Title
Precision and Robustness in the Interpretation of Morphogen Gradients

Author
Zhang, Peng Cheng

Publication Date
2015

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA,
IRVINE

Precision and Robustness in the Interpretation of Morphogen Gradients

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Biomedical Engineering

by

Peng Cheng Zhang

Dissertation Committee:
Professor Arthur D. Lander, Chair
Professor Kavita Arora
Professor Elliott Botvinick
Professor Enrico Gratton
Professor Qing Nie

2015

DEDICATION

To the best parents anyone could ask for

My Dad,
Who always supported me and sacrificed everything for me,
Even as he bravely battled ALS.
I will always remember everything you have taught me
And you will forever live in my heart

My Mom,
Who loves unconditionally
And always there for me.

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGMENTS

I would like to express my deepest appreciation to my committee chair, Professor Arthur Lander. Dr. Lander is the smartest person I have ever met and the most knowledgeable by far. His ability to distill a problem to its core and come up with insightful answers will inspire me to be a better problem solver for the rest of my life. Without his help and guidance, this dissertation would not have been possible.

I am grateful to all the friends I had the pleasure of knowing in graduate school. My appreciation goes out to everyone in Lander lab, past and present. Shaohua, who taught me most of what I know about flies, and built a solid foundation for the rest of the fly pushers in lab. Martha, who is always helpful and took care of everyone in lab. Karen and her team of wonderful staffs, that provided all the coffee and support to everyone under the CCBS umbrella. Sameeran, I'm glad to have roomed with you, and for helping me with the intangibles. Abed, Jimmy, Josh, Andrew, WanJong, Linda, Marcos, Peter, everyone made lab a more interesting place. I would also like to thank my extended fly family from Arora lab and Warrior lab, Jane, Maribel, Sopheap; people who have helped me so much with image segmentation and provided their software and support, Michael and Dr. Olivier Cinquin; math help and collaborations with Jon and Jeremy from Qing's lab.

I would like to thank my committee members, Professor Qing Nie, Professor Enrico Gratton, Professor Kavita Arora and Professor Elliot Botvinick. I know how busy every single one of my committee members are, but everyone always took the time out to talk to me about anything I needed for my research.

I want to thank my mom and my dad, who provided me with everything they had. I want to thank my cousin, when nothing seems to be going right; you were always just one phone call away.

Finally, I want to thank my girlfriend, Ying. Thanks for taking care of everything, especially during the last stretch of writing. Nothing can change what I have been through and all the memories I have, and it will always be a part of me.

Financial support was provided by the NIH R01-GM067247, NIH P50-GM076516.

CURRICULUM VITAE

Peng Cheng Zhang

Education

2015	Ph.D in Biomedical Engineering University of California, Irvine
2008	M.S. in Biomedical Engineering University of California, Irvine
2007	B.S. in Bioengineering University of California, San Diego

Experience

2008 – 2015	Graduate Student Researcher University of California, Irvine Center for Complex Biological Systems Advisor: Dr. Arthur D. Lander Research Interest: Pattern formation and quantitative stochastic gene expression in morphogen gradients
2006 – 2007	Undergraduate Student Researcher University of California, Irvine Advisor: Dr. Wayne Giles Research Interest: Mathematical Model of *Drosophila* Myocyte Action Potential
2006	Summer R&D Intern Scios Inc, Johnson & Johnson Company Research: Safety pharmacology experiments on cardiac arrhythmia drugs

PUBLICATIONS

ABSTRACT OF THE DISSERTATION

Precision and Robustness in the Interpretation of Morphogen Gradients

By

Peng Cheng Zhang

Doctor of Philosophy in Biomedical Engineering

University of California, Irvine, 2015

Professor Arthur D. Lander, Chair

Developmental pattern formation is orchestrated by diffusible signaling molecules, termed *morphogens,* that form gradients from which cells can determine positionally appropriate fates. Stochasticity in morphogen binding, signal transduction, and gene expression create local cell-to-cell variability in the readout of morphogen gradients. However, little is known about the actual levels of noise in morphogen gradient responses, or the mechanisms that might control it. To investigate this, I quantified the transcriptional activity noise, protein expression noise, and protein half-life of *optomotor blind* (*omb*), one of the downstream targets of the morphogen Dpp in the *Drosophila* larval wing imaginal disc. Using combined fluorescence *in situ* hybridization (FISH) with intronic probes, immunofluorescence, image segmentation and image analysis, I observed a very high level of transcriptional variability characterized by coefficients of variation (CV) as high as ~110%, in the cells in which *omb* plays a central role in specifying the location of wing vein primordium L5. However, the half-life of the Omb protein was found to be very long, ~ 6 hours, which would be expected to provide

significant temporal filtering of the transcriptional noise. I showed that the reduction in noise from transcript to protein is sufficient to account for the precision of the positional information that patterns vein L5.

I also investigated why the positioning of vein L5 is remarkably robust to genetic manipulation that change the shape of the Dpp morphogen gradient. I observed that patters of Dpp signaling and Omb expression are not constant during larval development, but change continuously, and not always in concert with each other. By taking into account the long half-life of Omb it was possible to build a model that explains both these movements and the remarkable robustness of L5 patterning to changes in Dpp gradient shape.

CHAPTER 1

INTRODUCTION

Morphogen Theory

Spatial patterns can be found in almost all biological organisms, from the evenly distributed spots on cheetahs, to the periodic strips or zebras, to a mix of shapes and patterns on butterfly wings. Even the parts that may not take the spotlight are incredibly interesting, from the repeating elements in vertebrate backbone, to the placement of limbs, to the veination of fruit fly wings. All these patterns arise through development from an embryo that's seemingly homogenous from the outside. How do biological systems provide information and trigger different developmental fates? Early studies in the late nineteenth and early twentieth centuries recognized that embryos somehow contain spatially-coordinated pattern information in the form of properties or substances (Lawrence, 2001). Furthermore, German embryologist Hans Spemann and his graduate student Hilde Mangold showed that this pattern information could be physically or chemically transmitted from one embryo to another in Xenopus (Hamburger, 1969).

The mystery of these chemicals and how they worked persisted for the first half of the twentieth century. It was not until 1952 that Alan Turing, a British mathematician, code-breaker and computer scientist, first coined the term "morphogens" and proposed the idea that diffusion of these hypothetical substance can provide the information needed to generate patterns (Turing, 1952). Turing formulated a set of reaction-diffusion equations with a slowly diffusing local activator and a fast diffusing long-range inhibitor,

which reached steady state concentrations that formed different patterns as a result of their interactions with each other. This theory of morphogenesis remained largely obscure until 1969 when Lewis Wolpert re-introduced the idea of the morphogen through the “French Flag Model” (Wolpert, 1969). Wolpert proposed a simple system where molecules of morphogens were synthesized at a localized source, from which they spread away, and formed a concentration gradient. Cells in the receiving region would then determine their position according to a threshold level of the morphogen concentration and respond by turning on different gene expression.

Almost 20 years after the theoretical framework for morphogens was defined, the first experimental evidence of a morphogen was found in *Drosophila* embryos in 1988. Bicoid (Bcd) is a maternal effect gene and was discovered through a screen for genes involved in development of the anterior-posterior axis of the embryo. The Bcd protein works in the syncytial embryo and act as a cytoplasmic morphogen; it is expressed most highly at the anterior end of the embryo and gradually decreases toward the posterior, forming a concentration gradient (Driever and Nusslein-Volhard, 1988b). Distinct cell fates are specified along the concentration gradient in a dose-dependent manner (Driever and Nusslein-Volhard, 1988a). Since the discovery of Bicoid, many other molecules of various sizes that also fulfill the criteria of a morphogen have been identified in different organisms. Secreted proteins that qualify as morphogens include Sonic Hedgehog (Shh) in chick neural tube development (Briscoe et al., 2001), the TGF-beta family protein Decapentaplegic (Dpp) in *Drosophila* dorsoventral development (Ferguson and Anderson, 1992; Nellen et al., 1996), and the Wnt family protein

Wingless (Wg) in *Drosophila* wing development (Neumann and Cohen, 1997; Zecca et al., 1996). There are also non-polypeptide morphogens, such as retinoic acid (RA, derived from vitamin A), which influences zebrafish hindbrain development (Cai et al., 2012; White and Schilling, 2008) and vertebrate organogenesis (Duester, 2008).

After more than a decade of extensive research, the idea of cell fate determination and pattern formation from morphogen gradients is now widely accepted. Both theoretical and experimental studies have provided a diverse number of explanations and ample evidence for the formation and interpretation of morphogen gradients. Despite these tremendous efforts, there are still many questions about actual mechanisms and functional details that remain to be answered. The mechanistic issue of how cells determine a threshold level of morphogen concentration and execute a set of distinct response to generate sharp boundaries seems to be far more complex than previously thought, with increasing evidence that the level of concentration plays a more limited role than once thought (Alexandre et al., 2014; Chen et al., 2012; Ochoa-Espinosa et al., 2009). Then there are issues involving natural variability of the gradient, starting from synthesis of the molecule, to its secretion, to the binding of receptors, to activating signaling pathways and gene regulation. Coupling these sources of variability with environmental fluctuations and genetic differences makes the problem of biological noise seem almost insurmountable (Lander, 2013). Given the remarkable abilities of biological systems to develop normally, with consistent outcomes, the question for developmental biologists really comes down to what mechanisms and strategies do

different patterning systems employ to be able to achieve their final objectives reliably and robustly.

