Scaling and regulation of gene expression in the developing fly embryo

Eric Michael Smith

A Dissertation
Presented to the Faculty
of Princeton University
in Candidacy for the Degree
of Doctor of Philosophy

Recommended for Acceptance
by the Department of
Physics
Adviser: Thomas Gregor

June 2015

Abstract

The early embryo of the fruit fly *Drosophila melanogaster* has long been a model system in which to investigate questions of how cells acquire information about their positions in developing organisms. Only in the past few years have technological developments in molecular biology, microscopy, and computing enabled the measurement of precise levels of gene expression *in vivo* in the developing fly embryo at the levels of single-cell protein concentrations and single sites of mRNA transcription. In this work I discuss two projects that take advantage of these recent advances in order to push the limits of our understanding, both in how cells receive precise information about position and how this information is regulated by the genes of the patterning gene network. The first of these projects uses measurements of protein concentration, both *in vivo* and *in vitro*, in order to determine whether a key maternally-supplied patterning gene is responsible for conveying all of the information that downstream genes need to scale their expression profiles with the length of the fly embryo. We find that this maternal gene does not scale with embryo length, implying the presence of additional inputs in enabling downstream genes to scale. The second of these projects uses *in vivo* imaging of single transcription spots in the developing embryo in order to understand how the pattern of binding sites in the regulatory region of a gene controls the level of mRNA production, given known concentrations of input protein. Contrary to hypotheses suggesting that mRNA production level is a complicated function of the strength and position of every binding site present on the DNA sequence, we find that the level of mRNA production and the domain along the length of the embryo in which high levels of mRNA are produced are both very well characterized by a linear relationship with the number of endogenous binding sites present. The results of these projects can be refined and extended by further analysis of scaling in mutant embryos and by measuring expression level in embryos in which greater perturbations have been made to sequences of protein binding sites.

Acknowledgements

First and foremost, I would like to thank my adviser, Thomas Gregor, for giving a first-year atomic physics graduate student a chance to explore the field of biophysics and for always taking the time to be invested in my project and to give useful feedback. I would also like to thank William Bialek for his valuable explanations of the theoretical underpinnings of my work, Eric Wieschaus for his unceasing enthusiasm and his unparalleled understanding of the field of fruit fly embryogenesis, and Joshua Shaevitz for his help conveying common-sense fundamentals of microscopy and for troubleshooting many silly technical issues. Credit also goes to Robert Austin for reviewing this dissertation and to David Huse for serving on my defense committee. This work would also not have been possible without the dedication of Stephan Thiberge and Evangelos Gatzogiannis in helping me to wrap my head around the challenges of building and operating a homebuilt two-photon point-scanning microscope. Moreover, I am indebted to Claude Champagne, Ted Lewis, Darryl Johnson, Steve Lowe, and the rest of the support staff of the Physics Department, without whom none of the science would have been possible.

I especially wish to thank Hernan Garcia, Feng Liu, and Allyson Sgro for all of their valuable insights and scientific advice over the years, as well as all other members of the Gregor Lab past and present for many wonderful conversations and collaborations: Jacques Bothma, Hongtao Chen, Bryan Chun, Mircea Davidescu, Heinrich Grabmayr, Rebecca Khalandovsky, Albert Lin, Shawn Little, Troy Mestler, Alex Morrison, Zoom Nguyen, Mariela Petkova, Net Ratchasanmuang, Michael Tikhonov, Danny Wells, Darvin Yi, and anyone else who I may have forgotten. I am eternally grateful as well for the support of my parents, family members, and the friends I have made during my time at Princeton.

Publications and preprints associated with this dissertation

1. **Eric M. Smith***, Jia Ling*, Hernan G. Garcia*, Hongtao Chen, Stephen J. Small, and Thomas Gregor. Binding site number controls enhancer-mediated activation independently of affinity or position. 2015.

2. Feng Liu*, Tiago Ramalho*, **Eric M. Smith***, Mariela D. Petkova, Martin W. Scheeler, Ulrich Gerland, and Thomas Gregor. Bicoid is not sufficient for the spatial scaling of *Drosophila* gap gene expression patterns. 2015.

I would like to thank my collaborators Hongtao Chen, Hernan Garcia, Ulrich Gerland, Thomas Gregor, Jia Ling, Feng Liu, Mariela Petkova, Tiago Ramalho, Martin Scheeler, and Stephen Small for these projects.

Materials from this dissertation have been publicly presented at the following conferences:

1. Mid-Atlantic Society for Developmental Biology Meeting, 2015, Princeton, NJ

2. International Physics of Living Systems 2014, Munich, Germany

3. Physics of Living Systems 2012, New Haven, CT

4. Mats Gustafsson Memorial Symposium at the Janelia Farm Research Campus, 2012, Ashburn, VA

To my parents

Contents

List of Tables

List of Figures

Chapter 1

Introduction

By studying the development of the body plan of multicellular embryos, we can understand how simple initial conditions give rise to complex biological structure. In the case of the embryo of the common fruit fly (*Drosophila melanogaster*), the simple initial conditions are a lattice of roughly six thousand undifferentiated cell nuclei, and the complex biological structure is the bewildering array of cell types that make up the organs and body segments of the adult fly [37]. The process of cell specification that turns this uniform lattice of cell nuclei into structured tissue, in which cells have identities specific to their locations in the embryo, is rapid and precise: the protein concentrations of the genes responsible for differentiating cells from each other shift dramatically over the span of several minutes, and the patterns of protein concentrations of these genes along the length of the embryo are often positioned with an accuracy of roughly the distance of one cell nucleus, or a few microns [18, 26, 41]. If we want to understand the biological processes that drive this rapid and precise development, our measurement tools must be quick enough to track the rapid expression of genes and sensitive enough to record precise levels of information as it is conveyed from one gene to another. This dissertation presents two projects that use powerful quantitative measurements of protein and mRNA

concentration in the fruit fly embryo to understand the biological mechanisms driving embryonic development.

In this chapter, I give an introduction to key concepts that the research in this work builds upon: the study of genes known as morphogens that convey information about position, the issues of precision and reproducibility in the transfer of positional information in the fly embryo, and the use of experimental techniques to measure gene transcription as a function of space and time in the developing embryo.

The role of morphogens

How do the cells of the early fly embryo, uniformly positioned and undifferentiated, distinguish themselves in a very short amount of time by forming specific cell types and folding themselves into precise structures? We can start at the beginning: a little over two hours after the embryo has been deposited by the mother, roughly six thousand cell nuclei have formed just underneath the surface of the embryo, where they are closely packed into roughly 60 rows along the anterior-posterior (AP) axis, from the head to the tail of the embryo [4, 37, 41]. Within the following hour, these nuclei are separated into individual cells by the formation of cellular membranes, and soon afterward the single layer of cells just underneath the outer embryonic membrane begins to fold inward, or gastrulate, in order to form multiple layers of cells that are separated along the AP axis into various body segments that will make up the head, thorax, and abdomen of the future adult fly [4, 37]. These gastrulation folds occur at very specific positions along the surface of the embryo when compared across a large population of embryos. For example, the cephalic furrow, a long and narrow indentation of cells perpendicular to the AP axis, delineates the boundary between the head and thorax segments of the developing fly, and the position of this furrow is defined across multiple embryos with a reproducibility of roughly one cellular diameter, or 1.2–1.5% of the length of the AP axis [22, 72, 83]. Because gastrulation

folds such as the cephalic furrow occur at such reproducible positions after only three hours since deposition by the mother [37], highly accurate information about the position of each cell along the AP axis must be conveyed to that cell within that short amount of time.

Researchers have long been interested in discovering the source of the information that cells use to learn their positions in developing organisms. In the first decades of the 20th century, Morgan hypothesized that gradients of "formative substances" were responsible for differing rates of regeneration at different positions in the flatworm [81, 107], and Boveri and Hörstadius surmised that opposing gradients of some substance were responsible for patterning the larvae of sea urchins [12, 28, 47]. Gradients provide a clear means for cells in an organism to learn their positions: the information of the local concentration level of whatever chemical comprises the gradient can be used by each cell to make a decision as to which cell type it will differentiate into and how it will fold as part of a complex tissue structure. Although no such position-encoding gradients had yet been discovered in the fruit fly embryo, scientists had begun to theorize about the precise mechanism by which these gradients might convey information to all of the cells about their relative locations in the embryo. Two competing hypotheses regarding this mechanism were proposed:

1. In 1952, Turing proposed that multiple chemical gradients, which he termed *morphogens*, could be used to create complex spatial patterns of concentration from which cells could learn highly specific information about their positions, if the morphogens diffuse at different rates and interact with each other [104]. Later computer simulations demonstrated that these models, in which one morphogen acts to activate a spatial pattern and another acts to inhibit it, could predict, for instance, stationary oscillatory patterns of activator and inhibitor morphogen concentration as a function of position [36].

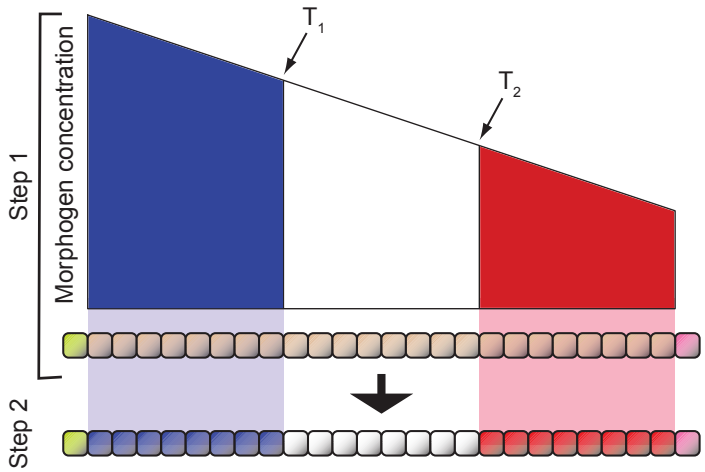

Figure 1.1: The French flag model posits that cell differentiation occurs via the creation of a morphogen gradient through diffusion and degradation (Step 1) and by cellular response to a simple read-out of the local morphogen concentration level (Step 2). T_1 and T_2 represent threshold morphogen concentrations delineating regions of distinct cell types. Taken from [57].

2. In 1969, Wolpert suggested a mechanism by which a single morphogen gradient could provide sufficient information for a cell to differentiate into one of a number of different cell types [109]: it posited a morphogen gradient that was produced by the diffusion of the morphogen chemical outward from one end of the organism, such as the anterior or posterior pole of the fruit fly embryo, and degraded upon reaching the other end of the organism. For such a morphogen, the concentration level of the morphogen at any point along the organism would be monotonically related to position, and so a cell could in principle learn its unique position by exactly reading the morphogen's concentration at that point. If the cell had several internal reference concentration thresholds that it compared the local morphogen concentration level against, it could use the knowledge of whether the local morphogen concentration exceeded none, some, or all of these thresholds in order to determine which cell type to differentiate into: it would thus adopt cell type A if the local concentration exceeded no reference threshold, cell type B if it exceeded only the lowest threshold, etc.

This is known as the "French flag" model because the resultant organism is divided into distinct regions, like the colors on the French flag, containing distinct cell types (Figure 1.1) [57].

A huge step forward in being able to test these two hypotheses of the role of morphogens in the *Drosophila* embryo was taken by Nüsslein-Volhard and Wieschaus in 1980 when 15 genes were discovered that, when absent from the fly genome, caused large sections along the head-to-tail axis of the fly larva to be completely absent [84]. Among these patterning genes is *hunchback* (*hb*), known as a *gap gene* because its absence in the genome leads to a gap in the normal body plan in developing fly larvae. It was later found that the patterning genes discovered by Nüsslein-Volhard and Wieschaus have high levels of expression in regions along the AP axis of the fruit fly embryo that largely correspond to the segments of the fly larva that are missing when each of the genes are removed from the embryo [1, 28]. However, since the discovered genes have complex expression profiles along the AP axis, the mechanism or the source morphogens by which such complex profiles could form was still unknown. This changed with the discovery of the *bicoid* (*bcd*) gene in 1986, whose absence in the genome causes the complete loss of all head and thorax segments in the anterior of the fly larva [32]. Antibody staining against Bcd protein revealed that Bcd concentration forms an exponential profile along the AP axis, peaking at the anterior of the embryo and declining to zero towards the posterior. This exponential profile is generated by diffusion of a clump of *bcd* mRNA that is deposited near the anterior pole of the embryo by the mother: as the mRNA diffuses outward, it is translated into protein and eventually degraded [23, 71]. From analyses of the expression patterns of genes such as those discovered by Nüsslein-Volhard and Wieschaus in mutant embryos lacking *bcd*, as well in as embryos in which the dosage of *bcd* was altered, *bcd* was discovered to be the hypothesized morphogen responsible for communicating information about position to these downstream genes [22]. Because the *bcd* expression profile is shaped

by diffusion and degradation and read out by downstream genes activated at certain target thresholds, *bcd* more strongly fits Wolpert's model of a morphogen than Turing's model of multiple morphogens whose expression profiles are shaped by mutual interaction. Other morphogens have since been found to influence downstream genes with expression patterns along the fly's AP axis: among them are *nanos* (*nos*), which is responsible for correct expression of *hb* in the posterior of the embryo and whose loss prevents formation of abdominal larval segments [49, 52, 97], and *torso* (*tor*), which produces a protein expression pattern localized near the anterior and posterior poles of the embryo and whose presence is necessary for correct production of anterior and posterior terminal structures [95].

Precise and reproducible transfer of information about position

Even if the specific genes that are responsible for conveying information about position along the AP axis to every cell in the first few hours of embryonic development have been identified, it still must be deduced whether the information about position encoded in gene expression levels is accurate enough to pattern the embryo correctly, as well as whether downstream genes can read in an upstream gene's concentration gradient with enough fidelity to make use of that information. In order to more concretely describe the properties that the information passed from gene to gene in the AP patterning system must possess so that cells can accurately differentiate, we distinguish from this point onward between the related but distinct notions of *precision* and *reproducibility* [41]:

1. Precision describes how accurately the information about the expression level of a gene is transferred to another gene. In the case of the input/output relation between Bcd and Hb, for instance, precision can be measured as the percentage

error with which a cell reads in the local Bcd protein concentration level, as given by the cell's output Hb protein concentration level. A very high precision indicates that the cell's Hb concentration at a given AP position is almost exactly at the level predicted by the Bcd protein concentration at that position.

2. Reproducibility is the extent to which information about the expression level of a gene is preserved throughout a population of embryos. For example, the reproducibility of the Bcd protein concentration at one particular AP position can be measured by the percentage variation in Bcd protein concentration at that location among a large number of embryos. A very high reproducibility indicates that there is almost exactly the same amount of Bcd protein at a given AP position across all embryos measured. Reproducibility can also be thought of as measuring how accurately specific markers, such as a certain Bcd concentration level, are positioned across a population of embryos: a high amount of reproducibility in this sense means that a given Bcd concentration threshold is crossed at very nearly the same AP position in every embryo in the population.

In order for cells in the developing embryo to receive accurate information about their positions along the AP axis, the transfer of information about gene expression levels in the patterning gene network must be both precise and reproducible: the expression level of each upstream gene must accurately be determined for use by downstream genes in the network, and the upstream gene's expression level must not vary to a large extent across embryos. Naïvely, we could assume that the precision and reproducibility exhibited by morphogenic markers such as the cephalic furrow are a result of each nucleus perfectly reading the protein concentration of upstream genes at its particular position and using that concentration information to determine the protein concentrations of downstream genes at that position. However, perfect read-in of input protein concentrations is impossible because the accuracy of an estimate of

local protein concentration is limited by the number of protein molecules that reach the nucleus in a given interval of time: for instance, given the measured concentration of Bcd protein halfway along the AP axis, reaching a level of precision of roughly 10% should theoretically take nearly two hours, not taking into account changes in the local Bcd protein concentration during that time [41, 102].

With such a limitation on the ability of genes to accurately transmit information about expression levels to each other, how can the embryo reliably form its body plan? Two competing hypotheses for how precision and reproducibility are introduced into the patterning network are as follows:

1. The earliest genes of the patterning network (such as Bcd) have a very low degree of reproducibility, and thus their protein concentration varies highly from embryo to embryo. Reproducibility is introduced progressively into the patterning gene network by cross-regulation of downstream genes: thus, downstream genes have more reproducible expression patterns than upstream genes, and the most downstream genes of the network have sufficient information about position to allow cells to accurately differentiate and morphogenic markers to accurately form based on their locations in the embryo [48, 58, 77, 99].

2. The transfer of information from one gene in the patterning network to the next is very precise, and as a result the earliest genes have the same level of reproducibility as later genes because expression patterns of upstream genes are faithfully read in by downstream genes, which produce their own reproducible expression patterns as a result. Thus, the variability of the protein concentration of an early gene such as Bcd from embryo to embryo should be roughly the same as that of a downstream gene such as Hunchback, which should have roughly the same level of variability as the cephalic furrow and other morphogenic markers [41, 72, 109]. This represents a conservation of information about position in the patterning gene network: because information in the system sufficient

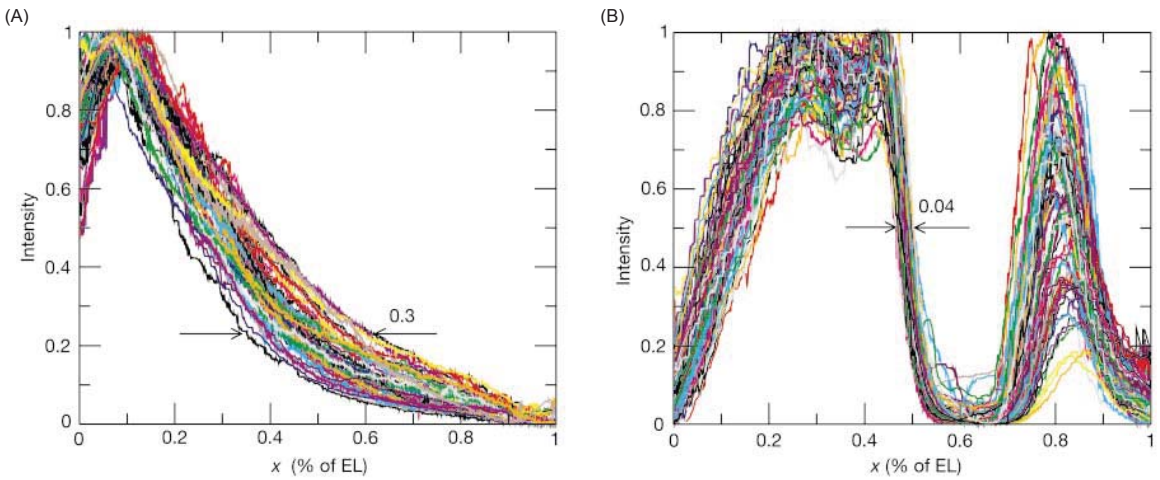

Figure 1.2: Hb is measured to be vastly more reproducible than Bcd. Each curve represents the measured Bcd (A) and Hb (B) protein concentration profile in one fixed embryo. Considering the mean Bcd concentration at the embryo's midpoint, Houchmandzadeh, *et al.* find that the standard deviation of the AP position at which each embryo's Bcd concentration profile matches this concentration is 7% of the embryo length. This is vastly greater than the equivalent quantity (1% of embryo length) in embryos' Hb concentration profiles. For each embryo and for both Bcd and Hb, the mean of the lowest 20 pixels and the mean of the highest 20 pixels along the measured curve of protein concentration are pinned to 0 and 1, respectively. Taken from Houchmandzadeh *et al.* [48].

> to pattern cells with observed precision and reproducibility is present in the earliest genes and never lost, it never has to be generated by later genes in the network.

In order to determine which of these hypotheses more correctly reflects how accurate patterning information is transferred in the patterning gene network, levels of gene expression must be measured with enough sensitivity and experimental rigor to make accurate calculations of precision and reproducibility possible. In the past 15 years, many such quantitatively-driven measurements of gene expression have been conducted, often with conflicting inferences about whether information sufficient to position all cells accurately can be found during the earliest stages of the patterning gene network [46, 56, 58, 61, 72, 99]. For example, in 2002 Houchmandzadeh, *et al.* [48], using fixed embryos stained for the presence of Bcd and Hb protein, found

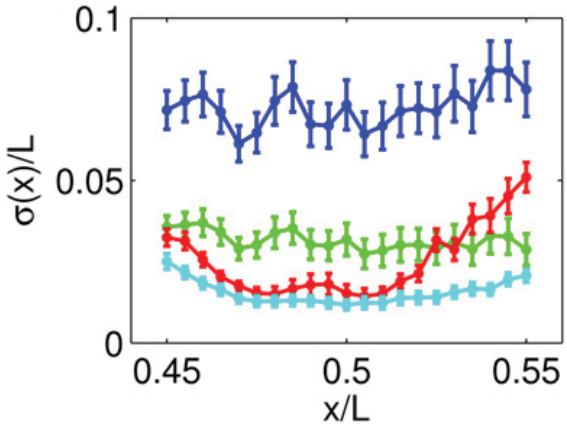

Figure 1.3: More subtle normalization of profile intensities shows that Bcd and Hb have similar levels of reproducibility. The y-axis represents the RMS error with which a certain threshold protein concentration level remains fixed at the same AP position from embryo to embryo, normalized by embryo length: lower errors correspond to greater reproducibility of profiles across the population of all embryos. The blue curve represents the RMS error in Bcd protein concentration profiles from Houchmandzadeh, *et al.*, using their normalization technique [48]; the green curve represents the RMS error in the same profiles, normalized to each other by χ^2 minimization of differences among all profiles along the AP axis. Red and cyan curves show the RMS error in Bcd and Hb expression profiles, respectively, as measured in fixed embryos by Gregor, *et al* and normalized through χ^2 minimization. Taken from Gregor *et al.* [41].

that the variability among embryos in the position along the AP axis at which the Bcd protein concentration gradient crosses a certain threshold level near the middle of the embryo is roughly 7 times as great as the variability in the position at which the Hb protein concentration gradient crosses an analogous threshold level near the middle of the embryo (Figure 1.2). This would support the first of the two reproducibility hypotheses above, as it would imply that information about position is made progressively more reproducible across embryos in a population as it moves from gene to gene in the patterning gene network.

However, questions were later raised in 2007 by Gregor, *et al.* [41] as to the validity of the normalization techniques applied to the data of Houchmandzadeh, *et al.* in order to arrive at that conclusion. By contrast, Gregor, *et al.*, applying more nuanced

methods of normalizing Bcd and Hb protein concentration levels in stained embryos, showed that the levels of reproducibility of Bcd and Hb are roughly equal (Figure 1.3). Furthermore, they found that the level of precision of the Bcd profile is roughly the maximum possible amount, given the number of Bcd proteins reaching the nucleus in a given time interval, and that that observed level of Bcd precision allows for position to be specified along the AP axis with an accuracy of one nuclear row. Thus, the information about position contained within the Bcd concentration profile should in principle be sufficient to communicate all necessary information about position along the AP axis to nuclei so that they may form cellular membranes and differentiate into different cell types with the observed accuracy of one row of cells. This is a strong confirmation of the second of the two reproducibility hypotheses listed above, because it implies that the necessary level of reproducibility present within the entire AP-axis patterning gene network can be found solely in one of the earliest genes in the network, and that that level of reproducibility is simply transferred from gene to gene in the network instead of progressively increasing as it moves from upstream to downstream genes.

Measuring transcription

However, solely measuring the levels of reproducibility and precision found in the protein expression profiles of the patterning genes leaves large gaps in understanding about the biological mechanisms by which these levels of reproducibility and precision are preserved throughout all stages of the patterning gene network. In particular, in order for the concentration of Bcd protein at one position in the embryo to serve as an input to determine an output concentration of Hb protein at that position, Bcd protein must reach and bind to a *hb* regulatory element on the genome, a certain amount of *hb* mRNA must be transcribed from the *hb* coding sequence as a result of that level of Bcd protein binding, and those *hb* mRNA molecules must be translated

into proteins in order to create the output Hb protein concentration. Thus, if we can measure spatiotemporal patterns of *hb* mRNA production with the same quantitative rigor and low experimental error with which protein concentration patterns have been measured, we will start to understand how the high level of reproducibility found in early genes in the pattering gene network is transferred from one gene to another via the processes of gene transcription and translation.

The reproducibility of *hb* mRNA transcription across a population of embryos can be compared to that of the Bcd protein concentration distribution that regulates it, as well as the Hb protein distribution that it produces, through the use of single-molecule fluorescence *in situ* hybridization (smFISH), a sample preparation and imaging technique that allows for highly quantitative measurements of mRNA transcription at single transcription sites in a cell nucleus [30, 59, 71, 94]. Using smFISH, Little, *et al.* [70] measured the fractional variation in the amount of *hb* mRNA production at individual transcription sites to be roughly 44%, compared to the roughly 10% fractional variation in the amount of Hb protein produced as a function of Bcd concentration level [41]: thus, levels of activity at single sites of *hb* transcription in the embryo are much less reproducible than the levels of Hb protein that result. Little, *et al.* argue that this discrepancy is a result of spatial averaging through mRNA diffusion and temporal averaging through mRNA accumulation [70].

One important limitation of fixed-tissue imaging techniques such as smFISH is that it only captures transcription in a particular embryo at one fixed point in time: thus, we lose any information about the dynamics of transcription in an embryo or how the variation among levels of transcription among different embryos changes over time. Garcia, *et al.* [33] refined a method for tracking transcription events in an embryo over time by labeling new mRNA molecules with fluorophores, fluorescent proteins, as soon as they are produced [31]: this method allows for the full dynamics of transcription, as opposed to transcription only at one fixed point in time, to be

compared across embryos. With access to dynamic information, fundamental quantities of the transcription process such as the rate at which mRNA is produced can be measured, which allows for more direct inferences to be made as to the mechanism by which genes in the patterning network regulate each others' expression levels in order to convey information about position.

Overview of dissertation

Drawing upon this previous research, this dissertation details two experimental and analysis projects that use quantitative measurements of protein and mRNA expression as a function of space and time to further understand the propagation of information about position in the early fruit fly embryo:

1. Previous work has studied whether levels of reproducibility of gene expression sufficient to pattern the entire embryo exist in all patterning genes. In Chapter 2, we ask a related question regarding variability in embryo size: given that the lengths of fly embryos along their AP axes are not all identical, are all patterning genes equally adept at scaling their protein concentration profiles to match the lengths of individual embryos, in order to achieve a high level of reproducibility? Our results show that, unlike the profiles of later gap genes such as Hb, the protein expression profile of the early patterning gene Bcd does not scale with embryo length, which hints that the scaling of gene expression profiles arises as a result of the collective input of multiple early morphogens in the patterning gene network as well as cross-regulation among multiple downstream genes.

2. With the advent of the technology to measure levels of mRNA transcription in live fly embryos, we not only get the chance to measure the dynamics of the transfer of information about position among patterning genes, but we also gain the opportunity to explore the biological mechanisms by which this information

is transferred through transcription. Chapter 3 discusses an experiment to characterize the effects that single regulatory elements have on the rate at which *hb* mRNA molecules are produced. Bcd proteins bind to 6 specific sequences near the *hb* coding region in order to increase the level of *hb* mRNA expression, and we can add or remove different combinations of these binding sites to see what effects this has on the resultant mRNA expression pattern. What we find is that the identities of the binding sites that are present are relatively unimportant in regulating the mRNA expression pattern, and that the level of mRNA expression and the width of the expression domain increase linearly as we increase the number of binding sites. This finding leads to several inferences about the roles of interactions between multiple Bcd proteins as well as between Bcd proteins and the transcriptional machinery.

Chapter 2

Bicoid is insufficient for the spatial scaling of *Drosophila* gap gene expression patterns

This chapter is an adaptation of a draft of a manuscript that will soon be submitted for publication. (Authors: Feng Liu, Tiago Ramalho*, Eric M. Smith*, Mariela D. Petkova, Martin W. Scheeler, Ulrich Gerland, and Thomas Gregor.)*

The maternal morphogen Bicoid (Bcd) is a primary source of positional information for the patterning of the early *Drosophila* embryo. If Bcd does not scale with embryo length, the positional read-out error from the Bcd profile should naively be up to the natural embryo length variability, e.g., 4% of the embryo length (EL) for the wild-type. However, the measured positional error of the downstream gap genes is just 1% EL. Given such low positional error in downstream genes, we ask if the Bcd gradient can actually provide sufficient information about position for *Drosophila* embryos of varying sizes. We present a method that characterizes the positional error of protein concentration profiles in a shape-independent manner, and we use this method to directly compare the error with which both maternal factors and gap genes

are positioned in embryos of varying egg lengths. We show that the positional error of Bcd exceeds that of the gap genes along the entire length of the embryo. In mutant embryos where Bcd is the only maternal positional input, the gap gene positional error is significantly larger in the posterior half than that of wild-type embryos. However, the gap gene positional error in *bcd* mutants matches wild-type levels. Therefore, our results imply both that precisely scaled gap gene expression patterns do not rely on Bcd alone and that scaling does not require Bcd. This agrees with a principal component analysis that demonstrates that the hypothesis of scaling with embryo length can be rejected for Bcd profiles, but generally not for gap gene profiles. Our results suggest that scaling of downstream gene expression patterns emerges as a collective network property, i.e., via a synergy between the combined maternal inputs and cross-regulation of the downstream genes.

