental system drift [183], in which the underlying genetic network changes over evolutionary time while maintaining a specific phenotype. However, no detailed mechanism has been proposed for that phenomenon. Our work suggests that canalization of mutant phenotypes compresses the width of their phenotypic distribution. This reduces the consequences of genetic variance, allowing otherwise dramatic mutations, such as the partial loss of the kni_+1 enhancer, to segregate in natural populations for a long enough time without being eliminated by selection to allow other compensatory mutations to occur. A computational model of enhancer evolution under stabilizing selection [17] predicts rapid turnover of binding sites, and in many cases deleterious mutations persist before positive selection takes place. Buffering mechanisms will greatly reduce the fitness cost of a deleterious mutation, and therefore greatly increase the rate of compensatory evolution [36]. Canalization, compensatory evolution, and developmental system drift may be multiple consequences of stabilizing selection acting to maintain the fidelity of developmental processes.

Variation in spatio-temporal patterns of gene expression in conserved developmental systems, as described in this work, is likely to be widespread as a consequence of buffering mechanisms that mitigate their developmental consequences on adult fitness. We note that strongly deleterious mutations, though rare in their individual frequency, are numerically abundant (though absent in the inbred DGRP lines). The *kni* deletion mutant may be an

example of a non-lethal deleterious allele exhibiting a dramatic molecular phenotype. Deleterious mutations with molecular phenotypes are expected to be much more common in natural populations than in the DGRP lines. We might also anticipate finding some of these variants to be common in populations if they co-occur with compensatory mutations. The continued development of models of transcription, such as [91], will be useful for predicting the effects of natural variation on gene expression.

CHAPTER 3

TESTING A MECHANISM PREDICTED BY A THEORETICAL MODEL: EFFECT OF *CAUDAL* ON SEPSID *EVE* STRIPE 2 ENHANCER

3.1 Abstract

Drosophila melanogaster (D. mel) eve stripe 2 enhancer (S2E) is one of the most well-characterized enhancers. Bicoid (Bcd) and Hunchback (Hb) act as activators to drive *D. mel* stripe 2 expression. Sepsid flies, which are about twice divergent as the distance from *D. mel* to the most distantly-related *Drosophila* species, whose stripe 2 enhancers are able to drive stereotypical expression in *D. mel*. Despite of the conservation in expression pattern, little sequence conservation is found in the enhancer sequences. To understand how these divergent enhancer sequences are able to drive similar expression patterns, we sought to use a computational model of transcription in the lab to understand possible mechanisms. For one sepsid species, *Themira putris (T. put)*, it is predicted its stripe 2 enhancer is activated by Caudal (Cad) by the computational model, rather than Bcd when driving expression in *D. mel*. We then set out to test this hypothesis by comparing the expression of the S2E from both *D. mel* and *T. put* in both wild-type and *cad* mutant backgrounds. Enzymatic *in situ* hybridization shows *T. put* S2E reduces expression much more than *D. mel* S2E, which provides support to the *in silico* prediction.

This chapter is in collaboration with Dr. Ah-Ram Kim. Ah-Ram initiated the project by doing all the computational model predictions (Figure 3.2, Figure 3.3, Figure 3.4). I performed all the experiments to test the role of Cad in activating the two enhancers that will be described below.

3.2 Introduction

Drosophila melanogaster (D. mel) eve stripe 2 enhancer is one of the most well-characterized enhancers [166, 7]. Bicoid (Bcd) and Hunchback (Hb) act as activators to drive *D. mel* stripe 2 expression [166]. Mutagenesis in footprinted binding sites within *eve* stripe 2 enhancer show that stripe 2 expression is sensitive to mutations in binding sites, such that mutations that delete Bcd-1 and Hb-3 sites abolish *eve* stripe 2 expression [166]. On the other hand, *eve* stripe 2 enhancers from other *Drosophila* species have many sequence changes (binding sites turnover); however, they can still drive a nearly identical expression pattern in *D. mel* [107]. Spatio-temporal expression of endogenous *eve* expression is also similar, if not identical, across these species [107]. Despite these obvious functional constraints on *eve* expression, binding sites turnover between species suggests that enhancers may utilize some "loose grammar" that generate similar outputs from different inputs.

Sepsid flies, which are about twice divergent as the distance from *D. mel* to the most distantly-related *Drosophila* species, whose stripe 2 enhancers are able to drive stereotypical expression in *D. mel* [71]. Sequence conservation falls off rapidly when moving away in the phylogenetic tree. 92% of binding sites in *D. mel eve* stripe 2 enhancer are conserved in the monophylogenetic *melanogaster* species subgroup that include eight close relatives of *D. mel*, whereas only 29% are conserved to *virilis-repleta* clade. Astonishingly, at most only 5% conservation of binding sites is observed for pair-wise comparison of *D. mel* versus sepsid flies [71]. Another way to visualize the conservation/divergence of sequences is through alignment of binding sites. Figure 3.1 shows the alignment of binding sites within *eve* stripe 2 enhancer from 12 species, including 6 *Drosophila* species and 6 sepsid species [71]. There are several clusters of conserved binding sites in *Drosophila* species, including three Bcd-Kr clusters and one Slp1-Gt cluster. However, this pattern is not observed in sepsid species. The only conservation in both *Drosophila* and sepsid species is one Bcd-Kr cluster at 5'. Another conserved cluster in sepsid species but not in *Drosophila* species is at 3', containing several Slp1 binding sites. Note that this alignment ignores the actual length of enhancers

by ignoring the space between binding sites, thereby maximizing possible alignment between the binding sites. For example, the stripe 2 enhancer from *D. mel* is around 500bp, while the stripe 2 enhancer in *T. put* is around 1700bp. The total alignment in Figure 3.1 is 3600bp. For this reason, even the few "aligned" binding sites may not be orthologous (and thus conserved) but instead may have been lost and subsequently reestabilishes in the enhancer.

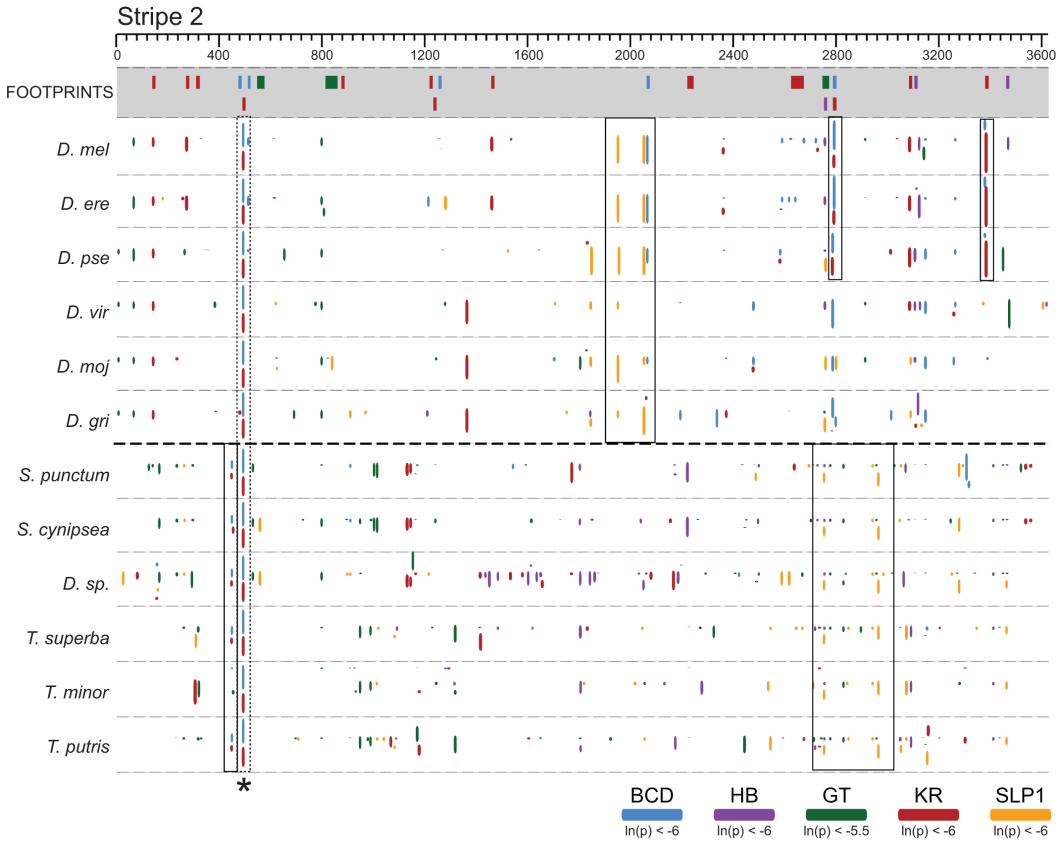


Figure 3.1: **Binding sites alignment in S2E across species [originally from [71](Fig. 6)]** Predicted binding site alignment of five known transcription factors in S2E from 12 species, including 6 *Drosophila* species and 6 sepsid species. The height of each binding sites indicates predicted binding affinity. The top gray panel shows all the experimentally validated footprinted binding sites. Conserved binding sites are shown in solid boxes, except one Bcd-Kr binding site cluster which has been conserved for all these species are marked with a dashed box.

3.3 Results

3.3.1 In silico model predicts the role of Cad in activating sepsid eve stripe 2 enhancer

In order to understand how sepsid S2E enhancers are able to drive stripe 2 expression in *D. mel*, a computational model of transcription was used to examine the detailed mechanism predicted in the model [91] (shortly described in Section 1.5). This specific model is trained on gene expression from several enhancer-reporter constructs from *D. mel* and shows good predictive power for enhancer expression from other *Drosophila* species in *D. mel* background. Using the fit parameters from this model, the analysis calculates binding sites that contribute to top 80% to activation in stripe 2 expression at a specific anteroposterior position. This analysis is referred to as Highly active sites Of TFBSs (HOT) analysis. HOT analysis was first performed on *D. mel* S2E. Figure 3.2A shows all the predicted binding sites in *D. mel eve* stripe 2 enhancer (S2E) and Figure 3.2B shows the binding sites after HOT analysis at 38% egg length (anterior border for *eve* stripe 2 expression; 0% being anterior tip). The number of binding sites reduces drastically with this analysis, and the binding sites contributing to activation are clustered, which is consistent with our current knowledge of enhancers. Clustering of binding site has been used successfully as a criterion in algorithms to predict enhancers in noncoding DNA [62, 115, 13]. According to the model, the activation of stripe 2 *eve* comes from two Hb binding site clusters—Hb-3 and Hb-2 (Figure 3.2C, D); the contribution of Bcd to activation in S2E is mostly through Bcd co-activation of Hb-3.

Core binding sites structures from 4 *Drosophila* species and 2 sepsid species were analyzed the same way using HOT analysis for comparison, including *D. mel, Drosophila yakuba (D. yak), Drosophila erecta (D. ere), Drosophila pseudoobsecura (D. pse), Sepsis cynipsea (S. cyn)* and *Themira putris (T. put)*. HOT binding sites are stacked for comparison by aligning to the putative Hb-3 site, which is present in all of the species. Only HOT activators are shown in Figure 3.3. There is a clear pattern for S2E in the four *Drosophila* species:

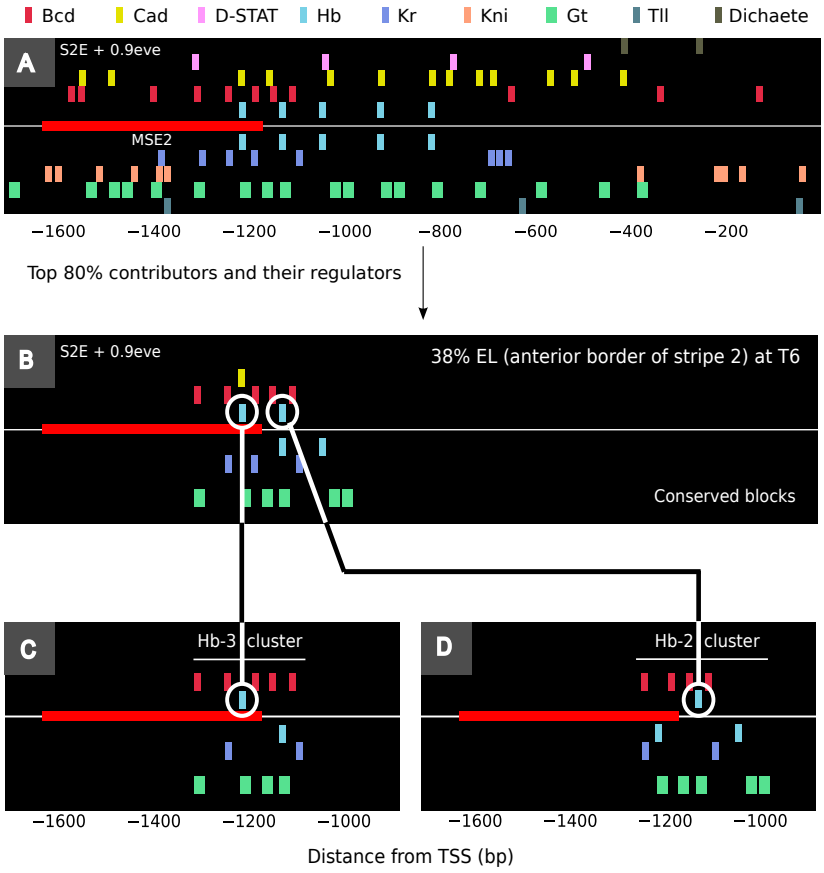


Figure 3.2: **HOT analysis for *D. mel* S2E [modified from Ah-Ram Kim thesis]** A. All of the predicted binding sites for *D. mel* S2E using a model of transcription described in [91]. DNA sequence is represented as a solid line in the middle. Red box shows the position of minimal stripe 2 enhancer (MSE2). Binding sites above this line are activators, while below the line are repressors. B. Binding sites after HOT analysis at 38% egg length (0% being the anterior tip). B-D. Two Hb functional clusters that contributes to $> 80\%$ activation of *D. mel eve* stripe 2 expression using HOT analysis. C. Hb-3 cluster. D. Hb-2 cluster.

conservation of the two Hb clusters that are essential in activating *D. mel*, despite changes in spacing and gain/loss of Bcd binding sites.

Strikingly, this pattern changes dramatically in the sepsids. Bcd binding site number reduces dramatically in *T. put*, with only one HOT Bcd binding site compared to 5 in *D. mel*. Since Bcd is essential in activating stripe 2 expression in *D. mel*, this poses a question of how sepsid stripe 2 enhancers are activated.

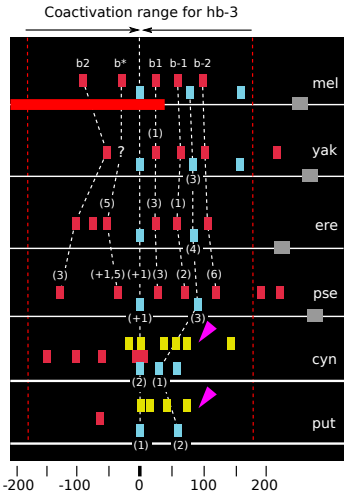


Figure 3.3: **Core binding site alignment for activators in S2E from different species. [from Ah-Ram Kim thesis]** Comparison of binding sites after HOT analysis for 4 *Drosophila* species (*D. mel*, *D. yak*, *D. ere* and *D. pse*)and 2 sepsid species (*S. cyn*, *T. put*). Binding sites are stacked by aligning to the putative Hb-3 binding site. The color notation of binding sites are the same as in Figure 3.2. Arrowhead show the gain of Cad binding sites after HOT analysis in *S. cyn* and *T. put* stripe 2 enhancers.

HOT analysis identified Cad binding sites, present in sepsid but not *Drosophila* sequences, in the same region in both *S. cyn* and *T. put* (Figure 3.3). Specifically, the model uses Cad as co-activator for Hb to activate stripe 2 expression in the sepsid S2E. This striking finding that Cad binding contributes to stripe 2 expression in the sepsid S2E is the basis for the hypothesis, tested here, that Cad has an activating role in the sepsid stripe 2 enhancer.

We first conducted *in silico* experiments to ask whether Cad is required to activate sepsid stripe 2 expression. Figure 3.4 shows the predicted expression driven by *D. mel, S. cyn and T. put* S2E in wild-type, removing all Cad binding sites or removing all Bcd binding sites. In the wild-type background, the model correctly predicts stripe 2 expression for all three enhancers, with some posterior shift for *S. cyn* and *T. put* enhancers, similar to a previous experimental report [71]. Cad is not an activator for *D. mel* S2E, and as expected, *D. mel* S2E drives stripe 2 expression in the absence of all Cad binding sites. One previous experiment also supports this view: endogenous *eve* stripe 2 is strongly expressed in $cad^{mat+zyg}$ mutants [132]. However, both *S. cyn* and *T. put* abolish expression when

Cad binding sites are removed, confirming the HOT analysis that the sepsid S2E is Cad-dependent.

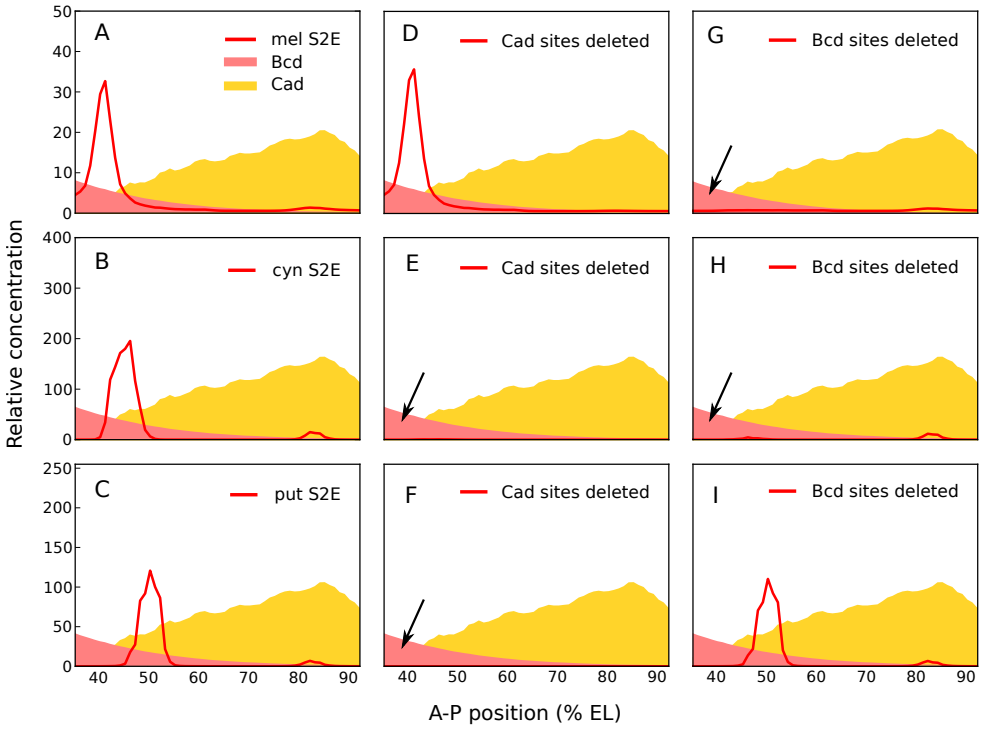


Figure 3.4: ***In silico* experiments for *D. mel, S. cyn, T. put* S2E in wild type, removing all Cad binding sites or removing all Bcd binding sites. [from Ah-Ram Kim thesis]** A-C: Wild-type expression of *D. mel, S. cyn, T. put* S2E predicted using a model of transcription [91]. D-F: Predicted *D. mel, S. cyn, T. put* S2E expression when all Cad binding sites are deleted. G-H: Predicted *D. mel, S. cyn, T. put* S2E expression when all Bcd binding sites are deleted. Arrows show the cases in which stripe 2 expression is abolished.

It is well known that *D. mel* S2E will not express a stripe when Bcd binding sites are removed *in vivo* [166], which our *in silico* model confirms. For, *S. cyn*, the model also predicts loss of expression when Bcd binding sites. For *T. put*, however, stripe 2 expression is not abolished in with all Bcd binding sites removed.

The *in silico* experiments, therefore indicate that sepsid S2E may utilize Cad as an activator or coactivator, this despite the fact that Cad is a posteriorly-expressed transcription factor in these experiments. Moreover, the analysis suggests that *S. cyn*, which requires both

Bcd and Cad binding sites, may be a functional (and evolutionary) intermediate between *D. mel*, which only requires Bcd binding sites, and *T. put* which only requires Cad binding sites. The same result holds when setting the *trans*-background of either Cad or Bcd expression to be zero.

3.3.2 Experimentally testing the role of Cad in activating T. put S2E

To test whether Cad has a role in activating stripe 2 expression for sepsid S2E, I compared S2E expression from *D. mel* and *T. put* in the wild-type and *cad* background since those two species gave the most contrasting behaviors predicted by the *in silico* model. Previous experiments show that anterior gap gene expression remains mostly intact in *cad* embryos, and there is an endogenous Eve stripe 2 [132], indicating that Cad has little impact on the anterior gap gene network. We chose to test this hypothesis by constructing *cad* null embryos rather than by removing Cad binding sites for two reasons: (1) identifying all Cad binding sites in the S2E for removal can not be guaranteed given the questionable quality of the Cad PWM quality, and (2) mutating Cad binding sites may create or disrupt overlapping TF binding sites. The model predicts, specifically, that the *T. put* S2E but not the *D. mel* S2E will drive transgene expression in a *cad*-dependent manner.

Cad is expressed both maternally and zygotically, and it is homozygous lethal for both maternal and zygotic *cad* null expression [132]. We therefore use Dominant Female Sterile (DFS) technique [27] to generate mosaic female flies which are zygotically heterozygous *cad*, and only eggs with homozygous *cad* are produced because germ cells which are heterozygous for the dominant female sterile allele arrest early in oogenesis. These female flies are crossed with homozygous males carrying an enhancer-reporter construct (either *P*(S2E*mel-lac*Z) or *P*(S2E*put*-CFP)). The embryos produced by this cross are all maternally homozygous *cad*. Zyogtically, half the embryos are heterozygous *cad* and half are homozygous *cad*, and all will carry one copy of the enhancer-reporter construct.

Because there is one copy of an enhancer-reporter construct in the embryos in a *cad*

background, control embryos are generated so that it also contains one copy of enhancer-reporter construct in the wild-type background. Control embryos were generated by crossing *P*{S2E*put*-CFP}attP2 virgins with *P*{S2E*mel-lac*Z}attP2 males. This cross has also been made in reserve direction, but no obvious difference was observed. To test staining conditions, embryos from crosses generated from either direction have been used. Only the described crossing direction is used for the final experiment, carried out at the same time with the *cad* embryos.

3.3.3 Expression comparison of D. mel S2E, T. put S2E in wild-type and cad mutant embryos

The final cross to get reporters to *cad* backgrounds was set up in bottles and then moved to a cage. Embryos were collected and fixed the same way as described in Chapter 2. Control embryos used in the final experiment were also fixed around the same time as *cad* mutant embryos to minimize effect of fixation on the result. Both fluorescence and enzymatic *in situ* hybridization were performed to visualize reporter expression.