Threshold Interpretation of Morphogen

The Wolpert model of morphogen theory assumes cells read and respond to thresholds of morphogen concentration in order to generate distinct cell fates. Experimental evidence has suggested between three to seven distinct thresholds of gene expressions in morphogen gradients from different organisms. For example, the Wingless (Wg) gradient in the *Drosophila* wing disc produces three thresholds (Tabata and Takei, 2004), the Activin gradient in the *Xenopus* blastula cells induces five different cell fates (Green et al., 1992), and Dorsal (Dl) can specify between four to seven different targets in *Drosophila* embryo (Stathopoulos and Levine, 2002). One piece of direct evidence that cells can simply read the level of morphogen concentration comes from the *Xenopus* embryo where Activin regulates *goosecoid* (*gsc*) and *Xenopus brachhyury* (*Xbra*) in a dose dependent manner. An absolute receptor occupancy of 100 molecules of Activin bound to a cell (~2% of total number of cell receptor) will activate the expression of *Xbra,* whereas 300 molecules (6% of total number of cell receptors) will induce *gsc* and repress *Xbra (Dyson and Gurdon, 1998; Green and Smith, 1990; Smith, 2009).* If all cells indeed read the level of morphogen concentrations, how do they interpret the information and translate continuous gradients into discrete gene expression changes?

One molecular mechanism for interpreting morphogen concentration is via different genes having different binding-site affinities to an intracellular morphogen (or signal produced in proportion to the amount of extracellular morphogen). Target genes with low-affinity binding can only be activated in the region of high morphogen signal, and target genes with high-affinity binding can be activated even if the morphogen signal is low, thus creating a staggered expression with other downstream targets of the morphogen. This mechanism of morphogen interpretation, through molecular control of transcription, has been demonstrated in the D/V patterning of the *Drosophila* embryo. The dorsal (Dl) gradient is responsible for initiation of the mesoderm, neuroectoderm, and dorsal ectoderm where specific enhancers from different genes have different affinities for Dl (Jiang and Levine, 1993; Stathopoulos and Levine, 2004). Computational analysis from a large set of enhancers of Dl target genes further confirms that their expression domain is related to Dl affinity (Stathopoulos and Levine, 2005).

Binding site affinity is a good starting point for a morphogen threshold readout, but it is usually insufficient to create multiple, discrete gene expression boundaries. In many cases, genes integrate information from the morphogen signal itself and other transcription factors that bind to regulatory elements to make cell fate decisions. In the *Drosophila* embryo, regulation of the AP axis depends on the morphogen Bcd, but most downstream targets of Bcd have enhancers that receive additional inputs from transcription factors such as Hunchback (Hb, activated by Bcd), Caudal (Cad, repressed by Bcd), and Krüppel (Kr) (Ochoa-Espinosa et al., 2005). Downstream targets can also integrate the signal with itself, creating positive feedbacks that can help refine

expression borders. In the AP axis of the vertebrate hindbrain development, a gradient of RA is responsible for setting the anterior limit of *Hoxb4*. While the initial anterior border during early development is diffuse, the *Hoxb4* locus is upregulated by the Hoxb4 protein itself. This auto-regulation helps refine the anterior limit and maintain its expression as development progresses (Gould et al., 1998; Gould et al., 1997).

Another mechanism for generating sharp boundaries is cross repression of two or more morphogen regulated genes. In dividing up the neuroectoderm of the *Drosophila* embryo into three columns along the DV axis, three homeobox transcription factors (Vnd, Ind and Msh) respond to a distinct threshold of Dl signal. But maintenance and creation of an abrupt switch in expressions depend on cross regulatory interactions, where genes expressed more ventrally repress genes closer to dorsal regions (Cowden and Levine, 2003).

It is apparent that interpreting morphogen gradients is not simple in most systems. In order to translate the level of a graded signal into regions of distinct gene expression, there may be a variety of regulatory mechanisms at work. In some cases, the actual level of the morphogen in a system has been shown to have a surprisingly small effect on where downstream targets are expressed. For example, two independent studies in *Drosophila* altered the shape of the *bicoid* morphogen gradient: one flattened the gradient by delocalizing *bcd* mRNA and relieving its transcriptional repression (Ochoa-Espinosa et al., 2009); the other triggered cytoplasmic flow within the embryo by creating a temperature gradient, which thereby changed nuclear locations and disturbed stability of the Bcd gradient (Lucchetta et al., 2008). Both studies found

bcd downstream anteroposterior gene expression domains (i.e. even-skipped and gap genes) were altered far less than expected for the amount of disruption of the gradient. These results suggest that the downstream cross-regulatory gene networks work in conjunction with morphogen gradients to produce discrete expression patterns. Indeed, it was found that such cross-regulation in the gap genes (Manu et al., 2009), working with a system of repressive gradients that antagonize Bcd mediated activation (Chen et al., 2012), contribute to the robust boundaries of gene expressions.

Spatiotemporal Interpretation of Morphogen

To reliably generate gene expression patterns from a morphogen gradient, developing systems are not limited to using instantaneous, single-cell reads of concentration level as their sole source of positional information. There are secondary properties inherently built into morphogen gradients, and cells can incorporate mechanisms that make use of these properties as a source of positional information.

For example, in the spatial domain, cells can make use of both the spatial integral or the spatial derivative of the gradient. To measure the spatial integral of the morphogen, cells can respond by polling (or averaging) information with their neighbors. This can happen purely due to the physical state of the developing organism, such as in syncytial embryos, where downstream effectors of morphogen signaling can diffuse from one nucleus to another (Gregor et al., 2007). Cell can also measure the spatial integral when a morphogen induces the expression of a diffusible factor. The diffusion range of this factor will be proportional to the range of spatial integration, and the level

will reflect the sum of the morphogen signal in that area. To measure spatial derivatives, cells can respond to differences in concentration level between themselves and their neighbors. This measurement of the spatial slope has been found to be involved in growth regulation in *Drosophila* wing discs. BMP signaling regulates growth in wing discs by two distinct mechanisms: one depends on the concentration level of Dpp, and the other depends on the slope of the gradient (Rogulja and Irvine, 2005; Rogulja et al., 2008). These two distinct mechanisms act in conjunction to create a relatively even distribution of cell proliferation, in which the gradient response mechanism promotes proliferation proportional to the slope of the gradient.

While cells can use physical mechanisms to extract spatial derivative or integral information from morphogen gradients, they can extract temporal information through downstream gene regulatory loops. To measure the temporal integral, or duration of exposure to the morphogen signal, cells can use a feed-forward loop downstream of the target. In this case, the morphogen signaling first activates an intermediate product that accumulates for a certain amount of time, and this intermediate product then activates the final target (Dessaud et al., 2010). The delay to activate the final target is proportional to the strength and duration of the morphogen signal, which is a measure of the temporal integral. One of the best-understood examples of this is vertebrate spinal cord development. Shh forms a ventral to dorsal concentration gradient that assigns neural progenitor identities in which the duration of exposure to Shh plays a crucial role for downstream target V3 interneuron marked by *Nkx2.2.* Shh initially activates *olig2* expression where *Nkx2.2* is repressed by Pax6. When Olig2 accumulates sufficiently it

represses Pax6 and allows Nkx2.2 to be activated. The signal accumulation of Olig2 in this feed-forward loop serves as a temporal integral readout of Shh signaling that turns on *Nkx2.2* (Dessaud et al., 2010; Dessaud et al., 2007; Lek et al., 2010).

Cells can also respond to the relative rate of change in the morphogen signal, which is a measure of the temporal derivative. It has been observed in *Xenopus* dorsal-anterior development that the level of β-catenin, an intermediary signal in the morphogen Wnt, is sensitive to perturbations in the pathway. However, downstream target gene expression responds to fold-changes in β-catenin and is robust to the final level. In this system, fold-changes in β-catenin are a more precise and important output of Wnt than their absolute levels (Goentoro and Kirschner, 2009). The authors of this work point out that this fold change detection can be achieved through an incoherent feed-forward loop, in which signaling activates a response but also represses the response through a parallel pathway with a different timescale (Goentoro et al., 2009). Fold change detection have also been found in the ERK2 response to EGF signaling in cell lines (Cohen-Saidon et al., 2009), and has been proposed to be a driving force in *Drosophila* wing disc growth control (Wartlick et al., 2011).

Robustness and Precision of Morphogen Gradients

It is not surprising that organisms undergoing development possess a variety of mechanisms to interpret both the levels and spatiotemporal dynamics of morphogens, since each system must achieve its own unique performance objectives. While it is impossible for systems to be precise and robust in the face of all possible perturbations,

the mechanisms in place usually reflect which types of fluctuations the system is capable of overcoming. Robustness of a biological system refers to its ability to consistently produce the same (or extremely similar) outputs with different types of environmental and genetic perturbations, which usually enter upstream of individual cells and affect the entire tissue. It can be quantified by a sensitivity coefficient, a unit-less number that can be defined as the fold change in the location of a positional value divided by a fold change in the upstream process. Precision is the magnitude of natural variation in the output from normal developmental perturbations, mostly from cell-to-cell variations due to stochastic biochemical processes. In a patterning systems, it can be quantified as a transition width, a physical range of "salt-and-pepper" gene expression wherein a predefined maximum proportion (e.g., 15%) of cells express incorrect genes (Lander, 2013).