2.1 Introduction

The final products of developmental patterning processes are strikingly reproducible from one individual to the next, and this reproducibility has been traced back to the earliest molecular patterning events [26, 41]. Natural variations in the intrinsic length scales of the developing organisms require specific scaling mechanisms to be established to maintain reproducible patterning. The classical model for developmental patterning is the interpretation of long-range signaling molecules (i.e. morphogens) by downstream patterning genes in a concentration-dependent manner [109, 110]. While gene expression patterns are observed to be proportional to the overall size of the system [34, 39, 48, 75], the upstream morphogen gradients are theoretically size independent if their setup is only driven by simple diffusion [40, 43]. Hence the origin of spatial scaling of the developmental patterns is not well understood, but two possible scenarios have been proposed: (1) spatial scaling of developmental patterns

stems from scaled morphogen inputs (whose setup involves mechanisms other than simple diffusion-driven); (2) scaling is a property that emerges via the interaction of the downstream gene network, with unscaled morphogen inputs being irrelevant for scaling. Several patterning systems are reported to follow the first scenario, e.g., bone morphogenic proteins (BMP) in *Xenopus laevis* embryos [6, 7, 51], Decapentaplegic (Dpp) in the *Drosophila* imaginal discs [45, 106], Notch in mouse embryos [66], but theoretically the second scenario also works as spatial pattern formation in reaction-diffusion systems can be scale-invariant without scaled inputs [85].

Anterior-posterior (AP) patterning during *Drosophila* embryogenesis has long been a model system for the study of scaling mechanisms [15, 21, 39, 40, 74]. The major focus has been the scaling of the maternal morphogen Bicoid (Bcd) [15, 21, 39, 48], which forms an exponential gradient along the AP axis and activates a series of downstream genes, such as the gap genes [23]. The length constant of the Bcd gradient was reported to scale with embryo length among different fly species [39, 40], and thus the scaling of Bcd had been proposed to account for the scaling of AP patterning across different species with egg sizes that varied by up to a factor of 5. However, for the scaling of AP patterning within a given species, the role of the Bcd gradient has been controversial. One early study argued that the gradient was too noisy to establish the precision and scaling of downstream gap genes, and that a "filter" mechanism must be present [48]. Subsequently, the scaling of segmentation patterning was suggested to be independent from Bcd gradients [46], and the repressive mutual cross-regulation between gap genes was reported to "canalize" patterning, improving the precision and driving the dynamic shifts of patterns [58, 77]. Therefore these results support the second scenario, that AP patterning is driven by the interaction of the downstream gene network.

On the other hand, recent experiments have demonstrated that Bcd gradients are highly reproducible such that they should be able to confer scaled information

about position with higher accuracy than the 4% natural egg size fluctuations [41]. This finding inspired several proposed mechanisms following the first scenario, i.e. a spatially scaled morphogen gradient from which downstream genes acquire spatial scaling. For instance, the nuclear trapping model [17, 42] predicts that Bcd gradients have a scaled length constant, but several lines of experimental evidence contradict such a prediction [44, 78]. It has also been suggested that scaled Bcd gradients are generated by volume-dependent modulation of the amplitude of the Bcd gradient [15], or by volume-dependent modulation of the length constant of the Bcd gradient [16], but no experimental evidence has conclusively shown that this would be sufficient to generate the scaling of the downstream gap genes. Another controversy arose form the suggestion that the positional precision and the scaling of Bcd-target genes could stem from a pre-steady-state decoding of Bcd gradients [20, 21, 54]. Here we address these issues by focusing on the underlying mechanism for the spatial within-species scaling of segmentation patterns in the early *Drosophila* embryo. In particular we ask whether the Bcd gradient is able to precisely provide scaling information for embryos of varying sizes.

Analyzing scaling mechanisms in eggs of a given fly species has been challenging since natural embryo size fluctuations from one individual to the next are small ($< 4\%$), even in cases where these fluctuations have been artificially increased [15, 78]. A great hurdle for progress has been the absence of a method that compares unequivocally the scaling of different protein expression profiles regardless of their shape. The traditional methods based on simple correlation analysis [15, 16, 34, 48] or on positional error calculation using standard error propagation methods [26, 41] are often limited to protein profiles with specific shapes, thus preventing a direct comparison between the scaling property of, e.g., the Bcd profile and its downstream gap gene profiles. Here we present a method named dynamic profile warping (DPW) that is based on dynamic time warping (DTW) [53, 82, 89, 90] to quantify the scaling behav-

ior of both maternal genes and gap genes by measuring local positional errors. We demonstrate that DPW overcomes the limitations of traditional methods for measuring scaling behavior, and that the positional error calculated with DPW provides a universal, shape-independent metric for characterizing the spatial scaling of gene expression profiles.

We applied this method to an extensive set of new data in order to understand the origin of gap gene scaling in the *Drosophila* embryo. We have collected various data sets for maternal gradients, including immunostained Bcd, Caudal (Cad) and Capicua (Cic) gradients from artificially selected inbred fly lines with large and small embryos (LE&SE, $N = 512$), live Bcd-GFP gradients from temperature-varied embryos ($N = 48$), and live Bcd-GFP gradients from embryos at one fixed temperature ($N = 1018$). We also collected several data sets of embryos in various genetic backgrounds, immunostained against the four gap genes Knirps (Kni), Kruppel (Kr), Giant (Gt) and Hunchback (Hb). Our results indicate that the positional error of Bcd is too large to position the gap genes to the observed level of precision, and the data of the mutant embryos reveal that Bcd alone is insufficient to position the gap genes to wild-type precision without input from other maternal factors such as Nanos and Torso. Furthermore, Bcd is unnecessary for precise positioning of gap genes because the gap gene positional error in *bcd* mutants matches that of wild-type levels, despite severe alterations in gene expression patterns. Our combined results suggest an alternative scenario where each maternal gradient only provides partial scaling to the downstream network, and final scaled patterns result from collective interactions within the gap gene network.

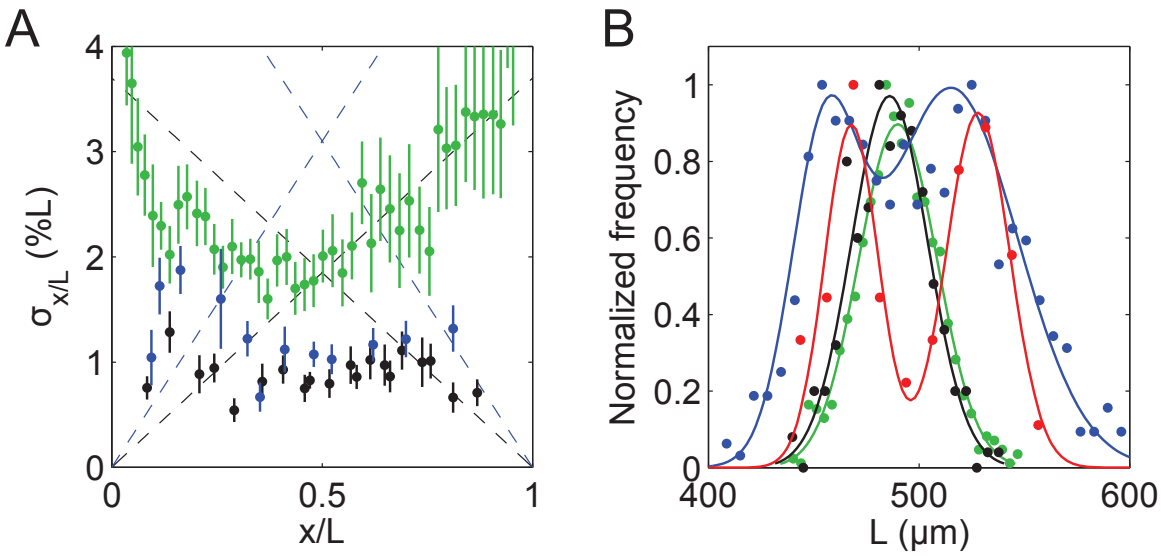

Figure 2.1: A scaling mechanism is needed. **(A)** Measured positional errors of gap gene expression boundaries from immunostained embryos at 42 ± 5 min into nc14 for OreR embryos (black, $N = 10$, [26]) and for the mixture of long and short embryos of artificial-egg-size-selected inbred fly lines (blue, $N = 15$, from [78]). Positional errors of live Bcd-GFP gradients (green, $N = 26$) measured as in [41, 72]. All error bars are standard deviations calculated from bootstrapping. Dashed lines represent the expected positional error due to egg length variation according to $\sigma_{\text{pos}}(x) = \sigma_L * x/L$ (see text) with $\sigma_L = 3.7 \pm 0.1\%$ for OreR (black), $\sigma_L = 6.2 \pm 0.2\%$ for the LE&SE mixture (blue). These lines are the graphical representation of the lower bounds of the positional error of completely unscaled profiles. The lines of positive (negative) slope represent the minimum positional error of unscaled profiles that completely coincide in absolute coordinate space with respect to the anterior (posterior) pole. **(B)** Egg length distributions of wild-type embryos (OreR, black, $N = 214$), Bcd-GFP-expressing embryos bred at 25 °C (green, $N = 968$), a mixture of Bcd-GFP-expressing embryos bred at 18 °C and 29 °C (red, $N = 46$), and a mixture of large and small embryos (blue, $N = 519$), from artificial egg-size-selected inbred fly lines LE and SE. Dots and lines represent the binned values and the Gaussian function based fitting results.

2.2 Necessity of a scaling mechanism in the early fly embryo

Embryo-to-embryo size fluctuations in natural populations of *Drosophila melanogaster* [15, 41] are relatively small. Therefore it is *a priori* unclear whether specific scaling mechanisms are needed to generate the proportionality of gene expression patterns

with egg size. The necessity for such mechanisms only becomes apparent when considering the precision of spatial gene expression markers, such as pattern boundaries or peaks along the AP axis. The precision of the exact positions of such a marker depends on how well the pattern scales with embryo length. If no scaling mechanism were in place, natural egg length variation would induce a positional error along the AP axis (Figure 2.1A). This inherent error from scaling thus sets a lower bound on the positional error that is found in expression patterns in the absence of a scaling mechanism. The magnitude of the inherent error in such "unscaled" expression patterns is position-dependent: if a given position along an expression profile has a fixed absolute distance from a reference point (e.g. the embryo's poles), its relative position with respect to embryo length should have a positional error σ_{pos} of at least $\sigma_{\mathrm{pos}}(x) = \sigma_L * x/L$, where $\sigma_L = \delta_L/\bar{L}$ is the ratio of the standard deviation δ_L of the length distribution of a population of embryos and the population mean $\bar{L}$; x is the absolute distance from the reference point. Therefore the linear function $\sigma_{\mathrm{pos}}(x)$ defines a lower bound for the positional error of unscaled expression patterns. If the positional error of a gene expression profile falls below these anteriorly (or posteriorly) anchored lines, which we will term the unscaled lines hereafter, the profile must contain more positional precision than what could be conveyed in a perfectly unscaled profile, and therefore such a profile must contain information about embryo size.

Given the measured natural length variation of *Drosophila* eggs, the lines of perfect unscaling provide a quantitative way to test whether a given gene's expression profile scales with embryo length (Figure 2.1A). Egg length variation in a wild-type (OreR) population at 25 °C is around 4% ($\bar{L} \pm \delta_L = 486 \pm 18$ µm; $N = 214$; $\sigma_L = 3.7 \pm 0.1\%$ from bootstrapping; Figure 2.1B), which is small in comparison to the variability of egg lengths between different fly species that differ by up to a factor of five [39, 40]. However, given the inherent positional error σ_{pos} in unscaled gene expression profiles, the natural wild-type egg length variation leads to a positional error of 2% L in

the central region of the egg and even larger in the posterior half of the egg (Figure 2.1A). This variation far exceeds the actual positional error of the gap gene expression boundaries or peaks, which is less than 1% L along nearly the whole AP axis (Figure 2.1A) [26]. We observe a similar scenario in inbred fly lines artificially selected for small and large egg lengths (SE and LE with length distributions of 451 ± 14 µm and 521 ± 25 µm, respectively; see Materials and Methods) [78], leading to an effective relative change in mean egg length between the two fly lines of 15.5% (Figure 2.1B), but no change in gap gene positional error is observed (Figure 2.1A). Thus, even the relatively small egg length variability within *Drosophila melanogaster* is large enough to conclude that gene expression profiles within the *Drosophila* segmentation network scale with egg length at the level of the gap genes. The question remains as to how this scaling emerges.

2.3 Measuring spatial scaling of gene expression profiles in a shape-independent manner

To test whether the upstream inputs to the gap genes, such as the maternal factor Bcd, are responsible for gap gene scaling, we ask whether the positional error of Bcd matches the level of the positional error of the gap genes, if so then gap gene scaling can be directly inferred from Bcd scaling, otherwise if the positional error of Bcd exceeds that of the gap genes, an additional source of positional precision is necessary. To answer this question, it is important to directly compare the positional error of both the maternal factor Bcd and the gap genes with the same method, regardless of the shape of their expression profiles. Furthermore, it is possible that Bcd is sufficient for conveying positional precision to the gap genes only in the central region of the AP axis but not along its entire length according to the volume-dependent production

model (Figure 2.1A, [15]), a method for measuring positional error has to be local, i.e. dependent on positions along the AP axis.

Using these criteria of shape-independence and locality, we assessed commonly used methods for measuring positional error in fly embryos to determine their suitability for reliably testing the sufficiency of Bcd to deliver positional precision to the gap genes (Figures 2.1A and 2.2). For instance, with the correlation analysis one measures the spatial scaling of the gap genes by finding the correlation between the absolute distances of the patterning markers (e.g., profile boundaries or peaks) from a reference point and the embryo length (Figure 2.2A-C). But this is unsuitable for measuring the positional error of Bcd because the Bcd expression profile has no well-defined boundaries or peaks. Instead, one often extracts a length constant λ or amplitude C_0 by curve fitting to the Bcd expression profiles. A method that combines shape-independence and locality is obtained by error propagating expression level variations among profiles into a variation in AP position [26, 41]. However, when applied to single gap genes this method leads to singularities at the profile's peaks and troughs, thus limiting its usefulness to partial regions of the AP axis (Figures 2.3D and 2.2).

Since different methods have different systematic effects, they could skew the difference of the positional errors and often lead to inconsistent results when comparing Bcd with gap genes. We measured the Bcd gradients (at 16 ± 2 min into nc14, bred at 25 °C) in a set of over 396 live embryos expressing Bcd-GFP. With correlation analysis, we find that λ is uncorrelated with embryo length L ($r = -0.01 \pm 0.04$, $p = 0.82$), whereas λ/L is correlated with L ($r = -0.42 \pm 0.04$, $p = 2.2 \times 10^{-18}$) (Figure 2.4A,B), implying that the shape of the Bcd gradient is independent of embryo length. On the other hand, we observe a weak correlation between C_0 and L (Figure 2.5B), suggesting that the volume-dependent production model might be valid [15] thus Bcd could be sufficient for the scaling of gap genes. However, this is inconsistent

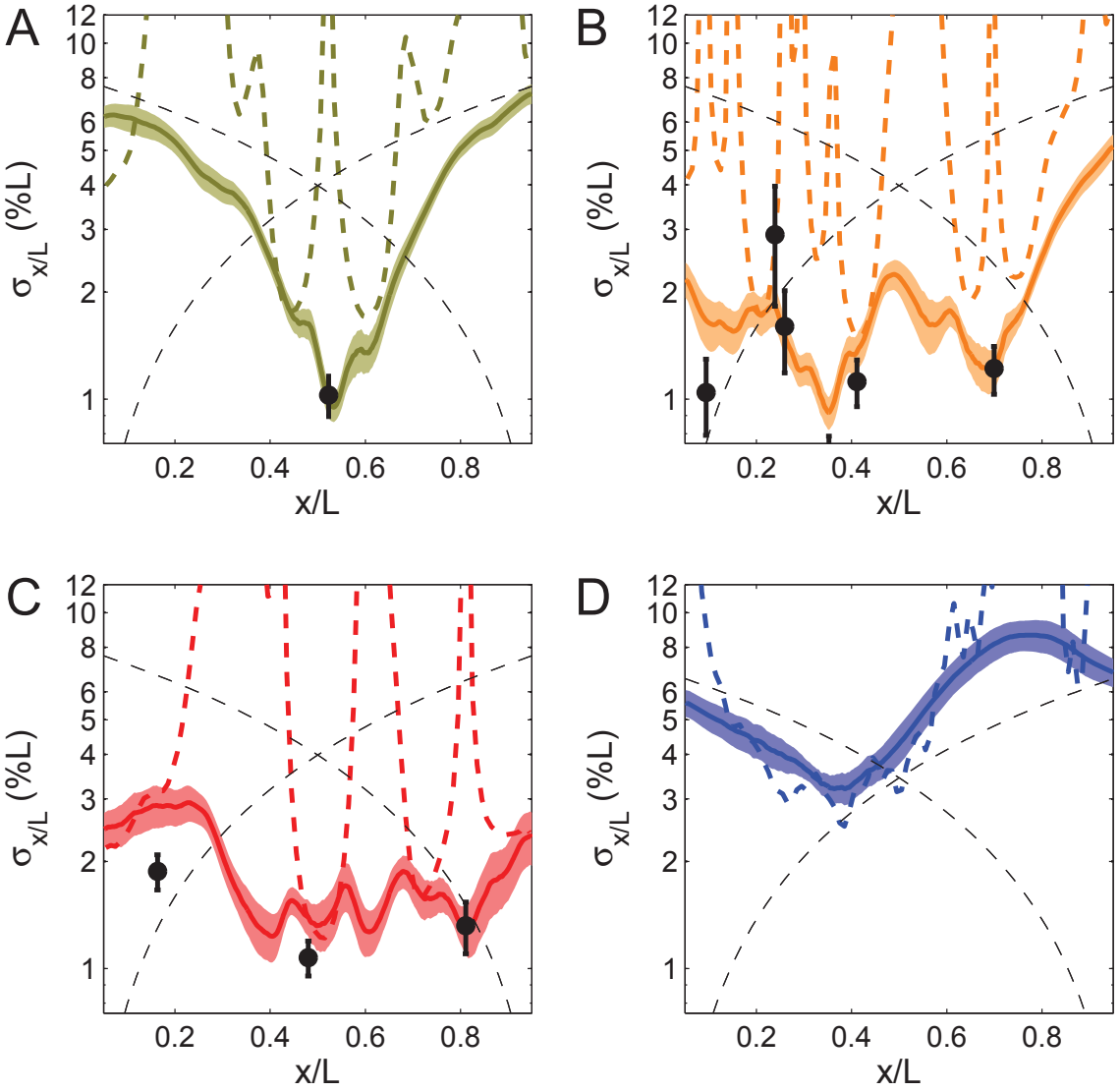

Figure 2.2: Comparison of the positional error of genes calculated using DPW (solid curves), the error-propagated intensity method (dashed curves, [26, 41]), and an analysis of the variability of individual peaks and boundaries (black error bars). Profiles are of Kr (A), Gt (B), and Hb (C) 35 ± 3 minutes into nc14 ($N = 26$), and Bcd (D) 16 ± 5 minutes into nc14 ($N = 20$). Shaded regions and error bars represent standard errors from bootstrapping. Black dashed lines represent the expected positional error due to 6.2% variability of the embryo length for LE&SE embryos.

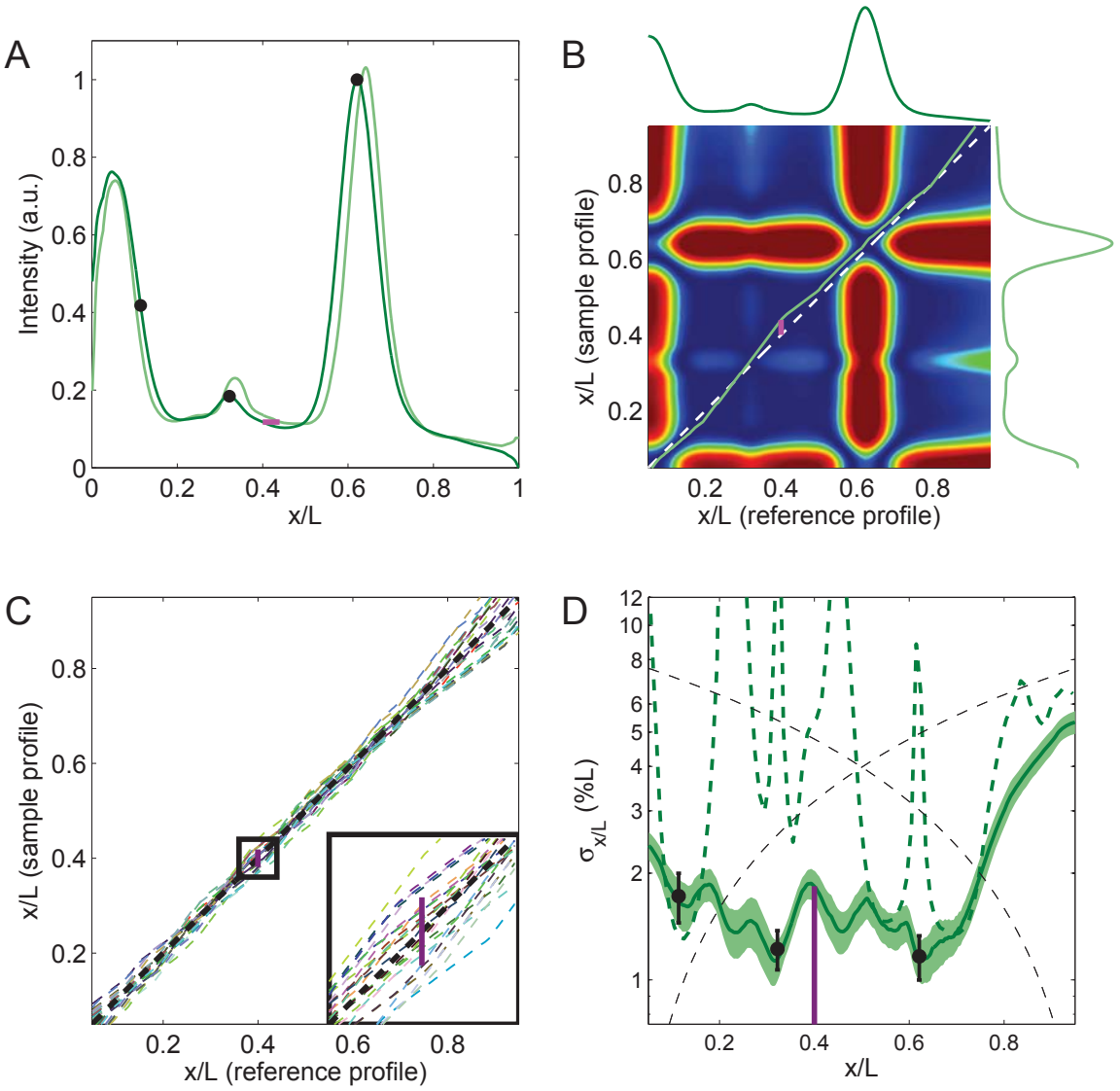

Figure 2.3: Positional error calculation with DPW. **(A)** Comparison between a sample Kni protein expression profile in an embryo 36 minutes into nc14 (light green line) and its reference profile (dark green line), which is the mean profile of Kni in a particular time class (35 ± 3 minutes into nc14, LE&SE embryos, $N = 26$). The magenta line represents the displacement between the sample and reference profile at $x/L = 0.4$. Three black dots mark representative peaks and boundaries. **(B)** The scaling path calculated by DPW of the sample profile (light green line in (A)) with respect to the reference profile (dark green line in (A)) is shown in green as a warped, nearly diagonal curve. The magenta line represents the displacement at $x/L = 0.4$ of the actual scaling path from the path (dashed gray line) that the sample profile would take if it perfectly scaled with the reference profile. The color map shows the relative intensity difference matrix between the sample profile and the reference profile (the intensity difference increases from blue to red). DPW calculates the scaling path by minimizing the sum of the squared intensity difference, normalized at each point

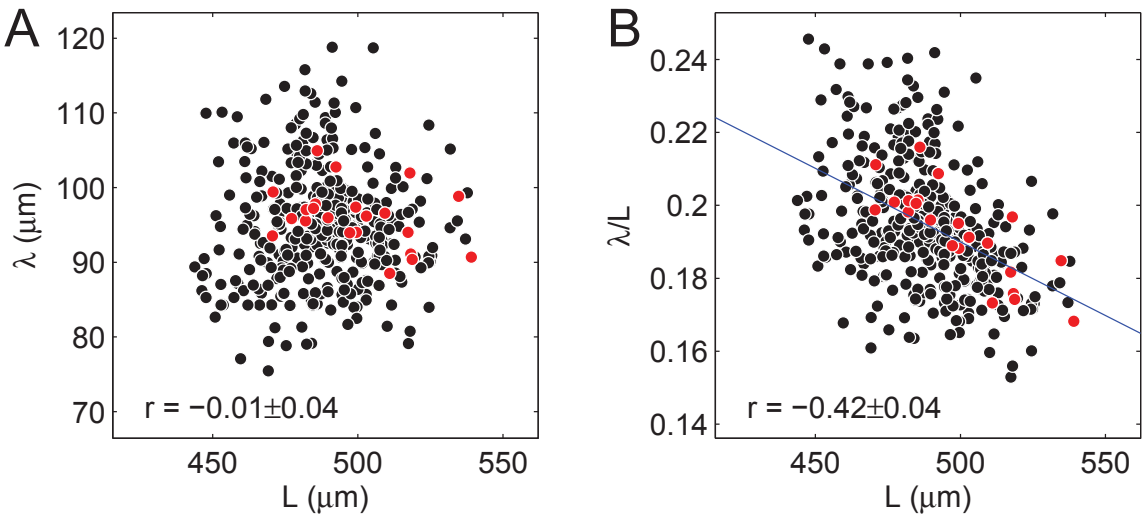

Figure 2.4: Regression analysis shows that the length constant of the Bcd profile does not scale with embryo length. **(A)** For Bcd-GFP embryos collected at 25 °C and imaged 16 minutes into nc14 ($N = 394$), no significant correlation exists between the embryo length L and the decay length λ of the Bcd profile, where λ was calculated for each embryo from a linear fit to the logarithm of that embryo's Bcd profile. **(B)** However, there is a strong negative correlation between the embryo length and the decay length when λ is normalized by L. The linear regression line is shown in blue. Bcd-GFP embryos were acquired from 48 different imaging sessions; the embryos corresponding to the imaging session plotted in Fig. 3B ($N = 22$) are highlighted in red. Pearson's r is shown in the lower left-hand corner of each plot [9]. Errors on r are the standard error from bootstrapping over embryos.

with the result that the positional error of Bcd calculated with the standard error prorogation method is higher than the positional error of the gap genes measured as the variability of the patterning markers (Figure 2.1A).

To overcome the limitations of the traditional methods for measuring positional errors, we have adapted a method known as dynamic time warping (DTW) [82, 89] to quantify the spatial scaling of gene expression profiles with different shapes. This technique has been extensively used for measuring the similarity between dynamic sequences such as audio recordings [53, 90]. The goal of DTW is to minimize the sum of the local intensity difference between two sequences by stretching or squeezing the sequences relative to each other at every point along the sequences' lengths.

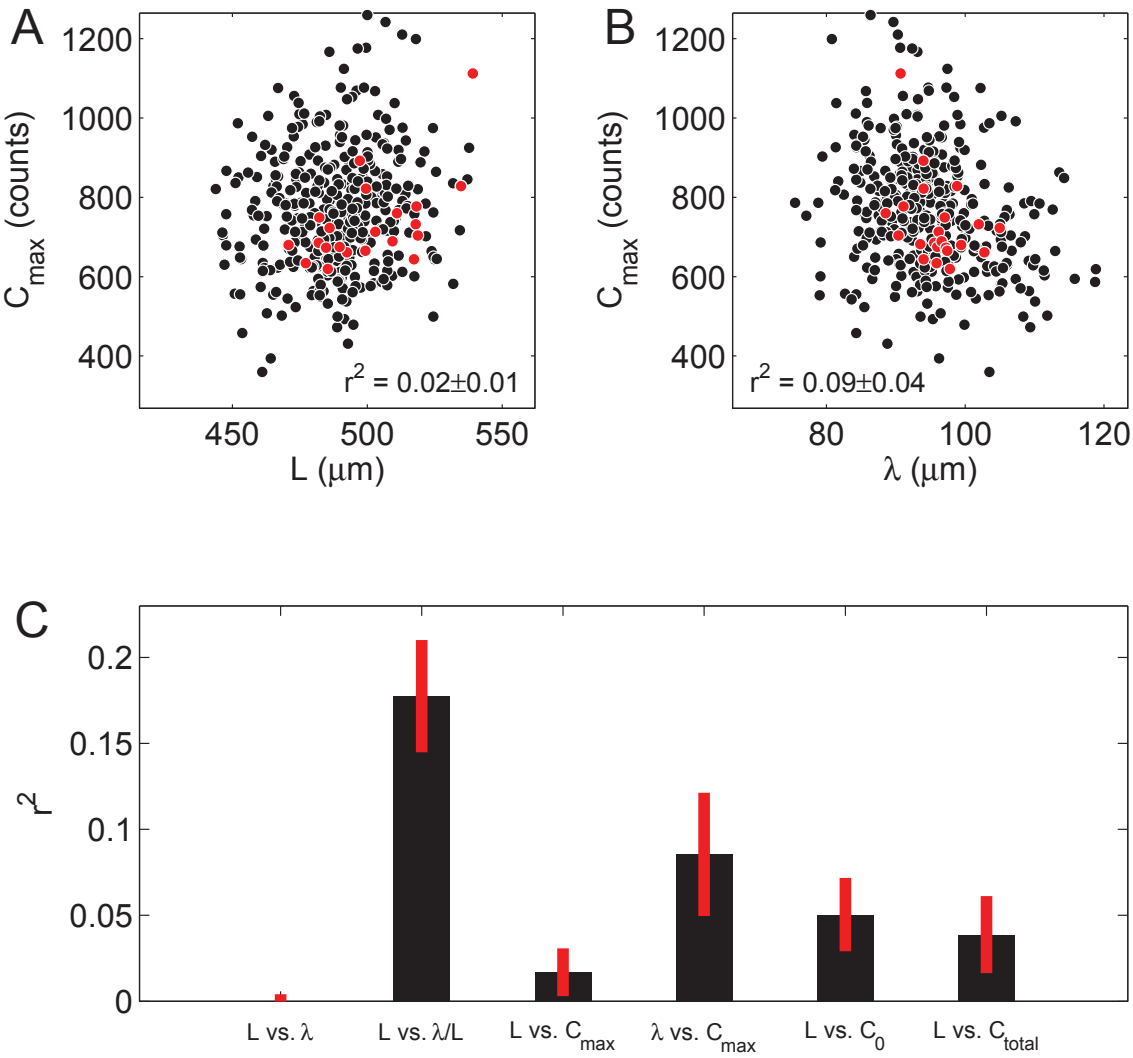

Figure 2.5: Correlations of the Bcd gradient amplitude with embryo length or the profile decay length are weak. **(A)** The correlation between the peak Bcd amplitude $C_{\max}$ and the length L of Bcd-GFP embryos imaged 16 minutes into nc14 ($N = 394$) accounts for only roughly 2% of the total variation in those two parameters. **(B)** The correlation between $C_{\max}$ and the decay length λ, as measured from a linear fit to the logarithm of the Bcd profile in Bcd-GFP embryos, accounts for roughly 9% of the variation in those two parameters. The Bcd-GFP populations plotted in red and black in (A) and (B) are as in figure 4. **(C)** Values of the coefficient of determination r^2 (where r is Pearson's r [9]) between L, λ, $C_{\max}$, the extrapolated anterior-pole intensity C_0 as calculated from a linear fit to the logarithm of the Bcd profile, and the total Bcd intensity C_{total} integrated along the AP axis from 10 to 90% AP. All errors on r^2 are from bootstrapping over embryos.