3.3.4 Fluorescent in situ hybridization

Figure 3.5 shows the reporter expression from *D. mel* S2E and *T. put* S2E in wild-type and *cad* backgrounds. Both *D. mel* S2E and *T. put* S2E strongly reduce their expression in the *cad* backgrounds. I will discuss the possibilities of the cause that reduce *D. mel* S2E in the discussion section. Since *T. put* S2E CFP embryos have overall higher background, it is not possible to conclude that there is a signal from *T. put* S2E CFP in *cad* mutants. *D. mel* S2E *cad* embryos, in contrast, display weak sparkling stripe pattern in *cad*.

Since we only observed the expression pattern from one copy of the enhancer-reporter construct, the expression in the wild-type background is already low. This suggests FISH may not be the ideal method to be used. Other more sensitive *in situ* hybridization methods

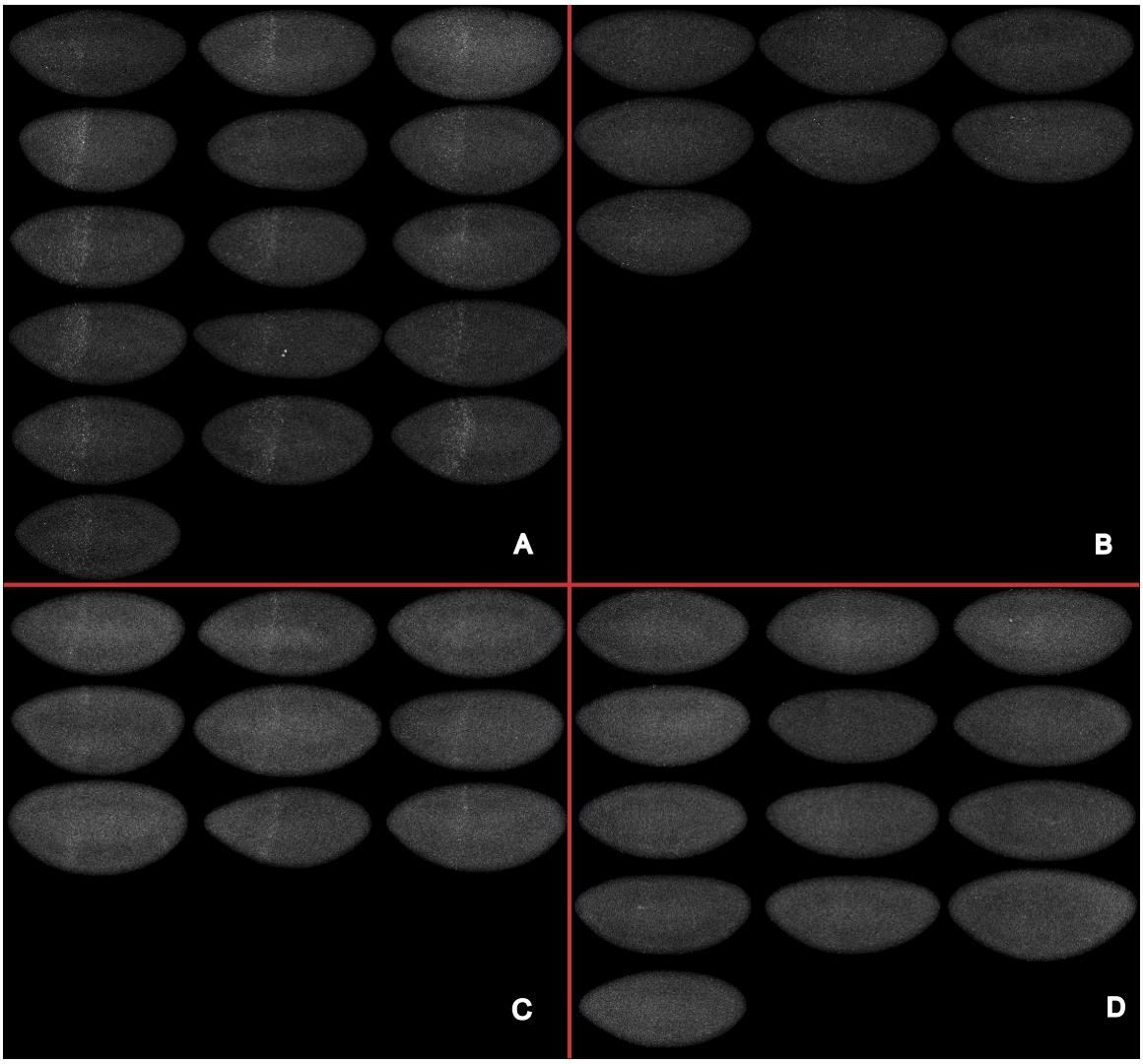


Figure 3.5: **Comparison of *D. mel* S2E versus *T. put* S2E in WT and *cad* mutant background using FISH.** A, B show the *D. mel* S2E expression in wild-type versus *cad* mutant background. C, D show the *T. put* S2E expression in wild-type versus *cad* mutant background. There is higher background for *T. put* S2E. Note the embryos in D are a mixture of zygotic heterozygous *cad* and homozygous *cad*.

could lead to a more definitive resolution of the hypothesis.

Also note that *T. put* S2E *cad* embryos imaged here are a mixture of heterozygous *cad* and homozygous *cad*, since embryos are mounted *en masse* on the slide, rather than individually, so no genotyping was performed to distinguish the two in this case. We should be cautious about drawing conclusions from these images.

3.3.5 Enzymatic in situ hybridization

We then use a more sensitive method to detect reporter expression, using enzymatic *in situ* hybridization. Specially, Alkaline Phosphatase (AP) is used. Since AP reactions have an arbitrary time of development, we first test the concentrations of probes and antibody using wild-type background embryos. After optimal conditions were determined, the experiment was performed with both wild-type embryos and $cad^{mat+zyg}$ embryos. Development of wild-type embryos and $cad^{mat+zyg}$ embryos are carried out at the same time, side by side. In each case, the sample of embryos are divided in half, to allow two stopping criteria. The first stopping criterion ("regular"), reactions were stopped when wild-type embryos showed evidence of saturation. For *P*(S2E*mel-lac*Z), reactions were stopped at 1:10h. For *P*(S2E*put*-CFP), reactions were stopped at 1:58h. The second stopping criterion aims to over-develop the reaction ("over-developed"), to allow examination of whether a low signal is present if no expression is observed in the regular development. For over-development, about twice the time of the regular development is used. For *P*(S2E*mel-lac*Z), reactions were stopped at 2:18h. For *P*(S2E*put*-CFP), reactions were stopped at 3:55h.

Genotyping was performed for *T. put* S2E *cad* embryos to distinguish zygotic heterozygous *cad* with homozygous *cad* [106]. All the embryos in enzymatic *in situ* hybridization shown below in *cad* were confirmed $cad^{mat+zyg}$.

Figure 3.6 shows typical embryos from enzymatic *in situ* hybridization experiment using regular stopping criterion. It is clear that although *D. mel* S2E has reduced expression in *cad* background, there is a stripe visible. However, for *T. put* S2E, which was stopped at the time point when the wild-type embryos first appear saturated, there is no visible expression. But if the reactions are over-developed, the *T. put* S2E in *cad* background now begin to show weak expression in several cases (Figure 3.7). These are consistent with the FISH experiment, but provide new insights about the results. We can take from the results that, qualitatively, *T. put* S2E reduces expression more in *cad* embryos than *D. mel* S2E does. Figure 3.7 shows all embryos from the enzymatic *in situ* hybridization for both the control

and the *cad* mutant background for both stopping times. Overall, the experimental results support the hypothesis that Cad co-activates Hb to drive stripe 2 expression in *T. put* S2E.

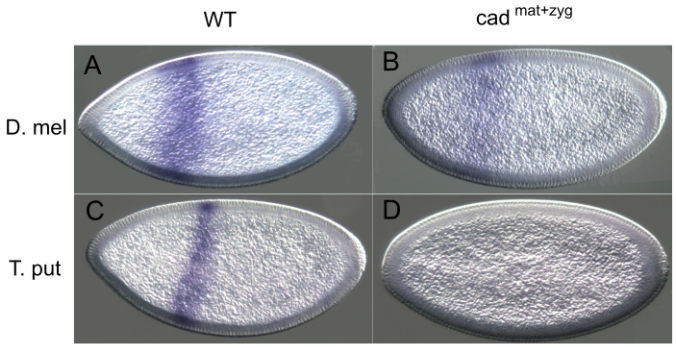


Figure 3.6: **Comparison of *D. mel* S2E versus *T. put* S2E in WT and *cad* mutant background.** A, B show the *D. mel* S2E expression in wild-type versus *cad* mutant backgrounds. Reporter expression is reduced in *cad* mutant for *D. mel* S2E, but there is still observable weak expression pattern of stripe 2. C, D show the *T. put* S2E expression in wild-type versus *cad* mutant backgrounds. The reporter is below detection threshold in *cad* mutant background for *T. put* S2E.

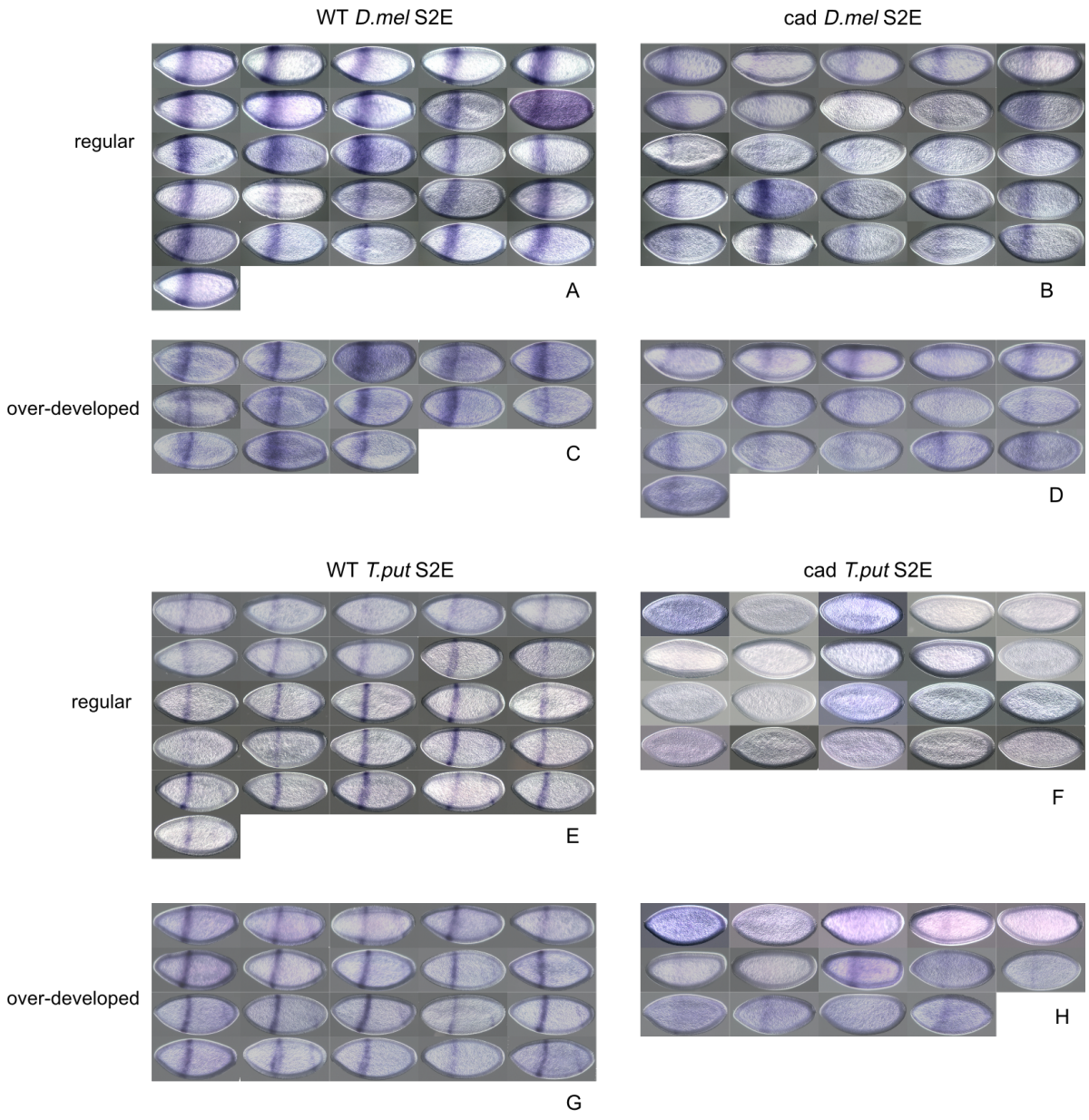


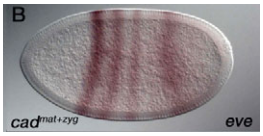
Figure 3.7: **Comparison of reporter expression from *D. mel* S2E and *T. put* S2E in wild-type and *cad* using enzymatic *in situ* hybridization (with AP development).** A-D show the *lac*Z reporter expression driven by *D. mel* S2E in wild-type and *cad* background. E-H show the CFP reporter expression driven by *T. put* S2E in wild-type and *cad* background. A, B, E, F reactions were stopped when the wild-type embryos are ready. ("regular") C, D, G, H reactions were over developed for twice the time of its corresponding "regular" time. For the "regular" reactions, reduced stripe 2 expression is observed in *cad* background from *D. mel* S2E; but no clear expression is observed from *T. put* S2E in *cad*.

Figure 3.7: Continued. For the "over-developed" reactions, weak expression from *T. put* S2E in *cad* was observed in some embryos.

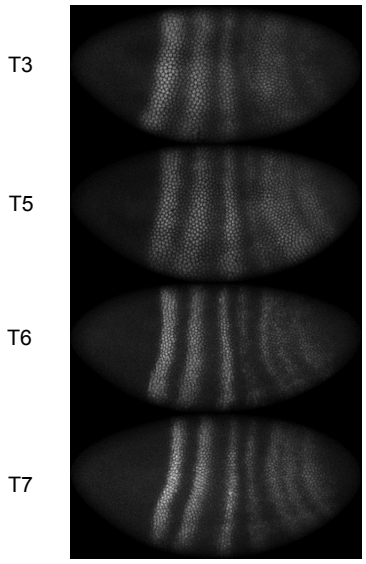
3.3.6 Eve expression pattern in cad embryos

We discovered dynamics in Eve expression in $cad^{mat+zyg}$ embryos that have not been described previously. Previous Eve expression in $cad^{mat+zyg}$ embryos was from enzymatic antibody staining exclusively, where anterior stripes are present but posterior stripes are missing [132] (stripe 1, 2, 3, 5 are present, Figure 3.8(a)). We discovered novel dynamics in Eve expression in $cad^{mat+zyg}$ embryos in our study, which are double stained with *D. mel* S2E using FISH. According to our Eve fluorescent antibody staining, the previous report on Eve expression in $cad^{mat+zyg}$ is an earlier developmental time point (similar to T3 embryos in our FISH experiments, Figure 3.8(b)). Late Eve pattern show weak posterior stripes (T6, T7 in Figure 3.8(b)), compared to anterior stripes, which could be driven by *eve* auto-regulation. Moreover, anterior stripes are shifted posteriorly in *cad*, and the interstripe distance between stripe 2 and 3 is larger (Figure 3.8(b)).

In contrast to the highly expressed endogenous Eve stripe 2 in $cad^{mat+zyg}$ embryos, the *D. mel* S2E expression is largely reduced (Figure 3.5B). This indicates that some sequences outside of *D. mel* S2E are responsible for the expression of stripe 2 in *cad* embryos.



(a) Eve expression in $cad^{mat+zyg}$, from [132]



(b) Eve expression in $cad^{mat+zyg}$ using FISH, from my study

Figure 3.8: **Comparison of Eve protein expression obtained using FISH in our experiment compared to previous data.**

3.3.7 No evidence shows T. put S2E is posteriorly shifted

I conducted enzymatic *in situ* hybridization with double stain of Eve antibody staining initially. But the color difference in enzymatic reactions are not distinct enough in the experiment to prevent background from Eve antibody staining to bleed into the reporter stain. For this reason, I performed experiments with *in situ* hybridization alone to explore reporter expression, as described above (Section 3.3.5). However, in the initial double stain experiment with Eve, we found *T. put* S2E expression is not posteriorly shifted as previously described [71]. Here we show several embryos with double stain in Figure 3.9. *T. put* S2E expression overlaps with *D. mel* endogenous Eve expression, and especially for the anterior nuclei in stripe 2. We also noticed that the reporter expression does not overlap 100% with the endogenous *D. mel eve* stripe 2, with occasionally some slight differences on the D-V axis. However, there is no evidence that *T. put* S2E is posteriorly shifted in *D. mel*. Such discrepancy in the previous result may be due to measurement errors due to comparing with embryos with slightly different stages, because previous conclusion was drawn from

measurements from *in situ* hybridization without double stain with Eve.

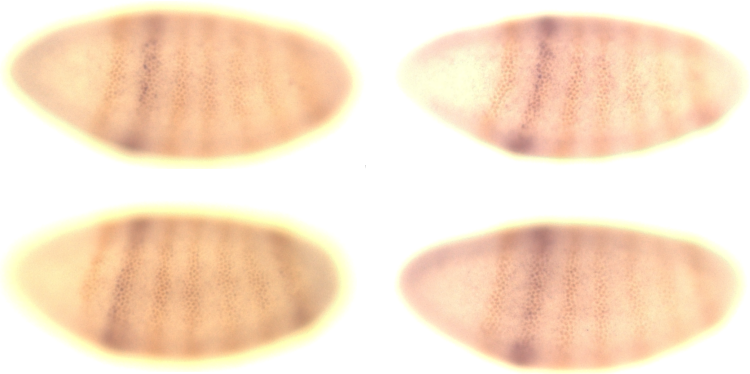


Figure 3.9: **Example embryos show double stain of *T. put* enhancer and Eve expression in control embryos (wild-type background).** Blue shows *T. put* S2E CFP reporter expression (AP development), and brown is the Eve expression (HRP development). *T. put* S2E expression overlaps with *D. mel* endogenous Eve expression quite well, and especially overlaps with the anterior nuclei in stripe 2, with occasionally some slight differences on the D-V axis.

3.3.8 Other immunofluorescence imaging attempts

Other immunofluorescence imaging methods have been attempted. One was Tyramide Signal Amplification (TSA) and the other was single molecule FISH (smFISH).

TSA method is a combination of enzymatic and fluorescence *in situ* hybridization. The initial development is enzymatic (anti-DIG HRP), followed by direct or indirect fluorescence labeling. It is sensitive to low signals and is semi-quantitative. The expression pattern is shown for control embryos with reporters (Figure 3.10A). However, some false positive signals are also observed for negative control embryos (Figure 3.10B), where no reporter is present. I further suppressed the endogenous H_2O_2 in case the false negative signals come from the embryo. By doing that, the signals from negative control reduced dramatically (Figure 3.10C), but there are still some residual signals compared to the negative control with no probe and no HRP (Figure 3.10D). No follow-up work was conducted to trouble-shoot this method; the enzymatic *in situ* hybridization method was pursued instead.

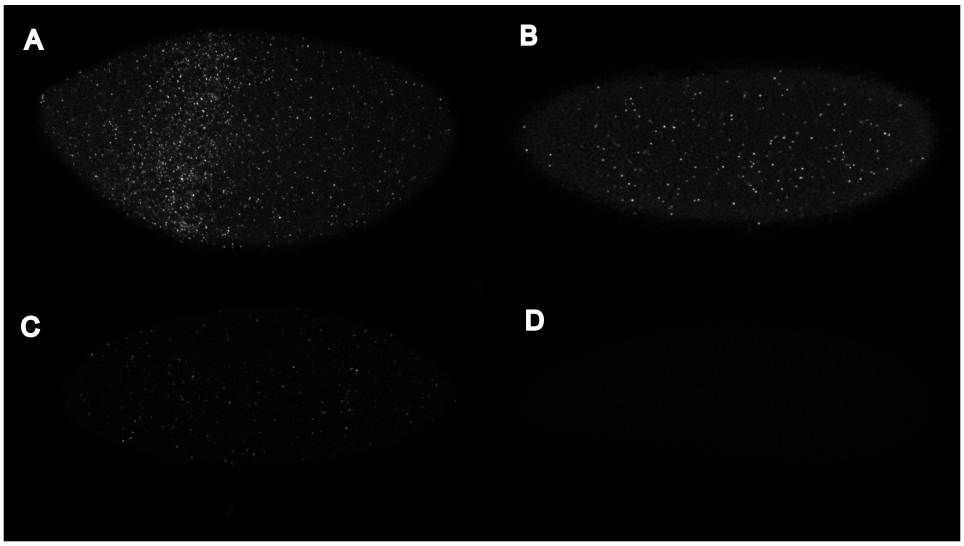


Figure 3.10: ***In situ* hybridization trials using TSA method.** A-D all use TSA method for *in situ* hybridization using control embryos. A. Positive control. Using *lac*Z-DIG probe, followed by TSA *in situ* hybridization. Clearly there are signals from this staining, indicates this positive control works. B. Negative control shows false positive signal. No probe added in the experiment, but some sparkling false positive signals show up using TSA method. C. After suppressing endogenous H_2O_2, the false positive signals reduce dramatically, but there are still some residual false positive signals. D. Negative control with no probe and no HRP. No signal is observed, as expected.

smFISH has also been attempted, follow protocol in [102]. However, I did not observe any signal from the experiment.

3.4 Materials and Methods

3.4.1 Getting D. mel S2E and T. put S2E to cad background

The full crossing scheme is shown in Figure 3.11. Since the crossing schemes for flies with *P*(S2E*mel-lac*Z) and *P*(S2E*put*-CFP) are performed exactly the same way, we only show the details for crosses with *P*(S2E*mel-lac*Z).

Figure 3.11: **Full crossing scheme** Red circles mark the parental lines needed for the cross.

The following part describes the crossing scheme from top down, left to right (Figure 3.11). Lines circled in red are parental lines needed for this cross, except the lines carrying reporter constructs. Parental fly line number 1 is a heat-shock Flipase line, courtesy of Ilaria Rebay's lab. Fly line number 2 and 3 are courtesy of Michael Ludwig. Fly line number 4 (Bloomingdale Stock Center: 5316), 5 (Bloomingdale Stock Center: 2121), 6 (Bloomingdale Stock Center: 7091). *P*{S2E*mel-lac*Z}attP2 fly is courtesy of Kenneth Barr (paper in review), reporter construct of *P*{S2E*put*-CFP}attP2 is from Michael Eisen's lab [71]. Reporter constructs were both injected into the *D. mel* attP2 landing site strain [67] and done by Rainbow Transgenic Flies, Inc.

ovo^{D1} is a dominant female sterile allele. Maternal female *cad* mutants were generated by crosses of cad^2FRT^{40A}/CyO virgins to *hs-Flp;* $ovo^{D1}FRT^{40A}/CyO$ males. 1/4 of the offspring will have $ovo^{D1}FRT^{40A}/cad^2FRT^{40A}$. Third instar larvae from this cross were heat shocked at 37°C water bath on two consecutive days after 72 hours for 2 hours each day for high recombination efficiency [27]. Flipase express under the activation of heat-shock promoter and drive recombination at FRT sites. Adult virgin flies with no *CyO* phenotype are selected for the right genotype. Only eggs that do not have ovo^{D1} allele will be fertile in these flies.