Several theoretical studies have shown how different mechanisms of gradient formation or interpretation affect robustness of morphogen gradients (Eldar et al., 2003; He et al., 2008; Lander, 2013; Lander et al., 2009). In morphogen gradients with self-enhanced degradation, where morphogen signaling leads to an increase in morphogen degradation, robustness to variations in levels of morphogen production increases (Eldar et al., 2003). Evidence for this strategy was found in the Wingless (Wg) gradient of the *Drosophila* wing and the retinoic acid gradient in the zebrafish (White et al., 2007). This strategy essentially increases morphogen uptake near the source where morphogen production increases and creates a shallower gradient away from the source—which unfortunately also leads to greater imprecision in the face of cell-to-cell

variability in the shallow parts of the gradient. Another strategy for reducing sensitivity to the morphogen production rate is by reading the pre-steady state gradient (Bergmann et al., 2007; Nahmad and Stathopoulos, 2009). While robustness to morphogen production is increased here, the system becomes fragile to the timing of when cells must make this measurement. In another example of pre-steady state interpretation, the Hh gradient defines the expression domain of *dpp* while it is undergoing a biphasic dynamic (Nahmad and Stathopoulos, 2009). Such a strategy becomes dependent on the rate of Hh spread relative to the rate at which its receptor, *ptc*, is upregulated.

Although there are tradeoffs for every mechanism, organisms have evolved to adapt to their specific circumstances. For example, smaller cells will likely place a greater emphasis than larger cells on robustness to levels of receptor fluctuation. One interesting example is in vertebrate hindbrain formation, where retinoic acid (RA) acts as a morphogen. RA is a derivative of dietary vitamin A, and since it is not encoded by the genome, its availability varies greatly from embryo to embryo depending on environmental factors. It has been observed that feedback regulation of RA biosynthesis, degradation and delivery all help to decrease sensitivity to variation in RA levels (White and Schilling, 2008). Large experimentally-induced fluctuations in RA level show minimal effect on the positional information of hindbrain cells. This strategy increases sensitivity to other components of the morphogen gradient, but it is likely that the other components have more reliable levels.

To improve precision in morphogen systems, spatial and temporal integration of morphogen signals can help. Temporal averaging can reduce short-lived disturbances

and produce a more stable readout, but this strategy requires time to be sufficiently long to average out such noise; consequently, positional information cannot be provided for such systems quickly. Spatial averaging can smooth out stochastic variations, but when systematic perturbations occur, patterns are unlikely to be placed in their correct locations. It is important to note that noise is not the only factor that determines precision. Transition width is calculated from the level of noise and steepness of the gradient, according to the formula $w=2cv \times \lambda$, where cv is the coefficient of variation of the noise, and λ is the length scale of the gradient. Because it is the product of noise level and gradient steepness that sets the transition width, systems can increase precision by either making a gradient steeper or reducing the variability between cells.

Spatial Precision or Noise in Morphogen Gradients

Morphogen gradients provide a solid framework for developmental systems to generate patterns of gene expression, but it is unavoidable that random noise can be found in every step of the process. Although developmental systems can employ a variety of mechanism to suppress or take advantage of fluctuations in molecular and cellular behavior, how much variability is present at different steps and how they are handled by morphogen systems is not fully understood.

The first step in a morphogen patterning is the production and formation of the gradient by transport of the morphogen. Morphogen gradients usually form through transport by diffusion (Muller et al., 2012; Schwank et al., 2011; Zhou et al., 2012), which relies on movement of molecules by random Brownian motion (Crick, 1970). This

process is obviously stochastic, but it is surprisingly reliable since diffusion coefficients are relatively insensitive to temperature fluctuations and furthermore independent of cellular mechanisms. Additionally, effective diffusivity is relatively insensitive to the tortuosity of the path even when cells are densely packed (Lander, 2013; Rusakov and Kullmann, 1998). Also, the rapidity of diffusion over short length scales means that a great deal of the randomness of diffusion time-averages much faster than the downstream processes that cells use to respond to diffusing molecules. These properties of diffusion make the transport step in the formation of diffusion gradients surprisingly predictable and reliable.

The next step in morphogen systems is translating the morphogen into a signal. It is at this step that the amount of noise that may be added into the process is not well studied, although some work has been done on the role of stochastic fluctuations in receptor occupancy (Lander et al., 2009). More attention has been focused, so far, on transcriptional noise downstream of morphogen signals, which is most likely a major source of input noise, given the stochastic nature of transcription. Whether a gene is transcribed depends on several factors: availability and binding of a regulatory protein to its designated sites on DNA, accessibility of the sites, and the necessity of multiple biochemical steps for initiating transcription (Coulon et al., 2013). Recent advances in technology have allowed researchers to quantitatively measure transcriptional noise in different organisms. The first quantitative study of transcriptional noise demonstrated single molecule mRNA detection in single cells using mRNA fluorescent *in situ* hybridization (FISH) (Femino et al., 1998). The technique have since then been

modified and improved and is still used in quantification of mRNA for a variety of developing organisms (Gregor et al., 2007; Paré et al., 2009; Raj et al., 2010; Raj and van Oudenaarden, 2009).

From the variety of systems studied, it has been found that transcription can be initiated either continuously or intermittently. In the continuous mode, random and uncorrelated binding of the necessary components determines transcriptional initiation, with a probability that is uniform over time. The Poisson process is a continuous time stochastic process and often used to count the number of events in a given time interval. The transcriptional noise associated with a continuous mode is thus considered "Poissonian" (Larson et al., 2011; Yunger et al., 2010). In the intermittent case, mRNA is synthesized in "bursts", with periods of high initiation followed by long periods of no initiation. The "burstiness" has been estimated using the mean expression level of a gene, as higher expression tends to be associated with higher or more frequent 'bursts' (Golding et al., 2005; Muramoto et al., 2012; Raj et al., 2006). The size of bursts correlate with the magnitude of the noise; larger burst sizes will have higher variance. Generally, bursty transcription will result in higher noise than the continuous transcriptional mode (Sanchez and Golding, 2013).

The amount of noise and the stochastic kinetics of transcription can be influenced by two mechanisms, gene-specific regulation and global genome-wide constraints. In gene-specific regulation, the specific promoter and enhancer architecture for individual genes determines the stochastic kinetics by controlling the binding and unbinding of different elements for transcriptional initiation (Blake et al., 2006; Carey et al., 2013;

Raser and O'Shea, 2005; Sanchez and Golding, 2013). Experiments from yeast *Saccharomyces cerevisiae* found no obvious trend between mean expression level and burst size from a large set of promoters (Bar-Even et al., 2006; Carey et al., 2013; Hornung et al., 2012; Newman et al., 2006). Additionally, different characteristic architectures were found between promoters with low and high transcriptional noise (Field et al., 2008; Tirosh and Barkai, 2008). These observations strongly suggest gene-specific regulation of transcription noise that is determined by the DNA sequence of the promoters. In gene-independent regulation, global constraints dominate transcriptional kinetics. Physiological or biophysical mechanisms such as cell-cycle-dependent regulation (Zopf et al., 2013) or cooperative recruitment of RNA polymerases (Guptasarma, 1996) limit the control that cis-regulatory regions have on transcriptional kinetics. For example, one proposed constraint is that all transcriptional events are bursty by nature. This would limit even fully active genes to adopting the same mechanism; all genes would follow the same trend between burst size and mean expression level (Dar et al., 2012; So et al., 2011). Evidence for global noise constraints has been found in *Escherichia coli* (Golding et al., 2005; So et al., 2011), cultured mammalian cells (Dar et al., 2012; Yunger et al., 2010), mouse (Itzkovitz et al., 2012), and *Drosophila* embryos (Little et al., 2013; Paré et al., 2009) in which a majority of transcription is bursty in nature (Raj et al., 2006). It is important to keep in mind that both mechanisms are not mutually exclusive, and transcriptional kinetics could be dominated by different mechanism during different developmental stages, or affected by a combination of the two depending on external factors (Sanchez and Golding, 2013).

Within a morphogen gradient system, measured variation (or spatial noise) of mRNA content does not come from just the noise introduced by transcriptional kinetics. It can be broken into two components, intrinsic and extrinsic transcriptional noise. Intrinsic noise is due to the process of transcription discussed above, and it refers to the moment-to-moment variability in transcriptional response to a signal. Extrinsic noise is the "input" noise that comes in upstream of the transcriptional process; it is the accumulated noise from all the steps leading up to transcriptional regulation and the cell-to-cell variability in perceived morphogen signal.

To uncover gene expression noise within a morphogen system, studies in the *Drosophila* embryo have attempted to quantify the variability of the morphogen Bcd and some of its downstream target genes. In the early syncytial embryo, the maternally-supplied transcription factor Bcd acts as a morphogen and triggers different gene expression patterns (Sauer et al., 1996). The Bcd protein concentration gradient across different embryos is remarkably reproducible (Gregor et al., 2007). Cell-to-cell variation in Bcd concentration at similar AP coordinates is also especially precise, varying less than 10% (Gregor et al., 2007). Furthermore, this level of precision in protein concentration is preserved downstream in zygotically expressed target genes such as Hb (Dubuis et al., 2013; Gregor et al., 2007). Interestingly, nascent transcriptional noise across these patterning genes downstream of *bcd* (including *hb*) is much higher. Quantification of FISH signals inside a nucleus revealed that transcriptional activities vary about 45% for each individual locus independent of specific cis-regulatory architecture (Little et al., 2013). The similarities in stochastic transcriptional noise across

different genes suggest that such variability is likely the result of a fundamental global constraint that limits the degree of precision. In order to recover precision from noisy mRNA production, additional regulation or an additional mechanism must be present in the embryos. Quantitative imaging of transcription in live embryos discovered that the rate of polymerase loading, the transcription time window, and the fraction of active nuclei essentially allow embryos to average expression over space and time to produce a more precise cytoplasmic mRNA level (Garcia et al., 2013).