Figure 2.5: *(continued)* by the sum of the intensities, when mapping the AP positions of the sample profile to those of the reference profile. **(C)** Colored dashed lines represent the scaling paths of all the Kni sample profiles in a particular time class (35 ± 3 minutes into nc14, $N = 26$). The dashed black line denotes the path of a perfectly scaled sample profile. The purple line marks twice the positional error at $x/L = 0.4$, which is calculated as the RMS of the displacement between the sample scaling paths and the path of a perfectly scaled profile. The region at $x/L = 0.35$–0.45 marked with the small black box is magnified in the large black box in the lower right corner. **(D)** Comparison of the DPW positional error (green solid line) with the error-propagated positional error (green dashed line [26, 41] and the positional error of the representative peaks and boundaries shown in (A) (black error bars). The error bars on the markers and the shaded regions on the DPW positional error curve represent standard error on the positional error calculated from bootstrapping. The purple line marks the DPW positional error at $x/L = 0.4$. Gray dashed curves represent the expected positional error due to 6.2% variability of embryo length for completely unscaled LE and SE embryos.

We have repurposed DTW for protein expression profiles (termed dynamic profile warping (DPW) hereafter) to find an ideal mapping between two profile shapes by locally warping the profiles along their coordinate axes (Figure 2.3A,B); specifically, we map the shape of an individual gene expression profile to the shape of the average expression profile of that gene in the population. DPW warps profiles by shifting each AP position in an expression profile a certain distance to overlap with its analogous AP position in the profile of the average embryo; the shifted distance represents the amount by which downstream genes would be expected to be spatially mispatterned if they could perfectly interpret the protein expression level of the read-out gene at that point. The positional error of a population of profiles at each AP position is calculated as the RMS of the shifted distance of all profiles at that AP position (Figure 2.3C, Materials and Methods).

Ideally, computing positional error with different methods should lead to the same answer. We confirmed that in regions of the AP axis where the existing methods are applicable, the positional error calculations using DPW are within measurement errors comparable to existing methods such as the defined boundaries and peaks of gap genes, as well as with the positional errors calculated with error propagation

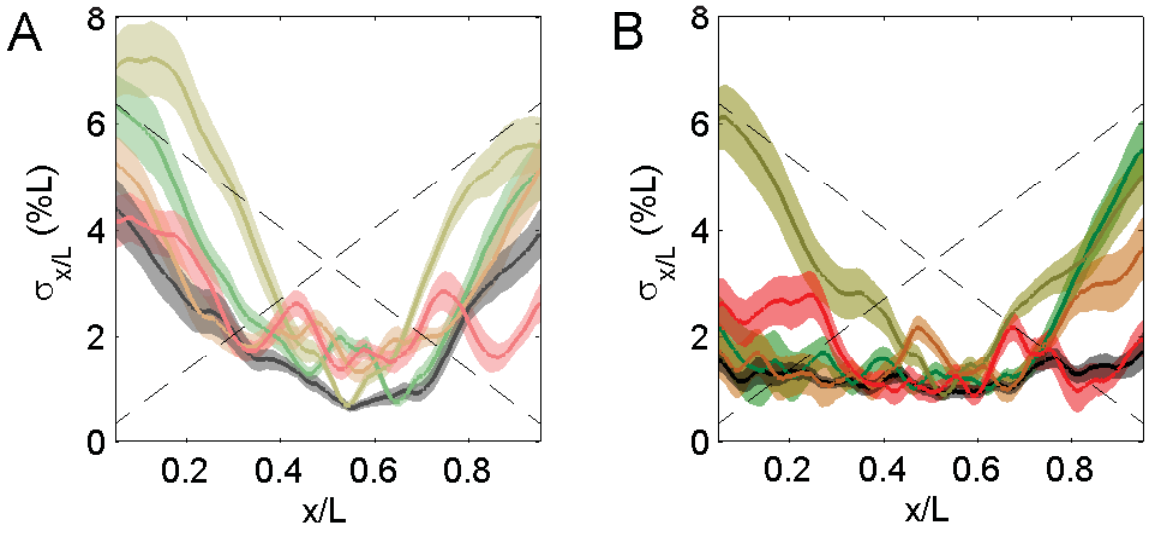

Figure 2.6: The combined positional errors of four gap genes. **(A)** Comparison of the individual positional errors of the gap genes Kni (green), Kr (yellow), Gt (orange), and Hb (red) with the combined positional error of the four gap genes (gray) measured in LE and SE embryos 16 minutes into nc14. $N = 20$ for all plotted data sets. Shadows around the positional error curves show the standard error of the positional error calculated from bootstrapping. Dashed lines are the positional errors of completely unscaled profiles with $\sigma_L = 6.2\%$, matching that of the mixed LE&SE embryos. **(B)** Comparison of the positional error curves of the same genes 50 minutes into nc14. The curve of the combined positional error of the gap genes evolves over nc14 from a "U" shape as in (A) to a flat line in (B) at roughly 1% L. Dashed lines and shadows around positional error curves are the same as in (A).

in the regions where the magnitude of the slope of the profile is sufficiently large (Figures 2.2 and 2.3D). This demonstrates that DPW matches the capabilities of these other methods. Furthermore, DPW avoids the pitfalls of previous techniques to measure scaling: it is a local measurement because it measures positional error along the entire AP axis, and it can be applied to profiles of any shape, allowing for direct comparison of the positional error of Bcd with that of the gap genes. The positional error of expression profiles calculated using DPW can additionally be compared to the unscaled lines to test whether expression profiles of any shape scale with embryo length.

DPW can be extended to measure the combined positional error of multiple genes' protein expression profiles (see Materials and Methods). The ideal way to weight different genes when calculating this combined positional error depends on the molecular details of how the genes regulate the downstream network. As an approximation, we weigh the four gap genes Kni, Kr, Gt, and Hb equally to obtain a combined positional error (Figure 2.6A,B).

2.4 The Bcd gradient is insufficient to establish the spatial scaling of gap gene expression profiles

In order to determine whether the degree of scaling inherent to Bcd is sufficient to establish the scaling of all gap genes using DPW, we need to measure and compare their respective positional errors. Such a comparison is influenced by three potential pitfalls that need to be addressed: the specific developmental time point of the comparison, the respective experimental errors of the different genes, and the overall embryo length distribution of the data sets. To address these pitfalls directly, we obtained three independent data sets:

i. Gene expression profiles as well as positional errors are well-documented to have a strong dependence on developmental time during the first three hours of development [26, 58, 72]. For wild-type embryos, Bcd's positional error is low during nc13 and in early nc14 whereas the combined positional error of the gap genes has a minimum roughly 40 min into the one-hour-long nc14. Therefore we obtained gene expression profiles of Bcd and of the gap genes using immunofluorescence on fixed embryos of the LE&SE fly lines: this data set includes embryos at different developmental ages during all of nc14 and thus the dynamics of the positional error difference

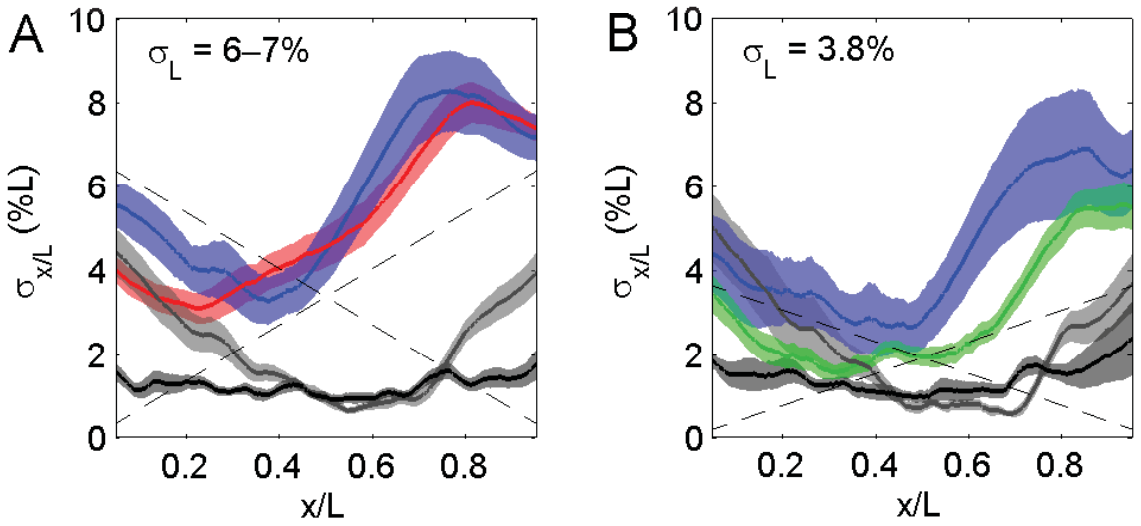

Figure 2.7: The positional error of Bcd is higher than that of the combined gap genes. **(A)** The DPW positional error of Bcd 16 min into nc14 measured in LE&SE embryos (blue, $N = 20$) and Bcd-GFP embryos with varied temperatures (red, $N = 46$) remains higher than the combined positional errors of four gap genes 16 minutes into nc14 (gray, $N = 20$) and 50 minutes into nc14 (black, $N = 20$) measured in LE&SE embryos. The dashed lines are the expected positional error of completely unscaled profiles given $\sigma_L = 6.2\%$ for the mixed LE and SE embryos. **(B)** The positional error of Bcd 16 minutes into nc14 measured in Bcd-GFP fly embryos collected at 25 °C (green, $N = 22$, $\sigma_L = 3.8\%$) has a minimum along the AP axis of roughly 2% L. The positional error of selected LE&SE embryos with $\sigma_L < 3\%$ (blue, $N = 8$) is smaller than that of the full distribution (shown in (A)) that they are drawn from. Note that this lower LE&SE positional error, as well as the Bcd positional error in Bcd-GFP embryos, is still larger than the combined positional error of the four gap genes 16 minutes into nc14 (gray, $N = 10$) and 50 minutes into nc14 (black, $N = 10$), measured in selected LE&SE embryos with a σ_L of roughly 4%. Dashed lines are the expected positional error of completely unscaled profiles given $\sigma_L = 3.8\%$ for the selected subset from the mixed LE&SE embryos. Shaded regions around positional error curves represent standard errors obtained through bootstrapping.

can be measured. We find that the positional error of Bcd is always larger than the combined positional error of the gap genes for all of nc14. In particular, the minimum combined positional error of the gap genes occurs at roughly 40 min into nc14 (agreeing with [26]), when it is roughly 1% L along the entire AP axis (Figure 2.7A). In contrast, Bcd's positional error is always position-dependent throughout all of nc14.

Its minimum value is near the middle of embryo (Figure 2.7A) and it only reaches approximately 3% L as the maximum Bcd expression reaches at 16 min into nc14, thereby significantly exceeding the minimum combined positional error of the gap genes. The combined positional error of the gap genes in the pole regions is higher during early nc14 than it is later in nc14; nevertheless, this early-nc14 combined gap gene positional error is still smaller along the entire AP axis than that of Bcd at the same time point (Figure 2.7A). Thus Bcd alone is insufficient to establish the level of scaling of the gap genes for LE&SE embryos during nc14.

ii. The accuracy that can be achieved when extracting gene expression profiles from immunostainings is limited by systematic errors due to the experimental setup, which is especially problematic here as we are interested not only in the mean protein concentrations but also in the variances [26]. Systematic errors are particularly high for immunostaining experiments to extract Bcd profiles, due to Bcd's low overall expression levels and its shallow exponential shape (Figure 2.8, Materials and Methods). We therefore obtained data from embryos expressing Bcd-GFP using live imaging, which greatly reduces the systematic errors for profile extraction [41, 72]. In order to maintain a broad egg length distribution comparable to that of LE&SE embryos, we bred flies at varying temperatures resulting in an egg length variability of roughly 6% (Figure 2.1B, Materials and Methods). Positional error measurements made at 16 ± 2 min into nc14 (when the overall positional error of Bcd is at its minimal level) in live embryos show a close similarity with the ones measured in fixed LE&SE embryos using immunofluorescence (Figure 2.7A). It is therefore implausible that the overall elevated positional error of Bcd compared to that of the gap genes can be explained by systematic errors.

iii. The previous two datasets have greatly skewed egg length distributions compared to that of wild-type (Figure 2.1B), which might systematically affect the positional error. In fact, we have some evidence from subsampling our embryo population

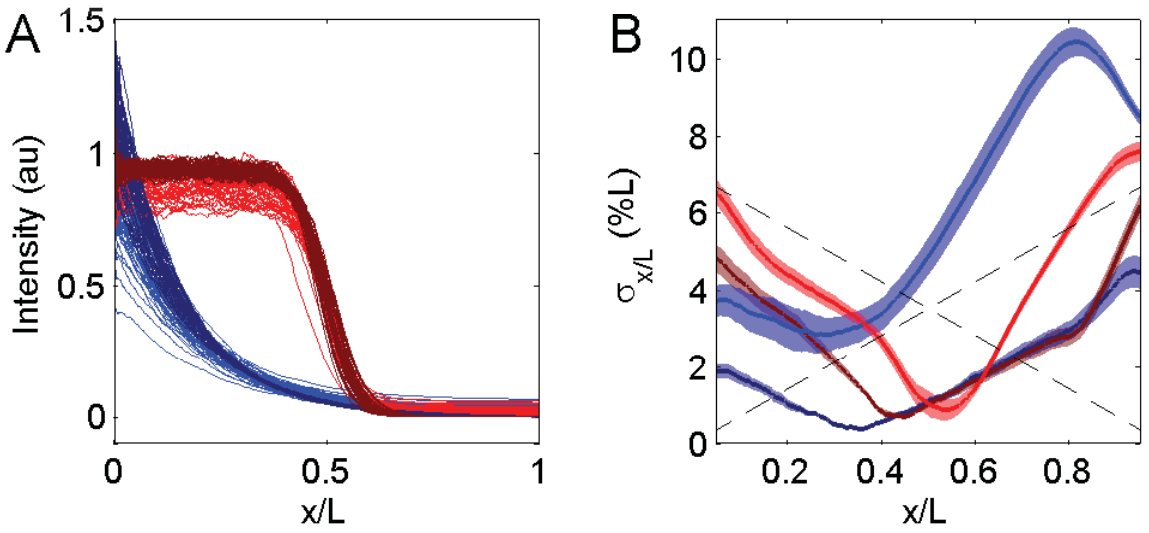

Figure 2.8: Systematic errors in positional error calculations. **(A)** Simulated exponential (blue) and step-function (red) profiles, with and without 5.5% and 1.5% noise in amplitude and baseline fluctuations, respectively (corresponding to the noise levels of Bcd measured in LE&SE embryos). Profiles without noise are plotted in darker colors and profiles with noise are plotted in lighter colors. **(B)** The shape-dependence of positional error calculations. The curve of the positional error of simulated noiseless exponential profiles (dark blue) overlaps with that of simulated noiseless step-function profiles dark red) except for in the region $x/L < 0.4$, implying that the positional error calculated by DPW is mostly independent of whether profile shapes vary quickly or gradually over space. However, the positional error of the exponential profiles (light blue) increases much more than that of the step function profiles (light red) after adding noise, suggesting that the former are more sensitive to overall intensity fluctuations than the latter. The dashed lines represent the minimum positional error of completely unscaled profiles with $\sigma_L = 7\%$, roughly that of the LE&SE embryos. The simulated exponential profiles ($N = 50$) were generated by assuming that the amplitude C_0 is correlated with embryo volume ($\frac{\delta C_0}{C_0} = 3\frac{\delta L}{L}$) and that the length constant $\lambda = 0.17L$ when measured in absolute coordinates. The step-function profiles ($N = 50$) have a transition point at $x/L = 0.5$ and were generated directly from the simulated exponential profiles using a Hill equation with Hill coefficient $n = 5$.

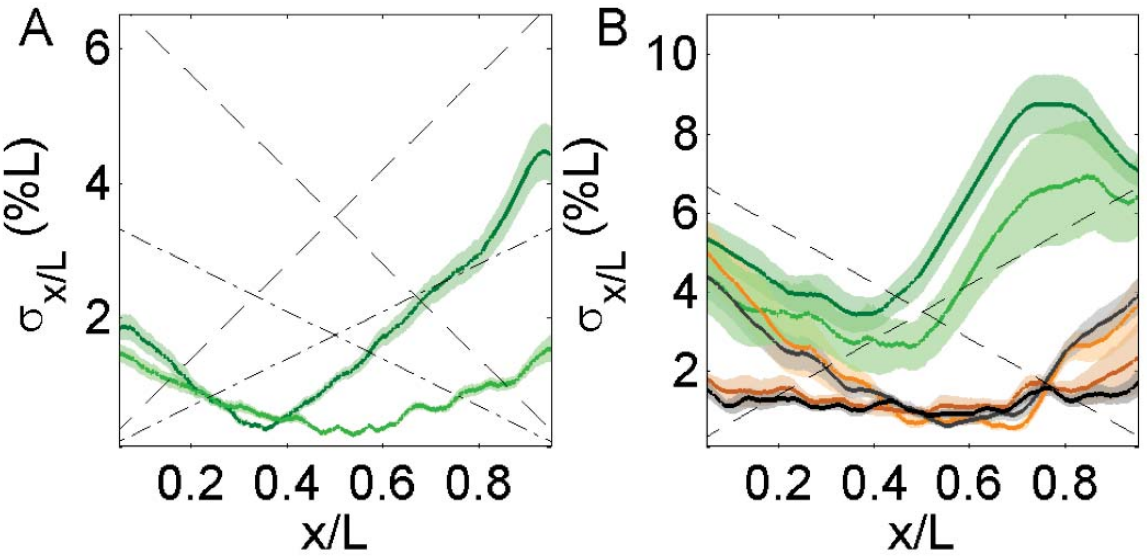

Figure 2.9: The positional error of Bcd gradients increases as the variability of embryo length increases. **(A)** The positional error of simulated exponential profiles (with no amplitude and baseline noise; see Materials and Methods) with volume-dependent amplitude increases as the variability of embryo length σ_L increases from 3.5% (light green) to 7% (dark green). Dashed lines and dash-dotted lines are the predicted positional error for $\sigma_L = 7\%$, and 3.5%, respectively. **(B)** The dependence of positional error on the variability of embryo length for measured Bcd and gap gene profiles. The positional error of measured Bcd profiles from LE&SE embryos (dark green, $N = 31$, $\sigma_L = 8\%$) decreases slightly if only those LE&SE embryos with $\sigma_L = 2\%$ (light green, $N = 9$) at the same time class (16 ± 5 minutes into nc14) are selected. However, the combined positional errors of four gap genes at 16 min into nc14 are almost the same regardless of whether we analyze all measured LE&SE embryos (gray, $N = 20$, $\sigma_L = 6.5\%$) or only selected LE&SE embryos with $\sigma_L = 2.7\%$ (orange, $N = 10$). The same is true for all LE&SE embryos (black, $N = 20$, $\sigma_L = 8\%$) versus selected LE&SE embryos (brown, $N = 10$, $\sigma_L = 2.4\%$) at 50 min into nc14. Dashed lines are as in (A). Shaded regions around positional error curves represent standard errors obtained through bootstrapping.

of the LE&SE data set that in the case of Bcd, the positional error has some egg length distribution dependence (Figure 2.7B), whereas we know that the positional error of the gap genes is independent of egg length distribution (Figure 2.1A, [26]). In order to test whether our conclusion regarding the insufficiency of Bcd remains true even with a wild-type embryo length distribution, we extracted nuclear Bcd concentrations from live embryos expressing Bcd-GFP, bred at 25 °C. The embryo length variability

of this dataset is comparable to that of WT (roughly 4%), whereas the embryo length variability of the other two datasets is roughly 6%. We confirm that the positional error of Bcd, but not the combined positional error of the gap genes, depends on the embryo length variability (Figures 2.1A and 2.9B). Moreover, the positional error of Bcd measured in these fixed-temperature Bcd-GFP embryos is higher than the minimal value of the combined positional error of the gap genes at the same level of embryo length variability (Figure 2.7B).

Overall our findings suggest that the scaling information that can be provided by Bcd to the system is incomplete, and that the scaling of the gap genes cannot be explained by the sole input of Bcd.

2.5 Bicoid is unnecessary to establish the spatial scaling of gap gene expression profiles

Bcd's insufficiency for conveying the low positional error observed in the gap gene profiles is not a proof that it is in fact unnecessary for gap gene scaling. The *bcd* gene, together with *nanos* (*nos*) and *torso* (*tor*) [38], is one of the primary maternal determinants in the early *Drosophila* embryo. All three set up gradients that are thought to provide most if not all of the positional information for patterning along the AP axis [110], including the setup, the precision and the scaling properties of the gap genes [26]. In one plausible scenario, the informational input of Bcd could be crucial for the gap genes to be positioned to our observed levels of precision, but other sources of positional precision could also be needed [72]. To test this hypothesis directly we repeated our analysis on embryos that are lacking one or two of these three maternal inputs.

First, to test whether *bcd* is necessary to establish the spatial scaling of the gap genes, we used DPW to measure gap gene positional error in in embryos mutant for

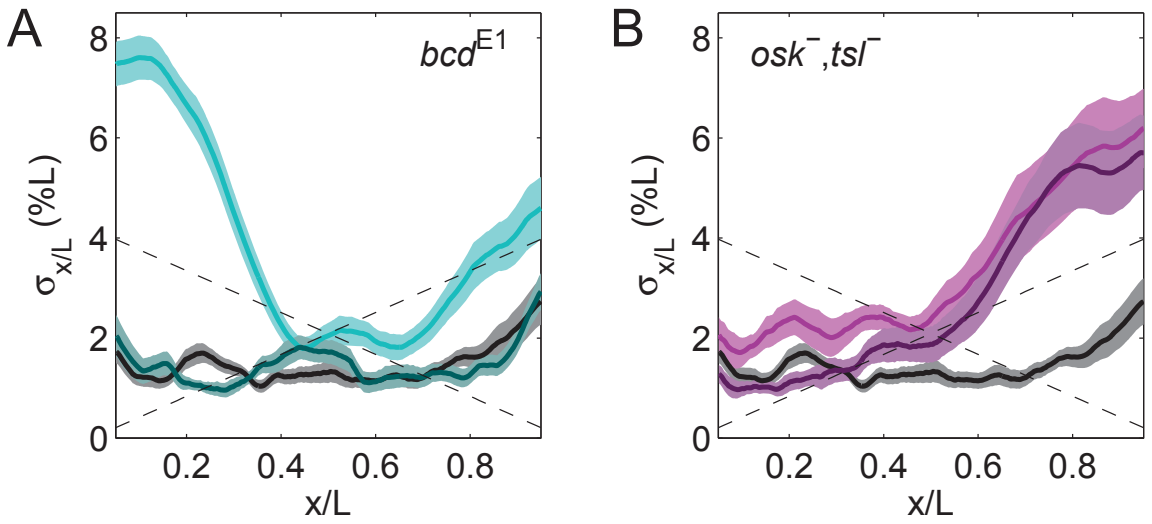

Figure 2.10: Bcd is neither sufficient nor necessary to establish the scaling of gap gene expression profiles. **(A)** Bcd is unnecessary for establishing the scaling of gap gene expression profiles in late developmental stages. The combined DPW positional error of the four gap genes in *bcd* mutant embryos (bcd^{E1}) is high in the anterior half of the embryo 15–25 minutes into nc14 (light cyan, $N = 18$), but it recovers to the same level as that of WT (black, $N = 26$, same as in Figure 2.7A) 40–50 minutes into nc14 (dark cyan, $N = 16$). Dashed lines represent the expected positional error of completely unscaled profiles with an embryo length distribution of $\sigma_L = 4\%$, matching that of the mutants. **(B)** Bcd alone is insufficient to establish the scaling of gap gene expression profiles. In the posterior half of the embryo, the combined positional errors of the four gap genes Kni, Kr, Gt and Hb in osk^{-},tsl^{-} mutant embryos remain high from 15–25 minutes (magenta, $N = 11$) to 40–50 minutes into nc14 (purple, $N = 16$), and they never approach the same level as that of the combined gap genes in WT embryos (black, same as in (A)). Dashed lines are the same as in (A). Shaded regions around positional error curves represent standard errors obtained through bootstrapping.

bcd (bcd^{E1}), which are lacking the Bcd gradient. Although the combined positional error of the gap genes is higher than in wild-type embryos in early nc14, it is nearly identical to that of wild-type embryos in late nc14 (Figure 2.10A). Thus, the lack of Bcd impairs the ability of the gap genes to be positioned precisely enough, but only temporarily. (Note that this early-nc14 impairment of positional precision is strongest in the anterior of the embryo, as would be expected from the loss of the Bcd profile.) This result implies that Bcd is not only insufficient, but also ultimately unnecessary, for the gap genes to achieve wild-type levels of positional precision. It also suggests

that the three independent maternal inputs Bcd, Nos and Tor seem to be redundant regarding the spatial scaling of the four gap genes in late developmental stages.

Second, to test whether the maternal inputs *nos* and *tor* are necessary to establish the observed positional precision and thus scaling, we measured the gap gene positional error in embryos mutant for *torsolike* (*tsl*) and *oskar* (*osk*) [38, 96], which are lacking the gradients of Nos and Tor. In these double mutant embryos the sole AP patterning input is provided by Bcd, and the combined positional error of the four gap genes Kni, Kr, Gt and Hb exceeds that of wild-type throughout all of nc14 in the posterior region (Figure 2.10B). This result agrees with our finding from the previous section in that Bcd cannot convey sufficient positional precision to the gap genes.