The right half of the crossing scheme show crosses of the enhancer-reporter construct onto a heterozygous *cad* background. In order to maximize the ratio of the final embryos having the correct genotype, males for the last cross carry two copies of enhancer-reporter construct.

Final *cad* zygotic mutants with *D. mel* S2E were generated by crossing *hs-Flp;* $ovo^{D1}FRT^{40A}/cad^2FRT^{40A}$ virgins (zygotic genotype) with cad^3/CyO *P(hb-lac*Z*);* *P*{S2E*mel-lac*Z}attP2 males. Note, the genotype of eggs from female is actually cad^2/cad^2. The final embryos are half heterozygous *cad* (cad^2/CyO *P(hb-lac*Z*);* *P*{S2E*mel-lac*Z}attP2/III), and half zygotic homozygous *cad* (cad^2/cad^3*;* *P*{S2E*mel-lac*Z}attP2/III). Note that embryos which are heterozygous zygotic *cad* also carry *hb-lac*Z on the second chromosome. Because

D. mel S2E uses *lac*Z as its reporter, when using *lac*Z probe to stain for *D. mel* S2E expression, embryos with *hb* expression are excluded. Therefore all the embryos without a *hb* expression for *D. mel* S2E are the desired genotype. However, for *T. put* S2E, since CFP probe is used, heterozygous *cad* embryos cannot be distinguished from homozygous *cad* embryos directly. Therefore we performed genotyping on individual embryos for *T. put* S2E after imaging of its expression patterns (only for enzymatic *in situ* hybridization).

3.4.2 in situ hybridization

Since the methods for *in situ* hybridization are similar to immunostaining staining for proteins, with one extra step of hybridizing probe to the mRNA at the beginning, I will introduce different immunostaining methods before introducing different *in situ* hybridization methods.

Immunostaining methods

Immunostaining (for proteins) can be categorized into either immunofluorescence staining or immunohistochemical staining. Immunofluorescence staining uses fluorescence as signal, while the traditional immunohistochemical staining uses the product of enzymatic reactions as readout. For enzymatic reactions, Alkaline Phosphatase (AP) and Horseradish Peroxidase (HRP) staining are commonly used. The enzymatic staining is usually more sensitive than immunofluorescence staining.

For immunofluorescence staining, depending on the abundance of the target protein product, different methods can be used [159]. If the target protein is abundant, primary antibody with conjugated fluorophores can be used. For proteins with medium to high level of abundance, usually a secondary antibody conjugate is used get modest amplification of the signal. The advantage of using either primary antibody alone or with secondary antibody is that this ensures the amplification of signal is linear in regards to the total protein levels. However, for low abundance targets, these methods are not sensitive enough to detect signals. In this case, enzymatic amplification methods are needed. The Tyramide Signal Amplification

(TSA) is a hybrid method of enzymatic amplification (with HRP), followed by fluorescent labeling. It can be used for high-resolution imaging of low-abundance targets.

Comparison of *in situ* hybridization methods

Below I compare four methods of *in situ* hybridization which have been widely-used (Table 3.1). Similar to the detection of protein expression, depending on the abundance of the target mRNA, different methods should be used.

method	experimental steps	pros	cons
FISH	hapten-labeled probe, primary antibody, secondary antibody conjugated fluorophore	amplify signal linearly with expression, quantitative	hard to get signal-to-noise ratio when expression is very low
smFISH	fluorescent conjugated probe	no amplification step, quantitative	hard to adjust experimental conditions, hard to get good signal-to-noise ratio in some cases
TSA	hapten-labeled probe, antibody conjugated with HRP, enzymatic development of HRP, tyramide conjugated with fluorophore	very sensitive, easy to adapt to multiplexing	nonlinear amplification of signal
AP	hapten-labeled probe, antibody conjugated with AP, enzymatic development	very sensitive	not quantitative, hard to choose cut-off time

Table 3.1: **Comparison of pros and cons of existing *in situ* hybridization methods**

Probe used for *in situ* hybridization

Different reporters are used for *D. mel* S2E and *T. put* S2E, because the major goal is to compare reporter expression in the presence or absence of *cad* for each enhancer. *D. mel* S2E uses *lac*Z as reporter, while *T. put* S2E uses CFP as reporter. Final embryos with *D. mel* S2E *lac*Z were hybridized with *lac*Z-DIG probe and *T. put* S2E CFP hybridized with eGFP-DIG probe. *lac*Z-DIG probe was courtesy of Michael Ludwig and eGFP-DIG probe was courtesy of Jackie Gavin-Smyth.

Fluorescent *in situ* hybridization

eGFP-DIG probe with 0.1 μl/100μl was used against CFP and *lac*Z-DIG probe with 0.25 μl/100μl was used against *lac*Z. FISH was done with Eve antibody staining. Primary antibody of 1:800 mouse anti-DIG, and 1:1000 GP anti-Eve were used, followed by secondary antibody 1:1000 anti-mouse Alexa Fluor 647 and 1:1000 anti-GP Alexa Fluor 555. Nuclei were stained with DAPI.

Enzymatic *in situ* hybridization

Enzymatic *in situ* hybridization is performed the same way as described in [108, 138]. Since CFP is shorter in sequence, its probe has higher background compared to *lac*Z probe when using similar concentrations (also observed using FISH, Figure 3.5). Therefore, we further lower the concentration of CFP in the enzymatic *in situ* hybridization experiment, but the reaction will take longer so as to give high signal to noise ratio. eGFP-DIG probe: 0.05 μl/100μl and *lac*Z-DIG probe: 0.25 μl/100μl, followed by anti-DIG AP 1:2500, for 1 hour.

Genotyping to distinguish homozygous zygotic *cad* for *T. put* S2E embryos

We used two sets of primers, one set as positive control and the other to distinguish the presence or absence of *hb-lac*Z. Positive control (*snail*): forward: GTGGACTAACG-

GCAGGAACA, reverse: CCAGATGGGGCTGATAGCTG. *hb-lac*Z: forward: CTGCCAGTTTGAGGGGACGACGACA, reverse: ACCAACGTAATCCCCATAGAAAA.

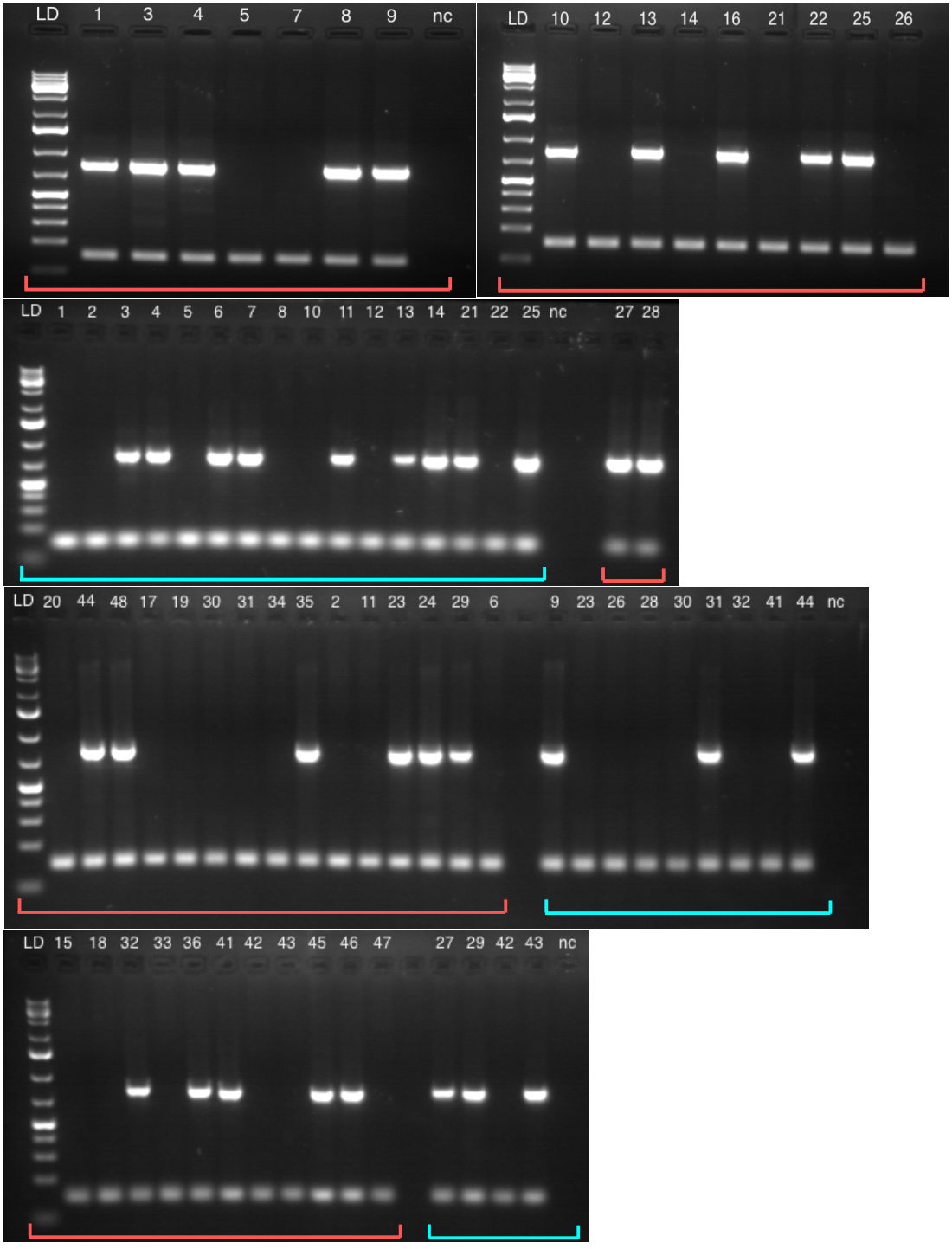


Figure 3.12: **Genotyping to distinguish zygotic heterozygous *cad* versus homozygous *cad***

Figure 3.12: Continued. The lower band is the positive control (*snail*), and the upper band shows the presence or absence of *hb-lac*Z. Red lines mark the embryos from regular stopping criteria. Blue lines mark embryos that are over-developed. Numbers represent individual embryo number using a specific stopping criterion.

3.5 Discussion

In summary, the much larger reduction in the expression level in *cad* embryos for *T. put* S2E compared to *D. mel* S2E from the enzymatic *in situ* hybridization indicates that Cad is playing a larger role in *T. put* S2E to activate its expression. These experimental results in part support the prediction from the *in silico* model, in which *T. put* S2E uses Cad as a co-activator for Hb to activate *eve* stripe 2 expression. However, the reduction of *D. mel* S2E is not predicted by the model. One possibility is that Cad is activating *D. mel* S2E, which has not been explored or characterized, and therefore not correctly represented in the model. Another possibility is that anterior gap gene expression levels are altered in *cad* embryos which are assumed to be constant in the model prediction. Previous study only shows the anterior gap gene domains are present [132], without showing quantitative comparison of expression levels. It is possible that anterior gap gene expression are reduced in *cad* embryos, and it causes a reduction in *D. mel* S2E. However, endogenous Eve stripe 2 show high expression compared to the *D. mel* S2E, indicating that some sequences outside *D. mel* S2E must have compensated this effect for the endogenous locus.

CHAPTER 4

GENERALIZATION: THE RELATIONSHIP OF ROBUSTNESS TO EVOLUTION

4.1 Abstract

Developmental robustness, as an important aspect of canalization, is a common feature of traits in multi-cellular organisms. High robustness ensures the reproducibility of phenotype within each individual. On the other hand, it is evident that phenotypic evolution has taken place in the past, causing phenotypic differences between species. To understand the evolutionary process of how canalized traits may evolve when adapting to a new environment, a Boolean model which has a simplified genotype to phenotype map is used to explore the relationship of different degrees of robustness to evolution under different environmental regimes. Fitness is based on a distance function between a phenotype and the optimum one for the environment. Genetic robustness is an explicit parameter in the genotype to phenotype map. Under a stable optimal environment, high robustness results in a equilibrium population fitness closer to the optimal fitness value than low robustness, suggesting high robustness has selection advantage under constant optimal environment. This reduction in population fitness under low robustness is attributed to stronger genetic drift caused by enhanced background selection. The situation reverses, however, when there is a sudden shift in the environment. In this case, low robustness populations adapt to the new environment faster than high robustness populations and reach higher equilibrium mean fitness. This reversal occurs because under the model, low robustness populations can access a larger phenotypic space than high robustness populations. Limitations in genetic accessibility to an adaptive phenotype may be a general property of robustness, which can favor low robustness in a changing environment. We further encode the robustness parameter by a genetic locus that can co-evolve along with the phenotype under selection. For a population with a mixtures of robustness genotypes, the low robustness genotype wins out, as expected.

But a high robustness genotype then replaces it early in the process of phenotype evolution, well before maximum fitness is achieved, and it appears to prevent further invasion into the population of a low-robustness genotype. This unexpected behavior of robustness evolution highlights the value of the novel modeling approach employed here. Additional work with the model is expected to produce further insights about the evolutionary balance between robustness and the ability to evolve, and how genetic assumptions of the model may play into its behavior.

4.2 Introduction

Most developmental processes, and the resultant morphologies they regulate, are specified by the activities of a large number of genes. Many, if not all, multi-locus traits are often subject to some form of stabilizing selection to maintain an optimum trait value in the population [87]. Stabilizing selection is capable of maintaining trait constancy over long periods of phylogenetic evolution, as evidenced by morphological stasis in the fossil record, though traits can also change rapidly in response to environmental change (Punctuated Equilibrium: [64, 38]). Rapid change of a trait under stabilizing selection is possible under the conventional neo-Darwinian framework [78, 164, 33, 118, 52, 119], driven by natural selection when the optimum trait value changes [25].

This evolutionary process is contingent on organismal development. We are interested in the reproducibility of traits, termed canalization or alternatively high robustness, a universal property of development in both single and multi-cellular organisms [185, 119, 100]. We are also interested in gaining a better understanding of the constraints canalization imposes on the subsequent ability of a population to evolve towards a new phenotypic optimum, since a more canalized phenotype, by definition, will exhibit less phenotypic variation, the raw material on which natural selection acts. There are seemingly competing trade-offs between natural selection for robustness and the rate at which a population can evolve in response to the demand for a new optimum phenotype. We narrow our focus to genetic (mutational)

robustness in this chapter.

One of the challenges in gaining a better understanding of how canalization may constrain or facilitate phenotypic evolution has been a lack of good understanding of the genetic and mechanistic basis of canalization itself. C. H. Waddington was the first to propose a mechanism for how canalization might evolve—genetic assimilation [185, 186]. The first stage of this process involves gene-by-environment (GxE) interaction, utilizing physiological responses to environment changes. Selection then acts on this response, and the third step internalizes the physiological response into a hardwired genetic mechanism for the trait.

The first hint that there may be global molecular regulators of canalization came from investigations by Rutherford and Lindquist [156] of the heat-shock protein, Hsp90, which is a protein that assists other proteins to fold correctly and participates in additional protein processing pathways. Hsp90, they believed, might adaptively regulate "expressed" genetic variation in other proteins segregating in a population through its own sensitivity to environmental stress. Such process of releasing phenotypic variations at onset of environmental stress is also called de-canalization [57].

There are some differences between Waddington's genetic assimilation theory and the canalizing activity of Hsp90. First, the genetic assimilation theory involves internalizing a specific non-genetic physiological response into a fixed genetic response, whereas Hsp90 is hypothesized to be an evolved global regulator of canalization that acts on pre-existing mutations in many genes and genetic pathways [185, 186]. Second, Hsp90 can act as a gatekeeper of evolution by either suppressing or enhancing the realization of genetic variation. In contrast, the GxE interactions proposed by Waddington can be restricted to a particular pathway or even impose directionality of GxE interactions. Unfortunately, the otherwise attractive theory that Hsp90 acts as a global regulator of canalization and evolution, finds little experimental support of its role contributing to evolution. The most convincing case involves selection on Hsp90 for eye size in cavefish [152]. More recently, Geiler-Samerotte et al. [56] have found that although Hsp90 buffers standing genetic variations in yeast de-

rived from natural populations, it reveals more phenotypic variations for *de novo* mutations, indicating natural selection is selecting for mutations with high robustness.

Besides genetic assimilation and Hsp90, a number of other genetic mechanisms contributing to canalization and adaptation have been investigated empirically. One popular theory is the co-option of gene regulatory networks [142, 123, 127, 93, 90]. This idea is based on the assumption, supported by observations, that most gene regulatory networks are modular [137, 101, 141] so that new functions can arise by a small number of genetic steps [60, 90]. New functions can also arise by gene duplication followed by neo-functionalization [44]. Many other findings identify molecular components which contribute to canalization, but their contribution to adaptation is less clear. These examples include shadow enhancers in *Drosophila* embryos [139, 140], network motifs [2] and mutual repression in gene regulatory networks [114, 113]. In addition, studies of miRNAs provide molecular evidence of their role in canalization. one example of which is mir-9a RNA in *Drosophila*. Knocking down this gene produces variation in bristle number [23, 144].

As alternatives to experiments, mathematical and computational models can shape our understanding by making explicit assumptions and formulations [188, 161, 162, 31, 189, 4]. Some models support the conventional belief that there is an antagonistic relationship between robustness and evolvability [4], while others illustrate conditions under which robustness can actually promote evolvability [190, 191] while other find that an intermediate level of robustness is optimal for adaptation [34]. Conflicting ideas about the relationship between robustness and evolvability are clearly dependent on the specification of models and the assumptions underlying them, indicating that this issue is far from settled. And, some models bear little resemblance—or cannot be easily analogized—to known biological processes or mechanisms of evolution. There are also definitional issues: "evolvability", for example, does not have a strict definition, and is measured in different ways across studies. We avoid using this word in the current chapter.

Models to explore robustness and evolution can be divided into two general classes,

depending on whether or not they consider the GxE interaction. Without GxE, the genotype to phenotype map is fixed. Many of these non GxE models use the idea of a "neutral network", originally defined as a network in which changes in genotype do not produce changes in phenotype [189]. These models are also neutral in the sense that they tend to make use of the idea of fitness in a very limited way. For example, Wagner [190, 191] showed that robustness can promote evolvability by demonstrating that more phenotypes are accessable from a network with both neutral and non-neutral genotype changes than from one having only non-neutral genotype changes. This is not the whole story, however, because increasing robustness to its maximum level in fully neutral network with no alternative phenotype reduces evolvability to zero. Another study considered a neutral network with a universe of P possible phenotypes of which exactly K phenotypic neighbors are accessible by a single mutational step from each node in the neutral network. The authors find that intermediate levels of robustness maximize evolvability when $K < P$ [34]. This model ignores mutational effects arising from the contingency of mutations in different genetic backgrounds, which has been shown to be important in protein evolution [171, 5]. The lack of an explicit genotype to phenotype map removes important biological constraints. Moreover, adaptation is generally a multiple-step process, but here evolvability (adaptation time) is dependent on the waiting time for a beneficial mutant (phenotype) to arise. Understanding what happens to the relationship between robustness and evolvability when one or more of these assumptions are relaxed has not been explored.

Other types of models assume an explicit form of GxE interaction following an environmental shift, known as plasticity [74, 3, 146]. The general topic of plasticity is beyond the scope of current chapter, but we mention its interface with canalization. With GxE, a mutant allele that might be selectively neutral or even deleterious in one environment can become adaptive in a new environment. Under stabilizing selection the population can build up a reservoir of such genetic variants, including ones potentially adaptive in another environment. These genetic variants, called cryptic genetic variants because their effects are

masked under the current environment, are a realization of canalization [117, 135]. Following an environmental shift, in the presence of GxE interactions, cryptic genetic variants can produce new phenotype that could be selected upon.

Many previous models that only focus on exploring the relationship of robustness to stabilizing selection suggest that stabilizing selection results in high robustness [192, 40, 188]. A Boolean model which, in contrast, focuses on evolvability suggests long-term evolution of gene regulatory networks in changing environment can facilitate the efficiency of generating beneficial mutations [31] .

To better understand both how a trait evolves when it is subject to canalization, and how canalization evolves in relation to the intensity of selection pressure imposed by novel environments on a trait, quantitative models are needed that capture robustness or even allow both traits and robustness to evolve simultaneously as a coupled process. Formulation of the models should also be motivated by traits governed by known biological mechanisms and processes. Here we have devised a relatively simple, computationally tractable, Boolean model to explore the relationship between robustness and evolution. Our model falls into the class that does not assume GxE interaction and considers cell type evolution with a single optimum mix of cell types, or a shifted optimum. The model considers the effect of multiple genes on multiple end phenotype. Epistasis and pleiotropy are explicitly represented in the model. Here we only consider genetic robustness, though the model can be extended to investigate other forms of robustness, such as environmental robustness or physiological plasticity.

Because we have attempted to construct a multi-locus model that captures essential features of an explicit biological process—cell type specification—we first describe the model in terms of this biological process. We then proceed to describe technical details of the model. The first part of the model examines the effect of different robustness under either an initial optimum or a shifted optimum, with constant robustness. Later, robustness is introduced into the model as a second multi-locus trait, evolvable in the same manner as the cell type

trait itself, and coupled to cell type specification with a biologically explicit interpretation. The model is explored by forward simulations of Wright-Fisher populations evolving under mutation and selection for a fixed optimum target of cell types. This allows us to investigate not only the adaptive evolution of robustness—both its increase and its diminution— as a function of constancy or change in trait optimum, but also the rate at which robustness changes in relation to the rate of advance of cell types towards the optimum.

4.3 Results

Our model attempts to capture essential features of multi-locus gene regulation for a large number of traits. In order to follow evolutionary dynamics, we coarse-grain developmental dynamics into a genotype to phenotype map. Genotype is the set of all loci $\boldsymbol{v}$, each element v_i of which represents a gene expression ON/OFF ($\{1, 0\}$) state. A specific trait z_i will be present or absent ($\{1, 0\}$) in an individual in a manner that is determined by the genotype $\boldsymbol{v}$, and the whole collection of traits $\boldsymbol{z}$ constitutes the phenotype (Table 4.1). We consider a haploid organism with no recombination (*i.e.*, complete linkage of genes on a single chromosome).

vector	element	element definition
$\boldsymbol{v}$	v_j	jth genotype
$\boldsymbol{z}$	z_i	ith phenotype
$\boldsymbol{T}$	t_{ij}	functional contribution of genotype j on cell type i
$\boldsymbol{h}$	h_i	threshold for cell type i
$\boldsymbol{b}$	b_i	ith element in the environment, representing desired phenotype for z_i under current environment

Table 4.1: **List of variables**

A second feature of the model concerns robustness, by which we mean the sensitivity

of phenotype to changes in genotype. We consider two possible situations. In one, a level of robustness is imposed on the model and we monitor its selective consequences. In the alternative formulation, we allow robustness itself to become a trait subject to selection.