It is not surprising that morphogen gradients at the earliest stages of embryo development display precise spatial gene expression patterns, as this is the first step to ensure correct patterning in development. Although some studies have questioned whether such precision or even the presence of a morphogen gradient is necessary for early development (Lucchetta et al., 2008; Ochoa-Espinosa et al., 2009), experimental evidence supports the view that morphogen gradients can, when necessary, achieve a very high level of precision.

How do other morphogen systems compare to the *Drosophila* embryo in terms of noise and does the level of morphogen play a major role in interpretation of another often studied morphogen system? In this work, I have decided to look at the development of *Drosophila* wing imaginal disc.

Drosophila Wing Disc Development

The *Drosophila* wing imaginal disc is a classical model for studying growth and patterning. This epithelial structure grows from about 20-30 cells to 50,000 cells during

the three larvae stages, adopting the characteristic shape of a flattened sac with two distinct surfaces, by the third-instar stage. On one face of the sac is a thin layer of peripodial membrane, which only makes minor contributions to the adult structure of the wing. On the other surface, the thicker folded columnar cell layer becomes the wing blade, hinge and much of the thorax of the adult fly. The wing blade forms during prepupal and pupal development from a region of the disc referred to as the wing pouch, which everts, lengthens and folds over on itself to form the dorsal and ventral epithelia of the wing blade (Milan et al., 1996). The most noticeable pattern elements of the *Drosophila melanogaster* wing are the five longitudinal and two transverse veins. These veins serve as structural support for the wing and as vessels for the delivery of hemolymph, nerves, and blood cells. The arrangement of veins is stereotyped and robust, studies show vein locations vary around one single cell width across different genetic backgrounds and growth temperatures (Abouchar et al., 2014), providing an intriguing puzzle in pattern formation.

Starting from early development, the wing disc is bisected into anterior and posterior compartments, with certain genes being expressed only in the posterior compartment. These include Engrailed, Invected (Inv) and Hedgehog (Blair, 2007). Hh is a short-range secreted morphogen that forms a concentration gradient within the anterior compartment. The receptor for Hh, Patched (Ptc), is constitutively expressed in the anterior compartment and represses Hh downstream targets through inhibition of the transmembrane protein Smoothened (Smo) (Denef et al., 2000). Binding of Hh to Ptc induces its internalization and degradation, which allows Smo to be phosphorylated.

The phosphorylated Smo inhibits cleavage of the transcription factor cubitus interruptus (Ci), which is expressed only in the anterior compartment, and permits full-length Ci to activate Hh downstream targets by entering the nucleus (Apionishev et al., 2005; Jia et al., 2004; Methot and Basler, 1999). This results in activation of transcription of *Dpp*, a homolog of Bone Morphogenic Protein (BMP) 2/4, in a stripe of cells on the anterior side of anterior/posterior (A/P) boundary (Padgett et al., 1987). Dpp is a long-range morphogen protein that spreads to both sides of the disc through diffusion (Zhou et al., 2012), forming two morphogen gradients, one toward the anterior and one toward the posterior (Entchev et al., 2000; Teleman and Cohen, 2000a). Both Hh and BMP signaling are known to be critical factors that position the longitudinal veins (Blair, 2007). The zone of Hh reception directly controls the placement and spacing of veins L3 and L4, while BMP signaling is required to control the placement of L2 and L5.

The BMP activity gradient in the wing disc is actually formed from a combination of two different signaling ligands: Dpp and Glass bottom boat (Gbb, also known as 60A). Similar to other BMP4s, Dpp is synthesized as an inactive precursor and requires cleavage at two sites for proper function. Cleavage at site S1 adjacent to the ligand domain allows Dpp to function as a short-range morphogen in the embryonic midgut, while cleavage at both sites enables Dpp to act as a long range morphogen in the wing disc (Kunnapuu et al., 2009; Sopory et al., 2010). Unlike Dpp, Gbb is expressed uniformly in the wing and has phenotypic effects only on L5 (Ray and Wharton, 2001). Both BMP proteins bind to type I serine/threonine kinase receptors, Thickveins (Tkv), or Saxophone (Sax), and the type II receptor Punt (Put) (Nellen et al., 1994; Shimmi et al.,

2005). Dpp strongly binds to Tkv, whereas Gbb has a higher preference for Sax than Tkv (Khalsa et al., 1998). It has also been found that Dpp and Gbb can form heterodimers when synthesized in the same cells, and such heterodimers can signal through Sax-Tkv heteromeric complexes, apparently having a particularly strong effect on the placement of vein L5 (Bangi and Wharton, 2006a; Bangi and Wharton, 2006b).

BMP signal transduction in the wing disc occurs similarly to other members of the TGFβ family, starting with activation of the heteromeric receptor complex. Activated Tkv then phosphorylates a BMP specific Smad, Mothers against dpp (Mad), to form pMad. PMad forms a new complex with the *Drosophila* co-Smad, Medea (Med) and starts to accumulate in the nucleus. This Mad/Med complex participates directly in gene regulation through Smad-homology domain 1 which behaves as a DNA-binding domain that recognizes a CG-rich sequence motif (Kim et al., 1997; Massague et al., 2005). Even though Mad/Med can directly regulate many Dpp-responsive genes, two other proteins, Schnurri (Shn) and Brinker (Brk), turn out to be crucial in the readout of the Dpp morphogen gradient. Schnurri (Shn) is a transcription factor with eight zinc-finger DNA binding domains that forms a Shn/Mad/Med (SMM) complex in cells where BMP signaling is present (Arora et al., 1995). Brk functions as a constitutive repressor of Dpp target genes by binding to enhancers via the sequence GGCGYY (Sivasankaran et al., 2000; Winter and Campbell, 2004; Zhang et al., 2001). The SMM complex represses *brk* transcription by binding to the silencer element of the *brk cis*-regulatory region, which contains a GRCGNC(N5)GTCTG motif (Gao et al., 2005; Muller et al., 2003; Pyrowolakis et al., 2004). Thus, Dpp controls target gene expression through two

mechanisms: directly activating gene expression through Mad/Med complex and indirectly by relief of repression from Brk.

Since pMad in the wing disc is an important component for both gene control mechanisms, it is often considered the most direct readout of the BMP activity gradient. However, the shape of the pMad gradient is not just the result of a linear signal transduction. Dpp down-regulates its receptor Tkv, and this negative feedback helps to increase the range of Dpp by creating a Tkv expression profile that is low near the central region and high in the lateral region of the disc (Lecuit and Cohen, 1998; Tabata, 2001). At the same, Tkv downregulates Dpp by receptor mediated endocytosis that quickly internalizes BMPs (Jortikka et al., 1997). Another feedback loop in the system comes from the inhibitory Smad, Daughters against dpp (Dad), which decreases *dpp* expression. Dad is induced by BMP and prevents accumulation of pMad by feeding back on Mad interaction with its receptors (Ogiso et al., 2011). This feedback loop helps maintain a reliable signal amplitude by preventing an increase in steady-state intracellular response.

The exact reason for these feedback loops is not well understood, in terms of the eventual goal of pattern formation, but it is evident that the gradient of BMP activity and its inverse Brk gradient set important target gene expression domains in the wing disc. Two particularly important downstream target genes are *optimotor blind* (*omb,* as known as *bifid*) and spalt (*salm* for spalt major and *salr* for spalt related). The expression boundary of *Salm* has been found to be regulated by Brk but the transcriptional rate of *salm* is also regulated directly by pMad (Campbell and Tomlinson, 1999; Marty et al.,

2000), providing one example of gene regulation through two different mechanisms downstream of Dpp. In contrast to *salm*, the Omb expression profile is regulated by only Brk repression; activation of *omb* does not require Dpp signaling or Smad complexes and is turned on by unknown factors in the wing disc (Winter and Campbell, 2004).

Like *brk*, *salm* and *omb* are both transcription factors, and together they are involved in defining the location of the longitudinal vein primordia L2 and L5. Salm expression is centered around the A/P boundary in a broad region and borders the L2 provein (vein primordium) in the posterior direction. Two theories exist on how the L2 provein is induced. The first theory is based on experiments in which a clone of $salm^{-}$ cells inside the endogenous Salm region induce ectopic veins inside the $salm^{-}$ clone. One explanation is that cells expressing high levels of Salm produce a short-range signal X, to which they cannot respond, and which induces the expression of L2 organizing gene, *knirps* (*kni*) and *knirps related* (*knrl*), in neighboring cells (Lunde et al., 1998). The second theory comes from experiments where both *salm* and *salr* are removed from a clone, and L2 does not form. The authors suggest that L2 is induced only in low levels of Sal expression, typically found at the anterior limit of Salm/Salr domain, thus removing *salm* mimics the effect of lowering Sal expression level (de Celis and Barrio, 2000).

Salm and Salr may also play a role in formation of L5. Clones of *salm* and *salr* show expression of ectopic veination anterior to L5 and could suggest they repress L5 induction. Since Omb is expressed in a domain wider than Salm and Salr and since L5 forms in between Brk and Omb, Salm and Salr could represent the anterior limit of L5

provein. The induction of the L5 provein also depends on Brk and Omb. One theory proposes a similar scenario to L2. In the posterior compartment, Brk expressing cells produce a short-range signal, Y, to which they cannot respond to. This signal acts on neighboring *omb*-expressing cells and induces the L5 organizing gene, *abrupt* (*ab*), that only turns on where both signal Y and Omb are present (Cook et al., 2004). This combined requirement for a short-range signal, Omb expressing cells, and repression from Salm/Salr limits L5 initiation to a stripe of cells between Brk and Omb in the posterior compartment.