The results from the mutant data sets suggest that other maternal factors (in addition to Bcd) contribute to the establishment of the scaling of the gap genes. We further tested whether the combination of maternal factors is sufficient to establish the spatial scaling of gap gene expression profiles. That requires measurements of the scaling properties of the maternal factors Nos and Tor. However, neither Nos nor Tor are transcription factors, and thus they must provide their information about scaling indirectly via the transcription factors that they regulate, i.e. *capicua* (*cic*, regulated by Tor [60]) and maternal Hunchback (Hb^{mat}) (which is suppressed by Nos in the posterior of the embryo [52]). In addition, it is uncertain if the profile of the Bcd-regulated gene *caudal* (*cad*) contains extra positional information not bound in the Bcd profile. Thus we measured the positional error of the maternal factors Bcd, Cad and Cic in LE&SE embryos using DPW (as we were unable to measure the positional error of Hb^{mat} in LE&SE embryos). Just like Bcd, the positional errors of Cic and Cad alone are higher than the combined positional errors of the gap genes (Figure 2.11A). However, the combined positional error of the maternal inputs Bcd, Cad and Cic reduces the positional error along much of the AP axis to within a factor of 2 of

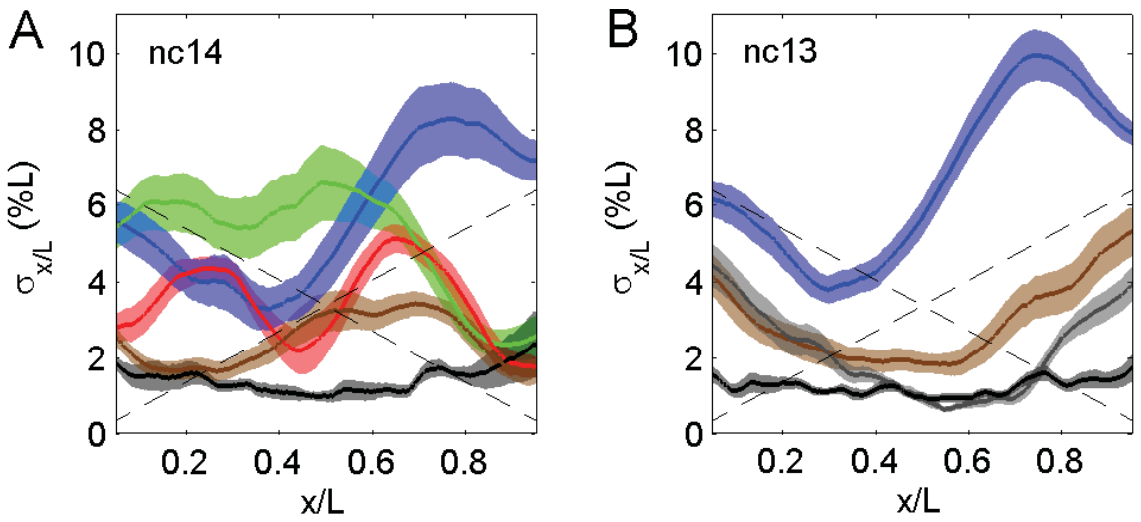

Figure 2.11: The combined DPW positional error of maternal genes approaches the same level as that of the four gap genes. **(A)** The combined positional error of the maternal genes Bcd, Cad, and Cic (brown) is reduced compared with the separate positional errors of Bcd (blue), Cad (green) and Cic (red) 16 minutes into nc14, as measured in LE&SE embryos. However, all four of those positional errors are higher than the combined positional error of the four gap genes Kni, Kr, Gt, and Hb 50 minutes into nc14 in LE&SE embryos. $N = 20$ for all data sets. Dashed lines and the black positional error curve are as in Figure 2.3A. **(B)** The combined positional error of the maternal genes Bcd, Cad, and Cad (brown, $N = 15$) has a similar shape in nc13 to that of the four gap genes in early nc14. By comparison, the positional error of Bcd in nc13 (blue, $N = 20$) is much higher than all of the other curves. The gray curve, black curve, and dashed lines are as in Figure 2.3A. Shaded regions around positional error curves represent standard errors obtained through bootstrapping.

that of the combined gap genes, and the functional shape of the combined positional error of Bcd, Cad and Cic during nc13 closely matches that of the combined gap genes at early nc14 (Figure 2.11B). The combined positional error of the maternal genes could potentially be further reduced by including maternal Hb, suggesting that the combined maternal factors might be sufficient to establish the spatial scaling of the gap genes in early nc14. Together our results argue that no single property of any of the maternal input factors is sufficient to explain gap gene scaling, but rather that

gap gene scaling emerges as a collective effect, an inherent property of the genetic network.

2.6 Discussion of scaling results with the dynamic profile warping method

By comparing the positional precision of patterning domains to the natural fluctuations of *Drosophila* egg sizes, we demonstrate that the *Drosophila* segmentation gene network must have solved the scaling problem at the level of the gap genes. To investigate the origin of gap gene scaling we have developed a unified, shape-independent metric for measuring the spatial scaling of gene expression profiles. Using this method we show that Bcd is neither sufficient nor necessary for gap gene scaling, and we conjecture that scaling emerges as a collective network property.

With DPW we have developed a unique method for characterizing the scaling properties of gene expression patterns that can be easily extended to systems other than the AP patterning system in *Drosophila* embryos. DPW has great advantages over traditional methods for measuring scaling in that it is applicable to genes of arbitrary profile shape for which neither specific landmarks (such as boundaries or peaks) nor specific functional forms can be identified. For example, the conventional correlation analysis that assesses the scaling of the Bcd gradient relies on fitting the Bcd gradient to an exponential shape. However, the Bcd profile might not be perfectly exponential [91], and reducing the Bcd gradient to a few parameters introduces fitting errors that are larger than what can be tolerated given the small natural egg size variation. Though one concern is that the magnitude of the positional error calculated with DPW might not directly represent how well profiles scale. The positional error includes natural variations in profile intensity as well as measurement errors in the profiles' intensities and baselines that are all unrelated to actual fluctuations due to

egg size variation. Carefully designed control experiments or simulations (as we did with Bcd gradients) will be necessary to clarify the contributions from the factors other than scaling (see Materials and Methods).

We also introduce lines representing the positional error of completely non-scaling profiles to set up an objective and practical criterion for determining whether a profile is spatially scaled. Compared with arbitrarily chosen thresholds in conventional correlation analyses that test for scaling [34], these lines represent a physically meaningful lower bound for the positional error of unscaled expression patterns. Namely, the line passing through the anterior pole is the theoretical minimum positional error for profiles that completely overlap when overlaid in absolute coordinate space with respect to the anterior pole, and likewise with the line passing through the posterior pole. For example, in the case of Bcd, the protein gradient originates from a source of localized *bcd* mRNA at the anterior pole and diffuses outward. Thus, it is only meaningful to compare the positional error of the Bcd gradient with the anterior-anchored unscaled line because we expect Bcd expression profiles across all embryos to be coincident at the anterior pole, regardless of whether or not they scale with embryo length. On the other hand, since the combined positional error of gap genes integrates the inputs from both poles, both unscaled lines should be used to determine scaling. By choosing the proper unscaled line(s), we are able to conclude that a gene must be scaled if its positional error is below its respective unscaled line as it is the case for the gap genes. One problem with the unscaled lines is that a gene cannot to be determined to be unscaled if its positional error is above the unscaling lines. This is because the magnitude of the positional error calculated with DPW has contributions from different sources including scaling, natural gene expression noise, measurement errors, etc., whereas the unscaled line only represents the positional error of perfect unscaled profiles without natural gene expression noise and measurement errors. Therefore we cannot conclude that Bcd is unscaled based on its

positional error lies above the anterior-anchored unscaling line (Figure 2.3) unless we can exclude the contributions from the factors other than scaling.

Our measurements challenge several previously proposed spatial scaling mechanisms for gene expression patterns in the early *Drosophila* segmentation gene network:

i. One hypothesis assumes that the length constant of the Bcd gradient is proportional to the embryo length, which would produce a scaled Bcd gradient along the whole AP axis if the amplitude of the gradient were fixed and the reproducibility of the Bcd gradient were high. However, based on SDD model [41, 43], the length constant of the Bcd gradient is $\lambda = \sqrt{D\tau}$, where D and τ are the diffusion constant and the degradation time of Bcd molecules, respectively. These two parameters are generally independent of embryo length. To solve this problem, τ was suggested to be proportional to L^2 if the protein degradation rate was tuned appropriately for the fly embryos in different species [39] or if Bcd degradation were dominated by nucleus-specific degradation and the nuclear number were fixed for the fly embryos within one species [17, 42]. In contrast, our measurements in embryos expressing Bcd-GFP demonstrate that the length constant (in absolute scale) is independent of L (Figure 2.4A). As previously noted, the length constant of the Bcd gradient of LE embryos is indistinguishable from that of SE embryos [15] and for WT embryos the AP position at which the Bcd gradient crosses at a given threshold is independent of the embryo length [48]. In addition, the formation of the Bcd gradient has been observed to be shaped independently of nuclei [44], and the nuclear number has been correlated with embryo size within one fly species [78]. Thus no evidence supports the hypothesis that Bcd is scaled via an adapted length constant of the gradient.

ii. Bcd has also been suggested to scale based on a volume-dependent production hypothesis. In this case, the length constant remains a constant, but the amplitude of the Bcd gradient is proportional to embryo volume [15, 105]. This model has the following three predictions. First, in the relative scale, the profiles of the longer

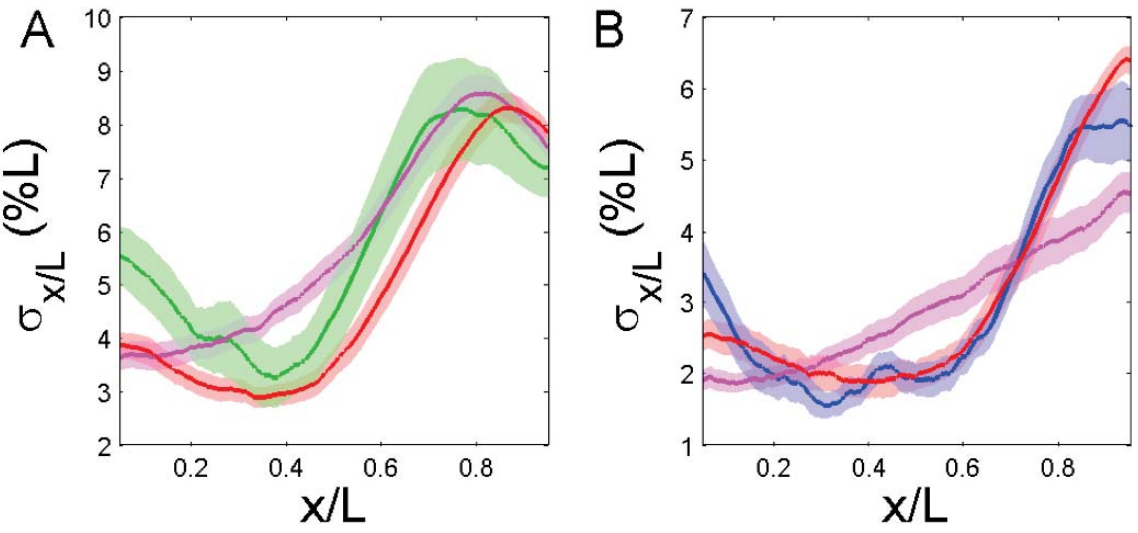

Figure 2.12: Comparison between the positional error of the simulated Bcd profiles and the measured Bcd gradients. **(A)** Compared to the positional expression error of the Bcd gradients measured in LE&SE embryos (green), the positional error of simulated Bcd profiles with an amplitude C_0 (red) extracted from the best fit ($A_{\text{noise}} = 0.1$, $B_{\text{noise}} = 0.01$, $k = 2.7$, see Materials and Methods) matches better than that of simulated best-fit Bcd profiles with fixed amplitude C_0 (magenta) ($A_{\text{noise}} = 0.14$, $B_{\text{noise}} = 0.01$, $k = 0$). **(B)** Compared to the positional error of Bcd gradients measured in Bcd-GFP embryos (blue), the positional error of simulated Bcd profiles with an amplitude C_0 (red) extracted from the best fit ($A_{\text{noise}} = 0.06$, $B_{\text{noise}} = 0.004$, $k = 3.1$) matches better than that of simulated best-fit Bcd profiles with fixed amplitude C_0 (magenta) ($A_{\text{noise}} = 0.08$, $B_{\text{noise}} = 0$, $k = 0$). Shaded regions around positional error curves represent standard errors obtained through bootstrapping.

embryos has short decaying constant and larger amplitude, they cross in the middle of AP axis with the profiles of the shorter embryos with longer decaying constant and smaller amplitude (A cartoon should help to illustrate this). Therefore the positional error of the Bcd gradient is position-dependent and has a minimum in the middle of the embryo (Figure 2.9A). Second, the positional error depends on the embryo length variability (Figure 2.9A). Third, the dependence between the amplitude of the Bcd gradient and the embryo length follows the formula $\frac{\delta C_0}{C_0} = k\frac{\delta L}{L}$, where $k = 3$, if the embryo volume is proportional to L^3. Consistent with these predictions, the positional error of the Bcd gradients of our three fly line datasets has approximately the same

curve shape as predicted with this model (Figures 2.7 and 2.12). Moreover, the positional error of Bcd gradients increases as the embryo length variability increases (Figure 2.9B). Last, based on the best fit of the positional error of simulated Bcd profiles to that of our measured profiles, C_0 is dependent on L and the coefficient k is between 2.7 to 3.1, which comes close to the predicted value of $k = 3$ (Figure 2.12, Materials and Methods). On the other hand, the correlation between C_0 and the embryo length L is very weak, and it is very hard to directly extract the exact dependence between C_0 and embryo size (Figure 2.5A). This implied that the expected change of C_0 is rather small compared with the gradient noise (including the biological noise and measurement errors), especially when the embryo variability of WT embryos is only roughly 4%. Thus the small modulation of C_0 is too weak to supply sufficient scaling information for Bcd gradients that could explain gap gene scaling, consistently the minimum positional error is still higher than the combined positional error of gap genes (Figure 2.7), limiting the effectiveness of volume-dependent modulation in delivering positional precision to the gap genes.

iii. The pre-steady-state decoding model suggests that the genes downstream of Bcd can be scaled if they read in the Bcd concentration before the Bcd gradient approaches its steady state, even if the steady-state Bcd gradient is not scaled with embryo length [8]. This model predicts that the scaling of the downstream genes' developmental patterns is position-dependent, with hyper-scaling close to the anterior pole and hypo-scaling close to the posterior pole [21]. However, our measurements do not support this hypothesis: we observe that the combined positional error of the gap genes in late nc14 is nearly evenly distributed along the entire AP axis (Figure 2.7). Moreover, developmental markers such as the cephalic furrow and individual stripes of the *even-skipped* expression pattern are also scaled with embryo length even though they are close to the anterior pole, contradicting the pre-steady-state decoding model. Given the fact that the Bcd gradient nearly stabilizes at about nc11 [42, 71],

which is when the transcription of zygotic gap genes also starts [55], it is plausible that the scaling information provided by Bcd is transferred to the gap genes not via an instant readout but rather through temporal averaging [70, 72]. This temporal averaging of the Bcd gradient would depend on the lifetime of the gap genes and the dynamics of the Bcd gradient, which starts to decay in nc14. Temporal averaging may explain why the specific developmental time for the gap genes to reach minimum positional errors is delayed by about 40 min compared to Bcd. Assuming that the volume-dependent production hypothesis is correct, it is quite possible that temporal averaging could help the system to filter the noise in the Bcd profile to obtain the scaling information encoded in a modulated C_0.

iv. A fourth model for spatial scaling would suggest that the Bcd gradient is irrelevant for the scaling of AP patterning, instead the scaling is driven solely by the interaction of the downstream gene network [46]. On the one hand, we find that the interaction between the gap genes plays a role in establishing gap gene scaling: although the combined positional error of the four gap genes is much higher than the unscaled line at early developmental stages in *bcd* mutant embryos, it recovers to the WT level at later developmental stages. Such a dynamic process might result from mutual cross-regulatory repressive interactions of the gap genes that are responsible for dynamic boundary shifts and boundary sharpening [58, 77]. One could imagine that the same cross-regulation helps gap genes recover scaling precision by adjusting their patterns based on the inputs from other maternal genes when Bcd is missing. A similar scenario was observed as a response of the network to Bcd dosage changes [72]. On the other hand, if the cross-regulation between gap genes does contribute to the establishment of the scaling of the segmentation pattern, it cannot achieve this goal by itself. As our double mutant data indicate, even at the end of nc14, the positional error of the four gap genes in the posterior half remains much higher than the WT level: all cross-regulation between the gap genes should still exist in these

mutants, but the crucial inputs of Nos and Tor do not. Thus the input from the maternal factors is imperative for scaled gap gene patterns.

In conclusion, our results indicate that the underlying mechanism of the scaling of segmentation patterns follows neither the scenario of initially scaled maternal morphogen gradients nor the scenario of scaling through the interaction of the downstream gene network. Further studies will be necessary to elucidate the exact origin of scaling in the patterning system, but our results project that the most likely scenario includes synergy between the combined maternal inputs and cross-regulation of the downstream genes, from which scaling emerges as a collective network property.

2.7 Principal component analysis affirms lack of scaling of Bcd

The DPW method is well-suited for comparing the positional error of genes as a function of position along the AP axis, but it has the limitation that it only provides an indirect measure of scaling. In this section we explore an alternative analysis technique that directly tests the hypothesis that a population of gene profiles scales with embryo length, considering the profile as a whole rather than investigating specific AP positions.

If we were to assume that a particular gene's expression profile completely scales with embryo length, then if we plotted the shape of the typical embryo's profile for this gene relative to its length (such that the anterior pole were fixed at 0 and the posterior pole were fixed at 1), we would expect that this shape would not vary as a function of embryo length: thus, if we observe any difference between the shape of the typical shorter embryo and the shape of the typical longer embryo when plotted in this manner, we would be able to attribute this shape difference to a lack of scaling. To illustrate this principle with real data, consider our population of temperature-

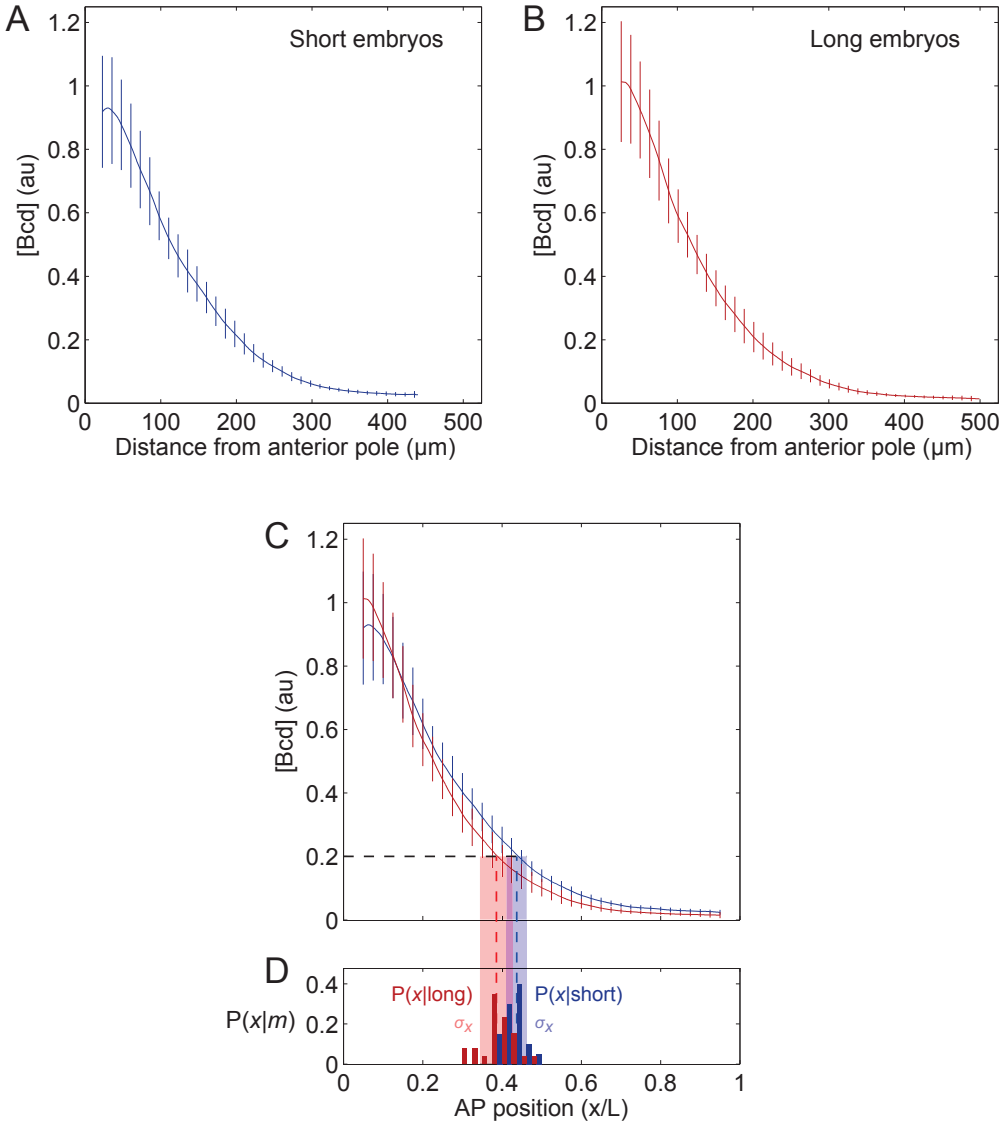

Figure 2.13: Differences in shape among differently sized embryos indicates a lack of scaling. **(A–B)** Bcd concentrations in a mixture of Bcd-GFP-expressing embryos bred at 18 and 29 °C ($N = 46$), plotted as a function of distance from the anterior pole, for all embryos shorter (A) and longer (B) than the median embryo length in the population. **(C)** When plotted in units relative to egg length, the mean Bcd profiles in short (blue) and long (red) embryos differ, most notably near the anterior pole and in the middle of the embryo. This pathological difference in shape is evidence of a lack of scaling. **(D)** As a result of this difference in profile shape between long and short embryos, the distribution of positions along the AP axis at which embryos cross a given Bcd concentration threshold (horizontal dashed line) is different for our population

Figure 2.13: *(continued)* of small embryos than for our population of large embryos. $P(x|m)$ is the probability density function for an embryo from population m to cross the threshold of a Bcd concentration of 0.2 (au) at an AP position of x. σ_x is the standard deviation of crossing positions x for the two populations of embryos, small and large. (A–C: error bars are the standard deviations of measured Bcd concentrations over all embryos in either the small-embryo-length or large-embryo-length populations.)

varied embryos (Section 2.4): we can split this population into two sets consisting of all embryos with lengths above and below the median (Figure 2.13A,B). If we plot the mean profile (relative to embryo length) in each of these two sets on top of each other, we find differences in profile shapes (Figure 2.13C,D). This is visual evidence of a lack of scaling in our population of temperature-varied embryos.

To make this analysis of scaling through visual comparison of profiles more scientifically rigorous, we can use the language of hypothesis testing to determine whether profile shapes at different egg lengths are different enough to constitute a likely lack of scaling. We take our null hypothesis to be the idea that profiles perfectly scale with egg length: the larger the difference in profile shapes in embryos of different lengths, the likelier it is that we will reject this hypothesis and conclude that these profiles do not actually scale. In order to quantify the difference in profile shapes in embryos of different lengths, we use principal component analysis (PCA) to algorithmically extract individual features in a set of profile shapes whose correlations with embryo length we can calculate [86, 101, 108].

If we think of our profiles as sampling the protein concentration level over p positions along the AP axis, the first principal component of PCA will be the direction in p-dimensional space in which the variability among profiles is the highest. The second principal component will be the direction, orthogonal to the first principal component, along which the remaining variability among profiles is highest, and so on [63]. In order to illustrate PCA with real data, Figure 2.14A presents Bcd profiles from our set of Bcd-GFP embryos imaged live in nc14, and Figure 2.14B,C presents

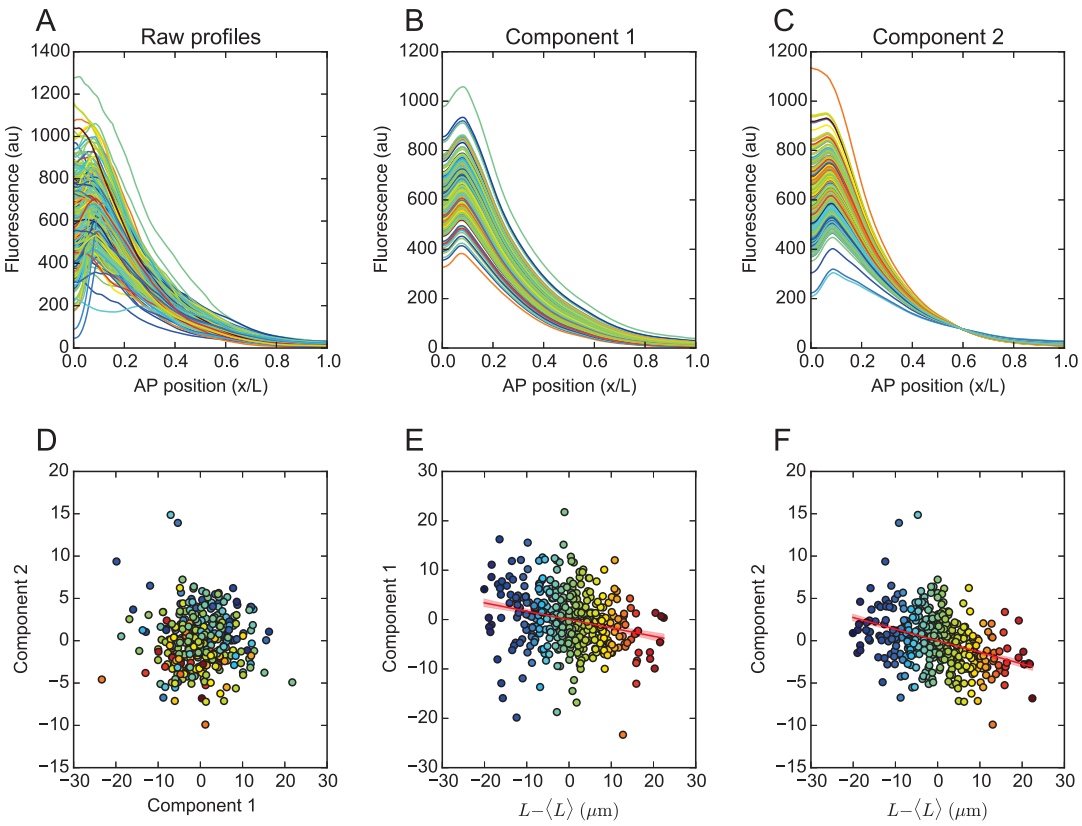

Figure 2.14: Principal component analysis tests the hypothesis that profiles perfectly scale with embryo length. **(A)** Bcd profiles from 394 Bcd-GFP embryos, imaged live 16 minutes into nc14, are plotted as a function of position along the AP axis and colored according to egg length, with red representing the longest embryos and blue representing the shortest embryos. **(B,C)** The first two principal components, accounting in this dataset for 89% of the total variance in the profiles, are extracted from the profiles. The projections of the profiles onto the vectors given by the first two principal components are shown. **(D)** Scatter plot of the magnitudes of component 1 vs. component 2 for all embryos, where color again represents embryo length. **(E,F)** The first two principal components plotted as a function of the deviation of egg length L from the mean egg length $\langle L\rangle$. Red lines indicate least-squares linear regressions to data, and error bands represent ± 1 standard deviation of uncertainty on the value of the slope of the regression.

the projections of the profiles onto the first two principal components as calculated using PCA. Typically, the first principal component of our Bcd datasets captures the variability in the overall magnitude of the measured protein concentrations among embryos, and the second principal component captures the variability in effective

protein concentration decay lengths. For each principal component, we find the magnitude and direction of the projection of each profile onto the vector represented by that component (Figure 2.14D). For the first two principal components, we perform a linear regression over the embryos in our dataset between the strength of each of the components and embryo length (Figure 2.14E,F): if profiles perfectly scale with embryo length, we would expect neither of these components to have a significant correlation with embryo length. Our null hypothesis, then, is precisely the hypothesis that neither of the first two principal components has a significant correlation with embryo length: more concretely, we reject this hypothesis if we can report with at least 95% confidence that the smaller of the two p-values of the correlations between the first two principal components and egg length is below 5%.

With this implementation of principal component analysis, we can test the hypotheses of whether Bcd and the gap genes, measured in the datasets described in the previous sections, scale with embryo length. First, we can apply PCA-based hypothesis testing to the datasets of live Bcd-GFP embryos and fixed wild-type embryos immunostained for the four gap genes Gt, Hb, Kni, Kr. We find rejection of Bcd scaling in individual imaging sessions of Bcd-GFP embryos if and only if the relative standard deviation $\sigma_L/\langle L\rangle$ of embryo length in the session is 3.7% or greater, which implies some sensitivity of the PCA method on the variability of embryo lengths in the population (Figure 2.15A). This rejection of the scaling of Bcd is further enhanced, as measured by the plotted p-value, when combining all imaging sessions, as shown by the largest circle in Figure 2.15A. These results clearly demonstrate a lack of scaling of the Bcd gene with embryo length.