We consider a picture where fitness is determined by an explicit interaction between the phenotype $\mathbf{z}$ and an environment $\boldsymbol{b}$, also described by a Boolean vector. In this work, we allow the optimal phenotype to be read off from the environment in an explicit manner, such that fitness decreases linearly with the distance of an individual's phenotype from the optimal phenotype.

These formulations of genotype, the genotype to phenotype map, robustness, and fitness allow us to carry out forward evolutionary simulation of a population as it adapts from its current state of phenotype towards the optimal state of phenotype. Our simulations include the case where the population begins in the fully optimal state of phenotype, and is challenged to maintain this optimal state in the face of mutation. The simulations also consider cases where a novel environment is introduced, and the population is challenged to evolve towards the new phenotypic optimum for this environment. Evolution is achieved by allowing mutation to operate on genotype by switching a single gene from ON to OFF (or the opposite) and considering the phenotypic consequences of each mutation in each individual. As mentioned above, robustness determines the sensitivity of phenotype to mutation. Initially we explore evolution of phenotype for fixed values of robustness. In a later section we allow robustness to evolve as well, to explore co-evolution of phenotype and robustness.

4.3.1 Model description

The model consists of two parts: the genotype to phenotype map, and the phenotype to fitness map, as indicated schematically in Figure 4.1. We now describe these sectors in detail.

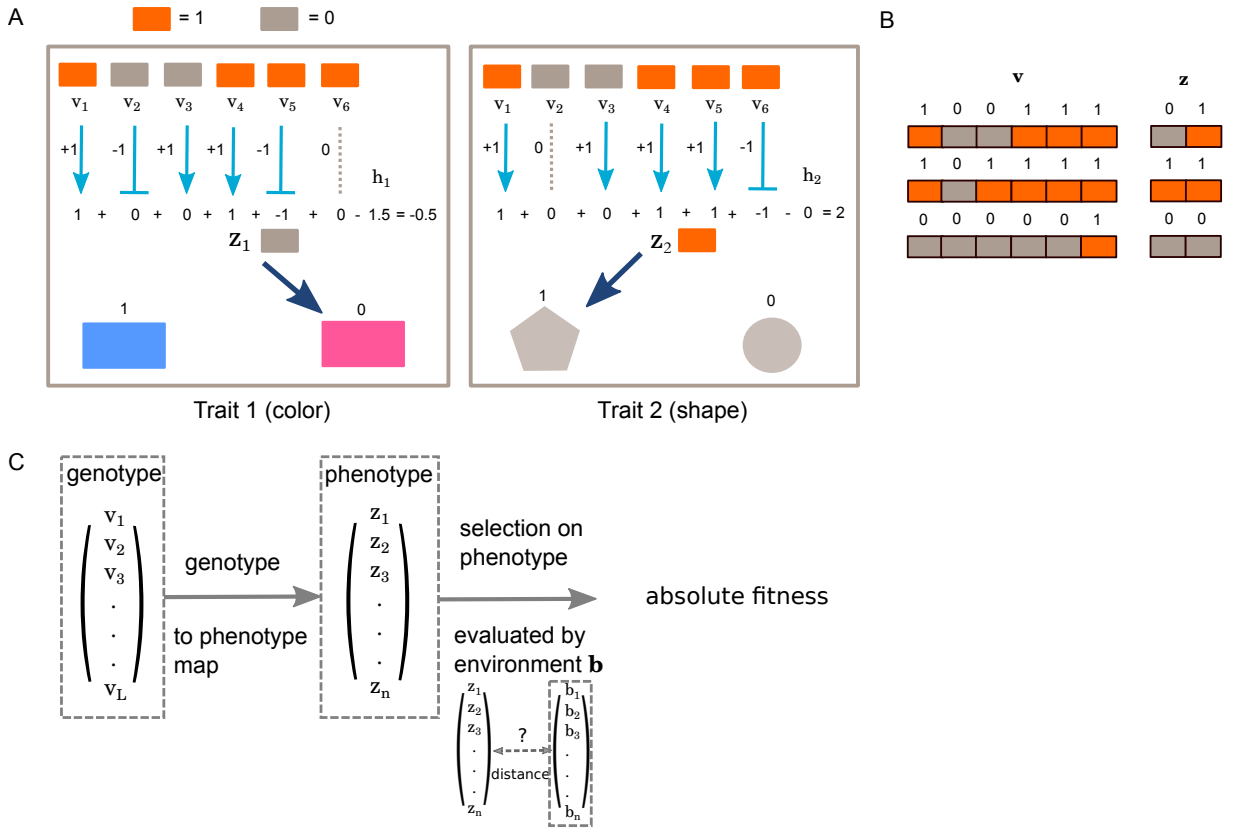


Figure 4.1: **Schematic example and overview of the model** A. Schematic figure showing two cell type determination by the genotype to phenotype map within one individual(equation 4.1). Phenotype z_1 determines color, z_2 determines the shape. In this example, the number of genes regulating the phenotype is 6, γ is set to 1. Interactions are marked by the types of arrow. Activation is represented by arrows with a triangle end, while repression is represented by a bar end. Grey dashed line means no effect. Under the current configuration, $z_1 = 0$ and $z_2 = 1$. B. Examples of different genotype configurations that give rise to different phenotype, given the genotype to phenotype map in A. Orange color indicates "1", and gray indicates "0" for each genotype or phenotype element. Note that under the genotype to phenotype map in A, phenotype vector $\boldsymbol{z}$ cannot take value of $(1, 0)$ by any configuration in genotype.

Genotype to phenotype map

We represent the genotype to phenotype map by a one-layer perceptron which maps $\boldsymbol{v}$ to $\boldsymbol{z}$ using the equation

$$z_i = \sigma(\sum^{j} t_{ij} v_j - h_i), \tag{4.1}$$

Figure 4.1: Continued. C. Overview of the full model. Genotype is $\boldsymbol{v}$, phenotype is $\boldsymbol{z}$. First part is a genotype to phenotype map, and the second part is phenotype to fitness map. Selection is on phenotype, evaluated by $\boldsymbol{b}$, representing environment.

where $z_i \in \{0, 1\}$ is the ith component of $\boldsymbol{z}$ and $v_j \in \{0, 1\}$ is the jth component of $\boldsymbol{v}$. Each component $t_{ij} \in \{-\gamma, 0, \gamma\}$ of $\boldsymbol{T}$ represents the action of terminal differentiation genes which produce a trait. h_i, a real number, is the ith component of the threshold $\boldsymbol{h}$. $\sigma(y)$ is a step function such that $\sigma(y) = 0 \,\text{iff}\, y \leq 0$, and 1 otherwise. $\boldsymbol{v}$ has L components while $\boldsymbol{z}$ and $\boldsymbol{h}$ have n components. In simulations, we assign elements of $\boldsymbol{v}$ a value of zero or one randomly in such a way as to ensure that the proportion of ones is a. Each v_i represents the gene expression state (ON/OFF) of gene i. If v_i equals 0, the gene is OFF (not expressed) and it cannot contribute to any trait. Alternatively, when v_i equals 1, the gene is ON (expressed), in which case it can have three possible roles in determining a phenotype: activator (v_5 for phenotype 2 [right panel] in Figure 4.1A); repressor (v_5 for phenotype 1 [left panel] in Figure 4.1A); no functional role (v_6 for phenotype 1 [left panel] in Figure 4.1A). The probability that each t_{ij} has a non-zero value is given by the parameter c, so that $Pr(t_{ij} = 0) = 1 - c$, and $Pr(t_{ij} = \gamma) = Pr(t_{ij} = -\gamma) = c/2$. $\sum^j t_{ij} v_j$ gives the difference between activating and repressing activities of genes in $\boldsymbol{v}$ for trait i. In order for phenotype i to be present ($z_i = 1$), this difference must be greater than the threshold value h_i.

Phenotype to fitness map

In our model, fitness is determined by an interaction of phenotype $\boldsymbol{z}$ and environment $\boldsymbol{b}$, rather than directly from genotype. In this application we restrict $\boldsymbol{b}$ and $\boldsymbol{z}$ to have the same number of components n, and then the absolute fitness w is given by

$$w = \frac{n - D(\boldsymbol{b}, \boldsymbol{z})}{n}, \tag{4.2}$$

where $D(\boldsymbol{b}, \boldsymbol{z})$ is the Hamming distance between the two vectors. This is a simple fitness function in which fitness is determined by the distance of an individual's phenotype from the optimal phenotype for a given environment.

4.3.2 Parameters controlling robustness

We represent robustness in our model by the sensitivity of a phenotype to genetic change, manifested by the probability that phenotype $\boldsymbol{z}$ changes given a mutation (bit flip) in genotype $\boldsymbol{v}$. Consider the argument to σ in (4.1), given by

$$y_i = \sum^{j} t_{ij} v_j - h_i. \tag{4.3}$$

The probability that a trait z_i changes as a consequence of of a mutation in v_i can be easily calculated because the t_{ij} are independently and identically distributed (i.i.d.), L is large, permitting the invocation of the Central Limit Theorem, and h_i itself is drawn from a Normal distribution $\mathcal{N}(0,\, acL)$ (see Methods). Then the probability P of a mutation causing a phenotypic change is

$$P = c[0.5 - \Phi(-\gamma/\sqrt{(1+\gamma^2)acL})], \tag{4.4}$$

where $\Phi(x)$ is the CDF of the standard normal distribution.

Genetic robustness, in this formulation, can be defined as $1 - P$. The parameters a, c, L and γ can all influence robustness; in the following analysis we only vary γ. We performed simulations to evaluate robustness changes by varying γ and calculating the number of trait changes in $\boldsymbol{z}$ given one mutation in $\boldsymbol{v}$. Figure 4.2A demonstrates that simulation results are consistent with (4.4). γ has a monotonic relationship with robustness—robustness decreases with increasing γ—but this relationship has upper and lower bounds.

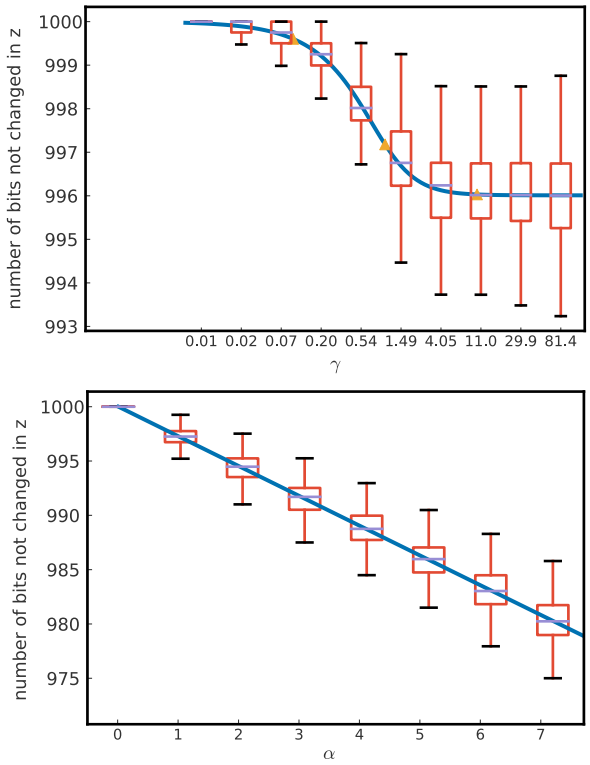


Figure 4.2: γ **and** α **controlling robustness** Y axis shows the number of bits that are not changed in phenotype z, which is a measure of robustness. A. γ controlling robustness. For the simulations, we first simulate cell types based on parameters. Then for each simulated cell type, every bit in the genotype is flipped and the new phenotype is calculated, the number of bits that are kept the same is recorded. The boxplot shows the simulation result. 1000 simulations for each parameter set. $c = 0.5, a = 0.5, L = 10000, n = 1000$. γ is varied as shown on the x axis. The solid lines shows the analytic result (from equation 4.4, showing the expected number of traits in phenotype that are not changed: $n(1 - P)$). Yellow triangles show three γ values: 0.1, 1, and 10, which are used for the following simulations. B. α controlling robustness. For the simulations, initial cell types are simulated the same way by parameters as A. $c = 0.5, a = 0.5, L = 10000, n = 1000, \gamma = 1$. Then α is re-scaled for each cell type, and number of bits that are kept the same in phenotype is evaluated for each α. 1000 simulations are done for each parameter set.

4.3.3 Robustness and evolution

Evolution in a constant, optimal environment

We first examine evolution under a constant optimal environment, a situation analogous to a population that is well adapted to its stable native environment. Models of robustness are often explored under a scenario such as this [192, 40, 188, 161, 17]. The simulation is initial-

ized for parameter values of c, a, n, L and γ given in Table 4.2 to generate a genotype vector and the phenotype vector it produces. We set the environment vector $\boldsymbol{b}$ to this phenotype vector. This initial genotype is thus optimal for this environment. All the individuals in the population initially have the same genotype. We consider a haploid population model with $N = 10,000$ individuals, complete linkage of genes, and no recombination. Standard Wright-Fisher forward simulations with mutation, selection and genetic drift were performed. We set $n = 1000, L = 10000, a = 0.5, c = 0.5$, the mutation rate per gene $\mu = 10^{-6}$, and population size $N = 10000$. We set L to be of the same order of magnitude as the number of genes in a typical eukaryote. Under this parameter regime, the genome-wide scaled mutation rate is $\theta = 200$. Unless stated otherwise, these parameter values were used for all simulations. We investigated three γ values: $0.1, 1, 10$.

parameters	definition	value
c	probability of t_{ij} not 0	0.5
a	proportion of 1s in $\boldsymbol{v}$	0.5
n	length of phenotype vector $\boldsymbol{z}$	1000
L	length of genotype vector $\boldsymbol{v}$	10000
γ	activation/repression coefficient of each gene on its target cell type	$0.1, 1, 10$
α	coefficient to scale robustness	$0.5, 1, 2$
μ	mutation rate per gene	10^{-6}
N	population size	5000
k	length of robustness genotype	2
μ_r	mutation rate on robust genotype	10^{-5}
I	lower bound when scaling robustness using α	-1
J	upper bound when scaling robustness using α	1

Table 4.2: **List of parameters**

Figure 4.3A shows the mean fitness changes for populations with three different γ values, under an optimal stable environment. The simulations are allowed to continue until a plateau in population mean fitness is observed. When γ is small (high robustness), the population mean fitness remains close to the optimal value ($\bar{w} = 1$), while when γ is large (low robustness), the population mean fitness equilibrates at slightly lower values. This confirms a previous finding that high robustness is favored under optimal stabilizing selection [188]. Because the populations start with optimal phenotype, most mutations are initially deleterious (Figure 4.3F). The greater reduction in mean fitness under our model when robustness is low is likely due to the effects of background selection against deleterious mutation [24, 26] caused by the greater initial variance in deleterious fitness effects of mutation (Figure 4.3F). It is also is similar to Muller's Ratchet [69]. As expected, mutations are purged from the population more effectively when robustness is low, resulting in lower mean population heterozygosity (Figure 4.3C).

We also investigated evolution under a constant environment for all combinations of values of n and $L \in \{100, 1000, 10000\}$ (Figure 4.4). In general, there is a decline in population fitness with decreasing robustness, though the sensitivity of population fitness to robustness is greater with larger number of loci. One combination of parameter values ($n = 10000, L = 1000$, other parameters being the same), however, produces an anomalous result in that this pattern reverses, with high robustness having severely reduced fitness ($\gamma = 0.1$) (Figure 4.4). This behavior may be a manifestation of Muller's Ratchet after the population exceeds a threshold number of deleterious mutations, initially hidden by robustness, at which point mean fitness begins to decline precipitously.

Evolution after a sudden environmental shift

What would the relationship of robustness and evolution be when there is a sudden environmental shift? We first explore this question by randomly selecting an environment $\boldsymbol{b}$ from the universe of possible environments, without regard to the phenotype $\boldsymbol{z}$ of individuals in

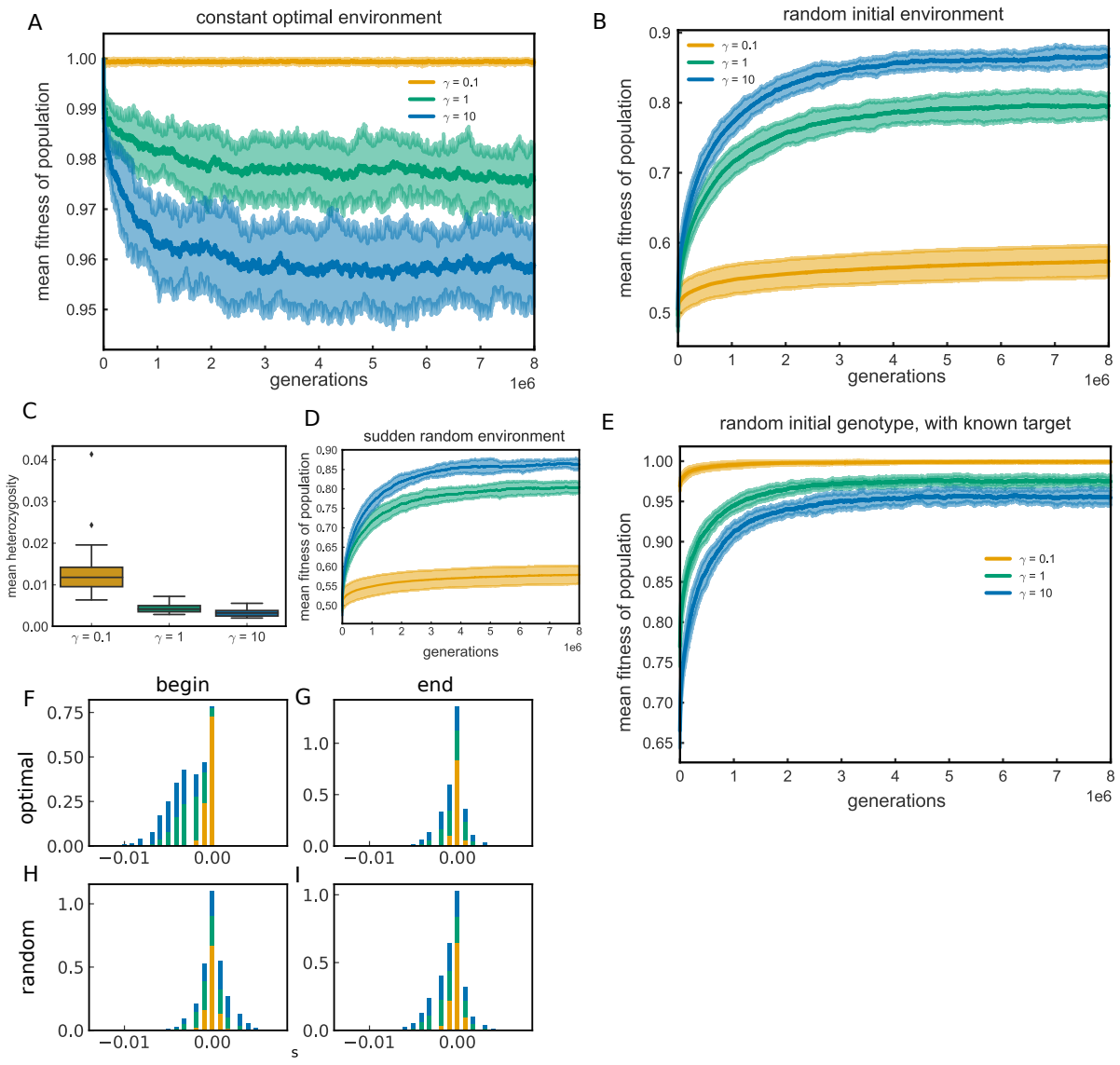


Figure 4.3: **Different evolutionary trajectories with populations under initial optimal environment and adapting to a new environment** A. Mean fitness changes over time for populations with different robustness levels under optimal stable environment. 20 simulations were done for each γ. The thick solid lines show the mean of the 20 simulations at each generation. The upper and lower boundaries are 1.5 standard deviation on each side. $\gamma = 0.1$ in yellow, $\gamma = 1$ in green and $\gamma = 10$ in blue. The color code is consistent throughout this figure. B. Mean fitness changes over time for populations with different robustness levels under a random initial environment. C. Boxplot of heterozygosity at the end of simulations under optimal stable environment. D. Mean fitness changes over time for populations that undergone a sudden environmental shift which were previously in optimal stable environment for 8000000 generations.

Figure 4.3: Continued. E. Mean fitness changes over time for populations initialized with a "attainable" random environment, where this new environment is initialized using a random genotype, and the same genotype to phenotype map to calculate the optimal environment $\boldsymbol{b}$. F-I show the distribution of selection coefficient s in different environments and at different times. F, G are for populations at optimal stable environment, with F at generation 0, and G at generation 8000000. H, I are for populations at random initial environment, with H at generation 0 and I at generation 8000000.

the initial population, so that the populations must adapt to this new environment. We found that populations with large γ (low robustness) respond to change in the environment more quickly than those with small γ (high robustness). We also found, however, that low robustness populations reached higher equilibrium mean fitness than high robustness populations (see Figure 4.3B). Both of these observations —the initial rate of adaptation and the ultimate level of population fitness—have monotonic relationships with robustness (Figure 4.5), in contrast to the findings in [34].