However, neither of the postulated short range signal proteins (X and Y) has been identified, and a recent study found that the endogenous Omb protein is actually expressed in a gradient with no sharp boundaries (Shen et al., 2010). Instead Omb was found to be expressed high in the center and to decline smoothly toward the lateral margin of the wing pouch in a linear fashion. These results complicate the current model of L5 formation and suggest that alternative theories should be explored.

Once the distinct organizing genes for L2 and L5, Ab and Kni/Knrl, have been induced, they act similarly, linking the positional information provided by BMP to activation of EGF signaling. Ab and Kni/Knrl promote expression of the vein gene *rhomboid* (*rho*) and repress expression of intervein gene *blistered* (*bs*) (Blair, 2007). Finally in the pupal stage, through a series of short-range activation and long range inhibition mechanisms involving EGF, Notch and BMP pathways, broad provein regions are refined to narrow veins (Blair, 2007; Yan and Lin, 2009).

CHAPTER 2

Materials and Methods

Fly strains

Fly strains were mostly obtained from Bloomington *Drosophila* Stock Center. Thickvein mutant allele tkv^{a12} was a generous gift from Dr. Hiroshi Nakato. Flies used to generate clones was a gift from Dr. Ken Irvine, listed below.

y[1] w[*]; P(s[+mC]=UAS-GFP.S65T)T2; P(w[+mC]=Act5C(FRT.y[+])Gal4.Switch.PR)3/TM6B, Tb[1], w hs-FLP; UAS-tkv[QD]/Tm6B, y[1] w[*] hs-FLP; UAS-ds(III), y[1] w[*] hs-FLP; UAS-fj(II), y[1] w[*] hs-FLP; UAS-brk.

RNAi stocks including *UAS-tkv-RNAi(KK), UAS-tkv-RNAi(GD), UAS-omb-RNAi(KK)* were obtained from Vienna *Drosophila* Resource Center. YFP fusion protein lines from Cambridge Protein Trap Insertion Project (CPTI) were created by UK laboratories using hybrid piggyback/P-element protein trap construct, tkv-YFP (*w[1118]; PBac[544.SVS-1]tkv[CPTI002487]*) were obtained from Kyoto *Drosophila* Genetic Resource Center.

Clone Generation and Transgene Induction

To generate Flp-out clones, flies of genotype *y[1] w[*]; P(s[+mC]=UAS-GFP.S65T)T2; P(w[+mC]=Act5C(FRT.y[+])Gal4.Switch.PR)3/TM6B, Tb[1], w hs-FLP; UAS-tkv[QD]/Tm6B, y[1] w[*] hs-FLP; UAS-ds(III), y[1] w[*] hs-FLP; UAS-fj(II), y[1] w[*] hs-FLP; UAS-brk* were crossed to *UASGFP; AyGal:PR[3]/TM6b* flies. Flp-out clones were then generated by heat shock for 6-10 minutes at 37 °C water bath.

To active the Gal4:PR, individual larvae were selected and transferred to instant *Drosophila* food (Connecticut Valley Biological) with addition of RU486 (mifepristone, Sigma). RU486 was dissolved in water to a concentration of 24 µg/ml for high dose experiments and 1.2 µg/ml for low-dose experiments. They are then added to dry instant food of various weight for a final effective concentration of 20 µg/ml and 1µg/ml.

Immunofluorescence (Antibody Staining)

Wing discs were dissected on ice cold phosphate buffered saline (PBS) and fixed in 4% paraformaldehyde (Electron Microscopy Sciences) for 25 minutes at room temperature. Fixed carcass were washed 3X 10 minutes with 1X PBT (PBS + 0.1% Tween-20) and blocked overnight at 4C or 2 hours at room temperature with 0.2 % BSA, 0.3% deoxycholate, 0.3% Triton X-100. Following 6 x 10 minute washes with 1X PBT, discs were incubated with primary antibody overnight at 4C and then washed 6X 10 minutes. Secondary antibody added with PBT for 1.5 hours at room temperature. If DAPI or Hoechst is performed, they are added after the 2nd antibody with PBT at 0.1% concentration, and followed by 5 washes of 10 minutes each. Primary antibodies used rabbit anti-abrupt from Dr. Ethan Bier, (1:500), rabbit anti-lamin Dm0 (1:1000), rabbit anti-lamin C (1:1000) courtesy of Dr. Paul Fisher, mouse anti-fj (1:500) from Dr. David Strutt, rat anti-knirps (1:500), guinea pig anti-knirps (1:500) gift from Dr. VijayRaghavan, rat anti-salm (1:150) courtesy of Dr. Tiffany Cook, rabbit anti-sal (1:500) from Dr. Adi Salzberg, rabbit anti-omb (1:1000) gift from Dr. Gert Pflugfelder, rabbit anti-tkv (1:500) from Dr. Michael O'Connor, mouse anti-Beta-Gal (1:1000, Promega), mouse anti-delta

(1:500, C594.9B, DSHB), mouse anti-wg (1:1000, 4D4, DSHB), mouse anti-lamin (1:500, ADL195 & ADL84.12, DSHB), mouse anti-Phospho-Smad (1:1000, Epitomics). Secondary antibodies all from Invitrogen0 used at 1:100 include Alexa 555-Goat anti-rabbit, Alexa 555 goat anti-mouse, Alexa 488 goat anti-rabbit, Alexa 488 goat anti-mouse, Alexa 647 sheep anti-guinea pig, Alexa 647 goat anti-mouse.

Wing disc is dissected from the carcass and mounted in between two thin strips of tape acting as a bridge in mounting medium. Images are acquired via confocal laser scanning microscope or Zeiss Apotome microscope.

Cell proliferation studies

Larvae were dissected on growth medium consisted of M3 medium supplemented with 2.5% Fetal Calf Serum, 2.5% fly extract, 10 µg/ml of insulin, 100 units/ml Penicillin G, 100 µg/ml Streptomycin. The samples were then incubated in fresh growth medium with Edu concentration of 10 µM for 30 minutes at 25 °C. This is followed by normal staining protocol replacing antibody detection steps with the Click-iT* imaging kit from Invitrogen.

To make fly extract that's added to the growth medium, healthy living flies were collected and froze at -80 °C until a minimum of 7.5 g total weight were reached. Next, I added 50ml of M3 medium to the flies and homogenize in small quantities using a loose-fit pestle until all 50ml are processed. Then the mush spun on a table-top centrifuge at 2600 RPM for 5 minutes. The supernatant of the product is transferred to a fresh tube and the oily layer that forms on top is discarded. The remaining extract is

heat inactivated in a 60 °C water bath for 10 minutes until a precipitate forms. This is then spun for 1 hour at 2600 RPM and the supernatant is filtered through a 0.22 µm filter. The final solution can be stored at -20 °C in small volumes for use.

Fluorescent In Situ Hybridization (FISH)

In situ probes are designed and gene synthesized from Genewiz and put into pUC57-Amp vector. It was then cloned into pBlueScript II vector flanked by T3, T7 or Sp5 promoter sequence (Roche). Template DNA is linearized and heated to 55°C for 2 minutes and rest of the steps performed with RNase-free reagents. Labeled RNA is then synthesized with different hapten-U NTP mix (DIG or DNP labeled) and the appropriate RNA polymerase, incubate at 37°C for 2 hours. The resulting mix is fragmentated with 2x carbonate buffer for 10-30 minutes and precipitated with sodium acetate and stored at -80°C.

All reagents used for FISH are RNase free. *Drosophila* larvae were dissected on cold PBS and fixed with 4% paraformaldehyde for 25 minutes at room temperature. This is followed by 4 quick washes in PBT for a total of 20 minutes, then rinsed with PBS once and then treated with 80% acetone at -20°C for 10 minutes. The carcass is not fixed for a 2nd time with 4% paraformaldehyde after quick PBT washes to get rid of any remaining acetone. The carcass are again washed with PBT and rinsed with 1:1 PBT:FHS (fluorescent hybridization solution). FHS consists of 50% formamide, 5X SSC (Saline Sodium Citrate), 100 µg/ml fragmented salmon testes DNA (Sigma) and 50 µg/ml heparin (Sigma).

Before hybridization of probes, carcass is pretreated with FHS for 90 minutes with 2 solution changes at 55°C, and probes are heated for 3 minutes at 80°C and chilled on ice for 5 minutes to denature secondary structure. For hybridization, 100ng of desired labeled probes are added with FHS to the carcass and incubated at 55°C for a minimum of 32 hours. Once hybridization is finished, a sequential wash method is used with FHS and PBT at 3:1, 1:1, 1:1, and 1:3 ratio to effectively remove FHS and all unattached probes, these washes last 20 minutes each at 55°C, and is finished with two PBT only washes. The sample is then cooled to room temperature, rinsed and blocked with 10% milk in PBT for 2 hours at room temperature.

To detect the hybridized probe(s) in single or double probe FISH experiment, I diluted primary detection reagent Biotinylated anti-DIG (Jackson Immuno Research Lab) at 1:100 in PBTB (PBT + 1% milk) and incubated overnight at 4°C. Next day, washed samples 6 times for a total of 2 hours with PBTB, then incubated with appropriate 2^{nd} detection antibody, in this case streptavidin-HRP (TSA kit, Invitrogen) at 1:00 dilution for 1 hour at room temperature. Samples are washed again for 2 hours changing solution every 20 minutes and incubated in dark for 1.5 hours with tyramide-alexa fluor (1:100) in amplification buffer (TSA kit Invitrogen). This is followed by more washes and if a 2^{nd} labeled probe needs to be detected, 1% hydrogen peroxide treatment follows the washes to clear all tyramide reaction.