On the other hand, an analogous principal component analysis of our dataset of fixed wild-type embryos stained for the gap genes reveals no rejection of the hypothesis of scaling for those genes (Figure 2.15B), either for early or late nc14: thus, for these two datasets with a wild-type embryo length distribution, our PCA results are

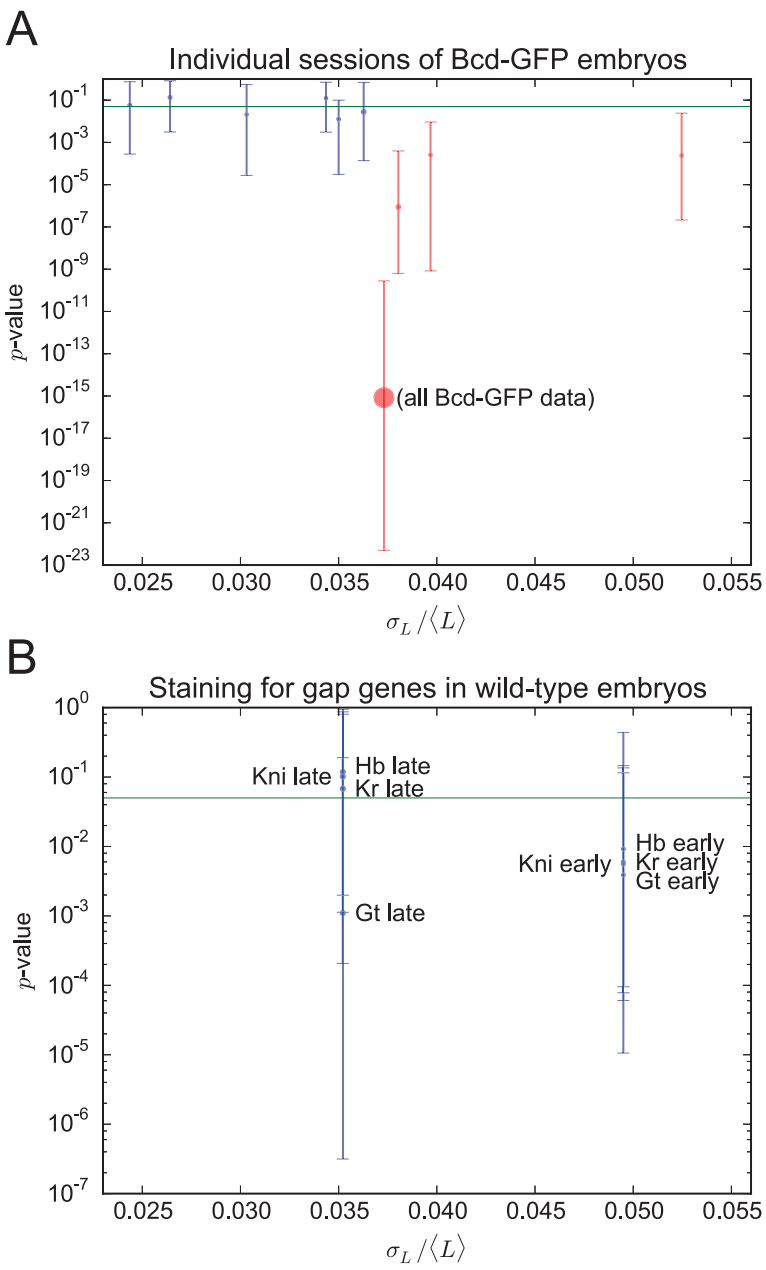

Figure 2.15: Principal component analysis of live and fixed embryos with wild-type length distributions rejects scaling of Bcd but not the gap genes. **(A)** Rejection (red) or non-rejection (blue) of the hypothesis of scaling in imaging sessions of symmetrically mounted Bcd-GFP embryos imaged live 16 minutes into nc14, as a function of the relative standard deviation $\sigma_L/\langle L\rangle$ of embryo length. The hypothesis of scaling for an imaging session is rejected if the 95% confidence interval (error bars) completely lies below a p-value of 5% (green horizontal line), where the p-value represents the lower of the two p-values of the correlations between the first two principal components and embryo length. Areas of circles are proportional to the number of valid embryos in each imaging session. Only imaging sessions with 12

Figure 2.15: *(continued)* or more valid embryos are shown. Large circle represents all imaging sessions combined. **(B)** Non-rejection of scaling in early- and late-nc14 wild-type embryos stained for the four gap genes. Early embryos were imaged from 15 to 25 minutes into nc14 and late embryos were imaged from 40 to 50 minutes into nc14. Colors, marker sizes, and error bars are as in top plot.

consistent with the hypothesis that the four gap genes scale with embryo length and that Bcd does not. Note that this finding controls for both relative embryo length variability and sample size in individual plotted points: first, Figure 2.15A implies that there is a certain minimum embryo length variability in a dataset necessary to reject the scaling hypothesis, and our early-nc14 gap gene dataset clearly exceeds this minimum (although our late-nc14 gap gene dataset may not). Second, the number of embryos in the Bcd and gap gene datasets are comparable, as shown by the areas of the circles in Figure 2.15A,B.

One potential difficulty in comparing the two aforementioned datasets with wild-type embryo length variability is the fact that they were acquired with different experimental techniques (live imaging of Bcd-GFP embryos vs. fixed immunostaining for gap gene profiles in wild-type embryos). Two alternative datasets, the LE&SE datasets discussed in Section 2.4, provide a way to compare the scaling of Bcd and the gap genes under more equal experimental conditions: both datasets are immunostainings of embryos bred for long and short embryo length, with a relative standard deviation of length of roughly 6 to 7% altogether. One dataset contains embryos stained for the maternal factors Bcd, Cad, Cic, and maternal Hb, and the other contains embryos stained for the gap genes Gt, Hb, Kni, and Kr. Both datasets contain embryos stained over several nuclear cycles before gastrulation, and we pick out specific time windows to test with PCA: early nc14 embryos, late nc14 embryos, and all embryos throughout all of nc14. Unlike the previously discussed datasets, which show a lack of scaling of Bcd in Bcd-GFP embryos and results consistent with scaling for the gap genes in immunostained embryos, this comparison of Bcd and the

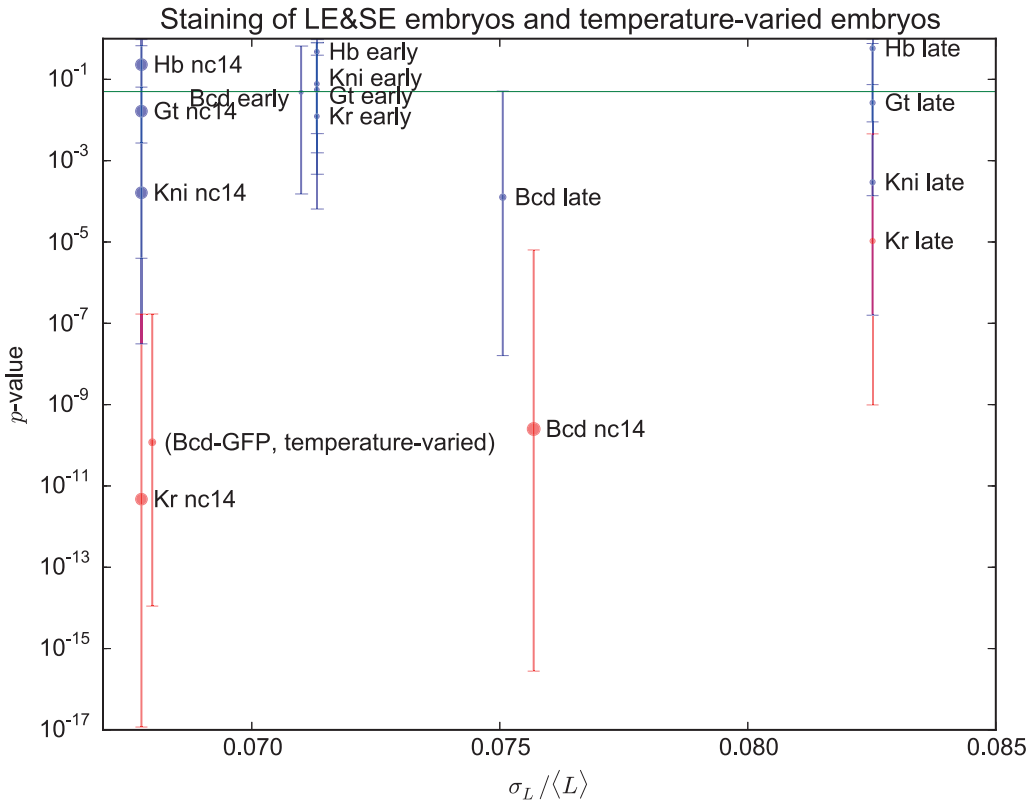

Figure 2.16: Principal component analysis of live and fixed embryos with enlarged length distributions rejects scaling for Bcd and Kr. p-values and 95% confidence intervals (calculated as in Figure 2.15) are shown for (1) LE&SE embryos stained for either Bcd or for the gap genes and (2) temperature-varied Bcd-GFP embryos (labeled as such). Stained Bcd and gap gene embryos are pooled into the age ranges of 15 to 25 minutes into nc14 ("early"), 40 to 50 minutes into nc14 ("late"), and all of nc14 ("nc14"). Axes, colors, green bar, marker sizes, and error bars are as in Figure 2.15.

gap genes in LE&SE embryos gives a slightly more nuanced result: the hypothesis of scaling is rejected for Kr in the time windows of late nc14 and all of nc14, for Bcd in the time window of all of nc14, and for no other combinations of gene and time window (Figure 2.16). Moreover, our dataset of Bcd profiles in temperature-varied Bcd-GFP embryos (discussed in Section 2.4) also fails the scaling hypothesis (Figure 2.16). While the lack of scaling of Bcd is to be expected, the implied lack of scaling of Kr is surprising, given that this finding is not corroborated by our PCA results in

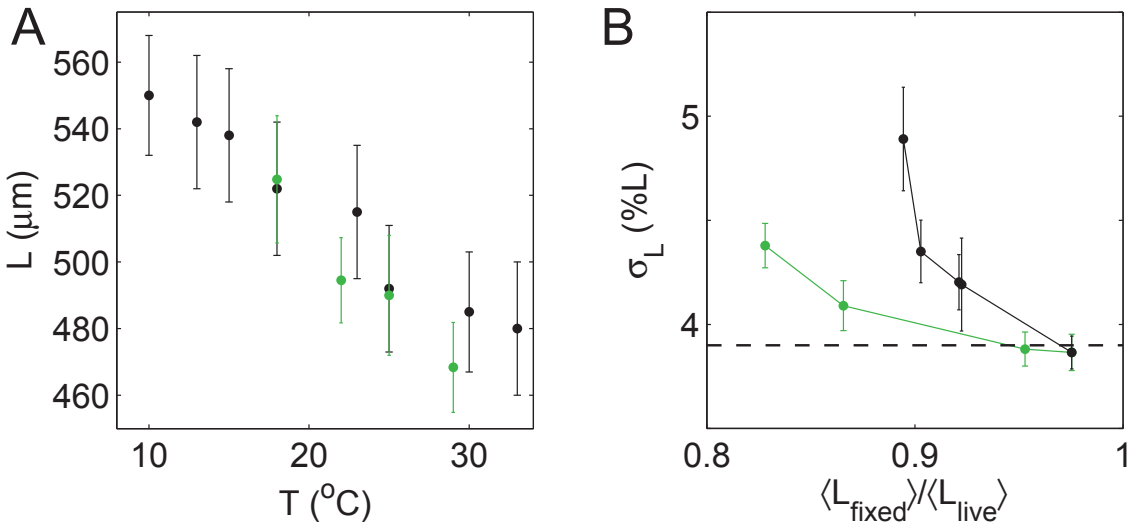

Figure 2.17: Embryo length varies with temperature and shrinkage rate from fixing. **(A)** Dependence of embryo length on temperature for OreR (black) and Bcd-GFP (green) embryos. Error bars are the standard deviation of the embryo length distribution. **(B)** Variability of embryo length vs. shrinkage rate of fixed embryos for OreR (black) and Bcd-GFP (green) populations. The dashed line represents the variability of the embryo length of live OreR embryos without fixing, $\sigma_L = 3.8 \pm 0.1\%$. Error bars are the standard deviation of the embryo length distribution.

immunostained wild-type embryos (Figure 2.15) or by our DPW results (Figures 2.2, 2.6). However, it is worth noting that of the four gap genes, DPW finds Kr to have the highest positional error as a function of position along the AP axis (Figure 2.6), and so it is plausible that the Kr expression profile may scale less precisely than the other gap genes.

2.8 Materials and Methods

Fly stocks and genetics

The artificially selected inbred fly lines with small and large embryos (SE and LE) are the fly lines #9.17 and #2.46, respectively, from Martin Kreitman's lab [15]. These stocks were kept at 22 °C. The Bcd-GFP fly line is $2X_A$ (*egfp-bcd*; +; bcd^{E1}) from our

Bcd-GFP fly line library [72]. One stock of Bcd-GFP was bred for embryo collection at 25 °C. Two additional Bcd-GFP fly stocks were bred at 18 °C and 29 °C environments for four months prior to imaging to make sure the embryo size adapted to the new temperatures. The Bcd-GFP flies bred at 18 °C had an embryo length of 524 ± 19 µm, and Bcd-GFP flies bred at 29 °C had an embryo length of 468 ± 14 µm (Figure 2.17A). All embryos from these two stocks were collected at 22 °C. The mutant fly lines Bcd$^+$ (+; +; *nos*$^-$*osk*$^-$) and Bcd$^-$ (+; +; *bcd*$^{\mathrm{E1}}$) are from Eric Weischaus's lab. All of these mutant stocks were bred for embryo collection at 22 °C.

Fixed imaging and live imaging

For fixed imaging, all embryos were collected at 25 °C, heat-fixed, immunostained, and imaged following the same protocol as described before [26]. The primary antibodies used for the maternal genes were rabbit anti-Bcd, rat anti-Cad, and guinea pig anti-Cic. All of the embryos were dechorionated to get rid of their outer eggshells before measurements were taken. One artifact of the immunostaining experiment was that it could introduce extra artificial variability of the embryo length due to differential embryo shrinkage. In our study, we only used the data sets with a shrinkage ratio $L_{\mathrm{fixed}}/L_{\mathrm{live}} > 0.9$, where L_{fixed} and L_{live} were the average embryo lengths of the fixed embryos and living embryos. We chose this cutoff based on our observation that when $L_{\mathrm{fixed}}/L_{\mathrm{live}} < 0.9$, the relative standard deviation σ_L of the fixed embryos is comparable with that of the living embryos (Figure 2.17B). This was critical if we wanted to investigate scaling in biologically relevant conditions. The live imaging with two-photon microscopy and imaging analysis were as described before [72].

Positional error calculation with DPW

All profiles were smoothed with 0.5% L (51 out of 1000 total data points). Only the data points in the region of 5–95% L were used in the analysis. For a given profile

set, the intensity of the individual profile was normalized by the global minimum and maximum intensity of all selected profiles. The reference profile was the mean of all profiles with an embryo length of $\langle L \rangle (1 \pm \sigma_L)$, where $\langle L \rangle$ was the average embryo length of all the embryos in the data set and σ_L was the standard deviation of the embryo length normalized by $\langle L \rangle$. This reference profile was the standard to which each individual sample profile in the population was compared using the DPW method. A relative intensity difference matrix was calculated as the pointwise intensity difference between points in the sample profile and points in the reference profile, squared and normalized by the pointwise sum of the intensities between the two profiles. A curve was found through this matrix that minimizes the sum of all of the matrix elements along the path: this represents the warping of the sample profile that best fits it to the reference profile. The core curve-finding algorithm was provided by Dan Ellis [27]. The positional error of the population of profiles of one gene at a particular position along the AP axis was calculated as the RMS distance between that position and the corresponding position in each of the individual profiles that that position was warped to. To calculate the combined positional error of multiple genes, the relative intensity difference matrix was calculated by summing the weighted matrices of the individual genes. For mathematical simplicity, the weight of each of the combined genes was assumed to be equal in this study. The magnitude of this positional error does not directly represent how well profiles scale. This is because the positional error includes natural variations in profile intensity unrelated to scaling and fluctuations in the profiles' intensity and baselines due to measurement errors. We systematically investigated the systematic errors of DPW: in the case that there are no amplitude or baseline fluctuations, the positional error calculation shows little shape dependence between exponential profiles and step-function-like profiles (Figure 2.8A,B). However, the positional error of more gradual profiles such as exponential profiles are more sensitive to intensity measurement errors (Figure 2.8A,B).

Simulations of positional errors

To understand the potential underlying mechanism for the scaling of Bcd, we also calculated the positional errors of simulated Bcd gradients using DPW. In the simulation, Bcd gradients approximated as exponential profiles $C(x) = C_0 e^{-x/\lambda}$. Based on the measurements, λ is not dependent on the embryo length L (Figure 2.7A), and thus it is fixed as 0.2 $\bar{L}$ for the gradients in the mid-coronal plane from the Bcd-GFP fly line and 0.14 $\bar{L}$ for the dorsal gradients from the LE&SE fly line. As suggested by the volume-dependent production model, C_0 could be modulated by the embryo length with the formula $\frac{\delta C_0}{C_0} = k\frac{\delta L}{L}$. The gradient noise at each position x is approximated as a Gaussian noise with the standard deviation of $C_{\text{noise}} = 0.15$ based on previous measurements on Bcd-GFP gradients. To take the measurement noise into account, we add a Gaussian noise with the standard deviation of A_{noise} to account for the overall intensity change due to either the offset of the focal plane, the fluctuation of the laser power and etc. We also add a Gaussian noise with the standard deviation of B_{noise} to account for the auto-fluorescence, the dark counts of the detector and etc. Taken all together, the observed profile $I(x) = (1 + A_{\text{noise}}g)\, C(x)\, (1 + C_{\text{noise}}g + B_{\text{noise}}g)$, where g is a random number drawn from the standard normal distribution. To fit the measured Bcd gradients, the positional errors of 100 simulated Bcd profiles were calculated with DPW to compare the positional errors of the measured Bcd gradients by searching on the parameter space. When the volume dependent production mechanism is on, there are three free parameters A_{noise}, B_{noise} and k; otherwise, k is set to be 0.

Chapter 3

Binding site number is the most important regulator of enhancer-mediated activation

This chapter reproduces a draft of a manuscript that will soon be submitted for publication. (Authors: Eric M. Smith, Jia Ling*, Hernan G. Garcia*, Hongtao Chen, Stephen J. Small, and Thomas Gregor.)*

Enhancers contain clusters of binding sites for transcription factors, but how the activities of individual sites are coordinated to direct gene expression is still unknown. Architectural features that are thought to affect expression output include binding site affinity, site number, and the spatial arrangements of adjacent sites. Here we use live imaging to investigate the impacts of single enhancer binding sites on two output parameters: mRNA expression levels and boundary positioning, which reflects enhancer sensitivity to transcription factor concentration. In contrast to previous studies, our measurements reveal that both parameters are linearly correlated with the number of activator binding sites, and not with measured binding site affinities or specific spatial arrangements of sites. The additive effects of individual sites reach a

saturation plateau that may reflect the maximum output of new transcripts from the basal promoter. These findings suggest a simple mechanism of transcriptional regulation in which binding site number is the most important input parameter driving gene expression output. We find that a cooperative model of protein-protein interactions, in which multiple enhancer-bound transcription factors interact with the promoter simultaneously, can largely explain this behavior. While we find the affinities and positions of endogenous binding sites to both be relatively unimportant in determining a binding site's effective role in regulating expression level, there exist synthetic binding site configurations in which both of these parameters dramatically influence expression level.

3.1 Introduction

Networks of transcriptional regulation are critical for establishing the body plans of multicellular organisms. The basic units of these networks are asymmetrically localized transcription factors and small DNA elements controlling target genes (enhancers), which contain between 5 and 20 transcription factor binding sites [2, 19]. Studies in many systems suggest that transcription factors can act as activators or repressors, and that the activities of individual enhancers are controlled qualitatively by the combination of binding sites in each enhancer [3]. Despite this extensive knowledge, the quantitative relationship between enhancer architecture, input transcription factor concentration, and transcriptional output is still largely unknown.

Here we focus on a detailed analysis of transcriptional activation mediated by a 270-bp proximal *hunchback* (*hb*) P2 enhancer, which is located upstream and adjacent to the *hb* P2 basal promoter. The P2 enhancer is activated by the maternal transcription factor Bicoid (Bcd), which is expressed in an anterior gradient [23, 71]. Bcd binds to 6 different well-characterized locations within the P2 enhancer [24], and

ID	Binding site combination	# Bcd sites	Orientation	*N*
1	__ __ __ __ __ __ __	0	1	5
2	__ __ __ __ Hb __ __	0	1	10
3	A1 __ __ __ Hb __ __	1	1	6
4	__ __ __ __ Hb A2 __	1	1	8
5	__ __ __ __ Hb __ A3	1	1	7
6	A1 __ __ __ Hb __ A3	2	1	8
7	__ __ __ __ Hb A2 A3	2	1	10
8	__ __ X2 X3 Hb __ __	2	1	6
9	A1 __ __ __ Hb A2 A3	3	1	12
10	__ X1 X2 X3 __ __ __	3	1	7
11	__ X1 X2 X3 Hb __ __	3	1	10
12	A1 X1 X2 X3 Hb __ __	4	1	6
13	__ X1 X2 X3 Hb A2 __	4	1	5
14	__ X1 X2 X3 Hb __ A3	4	1	6
15	A1 X1 X2 X3 Hb A2 __	5	1	7
16	A1 X1 X2 X3 Hb __ A3	5	1	9
17	__ X1 X2 X3 Hb A2 A3	5	1	6
18	A1 X1 X2 X3 Hb A2 A3	6	1	15

Table 3.1: Eighteen constructs were tested in which different combinations of Bcd and Hb binding sites were disabled via point mutation. All plots in Figure 3.1 were generated using embryos containing one of these 18 constructs, in which various combinations of binding sites were mutated out of the endogenous P2 enhancer. All listed constructs were inserted into the genome in the same orientation (labeled orientation 1). N represents the number of embryos imaged per construct.

drives gene expression throughout the anterior half of the early *Drosophila* embryo with an expression boundary at roughly 50% embryo length (EL) [25, 98]. Bcd is thought to control levels of *hb* RNA production in anterior nuclei and position the *hb* boundary, both via a threshold-dependent mechanism.

The six Bcd-binding sites in the P2-enhancer vary in sequence, with three high-affinity sites (A1–3, Figure 3.1A) and three tightly clustered lower-affinity sites (X1–3, Figure 3.1A) [24]. To quantitatively test the contributions of each site, we directly measured *in vivo* transcriptional outputs from 18 reporter genes containing precise

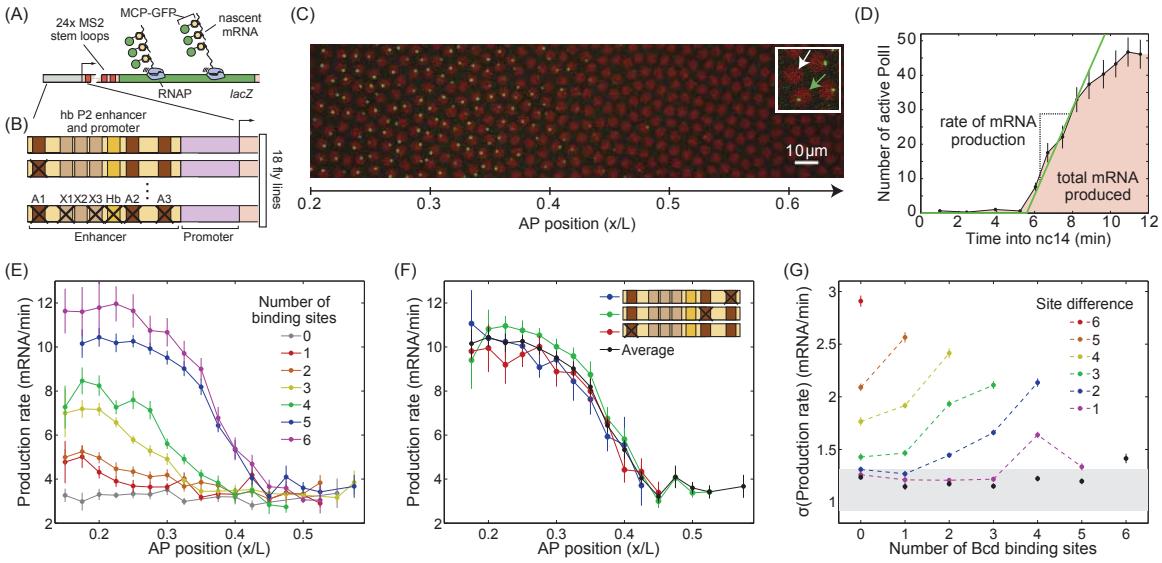

Figure 3.1: The rate of mRNA production is independent of binding site position or affinity. **(A)** Cartoon of template for reporter constructs. **(B)** Library of 18 P2 enhancer mutants with different combinations of deletions of Bcd and Hb binding sites. For the complete library list of mutations used refer to Table 3.1. **(C)** Individual sites of nascent transcript formation are visualized as fluorescent puncta allowing for quantification of the fraction of active nuclei and the rate of mRNA production. Inset: Magnified region of 14×14 μm showing an active (green arrow) and inactive nucleus (white arrow). **(D)** The mean fluorescence as a function of time for all active nuclei in a bin the size of 2.5% of the length of the AP axis reports on the average rate of mRNA production and the average number of total mRNA molecules produced. **(E)** Average rate of mRNA production as a function of AP position for different numbers of Bcd binding sites in the P2 enhancer. Numbers of Bcd binding sites reflect sites listed in [25]; more Bcd binding sites may be present in the *hb* P2 enhancer, as discussed in [76]. **(F)** Average rate of mRNA production as a function of position along the AP axis. Enhancers with the same number of binding sites yield statistically comparable rates of mRNA production. Here three examples of constructs with 5 Bcd binding sites each (colored lines) are shown, plus the average (blue line in E); for the complete set see Table 3.1. **(G)** The variability in the mRNA production rate among all embryos with a certain number of binding sites (black) is comparable to the embryo-to-embryo variability within a single fly line (grey bar), but lower than the variability among embryos bearing different numbers of Bicoid binding sites (colors). The black points represent the standard deviation of the rate of mRNA production among all embryos with a given number of Bicoid binding sites, and the gray band denotes the average (plus/minus one standard deviation) of the standard deviations of the rates of mRNA production over the individual collections of embryos sharing the same P2 enhancer mutant. Because the black points and gray band are at roughly the same height, we can consider different enhancers with the same number of binding sites to yield roughly the same rate of mRNA production. Each colored point is the standard deviation of the rate of mRNA production for all embryos with one of two numbers

Figure 3.1: *(continued)* of Bcd binding sites, where the lower of the two numbers for each point is given by its position along the horizontal axis. Because these colored points are above the black points that represent one given number of Bicoid binding sites, we can thus distinguish between embryos with different numbers of binding sites. (E–F: error bars are standard errors of the mean of the rates of mRNA production (fitted to raw fluorescence values) over multiple embryos; G, standard deviations are calculated as described in Figure 3.2 and error bars are obtained from bootstrapping over embryos. Absolute calibration error is not included and can be as high as 30% [33].)

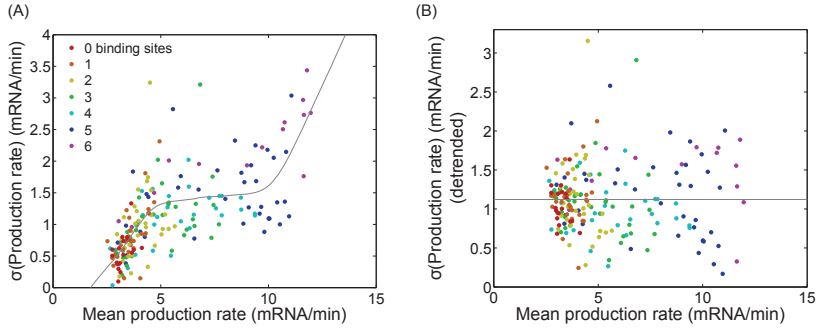

Figure 3.2: Detrending allows for a comparison in the standard deviations of the rate of mRNA production across different production rate levels. **(A)** Each point represents the mean and standard deviation of the rate of mRNA production in one 2.5% window along the AP axis for one of the 18 constructs listed in Table 3.1. Color represents the number of Bcd binding sites in the construct. A spline curve (gray) was fit to the standard deviation of production rate as a function of mean production rate. **(B)** The difference between each point in (A) and the spline fit, shown as a detrended standard deviation in production rate.

mutations in individual sites and combinations of sites (Figure 3.1A, B, Table 3.1). All constructs contained as a template the *hb* P2 enhancer, the *hb* P2 basal promoter, 24 MS2 stem loops, the *lacZ* coding region, and the alpha tubulin 3'UTR [33]. *In vivo* multi-photon fluorescence imaging of embryos containing single copies of each construct revealed individual sites of nascent transcription in each nucleus (Figure 3.1C). Intensity measurements of these sites over time allowed us to extract rates of transcriptional initiation and the total amount of produced mRNA molecules for each construct (Figure 3.1D).

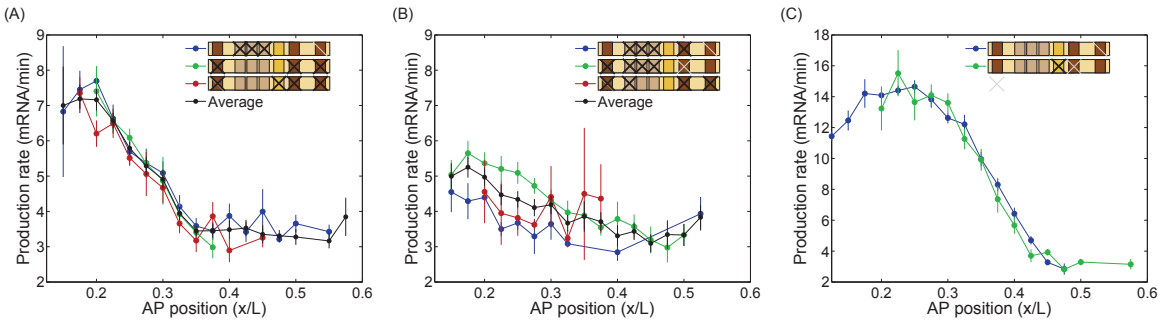

Figure 3.3: Enhancers with the same number of binding sites yield statistically comparable rates of mRNA production. **(A)** Average rate of mRNA production as a function of position along the AP axis. Here three examples of constructs with either only three "A" or three "X" Bcd binding sites (colored lines) are shown, as well as the average rate of mRNA production of embryos of all three constructs together (black line). **(B)** Same as (A), but three constructs with only two "A" or two "X" Bcd binding sites are compared. **(C)** Same as (A), but two constructs (#19 and 20 in Table 3.2) with all 6 endogenous Bcd binding sites that only differ in the presence of the Hb binding site are compared. (A–C: Error bars are standard errors of the mean of the rates of mRNA production (fitted to raw fluorescence values) over multiple embryos.)