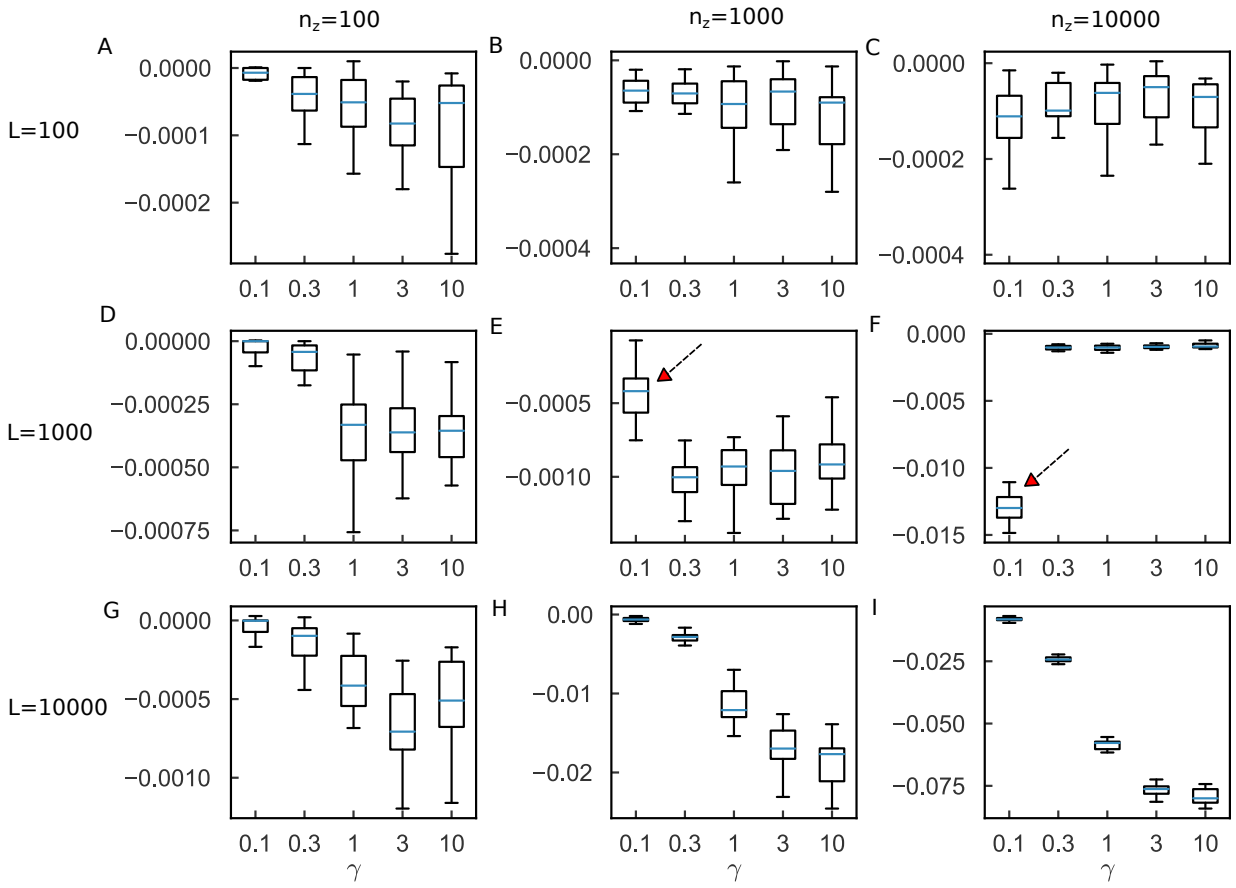


Figure 4.4: **Simulations under different parameter regimes** We changed n and L in the simulations, and in order to compare the mean fitness difference before and after, we set the simulations to run so that each simulation has the same expected number of mutations occurring (same mutation rate per gene. for $L = 100$, simulations run for 8×10^6 generations, $L = 1000$: 8×10^5, $L = 10000$: 8×10^4). Here plots all the mean fitness difference for different n and L. For $L = 10000$, which is what used in the main text, the mean fitness difference is as described in main text, where low γ (high fitness) is more close to optimal, and there is a monotonic relationship between robustness and the distance to optimal. There is an interesting behavior when $L = 1000$, where the relationship of robustness to the distance to optimal drastically shifts when n changes from 1000 to 10000. When n equals 10000, $\gamma = 0.1$, the deleterious effect of fixing individual mutation is so low that mutations get fixed, while for other γ values, mutations do not fix. The cumulative deleterious effect of fixing γ will be seen until it reaches a threshold, and the mean fitness of populations just keep decreasing, which is the case as Muller's ratchet. For $L = 100$, there is generally no effect of robustness on mean fitness change, because the per genome mutation rate per generation is so low that the deleterious mutations cannot be fixed at all, which the effect of robustness cannot be seen.

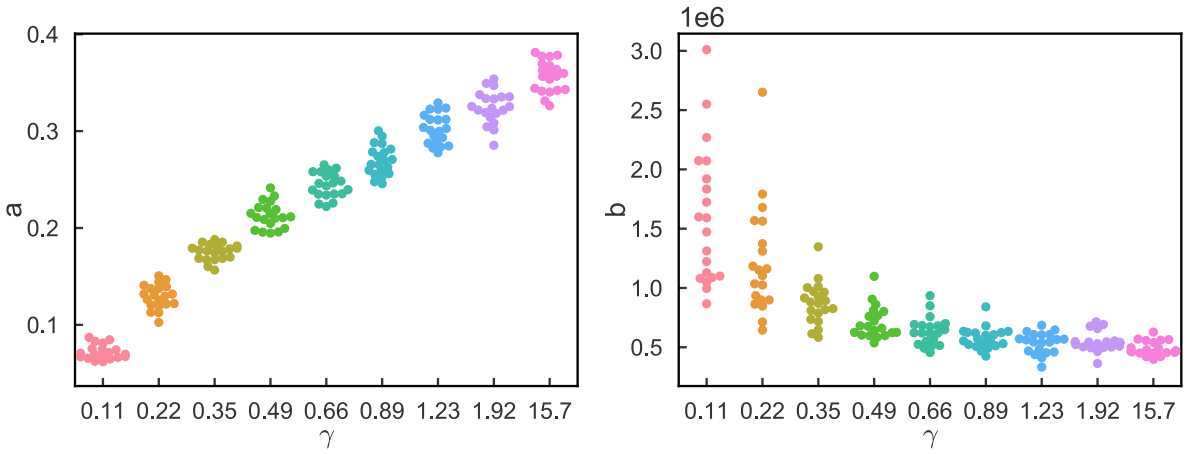


Figure 4.5: **Monotonic relationship of robustness to adaptation rate** To characterize the rate of adaptation for different robustness levels, we fit each mean fitness changes over time to the saturation function $y = at/(b + t) + c$. a describes saturation level, while b describes the time needed to reach saturation. Here we plot the Durant distribution of a and b with populations simulated with the that are equally separated in robustness.

In the previous simulations, the initial population started with a single genotype, whereas a population previously adapted to a stable environment will be genetically variable. We therefore investigated whether initial genetic variability affects the ability of a population to evolve to a new environment, noting that a more robust population will also be more genetically variable than a less robust population (Figure 4.3C). We initialized populations by subjecting them to 8×10^6 generations of stabilizing selection (Figure 4.3A) before applying a sudden change in environment, by assigning a random $\boldsymbol{b}$ at generation 8000001. We then allow the populations to adapt to their new environments for another 8×10^6 generations. We plotted the mean fitness change after this environmental shift in (Figure 4.3D), and found the same relationship between robustness and adaptation as for the homogeneous initial populations. This means that there is no initial advantage in highly robust populations to having higher genetic diversity in their ability to evolve to a new environment. This result runs counter to ideas stemming from models of genotype networks [190, 191].

We also calculated the selection coefficients s of new mutations at the beginning and end of these simulations (Figure 4.3F-I), and compared them to the distributions for evolution in

an optimal stable environment (Figure 4.3F-G). We find a large difference in the distributions of s under the two evolutionary scenarios. At the onset of simulations under an optimal environment, all mutations are either neutral or deleterious and, as expected, high robustness populations have a larger ratio of neutral mutations to deleterious mutations compared to low robustness populations. There is also a wider spread of deleterious mutations in low robustness populations. At equilibrium, the distribution of s becomes less skewed towards negative values, and high robustness populations have a narrower distribution than low robustness populations. In contrast, when adapting to a new environment, the distribution of s is initially symmetric (Figure 4.3H). Low robustness populations have a larger variance in fitness, which accounts for their more rapid initial increase in fitness. At equilibrium, however, the s distributions becomes skewed once again towards negative values, which is similar to the equilibrium distribution under an optimal environment (Figure 4.3G,I).

We were surprised by the relative inability of more robust populations to adapt to a new environment (Figure 4.3B) compared to less robust populations. We wondered whether the new, randomly generated environments are less genetically accessible for more robust populations. We therefore constructed environments known to be accessible by constructing a random genotype $\boldsymbol{v}_{\text{target}}$ and generating its phenotype $\boldsymbol{z}_{\text{target}}$ using the same threshold values and $\boldsymbol{T}$ matrix used throughout the simulation, and then setting $\boldsymbol{b} = \boldsymbol{z}_{\text{target}}$. After establishing an initial genetically uniform population population with a random $\boldsymbol{v} \neq \boldsymbol{v}_{\text{target}}$, we then challenged the population to evolve adaptively towards this new and accessible environmental target. We find that under all three robustness conditions, populations can now evolve to optimal values (Figure 4.3E), similar to the ones we found when populations evolve starting with the optimum genotype (Figure 4.3A). This indicates that robust populations are highly adaptable to environments within their accessible phenotypic space. It also demonstrates that new randomly generated environments are indeed less accessible, and supports the idea that the universe of possible phenotypes is smaller in populations subject to higher levels of robustness.

4.3.4 Evolution of robustness

In this section we investigate the dynamics of the evolution of populations when robustness can genetically co-evolve along with phenotype. The previous sections examined the relative advantages and disadvantages of robustness with respect to the evolution of a population in a constant or abruptly altered environment. The major questions we explore here are whether robustness can evolve adaptively under our model, and if it can, what is the magnitude and dynamics of this change in relation to evolution of phenotype?

We then sought to allow robustness itself to evolve under genetic control. Allowing mutations to change γ is not a suitable method for addressing this question because such changes will not only affect robustness, but can also change the phenotype directly. Instead, introduce a positive real number α which scales the phenotypic effect of mutation. Equation (4.3) is now written

$$y_i = \alpha \sum_j t_{ij} v_j - h'_i, \tag{4.5}$$

where $h'_i = (\alpha - 1) \sum_j t_{ij} v_j + h_i$. The multiplication of the first term by α scales the sensitivity of phenotype to changes in genotype in a manner similar to γ, but the modification to $\boldsymbol{h}$ keeps $\boldsymbol{z}$ the same for the current $\boldsymbol{v}$ when α is changed.

The probability P for changing the phenotype after a single mutational change, previously given in (4.4), now becomes

$$P = c[0.5 - \Phi(-\alpha\gamma/\sqrt{(1+\gamma^2)acL})] \tag{4.6}$$

Under this model, robustness is $1 - P$. We note that there are five variables that can influence P: a, c, L, γ and α. We set γ, which we previously varied, to be a constant, and only allow α to be the parameter that controls robustness. We set $\gamma = 1$ in this case unless otherwise stated. These values of a, c, L, and γ restrict Φ to a narrow region of its domain near $\Phi(0)$ which is almost linear. Then, to good approximation, α has a linear relationship

to robustness (Figure 4.2B): robustness decreases as α increases.

A population of individuals with differing robustness

A direct prediction from previous sections is that if a population consists of individuals having different degrees of robustness, high robustness will be favored under a stable optimal environment. Low robustness will be selectively favored, in contrast, when the environment changes. We set up simulations to test this hypothesis by allowing α to vary.

We initialized the $\boldsymbol{v}$, $\boldsymbol{z}$, $\boldsymbol{T}$ and $\boldsymbol{h}$, following the same procedure as the one described for investigating populations in a constant optimum environment (Section 4.3.3, above), with the parameter values given in Table 4.2. Individuals, initially all sharing the same genotype, were then assigned one of three α values: $0.5, 1, 2$. These three types are represented in the starting population with equal frequency. Figure 4.8A and B show a representative population simulation under an initial optimal environment and under an initial randomly chosen environment (compare to Figure 4.3A and B respectively). For populations with a stable optimal environment, in twenty trials individuals with the highest robustness became fixed in the population and the two lower robustness genotypes were eliminated (Figure 4.6). For random initial environments, 16 out of 20 simulations resulted in fixation of the most robust genotype, while in 4 cases the intermediate robustness genotype went to fixation.

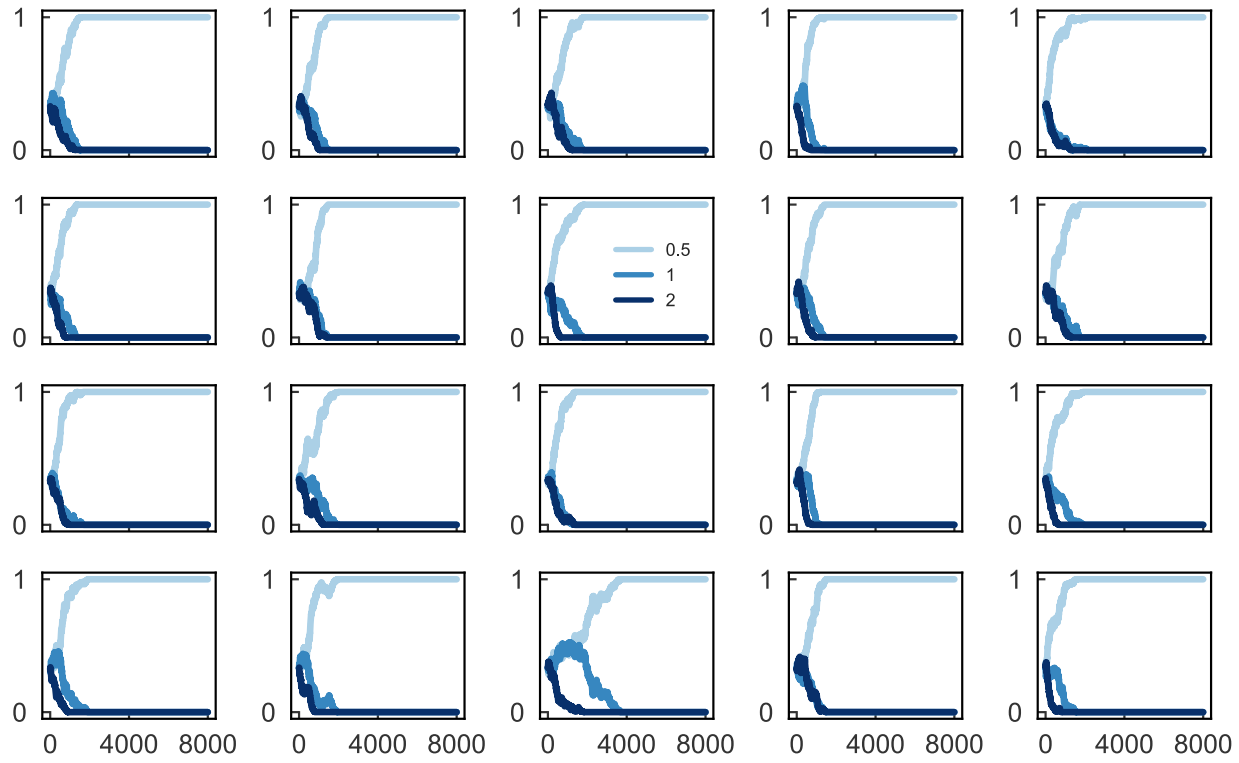


Figure 4.6: **Mixed population with three degrees of robustness, evolving under stable optimal environment** 20 simulations start with initially three different degrees of robustness, differing by α. In all the cases, lowest α (most robust) genotype takes over the population.

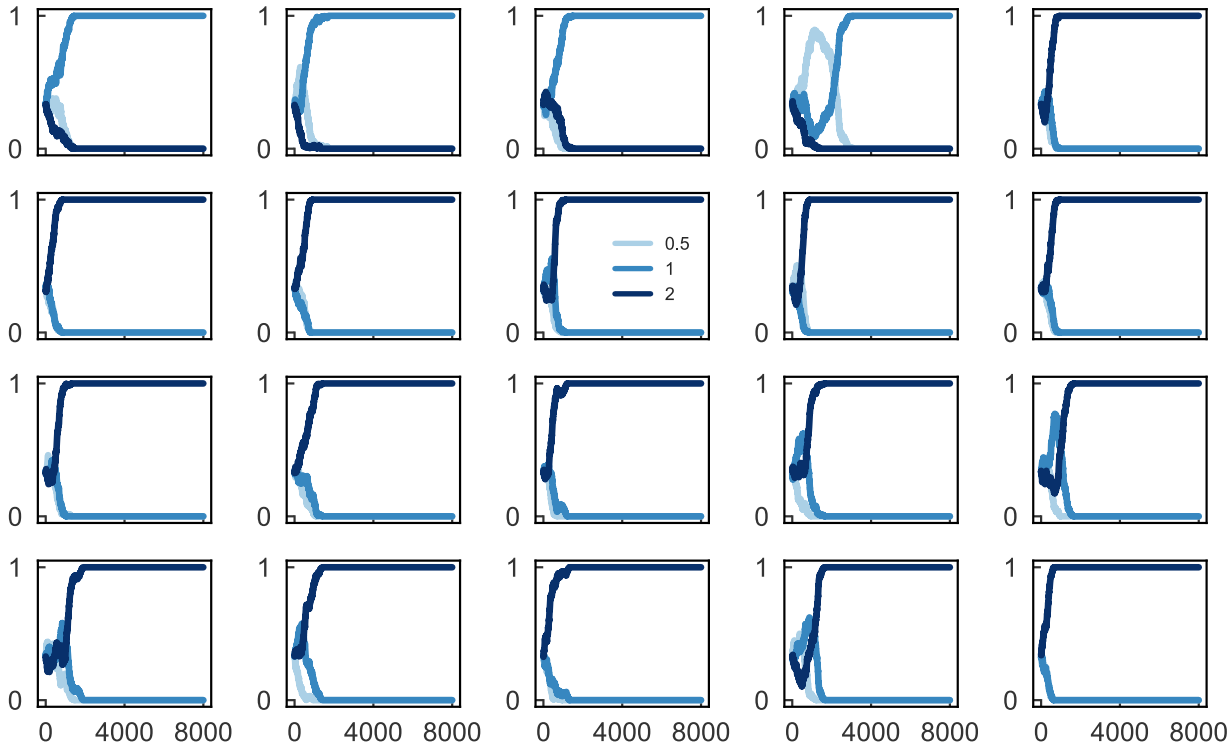


Figure 4.7: **Mixed population with three degrees of robustness, evolving under stable optimal environment** 20 simulations start with initially three different degrees of robustness, differing by α. In most cases, highest α (least robust) genotype takes over the population, but there is some stochasticity in the result. In several cases, intermediate robustness genotype takes over the population.

An evolving locus that controls robustness

To further investigate the evolution of robustness, we introduce a locus regulating robustness that can mutate to different values of α. This locus is completely linked to the other genotype loci $\boldsymbol{v}$. We use k bits of genotype to regulate robustness, and sum the bits to construct a map from discrete changes in genotype to real valued changes in α. This procedure yields $k+1$ different robustness levels. The robustness genotype has the value m, where $m \in \mathbb{Z} : m \in \{0, .., k\}$. It is useful to consider m to be the logarithm of robustness, so that

the scaling factor α is determined by

$$\alpha = 2^{(I+\frac{J-I}{k}m)}, \tag{4.7}$$

where m is the aforementioned robustness genotype, and I and J are the lower and upper bounds of robustness. We then use 2 bits to represent three different α genotypes. We take $I = -1, J = 1$ so that α will have the three possible values 0.5, 1, and 2. Each individual now has its own value of α, which can change by mutation. A mutation in the robustness locus alters a single bit, as is the case for the rest of the genotype. The mutation rate of the robustness locus is set to $\mu_r = 10^{-5}$. Fitness is calculated from equation (4.1), (4.2), and (4.5).

The initial ensemble of individuals were assigned three different degrees of robustness, as in the previous section. The only difference is that now robustness of these populations can evolve. We set up two environmental conditions, one in which the population starts in an optimal environment, and the other one in which it starts from a randomly selected initial environment. If starting in an optimal stable environment, the most robust genotype goes to fixation in all simulations that we examined (Figure 4.8C, D), as was the case when α was not allowed to evolve,

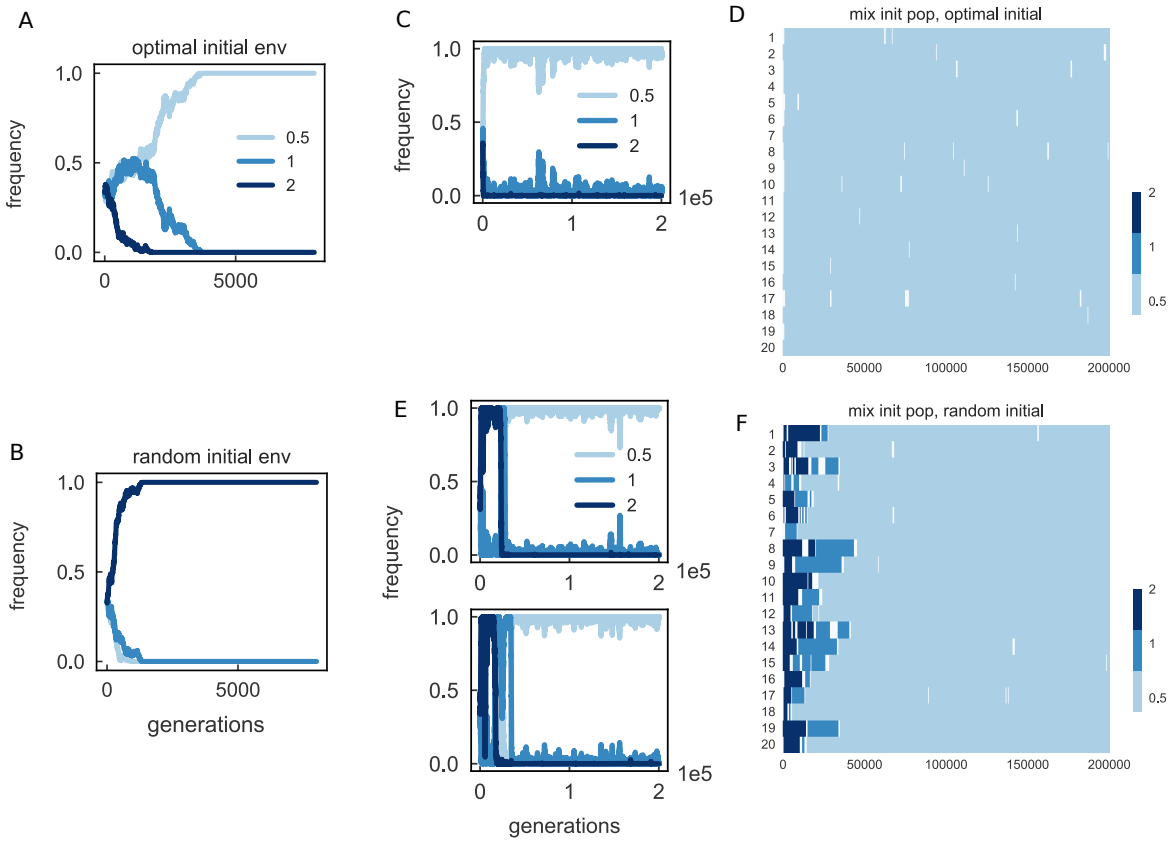


Figure 4.8: **Initial population with mixed robustness** Simulations of populations with the ability to evolve robustness. A. Introduce α parameter so that the initial population is a mixture of individuals with 3 different robustness. The three α values are 0.5, 1 and 2. C-F are results with evolved robustness, encoded by a second locus that can change α by mutation. They are all simulations starting from initial mixed 3 levels of robustness, each having equal frequency. C-D show the simulation results under optimal initial environment. C shows one typical simulation, in which the population is fixed with the most robust allele and keep it over time. D shows the fixation pattern of different robustness alleles from 20 simulations. Fixation is colored if that allele frequency is greater than 85% in the population. If none of the alleles pass that threshold, it was left blank. The fixation is evaluated at every 100 generations. We can see among all the simulations, the most robust allele fixes in the population from very early stage and is maintained fixed over time. E-F show simulation results under random initial environment. E shows two representative simulations. Initially the populations are fixed with the low robust allele, but over time, high robust allele becomes fixed in the population and is maintained. F shows the fixation plot of different robustness alleles under this random optimal environment. It is consistent from the 20 simulations that under a random environment shift, initially low robustness is preferred, but in the long-term, high robustness is selected for.