To detect the 2^{nd} probe, chose HRP conjugated antibody against the labeled hapten (HRP-anti-DNP, 1:100 in PBTB) and incubated overnight at 4°C. Similar to previous steps, samples were washed for 2 hours and then incubated with a different

tyramide-alexa fluor (1:100) in amplification buffer for 1.5 hours. This is washed and ready for mount unless antibody staining is required. To continue with immunofluorescent staining, block with 1% BSA and then follow normal antibody staining procedure.

This FISH protocol was modified from previously published methods (Kosman et al., 2004; Lecuyer et al., 2008).

Image segmentation and analysis for FISH

Segmentation of nucleus was done semi-automatically using a software plugin for imageJ developed by Dr. Cinquin (in preparation). Lamin or DAPI staining were used to mark the nucleus and nuclear centers were manually annotated. The images are then preprocessed for segmentation by normalizing, blurring and inverting the image. Next, an algorithm for running active contour initialized from the annotated nuclear center was used (De Solorzano et al., 2001) to generate segmentation masks. Borders of lamin signal or areas of low DAPI intensity in the original image along with neighboring cells were used to constrain the active contour boundaries.

To assess the impact of fluorescence attenuation along the z-axis of the image, I identified only nuclei that were on the “top layer” of the tissue and compared the results to that of using all nucleus in the z-stack image. Since signal attenuation comes mostly from the amount of tissue it has to traverse, the top layer of the tissue should have minimally attenuated signal regardless of its actual physical distance from the objective lens. This “top-layer” algorithm was part of the software plugin and identifies nuclei

whose projection to the top of the stack was at least some user-defined percentage (10-25%) free of overlap with nuclei closer to the top of the stack.

To quantify for Omb/GFP expression level, I calculated total and average pixel intensity in individual nuclear segmentation masks. The center of each nuclei was recorded as the midpoint of the maximum and minimum segmentation mask location in X, Y and Z direction.

To quantify *omb* FISH nascent transcription dots (NTDs), I start with the segmented nuclei and threshold the FISH signal to locate NTDs. Threshold level is determined manually but a range of threshold values will all gave similar results in the final NTD measurement. I identify the 3D local maxima in FISH channel within each nucleus, where the pixel with the highest fluorescent intensity above a threshold level is recorded. In cases where no maximum is found within a nucleus, a zero is recorded as the total NTD intensity for the corresponding nucleus. In all cases, the Euclidean center for the nuclei is calculated and stored along with the NTD values.

Using the local maxima as a starting location for the NTDs within each nuclei, a subset of the image enclosing the NTD in 3D is cropped out. Raw image from the FISH channel is used in the next step to avoid any preprocessing that may have affected accuracy of the intensity. I adapted a method from a recent publication by Mueller et al (Mueller et al., 2013) to quantify each NTDs in the cropped image by first fitting it to the 3D Gaussian function integrated over the voxel

$$I_{ijk} = B + A\frac{1}{x_{i,u}-x_{i,l}}\frac{1}{y_{j,u}-y_{j,l}}\frac{1}{z_{k,u}-z_{k,l}}\int_{x_{i,l}}^{x_{i,u}}\int_{y_{j,l}}^{y_{j,u}}\int_{z_{k,l}}^{z_{k,u}} G(x,y,z)\,dx\,dy\,dz .$$

Here, I_{ijk} is the modeled intensity of voxel *i,* $x_{i,l}$, $y_{j,l}$, and $z_{k,l}$ *are the lower border of the voxel,* $x_{i,u}$, $y_{j,u}$, and $z_{k,u}$ are the upper border of the voxel. A is the amplitude of the Gaussian, B is the background of the image and *G(x,y,z)* is the Gaussian function.

$$G(x,y,z) = e^{-\frac{(x-\mu_x)^2+(y-\mu_y)^2}{2\sigma_{xy}^2}} e^{-\frac{(x-\mu_z)^2}{2\sigma_z^2}},$$

where σ_{xy} and σ_z are the width of the Gaussian in xy and z, in which the Gaussian is assumed to be symmetric in the x-y direction. The center of the Gaussian in x, y and z is the terms μ_x, μ_y, and μ_z.

Matlab function "lsqcurvefit" is used to fit individual NTDs I_{image} to minimize the squared sum of residuals R

$$R = \sum_i \sum_j \sum_k (I_{image,ijk} - I_{ijk})^2 .$$

The result of the fit is an estimate of the following parameters: σ_{xy}, σ_z, μ_x, μ_y, μ_z, A and B. After each NTD is fitted, the integrated intensity under the Gaussian is calculated and recorded as the total intensity level for the NTD. Fitting results can also be selected if any of the parameters seems to be outside of a reasonable range, in cases where false positive was fit to a region with no real spots, σ_{xy} or σ_z would be much larger than size of the nucleus and a value of 0 is recorded for all parameters.

To assess the accuracy of the segmentation and the 3D Gaussian fit, 100 random nuclei from the FISH experiment images were selected and manually inspected. In cases where 3D Gaussian produced bad fits, over- or under- segmentation producing NTD that was not accurate, errors were recorded. I found an average accuracy rate of ~93% ± 1% across different images.

Using the nuclei center position, each image is broken up into stripes along the AP direction, and various statistics are calculated in Matlab to generate the final results on noise.

Image Analysis for Immunofluorescence (Antibody Staining)

All initial image analysis for various antibody stained disc starts in imageJ. Depending on the signal, one of two methods is used to acquire fluorescent intensity profile. For gradients that are sensitive to the curvature of the wing disc shape, I extract the profile from averages of cross sections in Anterior-posterior-Z section. Then manual line is draw along the curvature of the disc proper cells eliminating signals from peripodial membrane. For signals that are less sensitive to curvature, average of the z-stacks is created, sometimes only over z-stacks that contain the strongest signal. Region of interest is manually selected and fluorescent intensity profile is obtained. For size of the disc, polygon tool from imageJ is used circle the wing pouch area.

Once the intensity profile for gene of interest is saved, subsequent analysis are done in Matlab. Several Matlab function have been particularly helpful. For curve fitting, lsqcurvefit and fit can be used interchangeably in many situations. Savitszky-Golay smoothing filter, sgolayfilt, is great for smoothing fluorescent intensity profiles to preserve general shape and can be used for finding peaks, such as finding location of provein expressions from different markers.

CHAPTER 3

Noise Propagation in Spatial Patterning: From Stochastic Transcriptional Activity to Precision in a Stripe

INTRODUCTION:

Pattern formation in developing organisms has been a longstanding problem in developmental biology. To create a specific pattern in developing tissues, cells must adopt different cell fates according to their location. One way of achieving such a goal is to use gradients of secreted signaling molecules, called morphogens, which spread from a localized source and convey positional information to a field of cells (Wolpert, 1969). Cells can then interpret their positions based on the level (Green and Smith, 1990; Gurdon et al., 1994), or the dynamics (Dessaud et al., 2007; Nahmad and Stathopoulos, 2009) of the morphogen concentration to create spatial patterns. However, every step of this process involves stochastic components. Molecules move by random Brownian motion (Crick, 1970); ligand production and interactions are intrinsically stochastic (Raser and O'Shea, 2005); gene expression levels are subject to fluctuation in transcription and translation; and cell fate decisions are influenced by intrinsic cell history and extrinsic environmental conditions. These processes create noise (cell-to-cell variability) in the system, but it remains largely unknown how much noise (cell-to-cell variability) is present in every step of this process, how noise propagates through the process, and how noise affects the output (Lander, 2013).

Recent advances in quantitative experimental techniques have enabled acquisition of gene expression noise information. Most of this information concerns the

stochasticity of the transcriptional process. In single-cell organisms, such as bacteria and yeast, stochastic variation appears to be a universal feature of transcription (Raj et al., 2006; So et al., 2011; Stewart-Ornstein et al., 2012). There are two prevailing theories that attempt to explain stochastic kinetics of transcription. One view is that the process of gene regulation determines the kinetics; each promoter architecture individually influences the binding and unbinding of different elements such as histones and transcription factors (Blake et al., 2006; Carey et al., 2013; Raser and O'Shea, 2005; Sanchez et al., 2013), causing mRNA synthesis of some genes to be "Poissonian" (Larson et al., 2011; Yunger et al., 2010) while othes are "bursty" (Golding et al., 2005; Muramoto et al., 2012; Raj et al., 2006). A second view is that genome-wide (instead of gene-specific) constraints dominate variations in transcription kinetics; such biophysical and physiological constraints may include cell-cycle-dependent regulation of promoter activity (Zopf et al., 2013) or cooperative recruitment of RNA polymerases (Guptasarma, 1996). Studies using yeast *Saccharomyces cerevisiae* mostly point to gene-specific transcriptional noise for a large set of promoters (Carey et al., 2013; Hornung et al., 2012), while studies in *Escherichia coli* (Golding et al., 2005; So et al., 2011), cultured mammalian cells (Dar et al., 2012; Yunger et al., 2010), mice (Itzkovitz et al., 2012) and *Drosophila* embryos (Little et al., 2013; Paré et al., 2009) favor the presence of global constraints and "bursty" kinetics. The two theories above represent two limiting cases, and the transcription kinetics of most organisms are likely affected by a combination of gene-specific and genome-wide constraints (Sanchez and Golding, 2013).