3.2 Linearity of expression with binding site number

As expected, expression levels increase with the number of Bcd binding sites present in the constructs (Figure 3.1C,E). Surprisingly, however, mutations in strong A sites cause similar effects on mRNA production rates to those caused by mutations in weak X sites (Figure 3.3A,B). Additionally, mRNA production rates are independent of the positioning of the various binding sites and only depend on the number of binding sites present in a given enhancer construct (Figure 3.1F). The standard deviation of the rate of all embryos from all fly lines with the same number of binding sites is comparable to the standard deviations of multiple embryos of individual fly lines separately (Figure 3.1G). We thus grouped all the constructs by their binding site number, and collapsed the data for each group (Figure 3.1E). In contrast, when we group embryos with different numbers of binding sites together, the standard devia-

ID	Binding site combination	# Bcd sites	Orientation	*N*
19	(A1 X1 X2 X3 __ A2 A3)	6	2	8
20	(A1 X1 X2 X3 Hb A2 A3)	6	2	8
21	(A1 X1 X2 X3 Hb A2 A4 A3)	7	1	5
22	(A1 X1 X2 X3 Hb A2 A5 A3)	7	2	6
23	(A1 X1 X2 X3 Hb A2 A4 A5 A3)	8	2	3
24	(A2 A3) plus (A1 X1 X2 X3 Hb A2 A3)	8	2	3
25	(A1 X1 X2 X3) plus (A1 X1 X2 X3 Hb A2 A3)	10	?	4
26	(A1 X1 X2 X3 Hb A2 A3) plus (A1 X1 X2 X3 Hb A2 A3)	12	1	3

Table 3.2: Further constructs allow for investigation of additional effects on mRNA production. Constructs in the above table are not used in Figure 3.1 but feature in later figures. Among these are (1) constructs that were integrated into the genome in the opposite orientation (#19 and 20, labeled as orientation 2) to those in Table 3.1 and (2) constructs with additional Bcd binding sites not found in the endogenous P2 enhancer (#21 through 26), either by mutating binding sites into the endogenous enhancer or by integrating fragments of a second P2 enhancer next to the existing enhancer (Figure 3.4D). N represents the number of embryos imaged per construct.

tions of the rates in these groups are higher (Figure 3.1G), demonstrating that our measurements can distinguish between populations of embryos with different numbers of binding sites. Thus the specifics of the individual sequences and their relative position within the enhancer are much less predictive of transcription rate than the actual number and affinity of binding elements [14, 25, 98].

To systematically characterize the effect of binding site copy number on enhancer action quantitatively, we use a more convenient representation of our data in terms of input-output functions for the rate of mRNA production, where the position along the AP axis is replaced by its corresponding Bcd concentration (Figure 3.4A). When performing a fit of each input-output function to a sigmoidal curve, we obtain the maximum and basal rates of transcription, the dissociation constant, and the sensitivity of the curve (see Figure 3.4 caption). Despite the highly nonlinear nature of the input-output functions, we find that the rate of transcription increases linearly as binding site number is varied between 0 and 6, while the basal rate remains constant

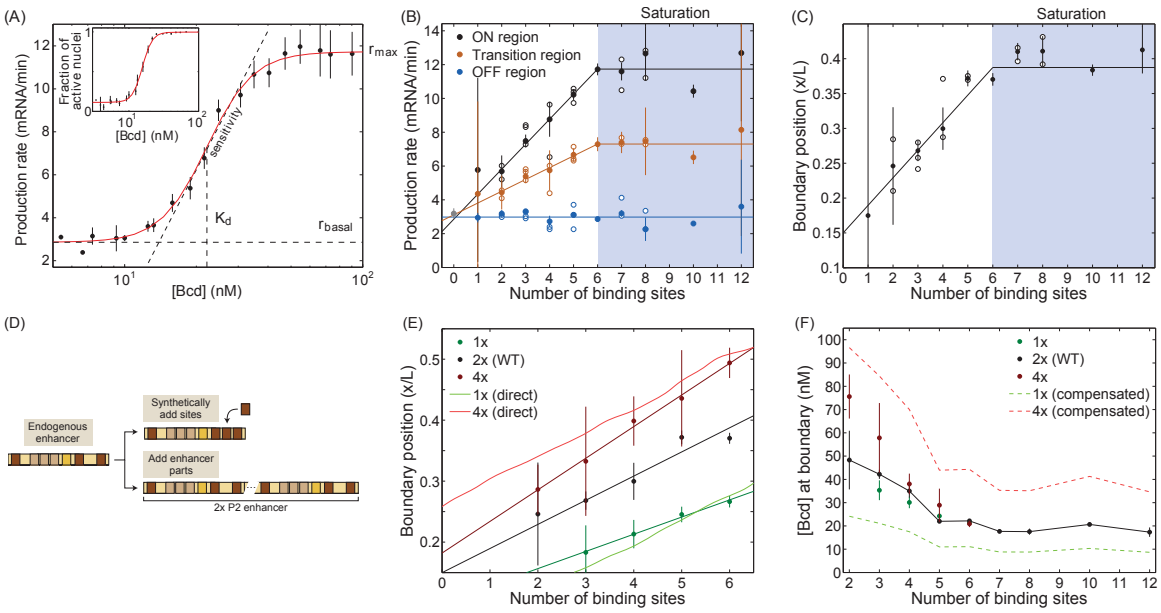

Figure 3.4: Rate of mRNA production and boundary position scale linearly with number of Bicoid binding sites. **(A)** Fit of the rate of mRNA production as a function of the Bcd activator concentration for the wild-type P2 enhancer (six Bicoid binding sites). Bcd concentration is inferred from the AP position (Figure 3.5A; [72]). The data are fitted to a sigmoidal curve with sensitivity 4.5 ± 0.8. Inset: Fit of the fraction of active nuclei as a function of Bcd concentration to a sigmoidal curve with a sensitivity of 8 ± 1. **(B)** Maximum (black), transition (orange) and basal (blue) rate of transcription as a function of the number of Bcd binding sites in the P2 enhancer display linear and saturated ranges. Gray point represents the mean rate of transcription for embryos with a P2 enhancer containing no Bcd binding sites. The slope in the increase of the maximum rate of transcription with binding site number is 1.5 ± 0.1 mRNA/min/binding site. The mean basal transcription rate is 3 ± 1 mRNA/min. Filled circles represent all embryos with a certain number of Bcd binding sites and open circles represent embryos from individual constructs listed in Table 3.1 and all constructs with more than 6 binding sites in Table 3.2. All data are from sigmoidal fits to the rates of mRNA production (Figure 3.5B). **(C)** Boundary position as a function of number of Bicoid binding sites. The position is inferred from the fitted K_d in Figure 3.5B following Figure 3.5A. **(D)** Two strategies were employed to add extra binding sites to the endogenous P2 enhancer: synthetically mutating Bcd binding sites into the endogenous enhancer (top) and adding parts of an extra enhancer beside and further away from the promoter than the existing one (bottom). **(E)** Boundary position as a function of number of binding sites shows a linear trend for different Bicoid dosages: for the wildtype level (2x) of Bicoid the slope of this trend is 0.040 x/L per binding site, for half the wildtype level (1x) the slope is 0.028 x/L per binding site, and for double the wildtype level (4x) the slope is 0.051 x/L per binding site. Light green and red lines represent the expected boundary positions for 1x and 4x dosage, respectively, if mRNA production was a direct readout of Bcd concentration for altered Bcd dosages. (See Figure 3.6A,B for analogous plots of the

Figure 3.4: *(continued)* fitted maximum production rate and boundary position of the fraction of active nuclei.) Expected boundary position curves calculated by interpolating from the known Bcd concentration profile with halved or doubled amplitudes, following Figure 3.5A. Dark green, black, and dark red lines represent regression lines to Bcd dosage points. **(F)** The Bicoid concentration at the boundary remains largely unchanged within experimental error independently of the activator dosage, supporting the claim that our reporter construct responds directly to Bicoid concentration. Light green and red dashed lines represent the expected boundary positions for 1x and 4x dosage, respectively, if mRNA production perfectly compensated for changes in Bcd dosage, i.e. if boundary positions exactly matched their 2x counterparts. (See Figure 3.6C for an analogous plot of the threshold Bicoid concentration at the boundary between high and low fractions of active nuclei.) (A: error bars are standard errors of the mean of the rates of mRNA production (fitted to raw fluorescence values) over multiple embryos; B, C, E, and F: error bars are 68% confidence intervals from fits of the rates of mRNA production to sigmoidal functions. C, E, and F: open and filled circles are as in B. Absolute calibration error is not included and can be as high as 30% [33].)

for all our constructs (Figure 3.4B). In addition, the Bcd concentration corresponding to each dissociation constant defines the boundary between high and low rates of expression along the AP axis. Interestingly, this boundary position also varies nearly linearly in the range of 0 to 6 Bcd binding sites (Figure 3.4C). Thus both the initiation rate and the boundary location depend additively on how many of the six endogenous Bcd binding sites are present, and they seem to be independent of binding site strength and site distance from the promoter.

3.3 Extensions to the endogenous system

To test the upper limit of these linear relationships, we increased the number of binding sites in our constructs beyond the six endogenous sites found in the P2 enhancer. Multiple cloning strategies were used, such as adding exogenous sites within the P2 enhancer, and splicing part of a second P2 enhancer next to the first (Figure 3.4D). However, for any number of possible binding sites larger than the endogenous six, both the maximum rate and the location of the boundary saturate (Figure 3.4B,C).

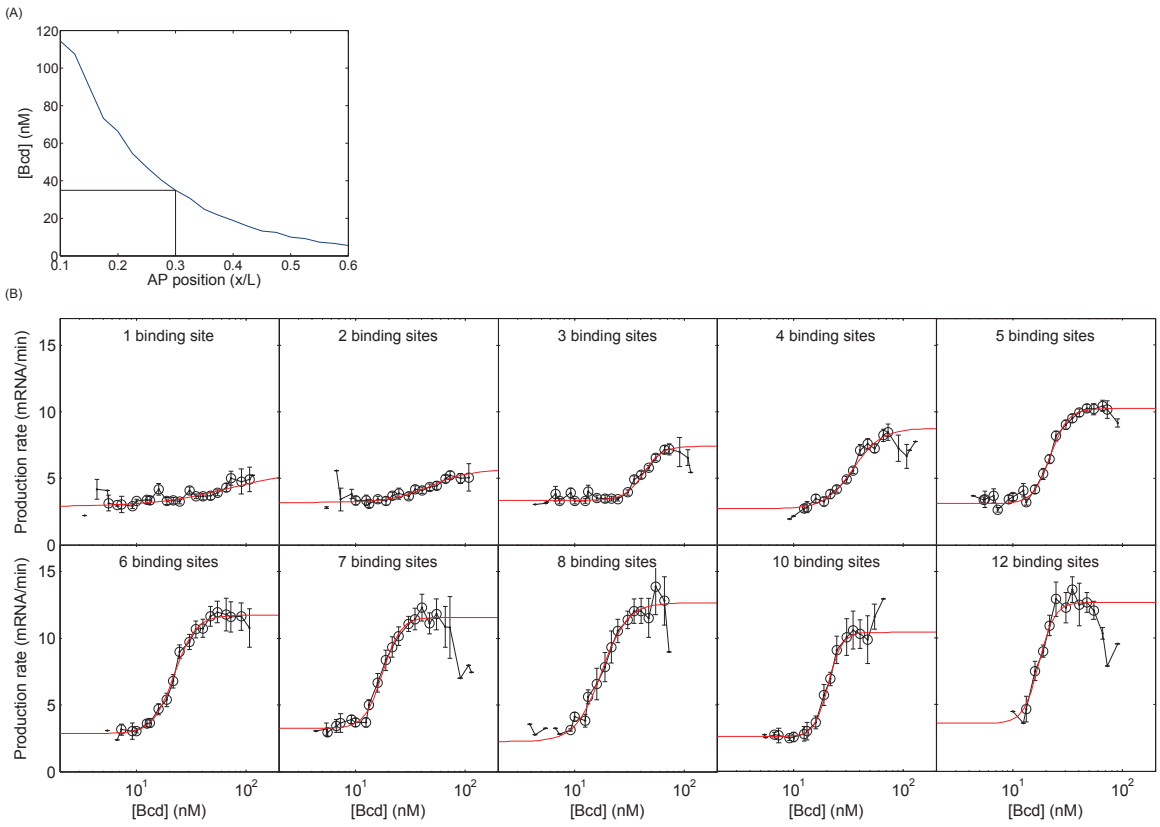

Figure 3.5: Extracting curves of mRNA production rate from raw data and known Bcd concentration. **(A)** The AP position can be extracted from Bcd concentration and vice versa. A curve of the Bcd concentration profile recorded live from Bcd-GFP embryos [72] is used to interpolate a Bcd concentration from an arbitrary AP position or vice versa, allowing for calculations of boundary positions from fits of K_d to rates of mRNA production, as in Figure 3.4A. The black vertical line represents 30% AP, and the black horizontal line represents the corresponding Bcd concentration, 35.0 nM, at that position. **(B)** Sigmoid functions (red) are fit to rates of mRNA production as a function of Bcd concentration (black). For each curve of N binding sites, the range of Bcd concentration to fit to was determined by eye (black circles). Error bars on rates of mRNA production are standard errors of the mean rate of mRNA production over all embryos with N binding sites.

Hence linearity is bound at the lower end by our detection limit (Figure 3.1G) and has an intriguing upper bound that coincides with the number of binding sites found in the endogenous enhancer.

To test whether Bcd is solely responsible for the observed linear relationships, we performed live imaging measurements on a subset of our constructs in different Bcd

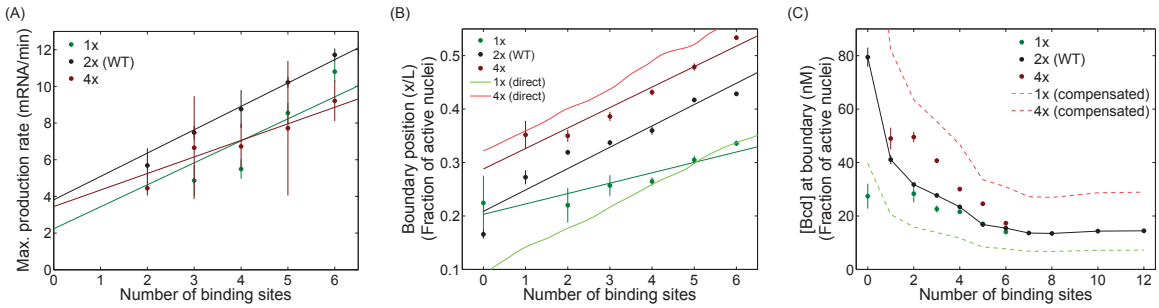

Figure 3.6: The maximum rate of mRNA production and the boundary position between high and low fractions of active nuclei vary as a function of Bcd dosage. **(A)** The maximum rate of mRNA production, as extracted from fits of sigmoid functions to the rates of mRNA production along the AP axis, is lower for embryos with half and twice the WT amount of Bcd (green and red respectively) than for embryos with WT amounts of Bcd, given the same number of Bcd binding sites. **(B)** The boundary position between high and low fractions of active nuclei (extracted according to Figure 3.4A inset) varies as a function of Bcd dosage and the number of Bcd binding sites in a way that suggests partial compensation for altered Bcd dosage. Light red and light green curves represent expected boundary positions if rates of mRNA production directly read out Bcd concentration with no compensation, as in Figure 3.4E. (A, B: Dark green, black, and dark red lines represent regression lines to Bcd dosage points.) **(C)** The Bcd concentrations at the boundary between high and low fractions of mRNA production vary according to Bcd dosage for a given number of Bcd binding sites. Light green and light red dashed lines represent the expected Bcd concentrations at the boundary if boundary positions along the AP axis stay fixed as the Bcd dosage changes, as in Figure 3.4F. Error bars for all plots represent 68% confidence intervals of fits of the rates of mRNA production (for A) or the fraction of active nuclei (for B and C) to sigmoid functions.

dosage backgrounds. Within the range of the six endogenous Bcd binding sites we see the same linear relationship for boundary position as a function of number of sites regardless of whether the measurements are performed with half or twice the wild-type Bcd dosage (Figure 3.4E). Thus, the placement of the mRNA expression boundary in embryos with altered Bcd dosage can be predicted solely (for constructs with many Bcd binding sites) or largely (for constructs with few Bcd binding sites) by the shift in Bcd spatial concentration caused by the altered dosage. These results suggest that the Bcd gradient is the main driver of the mRNA expression pattern directed by the P2-enhancer, and challenges the idea that it responds synergistically

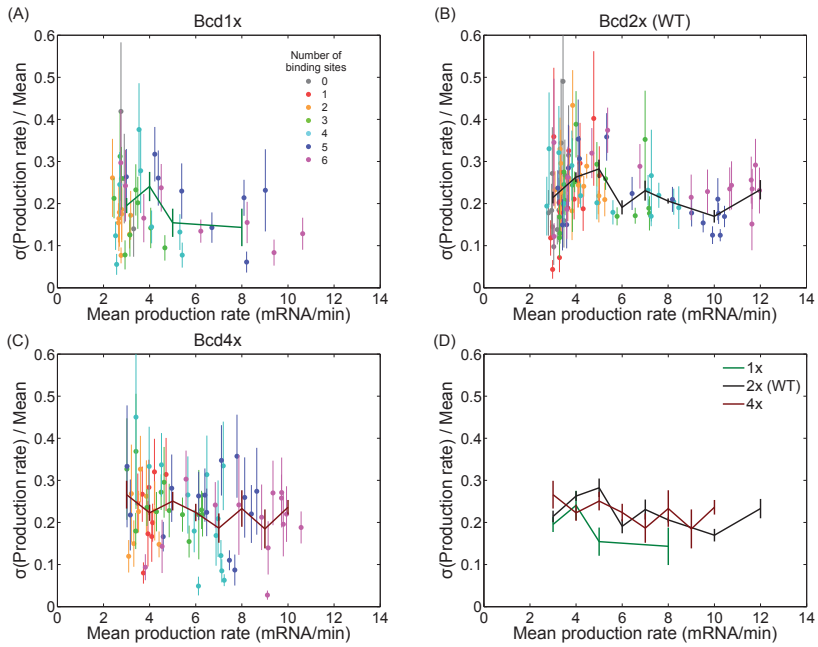

Figure 3.7: The variability of the rate of mRNA production is on the level of 20% for all Bcd dosages. **(A–C)** Each colored point represents the standard deviation of the rate of mRNA production divided by the mean rate of mRNA production in a 2.5% AP window for all embryos with a given number of Bcd binding sites, indicated by color, for embryos with half-WT (A), WT (B), and double-WT (C) levels of Bcd dosage. Error bars on points represent standard errors from bootstrapping over embryos, and only 2.5% AP windows in which rates of mRNA production were extracted from at least 3 embryos are included. Solid curves represent the plotted points collapsed into bins of width 1 mRNA/min; only bins with at least 3 points (indicated by the presence of an error bar) are plotted. Error bars on solid curves are standard errors over all points in the bin. **(D)** Each of the binned trendlines from (A–C) are plotted together, showing roughly 20% variability for each Bcd dosage.

to inputs from Bcd and Hb [87, 93]. To test this directly, we mutated the Hb site in the P2 enhancer, which had very little effect on enhancer activity (Figure 3.3A,C).

However, whereas constructs with high numbers of binding sites exhibit the same threshold Bcd concentrations regardless of Bcd dosage, we noticed that constructs with lower numbers of binding sites "compensate" for the change in Bcd concentration by shifting their threshold Bcd concentrations accordingly (Figure 3.4F). As a result, the boundaries between domains of high and low mRNA expression in constructs with few Bcd binding sites are closer to those found in embryos with WT Bcd dosage than

would be expected if the response to dosage is completely linear for all constructs (Figure 3.4E). The variability in the rate of mRNA production across embryos falls in the 20% region for all Bcd dosages (Figure 3.7). We speculate that the presence of many Bcd binding sites in the wild type P2 enhancer may override the effects of Hb in the placement of the mRNA expression boundary, regardless of Bcd dosage. In constructs with lower numbers of Bcd binding sites, however, Hb may tend to push the mRNA expression boundary in embryos of altered Bcd dosage towards where it would be if the Bcd dosage were unperturbed.

3.4 The role of protein-protein cooperativity in gene expression

To gain insight into the underlying molecular mechanisms of Bcd activation that have been further constrained by our observations, we invoke a thermodynamic DNA-binding model [10, 11]. Homotypic clusters of transcription factors play key roles in endogenous developmental programs [68], and they have been used extensively in synthetic contexts to increase the level of gene expression of minimal enhancers [29, 93, 100]. Despite extensive use of thermodynamic models describing protein-DNA binding of such constructs [29, 35, 92], no consensus exists about the effect of multiple activator DNA occupancy on the ensuing rate of transcription. Here we consider two scenarios that relate enhancer occupancy directly to actual levels of gene expression (Figure 3.8A).

In the first scenario, only a single activator bound to a single site within a homotypic cluster interacts with the transcriptional machinery [29]. In this *simple activation* case, the role of multiple binding sites in a regulatory element is to increase the probability of activator occupancy at the enhancer. In a second scenario, called *additive activation*, multiple activators can be bound simultaneously and inter-

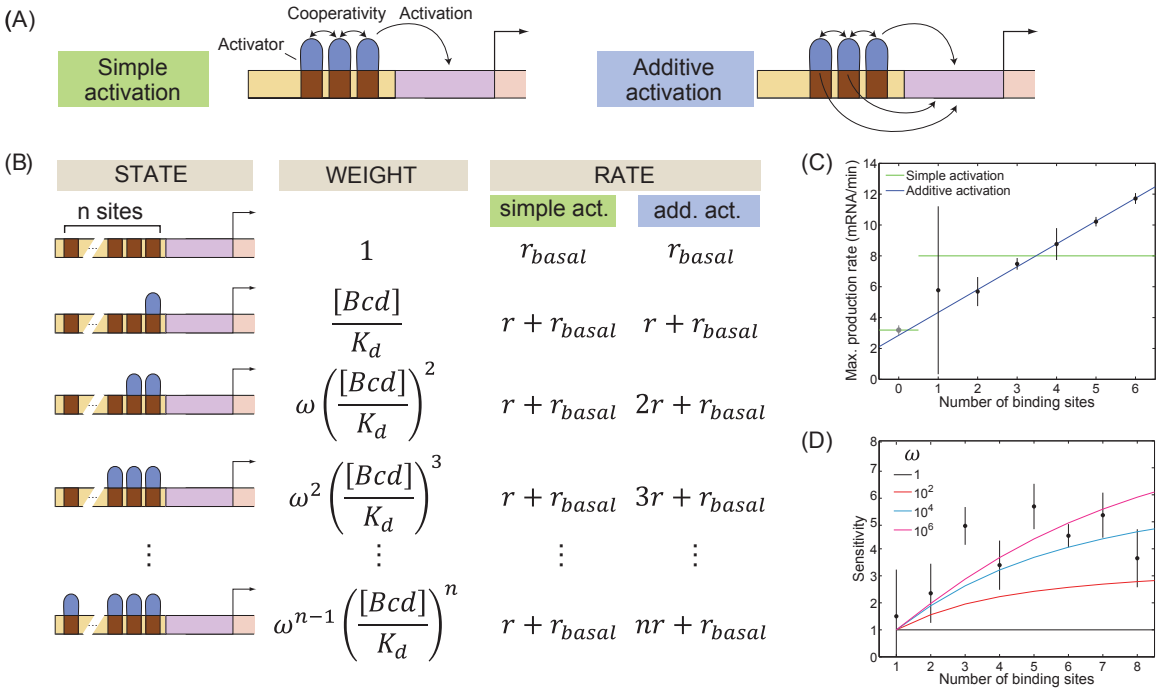

Figure 3.8: Multiple Bicoid molecules enhance transcription additively and are recruited to the DNA cooperatively. **(A)** Competing models of activator action. In simple activation models only one activator contacts the transcriptional machinery at a time. Here, the role of extra binding sites is, through cooperativity between Bcd molecules, to increase the probability of having one activator bound to the enhancer [29]. In a synergistic activation scenario each bound Bicoid molecule contacts the transcriptional machinery independently resulting in an additive contribution to the total rate of transcription. **(B)** Thermodynamic models accounting for simple and additive activation. The states, their degeneracies, and weights for both models are identical, but the rate of transcription corresponding to each state is unique to each model. Model parameters are the Bcd concentration [Bcd], the dissociation constant K_d corresponding to each binding site, and a parameter ω for protein-protein interactions between bound Bcd molecules. **(C)** The maximum rate of transcription in the simple activation scenario is independent of the number of binding sites. In contrast, the maximum rate of transcription for additive activation grows linearly with the number of activator binding sites and is consistent with our data. (Black and gray points are as in Figure 3.4B.) **(D)** Sensitivity of the rate of mRNA production as a function of the number of Bicoid binding sites. Sensitivity values larger than one suggest a cooperativity mechanism such as Bicoid protein-protein interactions reflected by values of the cooperativity parameter larger than one. (C and D: error bars are standard errors of the mean of raw fluorescence values over multiple embryos. Absolute calibration error is not included and can be as high as 30% [33]).

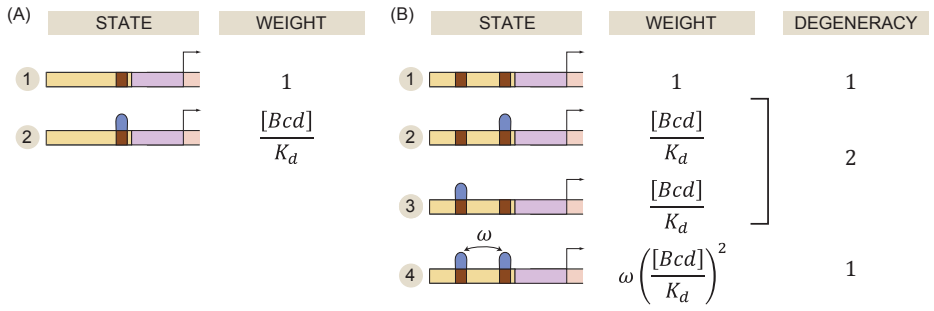

Figure 3.9: Hierarchy of models describing Bicoid binding. **(A)** Model of an enhancer with a single Bcd binding site. The dissociation constant is given by K_d. **(B)** Model of an enhancer with two Bcd binding sites. Here, the cooperativity parameter ω is introduced and the degeneracy of states bearing the same number of bound Bcd molecules is taken into account.

act with the transcriptional machinery independently, contributing additively to the ultimate rate of mRNA production [92]. The two models are identical in terms of the thermodynamic description of Bcd-binding, resulting in their having the same states and corresponding statistical weights (Figure 3.8B, see Materials and Methods, and Figure 3.9).

Despite identical statistical weights, however, each model predicts a different rate of transcriptional initiation. In *simple activation* all states with at least one Bcd molecule bound lead to a rate of mRNA production of r over the basal rate r_{basal}. In the *additive activation* model the rate of mRNA production scales linearly with the number of bound Bcd molecules. At saturating Bcd concentrations all binding sites are occupied, and the two models make polarizing predictions on how the rate depends on binding site occupancy, as seen in Figure 3.8C. Clearly, our data are consistent with linear scaling, and thus with the *additive activation* model. This suggests that multiple Bcd molecules can contribute independently to transcriptional activation and that this independent effect adds linearly.

The origin of the spatial non-linearity leading to a sharp posterior *hb* boundary has been repeatedly suggested to originate from Bcd activator protein-protein inter-

actions leading to DNA-binding cooperativity [14, 25, 41, 67, 73, 76, 98]. Indeed, *in vitro* and *in vivo* experiments have identified Bcd residues that alter the degree of such interactions [13, 67]. This cooperativity is reflected in our experiments by sensitivity values that are greater than one (Figure 3.8D). To account for these data, we incorporated cooperativity into our model using an interaction parameter ω (Figure 3.8B). A value of $\omega = 1$ denotes no cooperativity, while higher values correspond to cooperative interactions. Our data match values of ω that correspond to interaction energy values $\ln(\omega)\, k_{\mathrm{B}}T$ ranging between 9–14 $k_{\mathrm{B}}T$. These energies are significantly higher than those for pure protein-protein interactions, which are expected to be on the order of 3 $k_{\mathrm{B}}T$ [88]. The need for such high interaction energies to explain the sensitivity of observed boundaries suggests the presence of extra interactions leading to cooperativity (see SI). Despite these discrepancies, our measurements show an unequivocal direct measurement of cooperativity in Bcd binding to enhancers and present the potential to use our assay to quantify how this cooperativity changes for different Bcd mutants and enhancer architectures.

3.5 Discussion of enhancer dissection

Our model shows clear signatures of cooperativity leading to sharp boundaries along the embryo. The independence on Bcd binding site positioning together with the large measured sensitivities of the input-output functions suggest that Bcd protein-protein interactions are much more flexible than observed in classic examples such as the interferon gamma enhancer [64], Mat alpha 2 [62], and Lambda repressor binding [69]. In these examples cooperativity strongly depends on the relative position and orientation of nearby binding sites for transcription factors. The observed indifference of the effective cooperativity observed against changes in the position of binding sites could arise from global effects that are not based on direct protein-protein interactions

between Bcd molecules. For example, DNA binding proteins could indirectly interact with each other by aiding in the eviction of a nucleosome occluding DNA binding sites. Indeed, such mechanism of cooperativity has been demonstrated in yeast [79] and has been further theorized to be of broad relevance to transcriptional regulation [50, 80].