For populations starting in a randomly selected environment, low robustness is selectively favored initially, consistent with previous constant robustness results. However, as the adaptation process proceeds, a high robustness genotype will eventually take over the population (Figure 4.8E, F). This long-term behavior in the adaptation case is not seen in the case of fixed robustness.

4.4 Materials and methods

4.4.1 Forward evolutionary simulation

We use Wright-Fisher forward simulation to simulate the model. First of all, we want to simulate populations under constant environment. At each generation, only mutation, selection and genetic drift are considered. Genetic changes are on genotype, while selection is on the end phenotype. No recombination is considered. This is essentially a 2N haploid model.

The evolutionary scheme is as follows:

1. At each generation, certain number of mutations arise based on mutation rate.
2. Calculate mutant allele's fitness based on the new genotype.
3. Re-sampling the next generation alleles based on their fitness, and return to step 1.

Because we are simulation a large number of genotypes, and high genomic mutation rate, effective calculations are needed. We implemented population simulation under the c++ template framework of fwdpp model and made modification on top of that [182].

The code will be available on github.

4.4.2 Calculation of robustness

We define an extra parameter a: proportion of $\boldsymbol{v}$ which is 1. We let the elements t_{ij} in the T matrix to be i.i.d, and h_i is drawn from a Normal distribution $\mathcal{N}(0, acL)$, which is independent of γ. We can derive the probability of change in z_i when v_k mutates. From

equation 4.3

Before mutation, when L is large, using central limit theorem,

$$y_i \sim \mathcal{N}(0,\, (1+\gamma^2)acL)$$

When v_k mutates, the probability of flipping bits in z_i is:

$$\begin{aligned} P &= 2 \cdot P(y_i \text{ in the range } (-\gamma, 0)) \cdot P(\text{when } v_k \text{ mutates, } y_i \text{ increase } \gamma) \\ &= 2 \cdot [0.5 - \Phi(-\gamma/\sqrt{(1+\gamma^2)acL})] \cdot c/2 \\ &= c[0.5 - \Phi(-\gamma/\sqrt{(1+\gamma^2)acL})] \end{aligned}$$

where $\Phi(x)$ represent CDF of standard normal distribution.

4.5 Discussion

We constructed a Boolean model of phenotypic evolution to explore the influence of robustness on the evolution of traits under multi-locus genetic control. Our model features an explicit mapping of genotype to phenotype and of phenotype and environment to fitness. Robustness, the sensitivity of phenotype to changes in genotype, was described by either of two possible parametrizations, can be either be controlled explicitly or itself allowed to evolve. We employed forward population simulations to track trait evolution under both constant and abruptly changed environments.

We used a Boolean model for this investigation because it has a finite, albeit large, number of states. Moreover, only a finite possible set of rules govern its behavior. We explored two environmental regimes. The first is a constant environment favoring the current phenotype. Traits, under this regime, are maintained under mutation-selection balance, and evolutionary dynamics can be understood by nearly neutral theory, since the phenotype is close to optimum and most mutations are initially neutral or deleterious. One feature of the model is that it has a large number of loci contributing to many aspects of phenotype. Under this model, a single mutation rarely has a large effect on phenotype; a lethal mutation, for example, is not expected to occur. The effect of a mutation, if it is not completely neutral, is most likely nearly so. We found that under this scenario populations lose a small amount of mean fitness for all three values of γ, though high robustness populations equilibrate to a higher mean fitness than lower robustness populations (Figure 4.3A).

We attribute this reduction in mean fitness to a purely population genetic process: background selection against deleterious mutation [24, 26]. In particular, because our model has no recombination, a deleterious mutation anywhere in the "genome" (10,000 completely linked loci), dooms the entire genome on which it occurs to extinction, and the future ancestry of the population is governed by the sub-population of unloaded (or least loaded) genotype(s). With a broader distribution of selection coefficients, less robust populations undergo stronger background selection (Figure 4.3F-G). This background selection, in turn,

by reducing "effective" population size, also reduces the efficacy of selection against weakly deleterious mutations. Fixation of these weakly deleterious mutations by genetic drift reduces the mean fitness of the population. Finally, the population reaches a mutation-selection balance and maintains a certain mean fitness value.

To our knowledge, this is the first time a connection has been made between robustness and this population genetic process. One consequence of the reduction of population fitness with reduced robustness is that as the population genetic equilibrium is approached, selection will favor increased robustness simply to relieve the negative impact of background selection on population fitness. This will occur purely for genetic/population genetic reasons, independent of all other possible advantages of selection favoring robustness.

The second regime imposes a sudden environmental shift, sharply reducing mean fitness of the population. The selection coefficient distribution changes because the current phenotype is now far away from optimum, with a larger proportion of beneficial mutations (Figure 4.3H). This is a different evolutionary regime as compared to the stable environment regime because now adaptation is driven by fixations of beneficial mutations. It is closely related to Fisher's geometric model, where the rate of adaptation is a function of the distance to the optimal phenotype [41, 134, 72].

The initial rate of adaptation to a new phenotype, as expected, increases with decreasing robustness. This occurs because the variance in selection coefficients of new mutations is inversely proportional to robustness (Figure 4.3H), and the rate of adaptation is proportional to this variance (Fishers Fundamental Theorem of Natural Selection). Low robustness populations have greater mutational variance to fuel the adaptive process.

But we also discovered that robustness has consequences not only on the initial rate of adaptation but also on ability of selection to ever achieve an optimal phenotype (Figure 4.3B). Population fitness stalled in every case at a lower value than found under the constant environment regime for a given robustness. In addition, higher robustness resulted in lower net fitness, even after a very large number of generations of evolution. The different levels

of equilibrium mean fitness achieved by different degrees of robustness, we found, could be understood when we considered the dimension of the phenotypic space in our model, which increases as robustness decreases. Novel environments created new optimal phenotypes that were generally outside of the space achievable by any combination of genotype. To test this idea, we carried out simulations in which for each γ a randomly selected genotype was used to calculate the new optimal phenotype (rather than creating a random phenotype without regard to genotype, as we had done previously) so that this new optimum is genetically attainable by each population and robustness condition (Figure 4.3E). Notably, the initial mean fitness is higher with increasing robustness because genotypes map onto a smaller number of possible phenotypes, tightening the range of possible fitness values relating genotype to fitness. In this case, all of the simulations reached similar mean fitness levels as the corresponding ones starting from an optimal genotype, indicating that the optimal genotype was achieved in the course of adaptation. This result is in accord with the population genetic model of [65], which shows the existence of an attractor when mutation-selection balance is reached.

Phenotypic space in our model is smaller for high robustness than for low robustness. In the most extreme case—at 100% genetic robustness—all genotypes will map to a single phenotype, permitting no adaptive evolution. It is impractical to enumerate all genotype and phenotype combinations for our full model, so we illustrate this idea by visualizing all combinations in low dimension, specifically for $L = 6, n = 6$. We set up simulations using the genotype to phenotype map in our model, first simulating for a fixed $\boldsymbol{T}$ and $\boldsymbol{h}$, and use three γ values to re-scale $\boldsymbol{T}$ and enumerate all possible genotype and calculate the corresponding phenotype. In this particular simulation (Figure 4.9), when $\gamma = 0.1$, there is only one possible phenotype that can achieved from all possible genotypes; when $\gamma = 10$, eleven different phenotypes are possible in the same genotypic space.

The two environmental regimes investigated with our model — selection around an optimum and selection for a novel phenotype — may be applicable to patterns of phenotypic

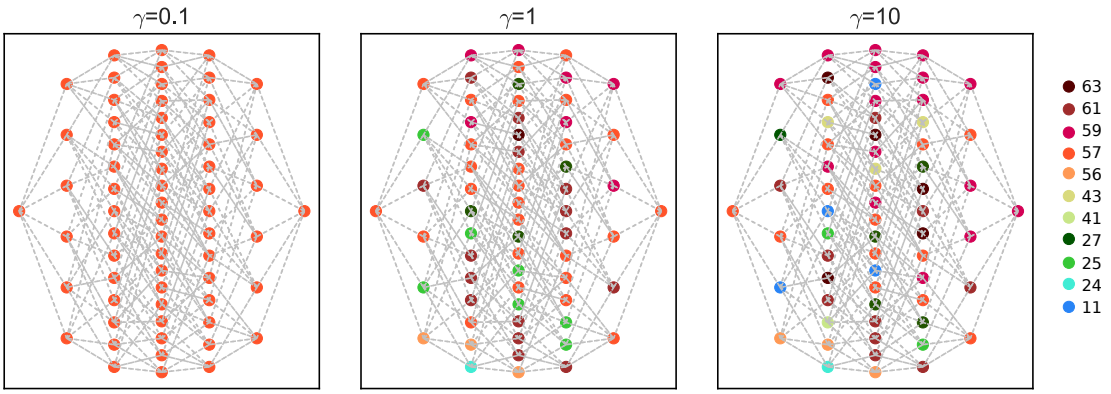


Figure 4.9: **Phenotypic space for all genotype combinations for different robustness** Simulation was done for $n = 6, L = 6$. Each circle represents a unique genotype, and each column indicates the total number of "1"s in the genotype vector (from left to right, zero "1"s to six "1"s). Each genotype is connected to its neighbor genotype that has 1 bit of difference. One set of randomly generated $\boldsymbol{T}$ and $\boldsymbol{h}$ was used for all γ values and for all combinations in genotype. For different γ values, the $\boldsymbol{T}$ was scaled to $\gamma\boldsymbol{T}$, according to the equation 4.1. The corresponding phenotype for each genotype was calculated. Each Boolean phenotype vector was then converted to an integer value, from which each unique value is assigned to a color.

evolution commonly observed in fossil record, where long-term stasis alternates with brief periods of rapid change (Punctuated equilibrium [64, 38]). But for canalized traits that are subject to long-term stabilizing selection with a fixed optimum, our model shows that genetic robustness against mutation can prevent adaptive evolution from achieving a new optimum in response to a large environmental change. In the absence of change in the robustness regime, genes that contributes to canalized traits may have limited opportunity to evolve. They may, however, be able to track small changes in the environment that select for phenotypes that are within the neighborhood accessible by mutation, as our further simulations show.

In the second part of the study, where we let robustness evolve along with the genotype that controls the phenotype (and completely linked to it), we observed initial selection for low robustness genotypes accompanied by rapid phenotypic evolution, followed by reestablishment of high robustness genotypes. Surprisingly, high robustness reappeared relatively

early in the process of adaptation, hindering further progress. As evidence, the equilibrium mean fitness reached when robustness was allowed to evolve was lower than the value reached with a fixed low robustness genotype (Figure 4.10). Currently we do not fully understand this process, and the extent to which it is dependent on assumptions of the model, such as complete linkage between genes controlling phenotype and genes controlling robustness. For our model, at least, selection favoring robustness occurs long before selection on phenotypes reaches its target (or population optimum), and it hinders further adaptive evolution of traits.

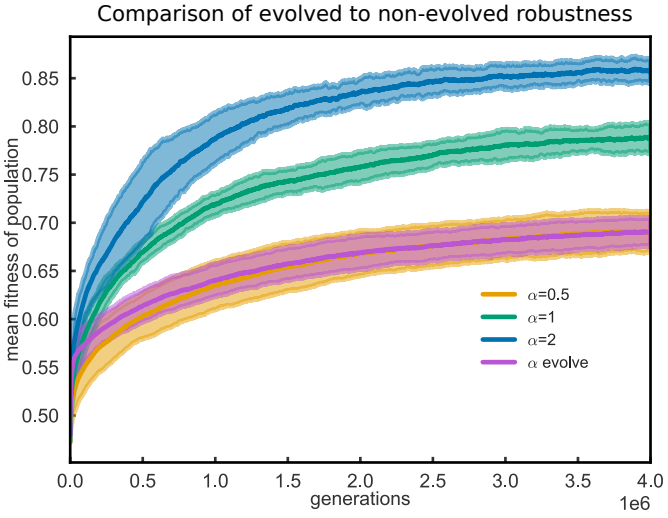


Figure 4.10: **Comparison of mean fitness change of population with evolved robustness and fixed robustness** The mean fitness curves with fixed robustness values are shown in blue, green and yellow, with α values $2, 1, 0.5$. The magenta curve shows the mean fitness change over time with evolved robustness, starting from a mixed population (with same number of individuals carrying the three different α values). The solid line shows the mean value of 20 simulations, and each upper and lower bound for each color is showing 1.5 standard deviations at each time point. When robustness is able to evolve, initially low robustness genotype is fixed in the population, and the slope of the initial mean fitness curve is steep, similar to the fixed low robustness curve (blue, $\alpha = 2$). However, when the high robustness genotype gets fixed in the population, the mean fitness grows slowly, of which trajectory falls similar to the high robustness one (yellow, $\alpha = 0.5$), resulting in a lower long-term mean fitness value compared to low robustness.

The computational efficiency of our Boolean model allowed us to replicate the evolution of simulated populations under selection for a large number of generations. In every case,

population fitness reached apparent asymptotes, for which a parabolic equation with saturation fit the data better than a power law equation (Figure 4.11). It is notable that power law trajectories do not have an upper limit, though they can decelerate. Wiser et al. [196] investigated whether the Long Term Experimental Evolution (LTEE) populations of *E. coli* appeared to be reaching an upper limit of fitness, but found that these populations have continued to improve absolute fitness for 50,000 generations, albeit with diminishing returns. The authors argue that in nature, small but steady improvement in *E. coli* fitness could continue essentially indefinitely, in contrast to our results with simulated populations, which suggest saturation. One possibility is that our simulations set a fixed target for populations to evolve towards, whereas selection in the LTEE populations is for relative competitive ability, which may not have a fixed target (even though the environment is constant). Another possibility is that in LTEE experiment, with only 50,000 generations of evolution, it simply may not be enough time for populations to get close an equilibrium. We show that, for our model, if taken fewer generations, such as the first 2.5×10^6 generations for random environment, it fit better with power law (Figure 4.11).

The period of low robustness at the onset of environmental change is usually observed in different systems, indicating low robustness is preferred for the initial stage of adaptaions. In bacteria, there is stress-induced mutagenesis, where under some stress *E. coli* cells increase mutation rates by activating the SOS pathway, changing to error-prone polymerase [14, 53], or resulting in DNA-damage, causing recombination [147, 37]; their behaviors are manifestation of lowering robustness. Other evidence supports that natural selection favors high robustness in natural population, which is similar to our stabilizing selection near optimum. Hsp90 was found to buffer standing genetic variation in yeast from natural populations, but reveals more variation (opposite effect) for *de novo* mutations in Mutation Accumulation (MA) lines [56], indicating natural selection is selecting for high robustness in the long-run. Similarly, selection on yeast TDH3 promoter was shown to reduce expression noise and thus favor robustness [120]. Polymorphisms, in this study, were shown to be less likely to increase

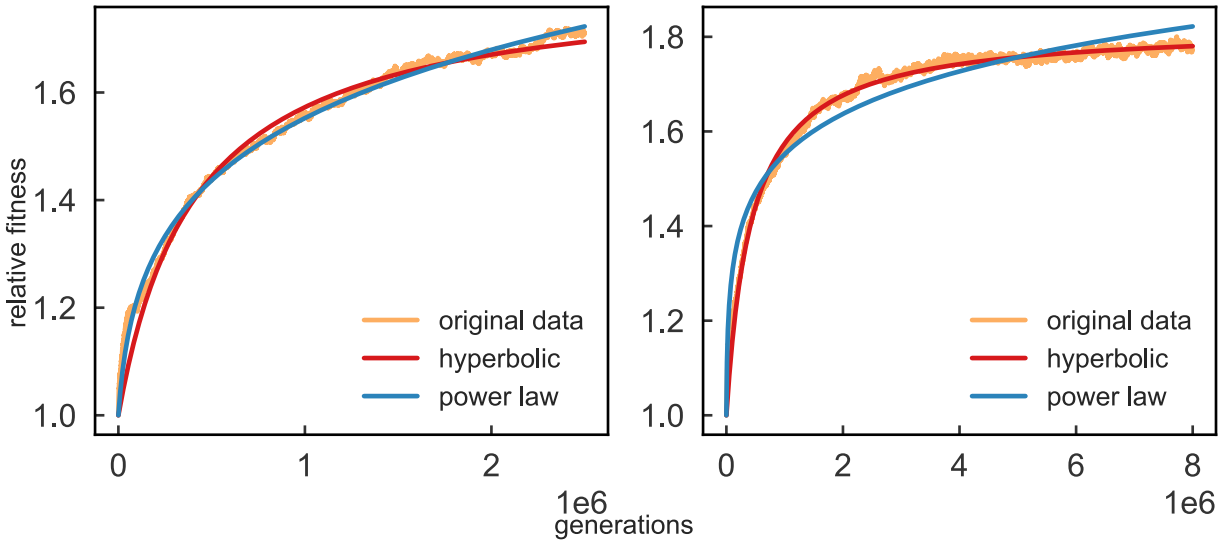


Figure 4.11: **Comparison of fitting mean fitness curve with hyperbolic and power law** Here shows fitting the same simulation that starts with a random initial environment with hyperbolic and power-law curves. The only difference between the two is the length of generations in simulations to fit with. If fitting is performed using the first 2.5×10^6 generations, power-law curve fits better (hyperbolic curve fit error: 0.020, power-law curve fit error: 0.013). If fitting is performed using 8×10^6 generations, a hyperbolic curve fits better (hyperbolic curve fit error: 0.013, power-law curve fit error: 0.033).

expression noise than random mutations, indicating that selection is favoring high robustness mutations under stabilizing selection.

CHAPTER 5

DISCUSSION AND CONCLUSIONS

In summary, I explored three questions related to canalization and evolution in my thesis. In the first study, I explored the natural variation in the conserved *eve* expression within species of *Drosophila melanogaster*. I found large deviation of gene expression could exist at the molecular level, even for this well conserved gene expression. However, such deviation of change is partially rescued in later developmental time, which results in a less severe alternation. Despite of that, the altered *eve* pattern does not fully rescue to wild type, and does have an effect on the downstream Engrailed expression. This is a good example of canalization, where a large deviation in phenotype get diminished during development. This also indicates that some rare alternations in gene regulatory network that pass through the sieve of development could potentially exist, which may allow compensatory evolution to occur.

In the second study, I followed up with a computational prediction to explain the mechanism of how two enhancers from two species with huge sequence divergence can drive similar expression patterns. By conducting rather complex genetic crosses, I compared the two enhancer reporter expression in the presence and absence of the predicted transcription factor. The experimental results are in support of the computational prediction.

Finally, I explored the general question of how canalization would evolve by using a computational model. I found under constant optimal environment, high robustness has a selection advantage, while in the face of a sudden change of environment, low robustness has a selection advantage. Low robustness has a larger phenotypic space compared to high robustness. When robustness is able to evolve with the phenotype and when adapting to a new environment, low robustness is initially selected for, but within relatively short period of time, high robustness get fixed in the population, much earlier than the population could reach for its potential mean fitness. Indicating that although high robustness genotype may impede with the population's long-term behavior to reach to a higher mean fitness value, it

still gets fixed relatively early during evolution according to our genetic model.

In the rest of this chapter, I further elaborate on issues pertaining to each chapter, and their implications for future studies.

5.1 Natural variation on *eve*

By careful quantification we show that gene expression variation in the pair-rule gene *eve* is present within a species, despite its conservation over long evolutionary time. Moreover, the observed deviation in pattern of *eve* expression, which we attribute to a *trans* deletion mutation in the *kni*, which itself has reduced early expression, is far greater than what is observed between different species. Mutational deviations in conserved traits, generally speaking, are not strong indicators of canalization because traits may be simply conserved by selection for a fixed optimum. A better indicator of canalization derives from the fact that weak early Kni expression in the mutant, and the lack of early cycle 14 Eve stripe 5, both are partially "corrected" by late cycle 14. This reduction in deviations in expression over development offer a possible definition of canalization. Yet, even this definition—a reduction in deviation of a phenotype through development—is insufficient because in this case there is also a spacing defect in Eve stripe 3/4 that is not corrected in cycle 14, and there is also a specific Engrailed patterning defect, which is regulated by Eve stripes, at a later stage of embryonic development. I believe that careful scrutiny and increased precision of measurement will reveal many subtle differences, some features of which resolve fully over time, and others of which do not.

These observations about the complexity of developmental variation highlight the challenge of defining canalization. It suggests to me the importance of using a dynamical system formulation to define canalization. The deviant *eve* pattern at the end of cycle 14 suggests that, rather than a single point attractor, when genotype changes, the underlying dynamics can change, which may result in a nearby attractor state or the deviant pattern can be a time-dependent transient state in the mutant. A complete and correct dynamical model

should be able to explain all of the observed complexity of a developmentally regulated phenotype.

Even though I investigated *eve* expression patterns for only a small sample of inbred wild-derived lines, I find phenotypic variation in expression within species. Its conservation between species is apparently not a good predictor of the extent of within-species variation. (I acknowledge that this could just be a case of the winner's curse.) It would be worthwhile to scale up this type of study to get a more comprehensive understanding about the distribution of variation in a tightly regulated developmental trait. However, imaging quantification across a developmental timecourse is a time-limiting step currently, where at most, only dozens of lines can be investigated with a reasonable effort. With the aid of micro-fluidic devices or other enabling technology, the throughput will need to increase another order of magnitude.

I showed that the intron in the *kni* may be responsible for the loss of stripe 5 at early cycle 14 for L2. To demonstrate this to be the causal effect, CRISPR technology can be employed to place the two alleles in a single isogenic genetic background. Construction of the two alternative alleles in a common genetic background will allow measurement of the possible fitness cost of this deletion allele.

The 448bp deletion of *kni* intron in L2 is not present in other lines derived from the same population, indicating that this deletion allele is probably deleterious in nature. Alleles that segregate at more intermediate frequency in the population probably will have smaller phenotypic effects. Identification of phenotypic variation in common genetic variants, therefore, will be much harder, because very careful quantification and larger sample sizes will be needed to distinguish small differences. The expression dynamics itself also change over time. A way to get around this challenge with the pattern formation system is to find a simpler phenotype that is canalized but has an easier to measure readout.

My results with *eve* also highlight the fact that a mutation of large effect, subject to likely elimination by selection, can occur in a canalized trait subject to stabilizing selection.

The mutant allele in *kni*, because it occurs at low frequency in the population, is indicative of mutation-selection balance, not stabilizing selection. Yet since we observe phenotypic conservation of *eve* patterning across all higher Diptera, I suspect that there is a general stabilizing selection with the same phenotypic target for this trait. The huge phenotypic diversity for the *eve* when looking at more basal lineages of Diptera shows that diversifying selection may have occurred earlier in evolutionary history. Subsequently, it is stabilizing selection that may be dominant for this trait in each species to maintain relative constancy of gene expression level and localization along the embryo.