In the development of the syncytial early *Drosophila* embryo, global constraints apparently dominate transcriptional kinetics. The maternally supplied transcription factor Bicoid (Bcd) acts as a morphogen, and triggers different gene expression patterns (Sauer et al., 1996). Its concentration gradient is remarkably reproducible and precise, varying by less than 10% among embryos (Gregor et al., 2007). On the other hand, nascent transcription activity across several patterning genes downstream of Bcd varies by about 45% for each individual locus, independent of specific promoter-enhancer architectures (Little et al., 2013). Despite the amount of transcriptional noise, these patterning genes display an overall mRNA and protein precision that is around 10%, similar to that of Bcd (Dubuis et al., 2013; Gregor et al., 2007). This precise gene expression from noisy transcription is most likely recovered by spatial and temporal averaging of the mRNA signal (Little et al., 2013).

Other morphogen systems also display qualities of precise phenotypic output including Decapentaplegic (Dpp) in the development of *Drosophila* wing disc. Dpp is a homolog of Bone Morphogenic Protein (BMP) 2/4; it acts over long ranges, creating gradients of BMP signaling that are critical to the placement of longitudinal veins L2 and L5 (Blair, 2007; Entchev et al., 2000; Teleman and Cohen, 2000b). Two downstream targets of Dpp, Spalt (splt) and Optimotor-blind (Omb), are responsible for the induction of the L2 and L5 vein primordium respectively (Cook et al., 2004; Lunde et al., 1998). However, little is known about either the spatial noise (cell-to-cell variation) in the different steps of this pattern formation system, or what possible strategies the organism employees to keep the effects of noise within tolerable limits. Here, I set out to measure

the level of noise in different parts of this system and determine if the positional information provided by the morphogen gradient is precise enough to define the positioning of a stripe of gene expression referred to as a pro-vein, the primordium of a longitudinal vein.. I start by quantifying the precision of the L5 pro-vein at the end of the larval stage, since correnctly positioning this pro-vein may be seen as a goal of the Dpp morphogen gradient. I then measure the transcriptional noise of *omb* using FISH, its protein noise using an antibody, and its half-life using RNAi to block translation. Finally, I connect how transcriptional noise of *omb* is reduced through temporal averaging by protein stability, resulting in protein expression that is just precise enough to produce the observed output.

RESULTS:

Pro-veins precision in the developing *Drosophila* wing

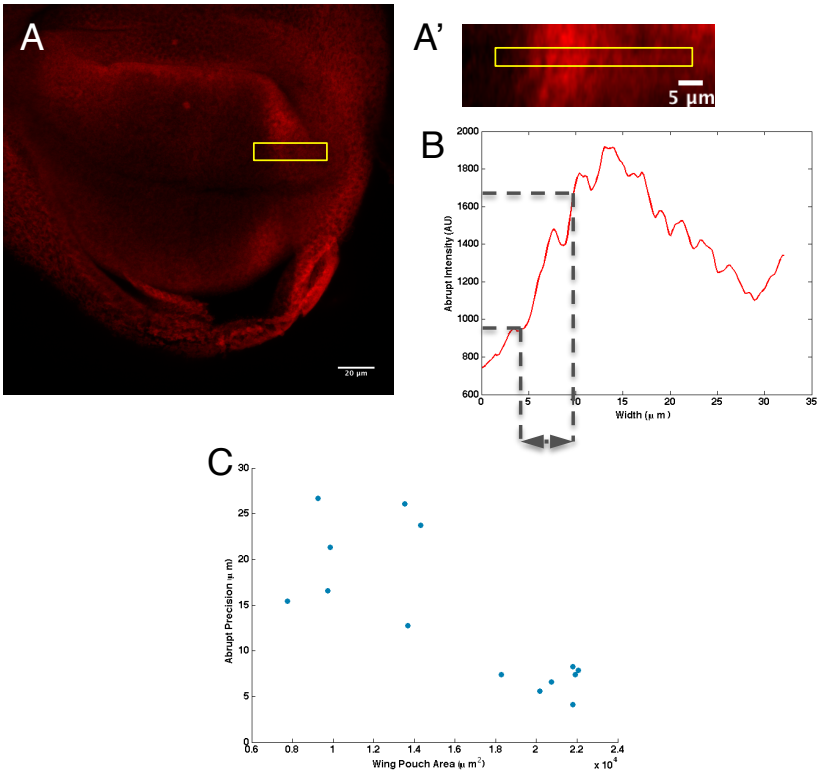
Wing vein formation in *Drosophila* is a multi-step process, starting from the interpretation of morphogen gradients to create specific gene expression patterns during the larval stage, to specification of a vein primordium domain, to vein-intervein refinement throughout pupal development (Blair, 2007; de Celis, 1997; Huppert et al., 1997). The eventual adult wing vein pattern is extremely reproducible and precise: Previous work has reported vein pattern variations of around one single cell width across different genetic backgrounds and growth temperatures (Abouchar et al., 2014). However, the level of precision in the vein primordium border before the pupal refinement process have not been quantified. To visualize the L5 vein primordium in mid

to late third instar larvae, I performed immunofluorescence using an antibody against one of the earliest known L5-specific organizing genes, *abrupt* (Cook et al., 2004). Confocal stacks were collected and the fluorescent intensity measured just below the dorsal-ventral (DV) boundary on the dorsal side (Fig 3.1 A, B). To quantify the noisiness of the Ab border (i.e. the anterior border of the domain of Ab expression), I analyzed the normalized fluorescence intensity profile and found the "transition width", where average pixel intensity is between 15 – 85% of the maximum intensity. The average transition width for Ab protein in late third instar is around 7μm±1μm, which is around 6-7 cells wide, far less than the precision observed in adult wing structures. Comparing discs from mid to late third instar, Ab expression was observed to sharpen over time. Using disc size as a measure of developmental stage, I noticed a linear decline in the transition width at the Ab expression boundary (Fig 3.1C).

Early studies of the L5 vein primordium showed that Ab is expressed along the border of the expression domains of omb and brinker (brk) and it was proposed that its establishment requires the presence of a short-range induction signal from Brk-expressing cells to neighboring omb expressing cells (Cook et al., 2004). More recent studies using anti-omb antibody showed that endogenous Omb expression is actually smoothly graded, in a nearly linear profile with no clear border (Shen et al., 2010). This discrepancy comes from the use of an omb-LacZ enhancer trap line in early studies, where homotetrametric protein beta-galactosidase is only active in its oligomeric form and will cause a sigmoidal dependence of the protein concentration (Shifrin and Steers, 1967). Although there is no clear omb border and the putative short-range induction

signal is still unidentified, observations from studies of *omb*$^{-}$ clones indicate an importance of omb expression for the establishment of the L5 primordium (Cook et al., 2004). One alternative to the current model is one in which the Omb concentration gradient sets the location of *ab*. I co-stained Omb and Ab to examine the location of L5 in relation to the Omb gradient (Fig 3.1D). Omb and Ab fluorescent intensities were obtained from the same region, plotted on the same graph and normalized to their respective maximum values. Analyzing 5 discs, I found that *ab* is located where *omb* expression is at about 30% of its maximum value (Fig 3.1E).

Figure 3.1 Precision of the L5 organizing gene, Abrupt, and its location in relation to Omb. (**A**) Wild type late 3rd instar discs immunostained with Abrupt antibody (red). Abrupt is a L5 specific organizing gene expressed in the late 3rd instar as a stripe near the lateral posterior end of the wing disc. (**A'**) is the averaged orthogonal view in the anterior-posterior-Z direction of the yellow-boxed region in (A). (**B**) A fluorescent intensity profile of Abrupt expression is taken from the orthogonal view, yellow box in (A'). (**C**) The transition width of Abrupt in μm is plotted against the overall wing pouch area for individual quantified discs. Abrupt's transition width is measured as the distance of 15-85% of the maximum fluorescent intensity in the expression border facing the anterior direction. (**D**) Wild type expression of Omb (visualized with antibody, rabbit-Omb, red) and Delta (antibody, green). Delta marks several provein locations including L5. (**E**) Normalized fluorescent intensity profiles of Omb and Delta from the blue box in (**D**) show that Omb (red) declines gradually and Delta is expressed at around 30% of Omb's maximum intensity.


A
A’
5 µm
B
Abrupt Intensity (AU)
Width (µ m)
C
Abrupt Precision (µ m)
Wing Pouch Area (µ m²)
x 10⁴
D
D’
D’’
Omb
Delta
E
Omb
Delta
Normalized Fluorescent Intensity (AU
Distance along A/P (µ m)

Measuring Transcriptional Activity of Omb in the Late Third Instar Wing disc

Given that *ab* expression is reasonably precise, with a transition width of around 7µm, one might expect *omb* expression to display similar level of spatial variance. I directly measured variation in the transcriptional activity of *omb* by combining FISH, immunofluorescence, and image analysis (Fig 3.2). FISH enables transcripts or collections of transcripts to be visualized as distinct fluorescent dots in cells, and variations of FISH have been used to quantify nascent transcriptional activity in the *Drosophila* embryo (Kosman et al., 2004; Paré et al., 2009; Xu et al., 2015). Most studies use probes complementary to the open reading frame of interested genes. Nascent transcription events are distinguished from spliced mRNAs by virtue of location (nascent events are in the nucleus) and signal intensity (multiple nascent transcripts are usually on a single gene at the same time). A more specific method to quantify nascent transcriptional activity is to use probes complementary to intronic regions. This will limit the signal to transcripts in the nuclei that are in the act of being elongated, given that splicing is usually co-transcriptional, with introns being released as soon as transcription reaches the next exon (Brody et al., 2011; Wada et al., 2009). I refer to these FISH signals as “nuclear transcription dots” (NTDs), the intensity of which quantifies the total number of transcripts undergoing elongation, and provides a quantitative estimate of transcriptional rate.