This study presents a systematic and quantitative dissection of the design features of a simple regulatory sequence element within a developmental program. Two different models for enhancer function have recently been put forth [65], the *enhanceosome* model and the *billboard* model. The *enhanceosome* model determines enhancer function by the precise placement of multiple transcription factors along the enhancer allowing for specific interactions between each other and with the transcriptional machinery. In the *billboard* model transcription factors interactions are flexible but the identity and number of the different molecular players bound to an enhancer rather than their particular spatial distribution determine enhancer action. Our data on the *hunchback* P2 enhancer is clearly consistent with the latter. The only parameter relevant for determining the output level of gene expression is its number of Bcd binding sites, with each bound Bcd molecule contributing additively and independently to the rate of mRNA production. This work exemplifies the power of quantitative analyses to constrain models of the microscopic action of transcription factors that determine the body plan of a multicellular organism and calls for the further systematic and quantitative dissection of more complex regulatory decisions in development.

3.6 Modifications of binding site position and affinity reveal additional regulatory effects

In Section 3.2, we removed binding sites from the endogenous proximal *hb* enhancer in order to demonstrate that all binding sites in the endogenous enhancer have roughly

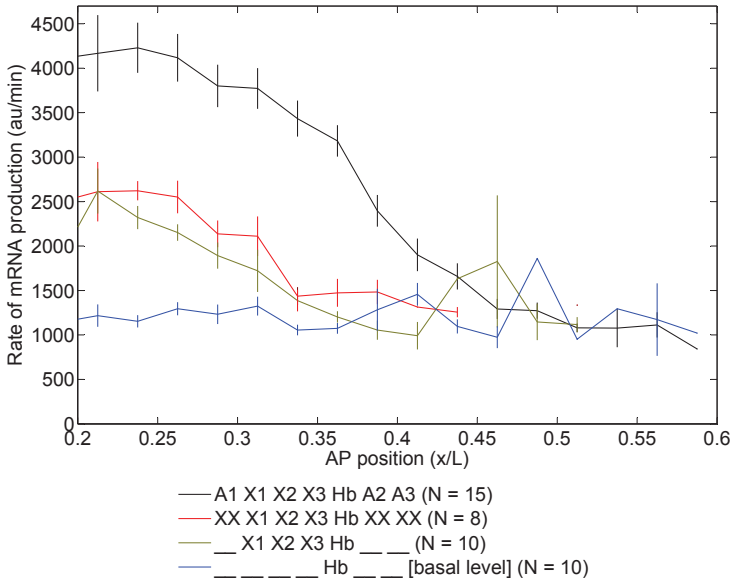

Figure 3.10: Mutating the affinity of strong sites to weak reduces expression level and domain. The rate of mRNA production in nc14 is shown for embryos containing constructs in which transcription is regulated by the WT *hb* P2 enhancer (black), a P2 enhancer in which the three strong sites have been mutated into weak sites (red), a P2 enhancer in which the strong sites have been disabled completely (yellow), and a P2 enhancer in which all six strong and weak sites have been disabled (blue). Error bars are standard errors of the mean rate of mRNA production over all embryos.

equal contributions to expression level, despite differences in binding site affinity and position: however, this does not prove that binding sites are always functionally identical *in vivo* regardless of affinity or position. This section details two experiments in which *alterations* of binding site affinity and position reveal both parameters to have immense regulatory power under certain conditions foreign to the endogenous enhancer.

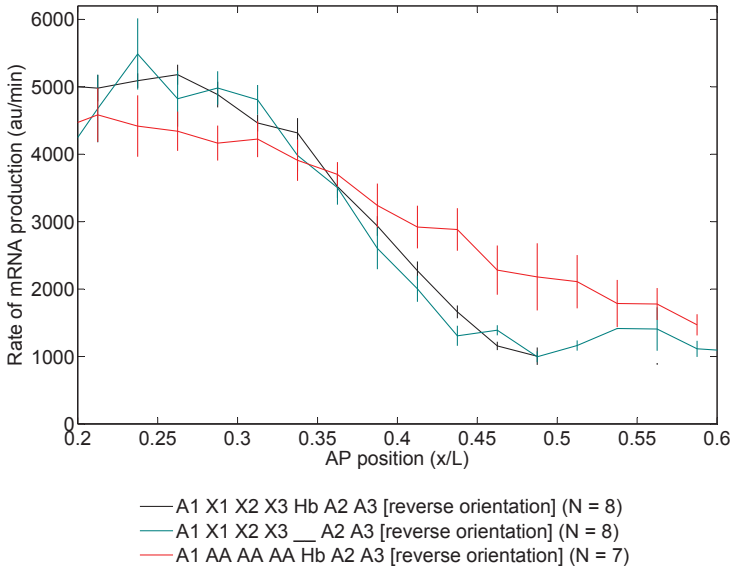

Figure 3.11: Mutating the affinity of weak sites to strong alters the expression pattern. The rate of mRNA production in nc14 is shown for embryos containing constructs in which transcription is regulated by the WT *hb* P2 enhancer (black), a P2 enhancer in which the Hb autoregulatory site has been disabled (teal), and a P2 enhancer in which the weak sites have been mutated into strong sites (red). All three constructs are in orientation 2 (Table 3.2). Error bars are standard errors of the mean rate of mRNA production over all embryos.

Changes in the affinity of endogenous binding sites change expression level and domain

Another way to test the affinity of binding sites, in addition to removing them completely via point mutations, is to change the affinity of endogenous binding sites by altering the binding site sequences to convert the strong sites to weak sites or vice versa. We previously found that the expression level and domain, under regulation by the three strong sites A1, A2, and A3 alone, is the same as when they are regulated by the three weak sites X1, X2, and X3 alone (Figure 3.3A): this implies that the contributions of the strong and the weak sites to the overall expression pattern are the same. One plausible hypothesis, then, would be that altering weak sites to strong

sites or vice versa should not have an effect on the expression pattern. However, measurements of such constructs completely betray this hypothesis: altering the three strong sites to weak sites reduces the level of gene expression to roughly that found with three weak sites alone (Figure 3.10), and altering the three weak sites to strong sites makes the expression domain larger and the boundary between high and low expression regions much shallower (Figure 3.11).

These results indicate that affinity plays a strong role in expression level *in vivo*: the challenge, then, is to reconcile this finding with the knowledge that the three strong and the three weak sites have the same effect when regulating expression by themselves. One possible hypothesis is that the cooperativity among binding sites, alluded to in Section 3.4, is dependent on the distance between neighboring binding sites: it is conceivable that the weakness of the weak binding sites in binding protein and driving increased expression may be counteracted in the wild-type *hb* P2 enhancer by the fact that the three weak binding sites are roughly 16 and 13 base pairs apart, end-to-end, whereas the strong binding sites are roughly 110 and 105 base pairs apart [25]. Thus, binding to a weak binding site may be made more likely when proteins are already bound to neighboring, nearby weak binding sites, whereas strong binding sites may be strong enough that cooperativity among those sites is less crucial.

Changes in the position of the wild-type enhancer relative to the promoter change expression level

Just as we mutated the affinity of endogenous binding sites in order to directly measure the differences in the functions of binding sites of different affinity, we can also alter the spacing between the binding sites and the promoter in order to provide another test of the dependence of the expression pattern on binding site position. The nearest of the six known Bcd binding sites on the *hb* P2 enhancer is located only 59 base pairs away from the transcription start site in the wild-type fly [25]. Thus, if

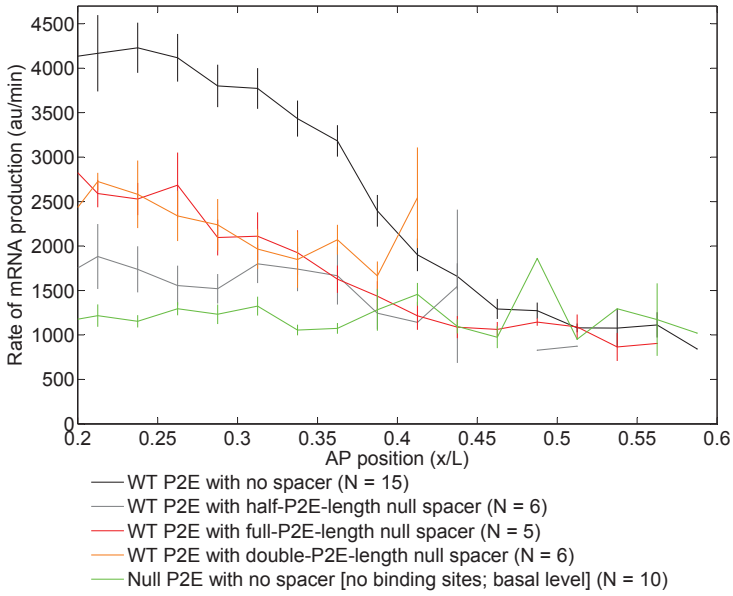

Figure 3.12: The expression level varies non-monotonically as a function of the distance between the *hb* promoter and the proximal enhancer. The rate of mRNA production in nc14 is shown for embryos containing constructs in which the wild-type *hb* P2 enhancer is placed in the wild-type position near the P2 promoter (black), as well as separated from the promoter by half of the length (gray), the entire length (red), and double the length (orange) of the P2 enhancer. The wild-type P2 enhancer, located in the wild-type position relative to the promoter but with all Bcd binding sites disabled via point mutation (green), is shown for comparison. Error bars are standard errors of the mean rate of mRNA production over all embryos.

there is any dependence of the expression pattern on binding site distance from the promoter, adding extra nucleotides between this closest Bcd binding site and the promoter (and thus moving all six binding sites further away from the promoter) should dramatically reveal this dependence by creating a large change in the expression pattern. We created three constructs in which the wild-type P2 enhancer is separated from the P2 promoter by a different fraction of the total length of the enhancer (either 0.5, 1, or 2 times the length of the enhancer), and we find that, surprisingly, the expression pattern varies non-monotonically as a function of the distance between the enhancer and the promoter (Figure 3.12). Specifically, the expression level drops

to close to the basal level when the enhancer is separated from the promoter by a distance of half of the length of the enhancer, but this level returns to roughly halfway between wild-type and basal levels of expression when the distance is increased to one or two times the length of the enhancer. (In all three constructs, the spacer between the enhancer and promoter is composed of duplicates of the wild-type *hb* P2 enhancer in which all known binding sites have been mutated out.) This counter-intuitive result clearly reflects a deeper level of complication in the regulation of expression level by binding site position, and may indicate the role of nucleosome positioning in affecting promoter-enhancer accessibility [50, 80].

It is important to note that both of the tests of the regulatory roles of affinity and position described in this section are in some way more invasive than the tests of affinity and position discussed in Section 3.2, where the linearity of expression level and domain with binding site number is described. Whereas those previous tests of affinity and position merely disabled binding sites in various combinations, the two tests described in this section create new enhancer configurations in which binding sites are placed in contexts exogenous to the wild-type fly embryo, which may not provide a fair assessment of the roles of binding sites in the wild-type enhancer. Still, both sets of perturbations to the *hb* P2 enhancer provide insights about the possible effects that changes in binding site position and affinity can have on transcriptional activity. Altogether, we see from these experiments that there are cases in which both affinity and position have very strong impacts on the resultant expression level, but that in the endogenous context these impacts are lessened.

3.7 Materials and Methods

Transgene construction and fly strains

A 472-bp fragment spanning the hunchback P2 enhancer and *hb* basal promoter (-321 to +151 bp) was either PCR-amplified (for #18 and #20) or chemically synthesized (for #1–#17, #19, and #21–#23) (Integrated DNA Technologies, Inc.), with HindIII added at 5' and NcoI added at 3'. A1 (TAATCC) was mutated to TCAACT, A2 and A3 (TAATCC) were mutated to CATACT, X1 and X2 (TAAGCT) were mutated to TAGACT, and X3 (TCATCCAA) was mutated to TACATACC, and Hb site (TTTTTTG) was mutated to ATGTTCA. None of the mutations introduce new Bcd or Hb binding sites according to JASPAR [111]. To create extra Bcd sites A4 and A5, hexamers 26-bp and 58-bp downstream of A2 resembling weak Bcd sites were converted to TAATCC. Amplified or synthesized fragments were digested with HindIII and NcoI and placed upstream of 24 MS2 stem loops and a *lacZ* reporter gene (Figure 3.1A) [33]. For other constructs with more than six Bcd sites, an intact or spliced P2 sequence was PCR-amplified, digested with HindIII and inserted 5' to the HindIII digested construct #18. The enhancer-promoter regions and the MS2 stem loops of all reporter constructs were verified by sequencing. Those constructs have two attB sites in the backbone and were inserted into the 38F1 landing site on chromosome II using ΦC31 integrase-mediated cassette exchange [5].

To identify the orientation of the transgene insertion in the genome, genomic DNA was extracted and amplified with two pairs of primers: sjs584 (AACAACACTATTATGCCCACCA) and P2TATAF (TTGGTGCTGCTTCTGTTG), sjs586 (CTGGAATTCGGCTTCGACT) and P2TATAF. sjs584 and sjs586 reside in the two attP sites flanking the transgene respectively and P2TATAF is inside the *hb* basal promoter. If the combination of sjs586 and P2TATAF produced an 800 bp band, the orientation of that transgene was designated as 1.

For measurements of embryos with WT Bcd dosage (Bcd2x), female virgins containing MCP-GFP and Histone-RFP fluorophores (yw;Histone-RFP;MCP-NoNLS(2) for embryos with WT Bcd dosage) were crossed with males containing the reporter constructs. For Bcd1x embryos, yw;sp/CyO;BcdE1,MCP-NoNLS(2) males were crossed with yw;His-RFP;MCP-NoNLS(2) females, and their female progeny were crossed with construct-bearing males. Bcd4x embryos were generated from a cross between construct-bearing males and Bcd-GFP;His-RFP/CyO;MCP-NoNLS(2) females.

Sample preparation and live imaging

Sample preparation and live imaging with two-photon microscopy was conducted following experimental procedures detailed in [33], with the following changes: embryos were imaged from mitosis 13 until 12 minutes into nc14 had elapsed, after which time the recording session was stopped. Each time point consists of 8 z-slices taken 1 µm apart, with 2 frames imaged per z-slice and aligned and averaged after imaging. Each frame has a resolution of 1024x256 pixels and extends to roughly 40% of the length of the AP axis in the anterior and middle of the embryo (typically in the range of 15 to 60% AP). Images were acquired on two custom-built two-photon microscopes [72] with a laser wavelength of 970 nm for excitation of both MCP-GFP and Histone-RFP. Fluorescence was collected using a Zeiss 25x (NA = 0.8) oil/water-immersion objective and two gallium-arsenide-phosphide photomultiplier tubes (H10770PA-40 SEL, Hamamatsu). Imaging power at the back aperture was 10–13.5 mW, with transcription spot brightness normalized among imaging sessions and between microscopes to correct for power differences (Figure 3.13A). Pixel sizes on the two microscopes were 220 and 240 nm, and time resolutions were 39 and 46 s, respectively.

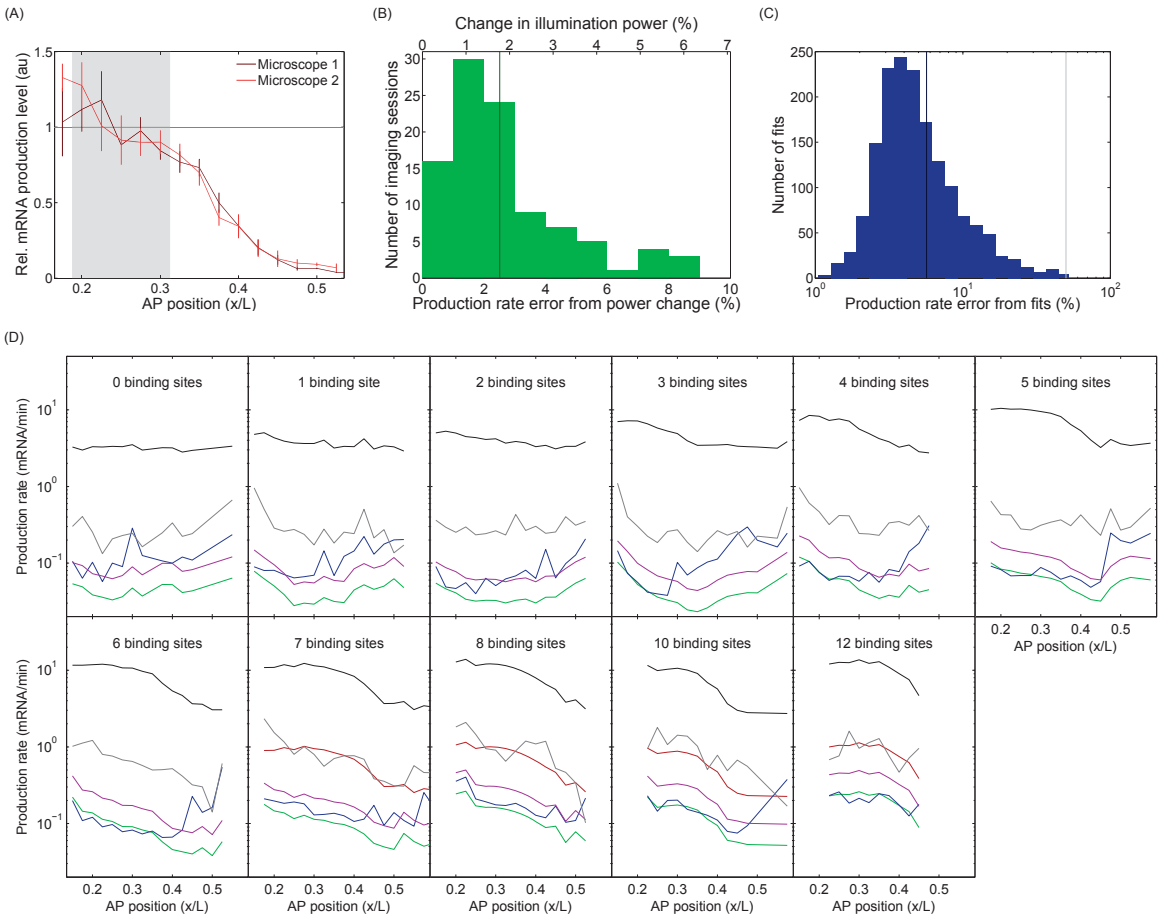

Figure 3.13: Analysis of measurement errors. **(A)** Calibration of transcription spot brightness between the two microscopes on which embryos were imaged for this work. Curves are of the relative amount of mRNA production, as measured by transcription spot brightness level, on each microscope, summed over the first 12 minutes after mitosis 13 and averaged among all transcription spots in 2.5% AP bins. Transcription spots were measured in embryos containing constructs with the *hb* P2 enhancer fully intact (#18 of Table 3.1). Normalization of fluorescence level (gray line: relative production rate = 1) was performed by averaging over the ratios of the values in the 5 shaded AP bins. Error bars are standard errors of the mean amount of mRNA production over embryos (8 embryos for microscope 1, 7 embryos for microscope 2). **(B)** Imaging power variability within each imaging session adds an average error of 2.5% to the measurement of the rate of mRNA production. For each of the 99 imaging sessions in which embryonic development was recorded for this work, the illumination power delivered by the objective was measured before and after imaging; the top axis of the plot records the percent change in illumination power between the beginning and end of each session. We find the standard deviation of each session's illumination power difference and propagate these standard deviations into corresponding percent errors in the rate of mRNA production in the embryos measured in each imaging session (bottom axis). The dark green vertical line represents the mean mRNA production

Figure 3.13: *(continued)* rate error over all imaging sessions, 2.5%. **(C)** Fits of rates of mRNA production to time traces of transcription spot brightness (Figure 3.1D) lead to an average error of 5.7%. Histogram shows the error (the half-width of the one-sigma confidence interval) in the mRNA production rate, as a percentage of the production rate, resulting from fitting the rate of mRNA production to the mean brightness time trace over all transcription spots in a bin of width 2.5% AP in a single embryo; 1627 errors from 167 embryos of normal Bcd dosage are shown. The mean percent error, 5.7%, is marked by a dark blue vertical line. All fits to rates of mRNA production with errors larger than 50% of the rate value (gray line) were automatically excluded by the code; 26 rates out of 1653 were excluded using this filter. **(D)** Comparison of measurement errors. Black curves represent the mean rates of mRNA production over all embryos of WT Bcd dosage with constructs containing a certain number of Bcd binding sites. Gray curves represent the standard errors on the mean rates of mRNA production resulting from averaging together multiple embryos: these are the errors shown in Figure 3.1F. Red, green, blue, and purple curves represent the measurement errors resulting from microscope calibration (A), power variability (B), fitting to rates of mRNA production (C), and imaging (Materials and Methods), respectively. Microscope calibration error only applies to embryos imaged on microscope 2 (last four panels), whose recorded fluorescence levels needed to be rescaled to match those of embryos imaged on microscope 1 (Materials and Methods). Curves are plotted in regions along the AP axis in which rates of mRNA production from at least 3 embryos could be fitted.

Analysis of movies from live imaging

Nuclei and particle tracking and transcription spot brightness detection in movies of embryonic development were conducted as in [33]. Rates of mRNA production were obtained for every embryo by a linear fit to the upward slope of the average fluorescence time trace for all detected particles in a 2.5% AP window (Figure 3.1D), with manual correction where necessary. The fitted values of the slopes are the reported rates of mRNA production. All embryos for which 4 or more 2.5%-AP-width windows (judged by either the rate of mRNA production, the fraction of active nuclei, or the total amount of mRNA production) lay outside 2 standard deviations of the average for that AP position and that construct were considered outliers and discarded: out of 430 otherwise usable embryos measured in service of this work, 11 embryos were discarded using this criterion.

Measurement errors

For this work, the rates of mRNA productions were calculated in 2.5%-AP-width windows in single embryos, which were then averaged over all embryos with a given construct or a given number of Bcd binding sites in order to produce the data shown in the main and supplemental figures. The reported rate of mRNA production r' in a single 2.5%-AP-width bin in a single embryo was calculated by

$$r' = M\left(\frac{p_0}{p}\right)^2 r, \tag{3.1}$$

where M is the scale factor to account for the difference in fluorescence between microscopes, p is the illumination power used to image the embryo (p_0 is a reference power, 10 mW), and r is the uncorrected rate of mRNA production as calculated by a linear fit to the mean rate of mRNA production in a 2.5%-AP-width bin in one embryo (Figure 3.1D). The errors associated with each of these quantities, as well as several others, are discussed below.

1. The scale factor M was calculated to correct for the different levels of fluorescence between the two two-photon microscopes used to image embryos for this work. M was produced by averaging over the ratios, in the five 2.5%-AP-width bins between 20% and 30% AP, of the mean amount of fluorescence produced in the first 12 minutes of nc14 for embryos imaged on either microscope (Figure 3.13A). By propagating the standard errors of the mean amount of fluorescence produced in the five named bins on both microscopes, we calculate the resultant error on the rate of mRNA production for embryos imaged on microscope 2 compared to microscope 1 to be 8.3% (Figure 3.13D). This microscope calibration error applies almost entirely to embryos with half- and double-WT Bcd dosage as well as to embryos with more than 6 binding sites, which constituted the vast majority of all embryos imaged on microscope 2.

2. The illumination power p was measured immediately before and after every imaging session in order to compensate for the differences in power among different imaging sessions. (The power correction factor applied to the uncorrected rate of mRNA production r is proportional to the square of the per-session illumination power because fluorescence level scales quadratically with illumination power on a two-photon microscope [103].) The standard deviation (using an unbiased estimator) of the illumination power measured before and after each imaging session was error-propagated using the above formula to find the mean per-session error on the rate of mRNA production arising from power variation over an imaging session, 2.5% (Figure 3.13B). Assuming no covariance between the rate of mRNA production and the power variability (for simplicity), the percentage error from power variability on the mean rate of mRNA production varies according to the number of embryos imaged per 2.5%-AP-width bin, but is generally (5th to 95th percentile) in the range of 0.7 to 1.9% (Figure 3.13D).

3. Each uncorrected rate of mRNA production r was calculated from a linear fit to the mean fluorescence time trace over all transcription spots in a 2.5%-AP-width bin in a single embryo (see Figure 3.1D, previous Materials and Methods section). By propagating the error in each fit to a resultant error in the corrected rate of mRNA production, we find that the mean error in the rate of mRNA production arising from fitting to a mean fluorescence time trace is 5.7% (Figure 3.13C). This error from fitting, when converted to a percentage error on the mean rate of mRNA production of all embryos with the same number of binding sites, varies by number of binding sites and AP position, but is generally (5th to 95th percentile) in the range of 0.8 to 7.0% (Figure 3.13D).

4. The error in the rate of mRNA production due to imaging noise was estimated by recording 3 movies of nascent transcription spots with roughly 2 to 3 times

the usual time resolution (14 to 23 s per time frame). splitting each movie into 3 separate sub-movies, each with a time resolution closer to that of a normal movie, and then calculating the standard deviation over the 3 sub-movies in the rate of mRNA production of all transcription spots in a bin of width 2.5% AP [33]. This variability in measured mRNA production rate within these movies of embryonic development, when converted to a percentage error on the mean rate of mRNA production resulting from imaging, varies according to the number of embryos imaged per 2.5%-AP-width bin, but is generally (5th to 95th percentile) in the range of 1.3 to 3.6% (Figure 3.13D).

5. The standard error of the mean rate of mRNA production, averaged over all embryos with a given number of binding sites, varies by binding site number and AP position, but is generally (5th to 95th percentile) in the range of 4.1 to 18.3% (Figure 3.13D).

From Figure 3.13D, we see that the largest of these measurement errors on the rate of mRNA production, the microscope calibration error and the error from averaging over embryos (#1 and 5, respectively, on the above list), can be as high as 1 mRNA/min or more, especially for embryos containing constructs with more than the 6 endogenous Bcd binding sites. However, in most regions along the AP axis and for most numbers of binding sites, our highest error only reaches to a few tenths of an mRNA/min. This is much lower than the observed variability in the rates of mRNA production among embryos both with the same and different numbers of binding sites (Figures 3.1G and 3.2), thus demonstrating our ability to detect very fine biological differences in rates of mRNA production.

Thermodynamic DNA binding model with cooperativity

Thermodynamic models of transcriptional regulation have been largely used to describe the interactions between transcription factors bound to regulatory regions and the transcriptional machinery [10, 11]. The main assumption of these models is that the time scales of transcription factor binding and unbinding to regulatory DNA are faster than the downstream rate of transcriptional initiation. Under this assumption, binding of transcription factors is in quasi-equilibrium such that the tools of statistical mechanics can be applied to describe the probability of occurrence of different binding configurations to the enhancer.

We invoke such a thermodynamic model to describe Bcd binding to the enhancer. For example, an enhancer bearing only one Bcd binding site can be either empty or occupied by one Bcd molecule (Figure 3.9A). For the empty state we assign a weight of 1. For the occupied state we assign a weight of $\frac{[\text{Bcd}]}{K_d}$, with [Bcd] being the Bcd concentration and the dissociation constant of Bcd to its binding site given by K_d. The probability of finding Bcd bound to the DNA is then given by

$$p_{\text{bound}} = \frac{\frac{[\text{Bcd}]}{K_d}}{1 + \frac{[\text{Bcd}]}{K_d}}, \tag{3.2}$$

which corresponds to dividing the weight of the bound state by the sum of the weights of all states the system can be found in.

As we increase the number of binding sites beyond one, we need to account for more possible Bcd configurations on the DNA: e.g. when two Bcd binding sites are present, the system can be found in multiple configurations (Figure 3.9B). Note that we assume that both binding sites have the same dissociation constant K_d. This assumption is consistent with our observation that binding site position and affinity do not affect the output level of gene expression significantly (Figure 3.1E,G). This observation also tells us that states 1 and 2, each corresponding to having only one

Bcd molecule bound to either of the two binding sites, should each have the same effect on the output rate of transcription initiation. As result of this indifference, we enumerate states based on the number of Bcd molecules bound to the DNA rather than their particular spatial configuration. This leads us to define the degeneracy of a state as the number of ways that m Bcd molecules can be arranged on n binding sites:

$$\text{degeneracy}\,(m,n) = \binom{n}{m} = \frac{n!}{m!(n-m)!}. \tag{3.3}$$

Finally, the fact that the values obtained for the sensitivity of the input-output functions are greater than 1 (Figure 3.8D) suggests the presence of protein-protein interactions between bound Bcd molecules leading to cooperativity. We account for this cooperativity through the pairwise interaction parameter ω, and we assume that these interactions only take place between neighboring Bcd molecules bound to the DNA. Note that an ω value of 1 corresponds to no interactions, while larger values correspond to cooperative interactions between Bcd molecules. With these results in hand, we can now write the states for an enhancer with n Bcd binding sites and their corresponding weights (Figure 3.8B).

A missing piece in our models is how to relate the enhancer occupancy state to the output rate of mRNA production. A simple mathematical description of the rate of mRNA production is

$$\frac{d[\text{mRNA}]}{dt} = \sum_i p_i r_i + r_{\text{basal}}, \tag{3.4}$$

where p_i is the probability of finding the enhancer in state i and r_i is the corresponding rate of mRNA production in that state. Additionally, we consider a basal rate of transcription given by r_{basal}. Note that we ignore mRNA degradation in this simple model. The probabilities p_i are calculated by dividing the weight for state i by the sum of the weights corresponding to all states that the enhancer can be found in.

In this paper we explore two extreme cases for r_i. In the *simple activation* model, only one bound Bcd molecule can interact with the transcriptional machinery at a time, leading to a rate of mRNA production r. In this model, the presence of extra Bcd molecules increases the overall occupancy of the enhancer but does not change the nature of the interaction between Bcd and the transcriptional machinery. In contrast, in the *additive activation* model, every bound Bcd molecule can interact with the transcriptional machinery and contribute a value r to the overall rate of mRNA production. Namely, the output rate of n bound Bcd molecules will be $nr + r_{\text{basal}}$.