5.2 More about *cad* experiments

In retrospect, the experimental design for the *cad* experiment could have been better. Originally, our goal was to test whether there is reduction or depletion of *T. put* S2E in *cad* embryos compared to *D. mel* S2E, and we did not anticipate finding a reduction of *D. mel* S2E in *cad* embryos. Therefore, we designed the experiments to compare each enhancer in its wild-type background and in a *cad* background. So, we did not think it would matter whether these two enhancers drove different reporters. However, to make quantitative comparisons across these two enhancers for the four combinations under the same conditions, the two enhancers should drive the same reporter, say *lac*Z. For *in situ* hybridizations, I noticed that it is easier to obtain a good signal to noise ratio with a *lac*Z probe than with a GFP probe. This difference is probably due to the fact that *lac*Z is a reasonably long sequence (∼3kb), while GFP is a relatively short sequence (∼500bp). If the two evolutionary versions of the enhancer had driven the same *lac*Z reporter, the experimental comparisons would have been more informative.

Another possible alternative experimental design to test the *cad* hypothesis, would be to make artificial enhancer constructs that mutate all the predicted Cad binding sites for *D. mel* S2E and *T. put* S2E. Through personal communication with Ben Vincent from Angela Depace's lab, I get to know that he conducted some experiments along this line. In his

preliminary data, when he mutated all the predicted Cad binding sites for *D. mel* S2E, he saw abolishment of expression. However when he deleted both Hb and Cad binding sites, he saw restoration of stripe 2 expression. His data indicates that Cad, rather than being a co-activator for Hb, may instead derepress Hb. Cad is a homeobox protein, which binds to a similar motif as many other homeobox transcription factors. It is worth exploring, in my opinion, possible effects from other homeobox transcription factors that express around the same time and location, and also the side effect of mutating Cad binding motifs on the binding strengths of other repressors.

Results in the literature led us to believe that anterior gap gene expression is largely preserved in the $cad^{mat+zyg}$ embryos. This provided a rationale for making *cad* mutant embryos. However, in retrospect, we cannot exclude the possibility that anterior gap gene expression levels are altered in *cad* mutants. If this is the case, the reduction of *D. mel* S2E in *cad* mutants may be caused by factors other than *cad*. The strong endogenous *eve* stripe 2 expression in *cad* mutants must be driven by sequences outside of the S2E. This makes me think about how, precisely, to define an enhancer. For the most accurate description, an enhancer should be both necessary and sufficient to drive a defined expression pattern as or a subset of the endogenous gene. However, most enhancers are identified as a minimal contiguous sequence element that can drive a pattern of expression that recapitulates wild-type expression [173, 166, 155]. Also, most enhancers are only being tested for their ability to drive expression in a qualitative way. Rarely are enhancers tested for their ability to rescue the endogenous expression level [107]. In our study, we found differences in the level of the enhancer expression compared with endogenous expression. In order to define a contiguous piece of DNA that can fully recapitulate the behavior of the endogenous locus, we should also test this DNA in different mutant conditions. If there exists such contiguous DNA that recapitulate wild-type expression under normal and mutant conditions, we should define it to be the enhancer. However, if we cannot find such a piece of contiguous DNA, the concept of modularity for an enhancer should be taken with caution. That being said, such experiments

would be unrealistically painstaking to perform exhaustively.

5.3 Robutness simulations

The framework of our model is quite general because it captures a specific genotype to phenotype map, and considers selection based on the idea that a change in environment shifts the phenotypic optimum to be selected upon. In a sense that when changing the environment in regards to the current phenotypic state of populations, different evolutionary dynamics occur and can be understood under a particular population genetic regime. For example, when the environmental optimum is close to the population's phenotypic optimum, the evolutionary process can be understood by near neutral theory. In contrast, if an environmental optimum is far away from the current phenotypic optimum, this process can be understood by a directional adaptation process (Fisher's geometric model of adaptation) where fixation of beneficial alleles dominate.

One interesting phenomenon that has been observed from this work which has not been fully understood is that when robustness is able to evolve at the same time as phenotypic evolution, high robustness evolves rather quickly in the adaptation process. As a result, the mean fitness reached by the population which has the ability to evolve robustness has a lower long-term stable fitness than the populations with low robustness. This indicates that the populations which evolve to high robustness are "trapped" in the local optimum, losing their long-term advantage to evolve to a global optimum. The underlying fitness landscape is quite rugged, so that this "trap" of local optima is possible. We currently do not have a good understanding as to why high robustness evolves so early in our model, and what triggers the initial selection advantage of fixing the high robustness allele. However, as the high robustness alleles increase in frequency under selection, though there are intermediate robustness segregating in the population, lower robustness cannot compete successfully against high robustness alleles in these cases. Perhaps the probability of getting a beneficial mutation is rare. It is worth further exploration of the model to determine if there is a way

to get populations out of the local optima.

There are some aspects or features that could be added to this model to further explore questions related to robustness and adaptation. One is to add a genotype to environmental interaction term. One outstanding question concerns the subtle difference of Waddington's genetic assimilation theory versus the Hsp90 mechanism of canalization. Hsp90 serves as a global regulator, which itself interacts with environment, controlling the phenotypic effects caused by mutations (aka robustness). However, Waddington's genetic assimilation theory involves specific gene-environment interactions, where a global regulator may be absent but genes in specific pathways respond to the environmental change. This process may also involve directional changes. Whether there is a directional change or not may depend on whether the organisms have experienced such changes in the past, in other words, whether those directional changes may have been the result of natural selection. It would be interesting to explicitly test the two scenarios by constructing a global regulator that controls robustness versus a pathway-specific regulator that interacts with the environment. Also it would be interesting to explore how natural selection may act on the directionality of response for different environments. Along the same line of thought, modularity may arise under this scenario, as many studies have already discovered. It would be very useful to investigate how different properties arise based on different assumptions in the model. With more emergent properties, evolving new traits from the current canalized state may not be as difficult to achieve as we initially thought.

REFERENCES

[1] Akam, M. (1987). The molecular basis for metameric pattern in the *Drosophila* embryo. *Development 101*, 1–22.

[2] Alon, U. (2007). Network motifs: theory and experimental approaches. *Nature Reviews Genetics 8*(6), 450–461.

[3] Ancel, L. W. (2000). Undermining the baldwin expediting effect: does phenotypic plasticity accelerate evolution? *Theoretical Population Biology 58*(4), 307–319.

[4] Ancel, L. W., W. Fontana, et al. (2000). Plasticity, evolvability, and modularity in rna. *Journal of Experimental Zoology 288*(3), 242–283.

[5] Anderson, D. W., A. N. McKeown, and J. W. Thornton (2015). Intermolecular epistasis shaped the function and evolution of an ancient transcription factor and its dna binding sites. *Elife 4*, e07864.

[6] Arnold, C. D., D. Gerlach, C. Stelzer, L. M. Boryn, M. Rath, and A. Stark (2013). Genome-wide quantitative enhancer activity maps identified by STARR-seq. *Science*, 1074–1077.

[7] Arnosti, D. N., S. Barolo, M. Levine, and S. Small (1996). The eve stripe 2 enhancer employs multiple modes of transcriptional synergy. *Development 122*, 205–214.

[8] Arnosti, D. N. and M. M. Kulkarni (2005). Transcriptional enhancers: Intelligent enhanceosomes or flexible billboards? *Journal of Cellular Biochemistry 94*(5), 890–898.

[9] Azpiazu, N. and M. Frasch (1993). *tinman* and *bagpipe*: Two homeo box genes that determine cell fates in the dorsal mesoderm of *Drosophila*. *Genes and Development 7*, 1325–1340.

[10] Banerji, J., S. Rusconi, and W. Schaffner (1981). Expression of a beta-globin gene is enhanced by remote SV40 DNA sequences. *Cell 27*, 299–308.

[11] Barrière, A., K. L. Gordon, and I. Ruvinsky (2011). Distinct functional constraints partition sequence conservation in a *cis*-regulatory element. *PLoS Genetics 7*, e1002095.

[12] Barrière, A. and I. Ruvinsky (2014). Pervasuve divergence of transcriptional gene regulation in Caenorhabditis Nematodes. *PLoS Genetics 10*, e1004435.

[13] Berman, B. P., Y. Nibu, B. D. Pfeiffer, P. Tomancak, S. E. Celniker, M. Levine, G. M. Rubin, and M. B. Eisen (2002). Exploiting transcription factor binding site clustering to identify cis-regulatory modules involved in pattern formation in the *Drosophila* genome. *Proceedings of the National Academy of Sciences USA 99*, 757–762.

[14] Bjedov, I., O. Tenaillon, B. Gerard, V. Souza, E. Denamur, M. Radman, F. Taddei, and I. Matic (2003). Stress-induced mutagenesis in bacteria. *Science 300*(5624), 1404–1409.

[15] Boyle, A. P., S. Davis, H. P. Shulha, P. Meltzer, E. H. Margulies, Z. Weng, T. S. Furey, and G. E. Crawford (2008). High-resolution mapping and characterization of open chromatin across the genome. *Cell 132*(2), 311–322.

[16] Buenrostro, J. D., P. G. Giresi, L. C. Zaba, H. Y. Chang, and W. J. Greenleaf (2013). Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, dna-binding proteins and nucleosome position. *Nature Methods 10*(12), 1213–1218.

[17] Bullaughey, K. (2011). Changes in selective effects over time facilitate turnover of enhancer sequences. *Genetics 187*, 567–582.

[18] Bullock, S. L., M. Stauber, A. Prell, J. R. Hughes, D. Ish-Horowicz, and U. Schmidt-Ott (2004). Differential cytoplasmic mRNA localisation adjusts pair-rule transcription factor activity to cytoarchitecture in dipteran evolution. *Development 131*, 4251–4261.

[19] Busturia, A. and P. A. Lawrence (1994). Regulation of cell number in *Drosophila*. *Nature 370*, 561–563.

[20] Cannavò, E., P. Khoueiry, D. A. Garfield, P. Geeleher, T. Zichner, H. E. Gustafson, L. Ciglar, J. O. Korbel, and E. E. M. Furlong (2016). Shadow enhancers are pervasive features of developmental regulatory networks. *Current Biology 26*, 38–51.

[21] Carreira, V. P., J. Mensch, E. Hasson, and J. J. Fanara (2016). Natural genetic variation and candidate genes for morphological traits in drosophila melanogaster. *PloS One 11*(7), e0160069.

[22] Carroll, S. B., A. Laughon, and B. S. T. . (1988). Expression, function and regulation of the *hairy* segmentation protein in the *Drosophila* embryo. *Genes and Development 2*, 883–890.

[23] Cassidy, J. J., A. R. Jha, D. M. Posadas, R. Giri, K. J. Venken, J. Ji, H. Jiang, H. J. Bellen, K. P. White, and R. W. Carthew (2013). mir-9a minimizes the phenotypic impact of genomic diversity by buffering a transcription factor. *Cell 155*(7), 1556–1567.

[24] Charlesworth, B. (2013). Background selection 20 years on the wilhelmine e. key 2012 invitational lecture. *Journal of Heredity 104*(2), 161–171.

[25] Charlesworth, B., E. Lande, and M. Slatkin (1982). A neo-darwinian commentary on macroevolution. *Evolution 36*(3), 474–498.

[26] Charlesworth, B., M. Morgan, and D. Charlesworth (1993). The effect of deleterious mutations on neutral molecular variation. *Genetics 134*(4), 1289–1303.

[27] Chou, T. B. and N. Perrimon (1992). Use of a yeast site-specific recombinase to produce female germline chimeras in *Drosophila*. *Genetics 131*(3), 643–653.

[28] Clark, A. G., M. B. Eisen, D. R. Smith, C. M. Bergman, B. Oliver, T. A. Markow, T. C. Kaufman, M. Kellis, W. Gelbart, V. N. Iyer, et al. (2007). Evolution of genes and genomes on the drosophila phylogeny. *Nature 450*(7167), 203–218.

[29] Consortium, . G. P. et al. (2010). A map of human genome variation from population-scale sequencing. *Nature 467*(7319), 1061–1073.

[30] Cramer, P., K. J. Armache, S. Baumli, S. Benkert, F. Brueckner, C. Buchen, G. E. Damsma, S. Dengl, S. R. Geiger, A. J. Jasiak, et al. (2008). Structure of eukaryotic rna polymerases. *Annual Review of Biophysics 37*, 337–352.

[31] Crombach, A. and P. Hogeweg (2008). Evolution of evolvability in gene regulatory networks. *PLoS Computational Biology 4*(7), e1000112.

[32] Davis, G. K. and N. H. Patel (2002). Short, long, and beyond: Molecular and embryological approaches to insect segmentation. *Annual Review of Entomology 47*, 669–699.

[33] Dobzhansky, T. (1937). *Genetics and the Origin of Species*, Volume 11. Columbia University Press.

[34] Draghi, J. A., T. L. Parsons, G. P. Wagner, and J. B. Plotkin (2010). Mutational robustness can facilitate adaptation. *Nature 463*(7279), 353–355.

[35] Driever, W. and C. Nüsslein-Volhard (1988). A gradient of Bicoid protein in *Drosophila* embryos. *Cell 54*, 83–93.

[36] Durrett, R. and D. Schmidt (2008). Waiting for two mutations: with applications to regulatory sequence evolution and the limits of Darwinian evolution. *Genetics 180*, 1501–1509.

[37] Echols, H. (1981). Sos functions, cancer and inducible evolution. *Cell 25*(1), 1–2.

[38] Eldredge, N. and S. J. Gould (1972). Punctuated equilibria: an alternative to phyletic gradualism. In T. J. M. Schopf (Ed.), *Models in Paleobiology*, pp. 82–115. Freeman, Cooper, San Francisco.

[39] Ephrussi, A. and D. S. Johnston (2004). Seeing is believing: The Bicoid morphogen gradient matures. *Cell 116*, 143–152.

[40] Eshel, I. and C. Matessi (1998). Canalization, genetic assimilation and preadaptation: a quantitative genetic model. *Genetics 149*(4), 2119–2133.

[41] Fisher, R. A. (1930). *The genetical theory of natural selection: a complete variorum edition*. Oxford University Press.

[42] Fisher, S., E. A. Grice, R. M. Vinton, S. L. Bessling, and A. S. McCallion (2006). Conservation of RET regulatory function from human to zebrafish without sequence similarity. *Science 312*, 276–9.

[43] Foe, V. E. and B. M. Alberts (1983). Studies of nuclear and cytoplasmic behaviour during the five mitotic cycles that precede gastrulation in *Drosophila* embryogenesis. *The Journal of Cell Science 61*, 31–70.

[44] Force, A., M. Lynch, F. B. Pickett, A. Amores, Y. Yan, and J. Postlethwait (1999). Preservation of duplicate genes by complementary, degenerative mutations. *Genetics 151*(4), 1531–1545.

[45] Fowlkes, C. C., K. B. Eckenrode, M. D. Bragdon, M. Meyer, Z. Wunderlich, and L. S. et al. (2011). A conserved developmental patterning network produces quantitatively different output in multiple species of *Drosophila*. *PLoS Genetics 7*.

[46] Fowlkes, C. C., C. L. L. Hendricks, S. V. E. Keränen, G. H. W. O. Rübel, M. Huang, S. Chatoor, A. H. DePace, L. Smirenko, C. Henriquez, A. Beaton, R. Weiszmann, S. Celnicker, B. Hamann, D. W. Knowles, M. D. Biggin, M. B. Eisen, and J. Malik (2008). A quantitative spatiotemporal atlas of gene expression in the *drosophila* blastoderm. *Cell 133*, 364–374.

[47] 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 and Development 1*, 981–995.

[48] 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*, 2527–2538.

[49] Fujioka, M., J. B. Jaynes, and T. Goto (1995). Early *even-skipped* stripes act as morphogenetic gradients at the single cell level to establish *engrailed* expression. *Development 121*, 4371–4382.

[50] Fujioka, M., H. Mistry, P. Schedl, and J. B. Jaynes (2016). Determinants of chromosome architecture: Insulator pairing in cis and in trans. *PLoS Genetics 12*(2), e1005889.

[51] Fujioka, M., G. L. Yusibova, N. H. Patel, S. J. Brown, and J. B. Jaynes (2002). The repressor activity of Even-skipped is highly conserved, and is sufficient to activate engrailed and to regulate both the spacing and stability of parasegment boundaries. *Development 129*, 4411–21.

[52] Futuyma, D. (2009). *Evolution*. Sunderland, Mass: Sinauer Associates.

[53] Galhardo, R. S., P. J. Hastings, and S. M. Rosenberg (2007). Mutation as a stress response and the regulation of evolvability. *Critical Reviews in Biochemistry and Molecular Biology 42*(5), 399–435.

[54] Gallo, S. M., D. T. Gerrard, D. Miner, M. Simich, B. Des Soye, C. M. Bergman, and M. S. Halfon (2011). REDfly v3.0: toward a comprehensive database of transcriptional regulatory elements in *Drosophila*. *Nucleic Acids Research 39*, D118–23.

[55] Gaul, U. and H. Jäckle (1989). Analysis of maternal effect mutant combinations elucidates regulation and function of the overlap of *hunchback* and *Krüppel* gene expression in the *Drosophila* blastoderm embryo. *Development 107*, 651–662.

[56] Geiler-Samerotte, K. A., Y. O. Zhu, B. E. Goulet, D. W. Hall, and M. L. Siegal (2016). Selection transforms the landscape of genetic variation interacting with hsp90. *PLoS Biology 14*(10), e2000465.

[57] Gibson, G. (2009). Decanalization and the origin of complex disease. *Nature Reviews Genetics 10*(2), 134–140.

[58] Gibson, G. and I. Dworkin (2004). Uncovering cryptic genetic variation. *Nature Reviews Genetics 5*(9), 681–690.

[59] Gilbert, S. F. (2003). *Developmental Biology* (Seventh ed.). Sunderland, MA: Sinauer Associates.

[60] Gompel, N., B. Prud'homme, P. J. Wittkopp, V. A. Kassner, and S. B. Carroll (2005). Chance caught on the wing: *cis*-regulatory evolution and the origin of pigment patterns in drosophila. *Nature 433*(7025), 481–487.

[61] Gordon, K. L. and I. Ruvinsky (2012). Tempo and Mode in evolution of transcriptional regulation. *PLoS Genetics 8*, e1002432.

[62] Gotea, V., A. Visel, J. M. Westlund, M. A. Nobrega, L. A. Pennacchio, and I. Ovcharenko (2010). Homotypic clusters of transcription factor binding sites are a key component of human promoters and enhancers. *Genome Research 20*(5), 565–577.

[63] 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*, 413–422.

[64] Gould, S. J. and N. Eldredge (1977). Punctuated equilibria: the tempo and mode of evolution reconsidered. *Paleobiology 3*(02), 115–151.

[65] Goyal, S., D. J. Balick, E. R. Jerison, R. A. Neher, B. I. Shraiman, and M. M. Desai (2012). Dynamic mutation–selection balance as an evolutionary attractor. *Genetics 191*(4), 1309–1319.

[66] Gregor, T., D. W. Tank, E. F. Wieschaus, and W. Bialek (2007). Probing the limits to positional information. *Cell 130*, 153–164.

[67] Groth, A., M. Fish, R. Nusse, and M. P. Calos (2004). Construction of transgenic *Drosophila* by using the site-specific integrase from phage phic31. *Genetics 166*, 1775–1782.

[68] Hahn, S. (2004). Structure and mechanism of the rna polymerase ii transcription machinery. *Nature Structural & Molecular Biology 11*(5), 394.

[69] Haigh, J. (1978). The accumulation of deleterious genes in a populationmuller's ratchet. *Theoretical Population Biology 14*(2), 251–267.

[70] 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*, 1205–1212.

[71] Hare, E. E., B. K. Peterson, V. N. Iyer, R. Meier, and M. B. Eisen (2008). Sepsid *even-skipped* enhancers are functionally conserved in *Drosopila* despite lack of sequence conservation. *PLoS Genetics 4*, e1000106.

[72] Hartl, D. L. and C. H. Taubes (1996). Compensatory nearly neutral mutations: selection without adaptation. *Journal of Theoretical Biology 182*(3), 303–309.

[73] Heather, J. M. and B. Chain (2016). The sequence of sequencers: The history of sequencing dna. *Genomics 107*(1), 1–8.

[74] Hermisson, J. and G. P. Wagner (2004). The population genetic theory of hidden variation and genetic robustness. *Genetics 168*(4), 2271–2284.

[75] Houchmandzadeh, B., E. Wieschaus, and S. Leibler (2002). Establishment of developmental precision and proportions in the early *Drosophila* embryo. *Nature 415*, 798–802.

[76] Huang, S. (2012). The molecular and mathematical basis of waddington's epigenetic landscape: A framework for post-darwinian biology? *BioEssays 34*(2), 149–157.

[77] Hughes, S. C. and H. M. Krause (2001). Establishment and maintenance of parasegmental compartments. *Development 128*, 1109–18.

[78] Huxley, J. (1942). *Evolution. The Modern Synthesis.* London: George Alien & Unwin Ltd.

[79] Jaeger, J., M. Blagov, D. Kosman, K. N. Kozlov, Manu, E. Myasnikova, S. Surkova, C. E. Vanario-Alonso, M. Samsonova, D. H. Sharp, and J. Reinitz (2004). Dynamical analysis of regulatory interactions in the gap gene system of *Drosophila melanogaster*. *Genetics 167*, 1721–1737.

[80] Jaeger, J., S. Surkova, M. Blagov, H. Janssens, D. Kosman, K. N. Kozlov, Manu, E. Myasnikova, C. E. Vanario-Alonso, M. Samsonova, D. H. Sharp, and J. Reinitz (2004). Dynamic control of positional information in the early *Drosophila* embryo. *Nature 430*, 368–371.

[81] Janssens, H., S. Hou, J. Jaeger, A. R. Kim, E. Myasnikova, D. Sharp, and J. Reinitz (2006). Quantitative and predictive model of transcriptional control of the *Drosophila melanogaster even skipped* gene. *Nature Genetics 38*, 1159–1165.

[82] Janssens, H., D. Kosman, C. E. Vanario-Alonso, J. Jaeger, M. Samsonova, and J. Reinitz (2005). A high-throughput method for quantifying gene expression data from early *Drosophila* embryos. *Development, Genes and Evolution 215*, 374–381.

[83] Jarvis, E. D., S. Mirarab, A. J. Aberer, B. Li, P. Houde, C. Li, S. Y. Ho, B. C. Faircloth, B. Nabholz, J. T. Howard, et al. (2014). Whole-genome analyses resolve early branches in the tree of life of modern birds. *Science 346*(6215), 1320–1331.

[84] Jiang, C. and B. F. Pugh (2009). Nucleosome positioning and gene regulation: advances through genomics. *Nature Reviews Genetics 10*(3), 161–172.

[85] Jiang, P., M. Z. Ludwig, M. Kreitman, and J. Reinitz (2015). Natural variation of the expression pattern of the segmentation gene *even-skipped* in *drosophila melanogaster*. *Developmental Biology 123*, 106–113. doi:10.1016/j.ydbio.2015.06.019; PMCID:PMC4529771.

[86] Johnson, D. S., A. Mortazavi, R. M. Myers, and B. Wold (2007). Genome-wide mapping of in vivo protein-dna interactions. *Science 316*(5830), 1497–1502.