These two models have been independently used to describe transcriptional regulation in flies and other model organisms [29, 35, 92]. However, their polarizing experimental predictions have not been contrasted. In particular, after some algebra, the rate of mRNA production in the *simple activation* case is given by

$$\left.\frac{d[\text{mRNA}]}{dt}\right|_{\text{simple act.}} = r\frac{\frac{1}{\omega}\left[\left(1+\omega\frac{[\text{Bcd}]}{K_d}\right)^n - 1\right]}{1+\frac{1}{\omega}\left[\left(1+\omega\frac{[\text{Bcd}]}{K_d}\right)^n - 1\right]} + r_{\text{basal}}, \tag{3.5}$$

whereas the rate of mRNA production for *additive activation* is

$$\left.\frac{d[\text{mRNA}]}{dt}\right|_{\text{add. act.}} = r\frac{\frac{[\text{Bcd}]}{K_d} n\left(1+\omega\frac{[\text{Bcd}]}{K_d}\right)^{n-1}}{1+\frac{1}{\omega}\left[\left(1+\omega\frac{[\text{Bcd}]}{K_d}\right)^n - 1\right]} + r_{\text{basal}}. \tag{3.6}$$

In the anterior region of the embryo, saturating concentrations of Bcd are present such that $\frac{[\text{Bcd}]}{K_d} \gg 1$, resulting in

$$\left.\frac{d[\text{mRNA}]}{dt}\left([\text{Bcd}]\to\infty\right)\right|_{\text{simple act.}} = r + r_{\text{basal}} \tag{3.7}$$

and

$$\left.\frac{d[\text{mRNA}]}{dt}\left([\text{Bcd}]\to\infty\right)\right|_{\text{add. act.}} = rn + r_{\text{basal}}. \tag{3.8}$$

Namely, at saturating concentrations of Bcd, the rate of mRNA production should not change with the number of Bcd binding sites n in *simple activation*, but it should grow linearly for *additive activation*. Note that at this concentration the cooperativity parameter ω plays no role in the predictions. These predictions are contrasted with our experiments (Figure 3.8C), which clearly support the *additive activation* model interaction between Bcd and the transcriptional machinery.

Given this model we can also calculate the sensitivity of the input-output function by finding the maximum of its log-log derivative, namely

$$\text{sensitivity} = \max \left(\frac{d \log \left(\frac{d[\text{mRNA}]}{dt} \right)}{d \log \left([\text{Bcd}] \right)} \right). \tag{3.9}$$

Fits to the sensitivity plotted as a function of n for different values of ω show how our model can account for the overall increasing trend of the sensitivity as a function of the number of Bcd binding sites (Figure 3.8D).

Chapter 4

Conclusion

The two projects discussed in the previous chapters deepen our understanding of the mechanisms by which patterning genes in the first developmental stages of the fruit fly embryo communicate information about position, and they do so by identifying the functions and limitations of two different ways by which positional information is conveyed in the patterning gene network. The study of the scaling of gene expression profiles with embryo length in Chapter 2 finds that the patterning gene Bicoid (Bcd) is unable on its own to create the scaled expression patterns of the downstream gap genes. In Chapter 3, the study of the DNA regulatory elements by which Bcd protein concentration determines the spatial pattern of *hunchback* (*hb*) mRNA expression finds that the level and domain of expression is a linear function of binding site number, and that the expression pattern is insensitive to the specificity and position of Bcd binding sites in the endogenous DNA sequence.

Our finding that Bcd is insufficient to convey scaled information to the gap genes can be extended by expanding our definition of what it means for a gene expression pattern to scale with embryo length, and by studying gap gene scaling under further perturbations to the maternal gene inputs. In Chapter 2, the dynamic profile warping method is used to assess the scaling of gene expression profiles by measuring the

variability of positions along the anterior-posterior axis, in a population of embryos, in which local features such as peaks and boundaries are found. By contrast, principal component analysis generalizes this definition of scaling to also consider scaling of the amplitude (or any other feature) of expression profiles with embryo length. Using this more holistic method, the scaling or lack thereof of any part of a gene's expression profile is taken into account through rigorous hypothesis testing to determine the overall scaling of the expression pattern with embryo length. A careful analysis of the scaling of gap gene expression profiles in embryos in which further combinations of the maternal factors Bcd, Nanos, and Torso are present or absent could help reveal the full extent to which these maternal factors work in tandem to supply information about embryo length to the gap genes.

Similarly, our finding in Chapter 3 of the linearity of *hb* expression level and domain with the number of Bcd binding sites is an important result on its own, but it could be extended by a careful analysis of further perturbations to the wild-type proximal *hb* enhancer, as discussed in Section 3.6. Our results from disabling combinations of binding sites show that mRNA expression level in this enhancer is independent of Bcd binding site position within the roughly 300-base-pair extent of the endogenous enhancer, but a steep drop in expression level exists when the entire enhancer is moved away from the promoter by several hundred base pairs. This may imply the presence of multiple length scales in the regulation of mRNA expression in this system: a short length scale on which expression is insensitive to position and a longer length scale over which expression is dependent on proximity to the promoter. Measurements of other spacing lengths between the promoter and enhancer will clarify this relationship between enhancer-promoter positioning and mRNA expression. Likewise, the changes in mRNA expression observed when mutating strong sites into weak sites and vice versa, and the accompanying hypothesis of distance-dependent cooperativity, can be tested by directly varying the distance between a pair of en-

dogenous strong or weak sites or by weakening strong sites and removing endogenous weak sites. Lastly, our result of the insensitivity of mRNA expression to the position and affinity of the endogenous binding sites in their wild-type configuration cannot necessarily be generalized to other enhancers in control of other genes: measurements of expression in response to perturbations of binding sites in systems other than the proximal *hb* enhancer may tell us how universal this phenomenon is across the fly genome.

Bibliography

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

[2] M. I. Arnone and E. H. Davidson. The hardwiring of development: organization and function of genomic regulatory systems. *Development*, 124(10):1851–1864, May 1997.

[3] David N. Arnosti. Analysis and function of transcriptional regulatory elements: insights from Drosophila. *Annual Review of Entomology*, 48:579–602, 2003.

[4] Michael Bate, editor. *The Development of Drosophila melanogaster*. Cold Spring Harbor Laboratory Press, Woodbury, N.Y., January 1994.

[5] Jack R. Bateman, Anne M. Lee, and C.-ting Wu. Site-specific transformation of Drosophila via phiC31 integrase-mediated cassette exchange. *Genetics*, 173(2):769–777, June 2006.

[6] Danny Ben-Zvi and Naama Barkai. Scaling of morphogen gradients by an expansion-repression integral feedback control. *Proceedings of the National Academy of Sciences*, 107(15):6924–6929, April 2010.

[7] Danny Ben-Zvi, Ben-Zion Shilo, Abraham Fainsod, and Naama Barkai. Scaling of the BMP activation gradient in Xenopus embryos. *Nature*, 453(7199):1205–1211, June 2008.

[8] Sven Bergmann, Oded Sandler, Hila Sberro, Sara Shnider, Eyal Schejter, Ben-Zion Shilo, and Naama Barkai. Pre-Steady-State Decoding of the Bicoid Morphogen Gradient. *PLoS Biol*, 5(2):e46, February 2007.

[9] Philip Bevington and D. Keith Robinson. *Data Reduction and Error Analysis for the Physical Sciences*. McGraw-Hill Science/Engineering/Math, 3rd edition, July 2002.

[10] Lacramioara Bintu, Nicolas E Buchler, Hernan G Garcia, Ulrich Gerland, Terence Hwa, Jané Kondev, Thomas Kuhlman, and Rob Phillips. Transcriptional regulation by the numbers: applications. *Current Opinion in Genetics & Development*, 15(2):125–135, April 2005.

[11] Lacramioara Bintu, Nicolas E Buchler, Hernan G Garcia, Ulrich Gerland, Terence Hwa, Jané Kondev, and Rob Phillips. Transcriptional regulation by the numbers: models. *Current Opinion in Genetics & Development*, 15(2):116–124, April 2005.

[12] T. Boveri. Die polarität von ovocyte, ei, und larve des strongylocentrus lividus. *Zool Jahrb Abt Anat Ont Thi*, 14:384, 1901.

[13] David S. Burz and Steven D. Hanes. Isolation of Mutations that Disrupt Cooperative DNA Binding by the Drosophila Bicoid Protein. *Journal of Molecular Biology*, 305(2):219–230, January 2001.

[14] David S. Burz, Rolando Rivera-Pomar, Herbert Jäckle, and Steven D. Hanes. Cooperative DNA-binding by Bicoid provides a mechanism for threshold-dependent gene activation in the Drosophila embryo. *The EMBO Journal*, 17(20):5998–6009, October 1998.

[15] David Cheung, Cecelia Miles, Martin Kreitman, and Jun Ma. Scaling of the Bicoid morphogen gradient by a volume-dependent production rate. *Development*, 138(13):2741 –2749, July 2011.

[16] David Cheung, Cecelia Miles, Martin Kreitman, and Jun Ma. Adaptation of the length scale and amplitude of the Bicoid gradient profile to achieve robust patterning in abnormally large Drosophila melanogaster embryos. *Development*, page dev.098640, November 2013.

[17] Mathieu Coppey, Alistair N. Boettiger, Alexander M. Berezhkovskii, and Stanislav Y. Shvartsman. Nuclear Trapping Shapes the Terminal Gradient in the Drosophila Embryo. *Current Biology*, 18(12):915–919, June 2008.

[18] Olivier Crauk and Nathalie Dostatni. Bicoid Determines Sharp and Precise Target Gene Expression in the Drosophila Embryo. *Current Biology*, 15(21):1888–1898, November 2005.

[19] Eric H. Davidson. *The Regulatory Genome: Gene Regulatory Networks In Development And Evolution*. Academic Press, Burlington, MA; San Diego, 1 edition edition, June 2006.

[20] Aitana Morton de Lachapelle and Sven Bergmann. Pre-steady and stable morphogen gradients: can they coexist? *Molecular Systems Biology*, 6:428, November 2010.

[21] Aitana Morton de Lachapelle and Sven Bergmann. Precision and scaling in morphogen gradient read-out. *Molecular Systems Biology*, 6(1), March 2010.

[22] Wolfgang Driever and C. Nüsslein-Volhard. The bicoid protein determines position in the Drosophila embryo in a concentration-dependent manner. *Cell*, 54(1):95–104, July 1988.

[23] Wolfgang Driever and Christiane Nüsslein-Volhard. A gradient of bicoid protein in Drosophila embryos. *Cell*, 54(1):83–93, July 1988.

[24] Wolfgang Driever and Christiane Nüsslein-Volhard. The bicoid protein is a positive regulator of hunchback transcription in the early Drosophila embryo. *Nature*, 337(6203):138–143, January 1989.

[25] Wolfgang Driever, Gudrun Thoma, and Christiane Nüsslein-Volhard. Determination of spatial domains of zygotic gene expression in the Drosophila embryo by the affinity of binding sites for the bicoid morphogen. *Nature*, 340(6232):363–367, August 1989.

[26] Julien O. Dubuis, Reba Samanta, and Thomas Gregor. Accurate measurements of dynamics and reproducibility in small genetic networks. *Molecular Systems Biology*, 9(1), January 2013.

[27] Dan Ellis. Dynamic Time Warp (DTW) in Matlab, 2012.

[28] Anne Ephrussi and Daniel St Johnston. Seeing Is Believing: The Bicoid Morphogen Gradient Matures. *Cell*, 116(2):143–152, January 2004.

[29] Jelena Erceg, Timothy E. Saunders, Charles Girardot, Damien P. Devos, Lars Hufnagel, and Eileen E. M. Furlong. Subtle Changes in Motif Positioning Cause Tissue-Specific Effects on Robustness of an Enhancer's Activity. *PLoS Genet*, 10(1):e1004060, January 2014.

[30] Andrea M. Femino, Fredric S. Fay, Kevin Fogarty, and Robert H. Singer. Visualization of Single RNA Transcripts in Situ. *Science*, 280(5363):585–590, April 1998.

[31] Kevin M. Forrest and Elizabeth R. Gavis. Live Imaging of Endogenous RNA Reveals a Diffusion and Entrapment Mechanism for nanos mRNA Localization in Drosophila. *Current Biology*, 13(14):1159–1168, July 2003.

[32] Hans Georg Frohnhöfer and Christiane Nüsslein-Volhard. Organization of anterior pattern in the Drosophila embryo by the maternal gene bicoid. *Nature*, 324(6093):120–125, November 1986.

[33] Hernan G. Garcia, Mikhail Tikhonov, Albert Lin, and Thomas Gregor. Quantitative Imaging of Transcription in Living Drosophila Embryos Links Polymerase Activity to Patterning. *Current Biology*, 2013.

[34] Mayra Garcia, Marcos Nahmad, Gregory T. Reeves, and Angelike Stathopoulos. Size-dependent regulation of dorsal–ventral patterning in the early Drosophila embryo. *Developmental Biology*, 381(1):286–299, September 2013.

[35] Jason Gertz, Eric D. Siggia, and Barak A. Cohen. Analysis of combinatorial cis-regulation in synthetic and genomic promoters. *Nature*, 457(7226):215–218, January 2009.

[36] A. Gierer and H. Meinhardt. A theory of biological pattern formation. *Kybernetik*, 12(1):30–39, December 1972.

[37] Scott Gilbert. Chapter 9: The Genetics of Axis Specification in Drosophila. In *Developmental Biology*. 8th edition edition, 2006.

[38] Scott F. Gilbert. The Origins of Anterior-Posterior Polarity. In *Developmental Biology*. 6th edition edition, 2000.

[39] Thomas Gregor, William Bialek, Rob R. de Ruyter van Steveninck, David W. Tank, and Eric F. Wieschaus. Diffusion and scaling during early embryonic pattern formation. *Proceedings of the National Academy of Sciences of the United States of America*, 102(51):18403–18407, December 2005.

[40] Thomas Gregor, Alistair P. McGregor, and Eric F. Wieschaus. Shape and function of the Bicoid morphogen gradient in dipteran species with different sized embryos. *Developmental Biology*, 316(2):350–358, April 2008.

[41] Thomas Gregor, David W. Tank, Eric F. Wieschaus, and William Bialek. Probing the Limits to Positional Information. *Cell*, 130(1):153–164, July 2007.

[42] Thomas Gregor, Eric F. Wieschaus, Alistair P. McGregor, William Bialek, and David W. Tank. Stability and Nuclear Dynamics of the Bicoid Morphogen Gradient. *Cell*, 130(1):141–152, July 2007.

[43] Oliver Grimm, Mathieu Coppey, and Eric Wieschaus. Modelling the Bicoid gradient. *Development*, 137(14):2253–2264, July 2010.

[44] Oliver Grimm and Eric Wieschaus. The Bicoid gradient is shaped independently of nuclei. *Development*, 137(17):2857 –2862, 2010.

[45] Fisun Hamaratoglu, Aitana Morton de Lachapelle, George Pyrowolakis, Sven Bergmann, and Markus Affolter. Dpp Signaling Activity Requires Pentagone to Scale with Tissue Size in the Growing Drosophila Wing Imaginal Disc. *PLoS Biol*, 9(10):e1001182, October 2011.

[46] David M. Holloway, Lionel G. Harrison, David Kosman, Carlos E. Vanario-Alonso, and Alexander V. Spirov. Analysis of pattern precision shows that Drosophila segmentation develops substantial independence from gradients of maternal gene products. *Developmental Dynamics*, 235(11):2949–2960, 2006.

[47] S. Hörstadius. Über die determination im verlaufe der eiachse bei seeigeln. *Publ Staz Zool Napoli*, 14:251–479, 1935.

[48] Bahram Houchmandzadeh, Eric Wieschaus, and Stanislas Leibler. Establishment of developmental precision and proportions in the early Drosophila embryo. *Nature*, 415(6873):798–802, February 2002.

[49] Martin Hülskamp, Christian Schröder, Christine Pfeifle, Herbert Jäckle, and Diethard Tautz. Posterior segmentation of the Drosophila embryo in the absence of a maternal posterior organizer gene. *Nature*, 338(6217):629–632, April 1989.

[50] O.A. Igoshin and J. Narula. Thermodynamic models of combinatorial gene regulation by distant enhancers. *IET Systems Biology*, 4(6):393–408, November 2010.

[51] Hidehiko Inomata, Tatsuo Shibata, Tomoko Haraguchi, and Yoshiki Sasai. Scaling of Dorsal-Ventral Patterning by Embryo Size-Dependent Degradation of Spemann's Organizer Signals. *Cell*, 153(6):1296–1311, June 2013.

[52] Vivian Irish, Ruth Lehmann, and Michael Akam. The Drosophila posterior-group gene nanos functions by repressing hunchback activity. *Nature*, 338(6217):646–648, April 1989.

[53] F. Itakura. Minimum prediction residual principle applied to speech recognition. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 23(1):67–72, 1975.

[54] Johannes Jaeger. A matter of timing and precision. *Molecular Systems Biology*, 6:427, November 2010.

[55] Johannes Jaeger. The gap gene network. *Cellular and Molecular Life Sciences*, 68(2):243–274, January 2011.

[56] Johannes Jaeger, Maxim Blagov, David Kosman, Konstantin N. Kozlov, Manu, Ekaterina Myasnikova, Svetlana Surkova, Carlos E. Vanario-Alonso, Maria Samsonova, David H. Sharp, and John Reinitz. Dynamical Analysis of Regulatory Interactions in the Gap Gene System of Drosophila melanogaster. *Genetics*, 167(4):1721–1737, August 2004.

[57] Johannes Jaeger, David Irons, and Nick Monk. Regulative feedback in pattern formation: towards a general relativistic theory of positional information. *Development*, 135(19):3175–3183, October 2008.

[58] Johannes Jaeger, Svetlana Surkova, Maxim Blagov, Hilde Janssens, David Kosman, Konstantin N. Kozlov, Manu, Ekaterina Myasnikova, Carlos E. Vanario-Alonso, Maria Samsonova, David H. Sharp, and John Reinitz. Dynamic control of positional information in the early Drosophila embryo. *Nature*, 430(6997):368–371, July 2004.

[59] Ni Ji and Alexander van Oudenaarden. Single molecule fluorescent in situ hybridization (smFISH) of C. elegans worms and embryos. *WormBook*, pages 1–16, December 2012.

[60] Gerardo Jiménez, Antoine Guichet, Anne Ephrussi, and Jordi Casanova. Relief of gene repression by Torso RTK signaling: role of capicua in Drosophila

terminal and dorsoventral patterning. *Genes & Development*, 14(2):224–231, January 2000.

[61] Jaeger Johannes, David H. Sharp, and John Reinitz. Known Maternal Gradients are not Sufficient for the Establishment of Gap Domains in Drosophila melanogaster. *Mechanisms of development*, 124(2):108–128, February 2007.

[62] Alexander D. Johnson and Ira Herskowitz. A repressor (MATα2 product) and its operator control expression of a set of cell type specific genes in yeast. *Cell*, 42(1):237–247, August 1985.

[63] I. T. Jolliffe. *Principal Component Analysis*. Springer, New York, 2nd edition edition, October 2002.

[64] Tae Kook Kim and Tom Maniatis. The Mechanism of Transcriptional Synergy of an In Vitro Assembled Interferon-β Enhanceosome. *Molecular Cell*, 1(1):119–129, December 1997.

[65] Meghana M. Kulkarni and David N. Arnosti. Information display by transcriptional enhancers. *Development*, 130(26):6569–6575, December 2003.

[66] Volker M. Lauschke, Charisios D. Tsiairis, Paul François, and Alexander Aulehla. Scaling of embryonic patterning based on phase-gradient encoding. *Nature*, 493(7430):101–105, January 2013.

[67] Danielle Lebrecht, Marisa Foehr, Eric Smith, Francisco J. P. Lopes, Carlos E. Vanario-Alonso, John Reinitz, David S. Burz, and Steven D. Hanes. Bicoid cooperative DNA binding is critical for embryonic patterning in Drosophila. *Proceedings of the National Academy of Sciences of the United States of America*, 102(37):13176–13181, September 2005.

[68] Alexander P. Lifanov, Vsevolod J. Makeev, Anna G. Nazina, and Dmitri A. Papatsenko. Homotypic regulatory clusters in Drosophila. *Genome Research*, 13(4):579–588, April 2003.

[69] J. W. Little, D. P. Shepley, and D. W. Wert. Robustness of a gene regulatory circuit. *The EMBO Journal*, 18(15):4299–4307, August 1999.

[70] Shawn C. Little, Mikhail Tikhonov, and Thomas Gregor. Precise Developmental Gene Expression Arises from Globally Stochastic Transcriptional Activity. *Cell*, 154(4):789–800, August 2013.

[71] Shawn C. Little, Gašper Tkačik, Thomas B. Kneeland, Eric F. Wieschaus, and Thomas Gregor. The Formation of the Bicoid Morphogen Gradient Requires Protein Movement from Anteriorly Localized mRNA. *PLoS Biol*, 9(3):e1000596, March 2011.

[72] Feng Liu, Alexander H. Morrison, and Thomas Gregor. Dynamic interpretation of maternal inputs by the Drosophila segmentation gene network. *Proceedings of the National Academy of Sciences*, 110(17):6724–6729, April 2013.

[73] Francisco J.P. Lopes, Alexander V. Spirov, and Paulo M. Bisch. The role of Bicoid cooperative binding in the patterning of sharp borders in Drosophila melanogaster. *Developmental Biology*, 370(2):165–172, October 2012.

[74] Susan E. Lott, Martin Kreitman, Arnar Palsson, Elena Alekseeva, and Michael Z. Ludwig. Canalization of segmentation and its evolution in Drosophila. *Proceedings of the National Academy of Sciences*, 104(26):10926–10931, June 2007.

[75] Susan E. Lott, Michael Z. Ludwig, and Martin Kreitman. Evolution and Inheritance of Early Embryonic Patterning in Drosophila Simulans and D. Sechellia. *Evolution*, 65(5):1388–1399, 2011.

[76] X. Ma, D. Yuan, K. Diepold, T. Scarborough, and J. Ma. The Drosophila morphogenetic protein Bicoid binds DNA cooperatively. *Development*, 122(4):1195–1206, April 1996.

[77] Manu, Svetlana Surkova, Alexander V Spirov, Vitaly V Gursky, Hilde Janssens, Ah-Ram Kim, Ovidiu Radulescu, Carlos E Vanario-Alonso, David H Sharp, Maria Samsonova, and John Reinitz. Canalization of Gene Expression in the Drosophila Blastoderm by Gap Gene Cross Regulation. *PLoS Biol*, 7(3):e1000049, March 2009.

[78] CM Miles, SE Lott, CL Luengo Hendriks, MZ Ludwig, Manu, CL Williams, and M Kreitman. Artificial selection on egg size perturbs early pattern formation in Drosophila melanogaster. *Evolution*, 65(1):33–42, January 2011.

[79] Joanna A. Miller and Jonathan Widom. Collaborative Competition Mechanism for Gene Activation In Vivo. *Molecular and Cellular Biology*, 23(5):1623–1632, March 2003.

[80] Leonid A. Mirny. Nucleosome-mediated cooperativity between transcription factors. *Proceedings of the National Academy of Sciences*, 107(52):22534–22539, December 2010.

[81] T. H. Morgan. Regeneration and liability to injury. *Science*, 14(346):235–248, August 1901.

[82] C. Myers, L. Rabiner, and A.E. Rosenberg. Performance tradeoffs in dynamic time warping algorithms for isolated word recognition. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 28(6):623–635, December 1980.

[83] R. Namba, T. M. Pazdera, R. L. Cerrone, and J. S. Minden. Drosophila embryonic pattern repair: how embryos respond to bicoid dosage alteration. *Development*, 124(7):1393–1403, April 1997.

[84] Christiane Nüsslein-Volhard and Eric Wieschaus. Mutations affecting segment number and polarity in Drosophila. *Nature*, 287(5785):795–801, October 1980.

[85] H. G. Othmer and E. Pate. Scale-invariance in reaction-diffusion models of spatial pattern formation. *Proceedings of the National Academy of Sciences*, 77(7):4180–4184, July 1980.

[86] K Pearson. On lines and planes of closest fit to systems of points in space. *Philosophical Magazine*, 2(6):559–572, 1901.

[87] Aude Porcher, Asmahan Abu-Arish, Sébastien Huart, Baptiste Roelens, Cécile Fradin, and Nathalie Dostatni. The time to measure positional information: maternal Hunchback is required for the synchrony of the Bicoid transcriptional response at the onset of zygotic transcription. *Development*, 137(16):2795–2804, August 2010.

[88] Mark Ptashne. *A Genetic Switch, Third Edition: Phage Lambda Revisited*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y, 3rd edition edition, April 2004.

[89] L. Rabiner, A.E. Rosenberg, and S.E. Levinson. Considerations in dynamic time warping algorithms for discrete word recognition. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 26(6):575–582, December 1978.

[90] H. Sakoe and S. Chiba. Dynamic programming algorithm optimization for spoken word recognition. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 26(1):43–49, 1978.

[91] Martin W. Scheeler. *Scaling in biological patterns: an experimental investigation of scaling in the Bicoid gradient*. PhD thesis, Princeton University, Princeton, NJ, 2011.

[92] Eran Segal, Tali Raveh-Sadka, Mark Schroeder, Ulrich Unnerstall, and Ulrike Gaul. Predicting expression patterns from regulatory sequence in Drosophila segmentation. *Nature*, 451(7178):535–540, January 2008.

[93] Marcia Simpson-Brose, Jessica Treisman, and Claude Desplan. Synergy between the hunchback and bicoid morphogens is required for anterior patterning in Drosophila. *Cell*, 78(5):855–865, September 1994.

[94] Alexander Spirov, Khalid Fahmy, Martina Schneider, Erich Frei, Markus Noll, and Stefan Baumgartner. Formation of the bicoid morphogen gradient: an mRNA gradient dictates the protein gradient. *Development*, 136(4):605–614, February 2009.

[95] Frank Sprenger, Leslie M. Stevens, and Christiane Nüsslein-Volhard. The Drosophila gene torso encodes a putative receptor tyrosine kinase. *Nature*, 338(6215):478–483, April 1989.

[96] Leslie M. Stevens, Hans Georg Frohnhöfer, Martin Klingler, and Christiane Nüsslein-Volhard. Localized requirement for torso-like expression in follicle cells for development of terminal anlagen of the Drosophila embryo. *Nature*, 346(6285):660–663, August 1990.

[97] G Struhl. Differing strategies for organizing anterior and posterior body pattern in Drosophila embryos. *Nature*, 338(6218):741–744, April 1989.

[98] Gary Struhl, Kevin Struhl, and Paul M. Macdonald. The gradient morphogen bicoid is a concentration-dependent transcriptional activator. *Cell*, 57(7):1259–1273, June 1989.

[99] Svetlana Surkova, David Kosman, Konstantin Kozlov, Manu, Ekaterina Myasnikova, Anastasia A. Samsonova, Alexander Spirov, Carlos E. Vanario-Alonso, Maria Samsonova, and John Reinitz. Characterization of the Drosophila segment determination morphome. *Developmental Biology*, 313(2):844–862, January 2008.

[100] P. Szymanski and M. Levine. Multiple modes of dorsal-bHLH transcriptional synergy in the Drosophila embryo. *The EMBO Journal*, 14(10):2229–2238, May 1995.

[101] M. E. Tipping and C. M. Bishop. Mixtures of probabilistic principal component analyzers. *Neural Computation*, 11(2):443–482, February 1999.

[102] Gašper Tkačik, Curtis G. Callan, and William Bialek. Information flow and optimization in transcriptional regulation. *Proceedings of the National Academy of Sciences*, 105(34):12265–12270, August 2008.

[103] Philbert S. Tsai, Nozomi Nishimura, Elizabeth J. Yoder, Earl M. Dolnick, G. Allen White, and David Kleinfeld. Chapter 6. Principles, Design, and Construction of a Two-Photon Laser-Scanning Microscope for In Vitro and In Vivo Brain Imaging. In Ron Frostig, editor, *In Vivo Optical Imaging of Brain Function*, pages 113–171. CRC Press, 2002.

[104] A. M. Turing. The Chemical Basis of Morphogenesis. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 237(641):37–72, August 1952.

[105] David M. Umulis and Hans G. Othmer. Mechanisms of scaling in pattern formation. *Development*, 140(24):4830–4843, December 2013.

[106] O. Wartlick, P. Mumcu, A. Kicheva, T. Bittig, C. Seum, F. Jülicher, and M. González-Gaitán. Dynamics of Dpp Signaling and Proliferation Control. *Science*, 331(6021):1154–1159, March 2011.

[107] Ortrud Wartlick, Anna Kicheva, and Marcos González-Gaitán. Morphogen Gradient Formation. *Cold Spring Harbor Perspectives in Biology*, 1(3), September 2009.

[108] Svante Wold, Kim Esbensen, and Paul Geladi. Principal component analysis. *Chemometrics and Intelligent Laboratory Systems*, 2(1–3):37–52, August 1987.

[109] L. Wolpert. Positional information and the spatial pattern of cellular differentiation. *Journal of Theoretical Biology*, 25(1):1–47, October 1969.

[110] Lewis Wolpert. Positional information and patterning revisited. *Journal of Theoretical Biology*, 269(1):359–365, January 2011.

[111] Xiaobei Zhao, et al. JASPAR 2013: An extensively expanded and updated open-access database of transcription factor binding profiles. 2013.