[87] Johnson, T. and N. Barton (2005). Theoretical models of selection and mutation on quantitative traits. *Philosophical Transactions of the Royal Society of London B: Biological Sciences 360*(1459), 1411–1425.

[88] Karolchik, D., G. P. Barber, H. Clawson, M. S. Cline, and M. D. et al. (2014). The UCSC Genome Browser database: 2014 update. *Nucleic Acids Research 42*, D764–D770.

[89] Kent, W. J. (2002). Blat–the blast-like alignment tool. *Genome Research 12*, 656–664.

[90] Keys, D. N., D. L. Lewis, J. E. Selegue, B. J. Pearson, L. V. Goodrich, R. L. Johnson, J. Gates, M. P. Scott, and S. B. Carroll (1999). Recruitment of a hedgehog regulatory circuit in butterfly eyespot evolution. *Science 283*(5401), 532–534.

[91] Kim, A. R., C. Martinez, J. Ionides, A. F. Ramos, M. Z. Ludwig, N. Ogawa, D. H. Sharp, and J. Reinitz (2013). Rearrangements of 2.5 kilobases of noncoding DNA from the *Drosophila even-skipped* locus define predictive rules of genomic *cis*-regulatory logic. *PLoS Genetics 9*, e1003243. PMCID: PMC3585115.

[92] Kirchhamer, C. V., C. H. Yuh, and E. H. Davidson (1996). Modular cis-regulatory organization of developmentally expressed genes: two genes transcribed territorially in the sea urchin embryo, and additional examples. *Proceedings of the National Academy of Sciences 93*(18), 9322–9328.

[93] Kopp, A. (2012). Dmrt genes in the development and evolution of sexual dimorphism. *Trends in Genetics 28*(4), 175–184.

[94] Kosman, D., C. M. Mizutani, D. Lemons, W. G. Cox, W. McGinnis, and E. Bier (2004). Multiplex detection of RNA expression in *Drosophila* embryos. *Science 305*, 846.

[95] Kosman, D., S. Small, and J. Reinitz (1998). Rapid preparation of a panel of polyclonal antibodies to *Drosophila* segmentation proteins. *Development, Genes and Evolution 208*, 290–294.

[96] Kuratani, S., S. Kuraku, and Y. Murakami (2002). Lamprey as an evo-devo model: Lessons from comparative embryology and molecular phylogenetics. *Genesis 34*(3), 175–183.

[97] Larkin, M. A., G. Blackshields, N. P. Brown, R. Chenna, P. A. McGettigan, H. McWilliam, F. Valentin, I. M. Wallace, A. Wilm, R. Lopez, J. D.Thompson, T. J. Gibson, and D. G. Higgins (2007). Clustal W and clustal X version 2.0. *Bioinformatics 23*, 2947–2948.

[98] Lee, C. K., Y. Shibata, B. Rao, B. D. Strahl, and J. D. Lieb (2004). Evidence for nucleosome depletion at active regulatory regions genome-wide. *Nature Genetics 36*(8), 900–905.

[99] Levine, M. (2010). Transcriptional enhancers in animal development. *Current Biology 20*, R754–R763.

[100] Li, F., T. Long, Y. Lu, Q. Ouyang, and C. Tang (2004). The yeast cell-cycle network is robustly designed. *Proceedings of the National Academy of Sciences of the United States of America 101*(14), 4781–4786.

[101] Lipson, H., J. B. Pollack, N. P. Suh, and P. Wainwright (2002). On the origin of modular variation. *Evolution 56*(8), 1549–1556.

[102] Little, S. C., G. Tkačik, T. B. Kneeland, E. F. Wieschaus, and T. Gregor (2011). The formation of the bicoid morphogen gradient requires protein movement from anteriorly localized mrna. *PLoS Biology 9*(3), e1000596.

[103] Lo, W. C., S. Zhou, F. Y. M. Wan, A. D. Lander, and Q. Nie (2015). Robust and precise morphogen-mediated patterning: trade-offs, constraints and mechanisms. *Journal of The Royal Society Interface 12*(102), 20141041.

[104] Lott, S. E., M. Kreitman, A. Palsson, E. Alekseeva, and M. Z. Ludwig (2007). Canalization of segmentation and its evolution in *Drosophila*. *Proceedings of the National Academy of Sciences USA 104*, 10926–10931.

[105] Ludwig, M. Z., C. M. Bergman, N. H. Patel, and M. Kreitman (2000). Evidence for stabilizing selection in a eukaryotic enhancer element. *Nature 403*, 564–567.

[106] Ludwig, M. Z., Manu, R. Kittler, K. P. White, and M. Kreitman (2011). Consequences of eukaryotic enhancer architecture for gene expression dynamics, development, and fitness. *PLoS Genetics 7*, e1002364. doi:10.1371/journal.pgen.1002364 PMCID: PMC3213169.

[107] Ludwig, M. Z., A. Palsson, E. Alekseeva, C. M. Bergman, J. Nathan, and M. Kreitman (2005). Functional evolution of a *cis*-regulatory module. *PLoS Biology 3*(4), e93.

[108] 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*, 949–958.

[109] Macdonald, P. M., P. Ingham, and G. Struhl (1986). Isolation, structure, and expression of *even-skipped*: a second pair-rule gene of *Drosophila* containing a homeo box. *Cell 47*, 721–734.

[110] Macdonald, P. M. and G. Struhl (1986). A molecular gradient in early *Drosophila* embryos and its role in specifying the body pattern. *Nature 324*, 537–545.

[111] Mackay, T. F. C., S. Richards, and E. A. S. et al (2012). The *Drosophila melanogaster* genetic reference panel. *Nature 482*, 173–178.

[112] Manu, M. Z. Ludwig, and M. Kreitman (2013). Sex-specific pattern formation during early *Drosophila* development. *Genetics 194*, 163–173.

[113] Manu, S. Surkova, A. V. Spirov, V. Gursky, H. Janssens, A. Kim, O. Radulescu, C. E. Vanario-Alonso, D. H. Sharp, M. Samsonova, and J. Reinitz (2009b). Canalization of gene expression and domain shifts in the *Drosophila* blastoderm by dynamical attractors. *PLoS Computational Biology 5*, e1000303. doi:10.1371/journal.pcbi.1000303 PMCID: PMC2646127.

[114] Manu, S. Surkova, A. V. Spirov, V. Gursky, H. Janssens, A. Kim, O. Radulescu, C. E. Vanario-Alonso, D. H. Sharp, M. Samsonova, and J. Reinitz (2009a). Canalization of gene expression in the *Drosophila* blastoderm by gap gene cross regulation. *PLoS Biology 7*, e1000049. doi:10.371/journal.pbio.1000049 PMCID: PMC2653557.

[115] Markstein, M., P. Markstein, V. Markstein, and M. S. Levine (2002). Genome-wide analysis of clustered dorsal binding sites identifies putative target genes in the drosophila embryo. *Proceedings of the National Academy of Sciences 99*(2), 763–768.

[116] Martinez, C., A.-R. Kim, J. S. Rest, M. Ludwig, M. Kreitman, K. White, and J. Reinitz (2014). Ancestral resurrection of the *Drosophila* S2E enhancer reveals accessible evolutionary paths through compensatory change. *Molecular Biology and Evolution 31*, 903–916. PMCID: PMC3969564.

[117] Masel, J. (2005). Evolutionary capacitance may be favored by natural selection. *Genetics 170*(3), 1359–1371.

[118] Mayr, E. (1942). *Systematics and the Origin of Species, from the Viewpoint of a Zoologist*. Harvard University Press.

[119] Meiklejohn, C. D. and D. L. Hartl (2002). A single mode of canalization. *Trends in Ecology and Evolution 17*(10), 468–473.

[120] Metzger, B. P., D. C. Yuan, J. D. Gruber, F. Duveau, and P. J. Wittkopp (2015). Selection on noise constrains variation in a eukaryotic promoter. *Nature 521*(7552), 344–347.

[121] Miles, C. M., S. E. Lott, C. L. L. Hendriks, M. Z. Ludwig, Manu, C. L. Williams, and M. Kreitman (2011). Artificial selection on egg size perturbs early pattern formation in *Drosophila melanogaster*. *Evolution 65*, 33–42.

[122] Mlodzik, M., A. Fjose, and W. J. Gehring (1985). Isolation of *caudal*, a *Drosophila* homeo box- containing gene with maternal expression, whose transcripts form a concentration gradient at pre-blastoderm stage. *The EMBO Journal 4*, 2961–2969.

[123] Monteiro, A. (2012). Gene regulatory networks reused to build novel traits. *BioEssays 34*(3), 181–186.

[124] Moris, N., C. Pina, and A. M. Arias (2016). Transition states and cell fate decisions in epigenetic landscapes. *Nature Reviews Genetics 17*(11), 693–703.

[125] Namba, R., T. M. Pazdera, R. L. Cerrone, and J. S. Minden (1997). *Drosophila* embryonic pattern repair: how embryos respond to *bicoid* dosage alteration. *Development 124*, 1393–1403.

[126] Negre, N., C. D. Brown, P. K. Shah, P. Kheradpour, C. A. Morrison, J. G. Henikoff, X. Feng, K. Ahmad, S. Russel, R. A. H. White, L. Stein, S. Henikoff, M. Kellis, and K. P. White (2010). A comprehensive map of insulator elements for the *Drosophila* genome. *PLoS Genetics 6*, e1000814.

[127] Noon, E. P. B. and N. Frankel (2015). Evolving genital structures: A deep look at network co-option. *Developmental Cell 34*(5), 485–486.

[128] Nüsslein-Volhard, C., H. G. Frohnhofer, and R. Lehmann (1987). Determination of anteroposterior polarity in *Drosophila*. *Science 238*, 1675–1687.

[129] Nüsslein-Volhard, C. and E. Wieschaus (1980). Mutations affecting segment number and polarity in *Drosophila*. *Nature 287*, 795–801.

[130] Nüsslein-Volhard, C., E. Wieschaus, and H. Kluding (1984). Mutations affecting the pattern of the larval cuticle in *Drosophila melanogaster*. I. Zygotic loci on the second chromosome. *Roux's Archives of Developmental Biology 193*, 267–282.

[131] Nüsslein-Volhardt, C. (1991). Determination of the embryonic axes of *Drosophila*. *Development (Supplement) 1*, 1–10.

[132] Olesnicky, E. C., A. E. Brent, L. Tonnes, M. Walker, M. A. Pultz, D. Leaf, and C. Desplan (2006). A caudal mRNA gradient controls posterior development in the wasp nasonia. *Development 133*, 3973–3982.

[133] Ong, C. T. and V. G. Corces (2011). Enhancer function: new insights into the regulation of tissue-specific gene expression. *Nature Reviews Genetics 12*, 283–293.

[134] Orr, H. A. (1998). The population genetics of adaptation: the distribution of factors fixed during adaptive evolution. *Evolution*, 935–949.

[135] Paaby, A. B. and M. V. Rockman (2014). Cryptic genetic variation: evolution's hidden substrate. *Nature Reviews Genetics 15*(4), 247–258.

[136] Palsson, A., N. Wesolowska, S. Reynisdóttir, M. Z. Ludwig, and M. Kreitman (2014). Naturally occurring deletions of hunchback binding sites in the *even-skipped* stripe 3+7 enhancer. *PLoS One 9*, e91924.

[137] Parter, M., N. Kashtan, and U. Alon (2008). Facilitated variation: how evolution learns from past environments to generalize to new environments. *PLoS Computational Biology 4*(11), e1000206.

[138] Patel, N. H. (1994). Imaging neuronal subsets and other cell types in whole-mount *Drosophila* embryos and larvae using antibody probes. *Methods in Cell Biology 44*, 445–487.

[139] Perry, M., A. N. Boettiger, J. P. Bothma, and M. Levine (2010). Shadow enhancers foster robustness of *Drosophila* gastrulation. *Current Biology 20*, 1562–1567.

[140] Perry, M., A. N. Boettiger, and M. Levine (2011). Multiple enhancers ensure precision of gap gene-expression patterns in the *Drosophila* embryo. *Proceedings of the National Academy of Sciences of the United States of America 108*, 13570–13575.

[141] Peter, I. S. and E. H. Davidson (2009). Modularity and design principles in the sea urchin embryo gene regulatory network. *FEBS letters 583*(24), 3948–3958.

[142] Peter, I. S. and E. H. Davidson (2011). Evolution of gene regulatory networks controlling body plan development. *Cell 144*(6), 970–985.

[143] Pool, J. E., R. B. Corbett-Detig, R. P. Sugino, K. A. Stevens, C. M. Cardeno, M. W. Crepeau, P. Duchen, J. J. Emerson, P. Saelao, D. J. Begun, and C. H. Langley (2012). Population Genomics of sub-saharan *Drosophila* melanogaster: African diversity and non-African admixture. *PLoS Genetics 8*, e1003080.

[144] Posadas, D. M. and R. W. Carthew (2014). Micrornas and their roles in developmental canalization. *Current Opinion in Genetics & Development 27*, 1–6.

[145] Poustelnikova, E., A. Pisarev, M. Blagov, M. Samsonova, and J. Reinitz (2004). A database for management of gene expression data in situ. *Bioinformatics 20*, 2212–2221.

[146] Price, T. D., A. Qvarnström, and D. E. Irwin (2003). The role of phenotypic plasticity in driving genetic evolution. *Proceedings of the Royal Society of London B: Biological Sciences 270*(1523), 1433–1440.

[147] Radman, M. (1974). Phenomenology of an inducible mutagenic dna repair pathway in escherichia coli: Sos repair hypothesis. In *Molecular and Environmental Aspects of Mutagenesis*.

[148] Reinitz, J. and D. H. Sharp (1995). Mechanism of *eve* stripe formation. *Mechanisms of Development 49*, 133–158.

[149] Rendel, J. M. (1965). Bristle pattern in scute stocks of drosophila melanogaster. *The American Naturalist 99*(904), 25–32.

[150] Richards, S., Y. Liu, and B. R. B. et al. (2005). Comparative genome sequencing of *Drosophila pseudoobscura*: chromosomal, gene, and *cis*-element evolution. *Genome Research 15*, 1–18.

[151] Rivera-Pomar, R., D. Niessing, U. Schmidt-Ott, W. J. Gehring, and H. Jäckle (1996). RNA binding and translational suppression by Bicoid. *Nature 379*, 746–749.

[152] Rohner, N., D. F. Jarosz, J. E. Kowalko, M. Yoshizawa, W. R. Jeffery, R. L. Borowsky, S. Lindquist, and C. J. Tabin (2013). Cryptic variation in morphological evolution: Hsp90 as a capacitor for loss of eyes in cavefish. *Science 342*(6164), 1372–1375.

[153] Rohr, K. B., D. Tautz, and K. Sander (1999). Segmentation gene expression in the mothmidge *Clogmia albipunctata* (Diptera, psychodidae) and other primitive dipterans. *Development, Genes and Evolution 209*, 145–154.

[154] Romano, L. A. and G. A. Wray (2003). Conservation of Endo16 expression in sea urchins despite evolutionary divergence in both *cis* and *trans*-acting components of transcriptional regulation. *Development 130*, 4187–4199.

[155] Rosée, A. L., T. Häder, H. Taubert, R. Rivera-Pomar, and H. Jäckle (1997). Mechanism and *bicoid*-dependent control of *hairy* stripe 7 expression in the posterior region of the *Drosophila* embryo. *The EMBO Journal 16*, 4403–4411.

[156] Rutherford, S. L. and S. Lindquist (1998). Hsp90 as a capacitor for morphological evolution. *Nature 396*, 336–342.

[157] Schneider, C. A., W. S. Rasband, K. W. Eliceiri, J. Schindelin, and I. A.-C. et al. (2012). NIH image to ImageJ: 25 years of image analysis. *Nature Methods 9*, 671–675.

[158] Schroeder, M. D., M. Pearce, J. Fak, H.-Q. Fan, U. Unnerstall, E. Emberly, N. Rajewsky, E. D. Siggia, and U. Gaul (2004). Transcriptional control in the segmentation gene network of *Drosophila*. *PLoS Biology 2*, e271.

[159] Scientific, T. Immunofluorescence imaging. `http://www.thermofisher.com/us/en/home/life-science/cell-analysis/cellular-imaging/immunofluorescence.html/`.

[160] Shorter, J., C. Couch, W. Huang, M. A. Carbone, J. Peiffer, R. R. Anholt, and T. F. Mackay (2015). Genetic architecture of natural variation in drosophila melanogaster aggressive behavior. *Proceedings of the National Academy of Sciences 112*(27), E3555–E3563.

[161] Siegal, M. L. and A. Bergman (2002). Waddington's canalization revisited: Developmental stability and evolution. *Proceedings of the National Academy of Sciences USA 99*, 10528–10532.

[162] Siegal, M. L. and J. Y. Leu (2014). On the nature and evolutionary impact of phenotypic robustness mechanisms. *Annual Review of Ecology, Evolution, and Systematics 45*, 495–517.

[163] Simcox, A. A. and J. H. Sang (1983). When does determination occur in *Drosophila* embryos? *Developmental Biology 97*, 212–221.

[164] Simpson, G. G. (1944). *Tempo and mode in evolution*. Number 15. Columbia University Press.

[165] Small, S., D. N. Arnosti, and M. Levine (1993). Spacing ensures autonomous expression of different stripe enhancers in the *even-skipped* promoter. *Development 119*, 767–772.

[166] Small, S., A. Blair, and M. Levine (1992). Regulation of *even-skipped* stripe 2 in the *Drosophila* embryo. *The EMBO Journal 11*, 4047–4057.

[167] Small, S., A. Blair, and M. Levine (1996). Regulation of two pair-rule stripes by a single enhancer in the *Drosophila* embryo. *Developmental Biology 175*, 314–324.

[168] Small, S., R. Kraut, T. Hoey, R. Warrior, and M. Levine (1991). Transcriptional regulation of a pair-rule stripe in *Drosophila*. *Genes and Development 5*, 827–839.

[169] Spitz, F. and E. E. Furlong (2012). Transcription factors: from enhancer binding to developmental control. *Nature Reviews Genetics 13*(9), 613–626.

[170] Stanojevic, D., S. Small, and M. Levine (1991). Regulation of a segmentation stripe by overlapping activators and repressors in the *Drosophila* embryo. *Science 254*, 1385–1387.

[171] Starr, T. N. and J. W. Thornton (2017). Exploring protein sequence-function landscapes. *Nature Biotechnology 35*(2), 125–126.

[172] Stern, D. L. and N. Frankel (2013). The structure and evolution of cis-regulatory regions: the shavenbaby story. *Philosophical Transactions of the Royal Society of London B: Biological Sciences 368*(1632), 20130028.

[173] Struhl, G., K. Struhl, and P. M. Macdonald (1989). The gradient morphogen Bicoid is a concentration-dependent transcriptional activator. *Cell 57*, 1259–1273.

[174] Sulston, J. E. and H. R. Horvitz (1977). Post-embryonic cell lineages of the nematode, caenorhabditis elegans. *Developmental Biology 56*(1), 110–156.

[175] Surkova, S., E. Golubkova, Manu, L. Panok, L. Mamon, J. Reinitz, and M. Samsonova (2013). Quantitative dynamics and increased variability of segmentation gene expression in the *Drosophila Krüppel* and *knirps* mutants. *Developmental Biology 376*, 99–112. PMCID: PMC3426564.

[176] Surkova, S., D. Kosman, K. Kozlov, Manu, E. Myasnikova, A. Samsonova, A. Spirov, C. E. Vanario-Alonso, M. Samsonova, and J. Reinitz (2008). Characterization of the *Drosophila* segment determination morphome. *Developmental Biology 313*(2), 844–862. PMCID: PMC2254320.

[177] Tautz, D. (1988). Regulation of the *Drosophila* segmentation gene *hunchback* by two maternal morphogenetic centres. *Nature 332*, 281–284.

[178] Tautz, D., R. Lehmann, H. Schnürch, R. Schuh, E. Seifert, A. Kienlin, K. Jones, and H. Jäckle (1987). Finger protein of novel structure encoded by *hunchback*, a second member of the gap class of *Drosophila* segmentation genes. *Nature 327*, 383–389.

[179] Tautz, D. and R. J. Sommer (1995). Evolution of segmentation genes in insects. *Trends in Genetics 11*, 23–27.

[180] Thom, R. (1969). Topological models in biology. *Topology 8*, 313–335.

[181] Thom, R. (1975). *Structural stability and morphogenesis*. Reading, Massachusetts, USA: W. A. Benjamin.

[182] Thornton, K. R. (2014). A c++ template library for efficient forward-time population genetic simulation of large populations. *Genetics 198*(1), 157–166.

[183] True, J. R. and E. S. Haag (2001). Developmental system drift and flexibility in evolutionary trajectories. *Evolution & Development 3*, 109–119.

[184] Turner, T. L., A. D. Stewart, A. T. Fields, W. R. Rice, and A. M. Tarone (2011). Population-based resequencing of experimentally evolved populations reveals the genetic basis of body size variation in drosophila melanogaster. *PLoS Genetics 7*(3), e1001336.

[185] Waddington, C. H. (1942). Canalization of development and the inheritance of acquired characters. *Nature 150*, 563–565.

[186] Waddington, C. H. (1953, June). Genetic assimilation of an acquired character. *Evolution 7*(2), 118–126.

[187] Waddington, C. H. (1957). *The Strategy of Genes*. London: George Allen & Unwin.

[188] Wagner, A. (1996). Does evolutionary plasticity evolve? *Evolution 50*, 1008–1023.

[189] Wagner, A. (2005). *Robustness and Evolvability in Living Systems*. Princeton University Press.

[190] Wagner, A. (2008). Robustness and evolvability: a paradox resolved. *Proceedings of the Royal Society of London B: Biological Sciences 275*(1630), 91–100.

[191] Wagner, A. (2012). The role of robustness in phenotypic adaptation and innovation. *Proceedings of the Royal Society of London B: Biological Sciences 279*(1732), 1249–1258.

[192] Wagner, G., G. Booth, and H. Bagheri-Chaichian (1997). A population genetic theory of canalization. *Evolution 51*, 329–347.

[193] Weigel, D., G. Jürgens, M. Klingler, and H. Jäckle (1990). Two gap genes mediate maternal terminal pattern information in *Drosophila*. *Science 248*, 495–498.

[194] Wheeler, M. R. (1987). Drosophilidae. In J. F. McAlpine (Ed.), *Manual of Nearctic Diptera*, Volume 2, pp. 1011–1018. Agriculture Canada, Research Branch Otawa.

[195] Wieschaus, E., C. Nüsslein-Volhard, and G. Jürgens (1984). Mutations affecting the pattern of the larval cuticle in *Drosophila melanogaster*. III. Zygotic loci on the X-chromosome and fourth chromosome. *Roux's Archives of Developmental Biology 1983*, 296–307.

[196] Wiser, M. J., N. Ribeck, and R. E. Lenski (2013). Long-term dynamics of adaptation in asexual populations. *Science 342*(6164), 1364–1367.

[197] Wittkopp, P. J. and G. Kalay (2012). *Cis*-regulatory elements: molecular mechanisms and evolutionary processes underlying divergence. *Nature Reviews Genetics 13*, 59–69.