The variability in cumulative fluorescence intensity from NTDs represents two sources of noise: the actual variability in the amount of transcripts, and noise introduced by the measurement process itself. Many recent studies have tried to limit the amount of

noise introduced from the experimental process by using DNA oligonucleotides that are directly labeled with fluorescent molecules (Little et al., 2013; Little et al., 2011; Raj et al., 2008). I took a different approach in the experimental design; rather than trying to eliminate the measurement noise before image acquisition, I designed a "two probes setup" so that measurement noise could be subtracted after quantification.

To estimate and subtract the measurement noise from the actual noise of *omb* transcription, I designed two probes of equal length (550bp) that are located only 50bp apart, both targeted against the 5' end of the 2nd to last intron (Fig 3.2A). These two probes are labeled with two different haptens, Digoxigenin (DIG) and dinitrophenol (DNP), and detected separately with Alexa Fluor® Tyramide signal amplification (See Materials and Methods). Since these two probes are located very close to one another, they are expected to bind to the same transcripts in the same cell. I assume that the true signal from both probes is essentially the same, so that the only difference between what I detect using them must represent measurement noise. Because measurement noise should be independent for the two probes, measurements may be treated as independent random variables with a common distribution.

If I let S_1 and S_2 stand for the measured signals from the two FISH probes, B for the true biological signal, and Δ_1 and Δ_2 for the measurement noise that comes from two independent common distributions, then

$$\begin{aligned} S_1 &= B + \Delta_1 \\ S_2 &= B + \Delta_2 \end{aligned} \qquad [1]$$

Since variances add, the measured variance of S is just the sum of the variance from B and its respective Δ.

$$Var[S] = Var[B] + Var[\Delta] \quad [2]$$

The difference between the measured signals in a single cell, $Q = S_1 - S_2$, is related to the measurement noise by $Q = \Delta_1 - \Delta_2$, and the variance of Q is the sum of the variances from the Δs (Bland and Altman, 1999; Bland and Altman, 2007).

$$Var[Q] = Var[\Delta_1] + Var[\Delta_2] \quad [3]$$

To get the variance of the true signal, I set

$$Var[B] = Var[S_1] - Var[\Delta_1], \quad [4]$$

and substitute in equation [3] for Var[Δ_1] to get

$$Var[B] = Var[S_1] - Var[Q] + Var[\Delta_2]. \quad [5]$$

which can be rewritten with equation [2] to get

$$2Var[B] = Var[S_1] + Var[S_2] - Var[Q]. \quad [6]$$

Therefore, the variance of the true signal is simply

$$Var[B] = \frac{Var[S_1] + Var[S_2] - Var[Q]}{2}. \quad [7]$$

To minimize off target signals from outside the nucleus, I coupled FISH with antibody staining (immunofluorescence) for *Drosophila* nuclear lamin, a major structural protein of the nuclear lamina, which lines the nucleoplasmic surface of the inner nuclear membrane (Riemer et al., 1995). Similar to a recent publication of labeling protocols in *Drosophila* embryo (Xu et al., 2015), I also found that performing immunofluorescence labeling after the FISH protocol gave the best results in preserving signals from both protein and RNA detection. All discs were mounted between a bridge to preserve the structural integrity of the disc, and images were then acquired with a scanning confocal

microscope. I found that the nucleus of the pseudostratified columnar epithelial cells of the disc are located apically, and appear in multiple staggered layers. *Omb* FISH signals appear as dots of various sizes in the nuclei, and can be seen in most (but not all) cells in both channels. Unlike the embryo (Little et al., 2013; Wilkie et al., 1999), I detect only one single NTD within each nucleus, a sign that homologous chromosome pairs remain in close proximity during transcription.

To quantify the *omb* NTD signals, I used an active contour algorithm started from a 3D segmentation of nucleus based on the Lamin stain (see Materials and Methods). This step served two purposes: it helped to eliminate any nonspecific signals that may appear in the cytoplasm, and it identified all nuclei in the wing pouch even if no NTD was detected within them. Once each nucleus had been segmented, I identified the presence and location of an NTD by thresholding the nuclear mRNA fluorescence intensity. Using the identified location, I fitted the NTD to a 3D Gaussian, and then calculated the integrated intensity under the Gaussian from the result of the fit (Mueller et al., 2013). This gave a single measure of total fluorescence for the NTD, whereas simple methods of NTD quantification using maximum intensities or average intensities over a set number of z-slices ignored the 3D extent of the transcription site (Boireau et al., 2007). This method of using a fitted Gaussian accounted for signals in the Z direction (Mueller et al., 2013).

To assess the accuracy of the segmentation and ability to correctly detect a FISH dot using this method, 100 nuclei were randomly selected for manual inspection. A sub-region of the z-stack image from a single segmented nucleus was displayed for all

channels; the maximum projected image of the mRNA signal was plotted along with its 3D Gaussian fit and residuals from the fit for both channels (Fig 3.2C). I found that this method had an accuracy of about 93 ± 1% across all sampled images to correctly identify and fit NTDs. Based on the result of the 3D Gaussian fit parameters, I further eliminated any false positives if their σ values were much larger than a fraction of the nucleus (NTD can not be physically bigger than size of the nucleus). Grossly misfit signals that had a ratio of the sum of the residuals to the original intensity higher than 30% were also detected and removed. Both of the above removals happened less than 0.1% of the time and showed that the 3D Gaussian fit works efficiently. Furthermore, I found a minimum of 78% of the NTDs detected appeared in both channels, indicating that ~88% of the signal was detected in at least one channel. Consistent with expectations, the probe that was located closer to the starting point of the intron had a slightly larger mean integrated NTD intensity and was also more likely to be detected in only one of the channels.

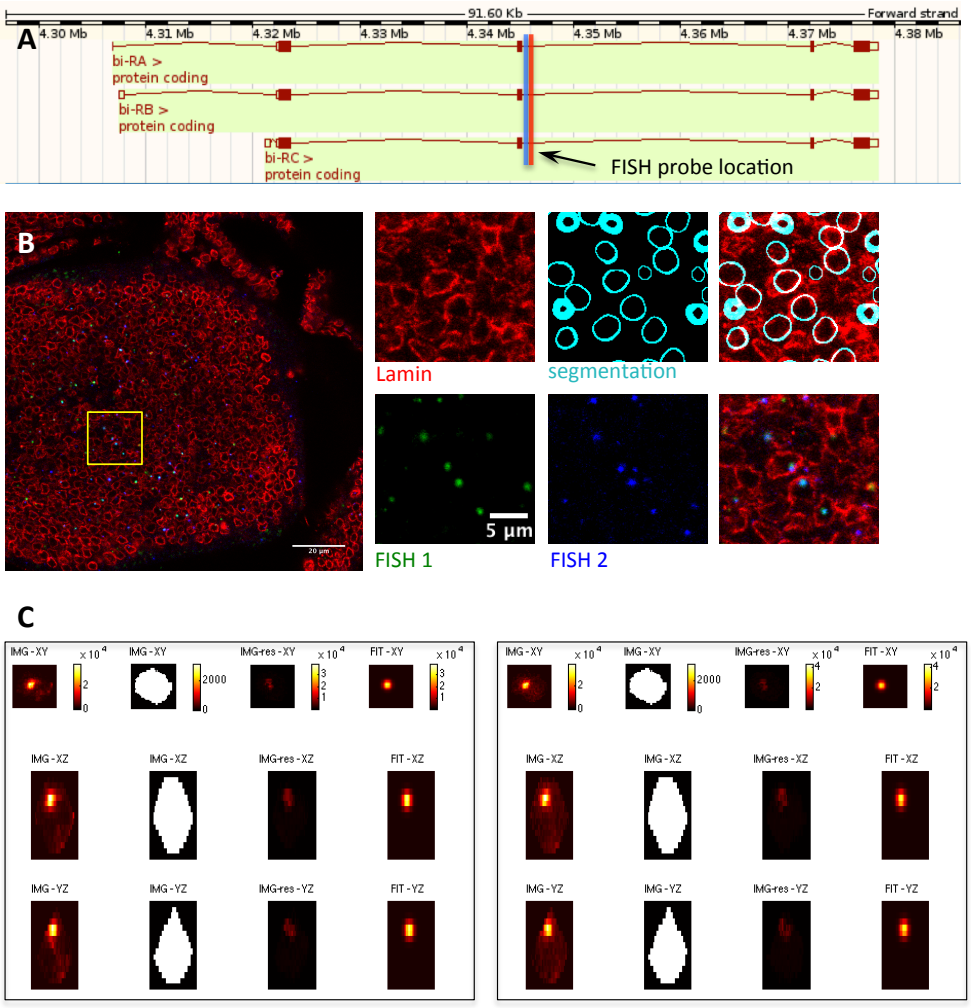


Figure 3.2 *omb* FISH experiment detected nascent transcription dots (NTDs). (**A**) Two FISH probes (red and blue bar) were designed to target the 2nd to last (also longest) intron of *omb*, closest to the 5' end. (**B**) A single z-section of FISH showing detected lamin (red, marking nuclear membrane), FISH signal 1 (green), FISH signal 2 (blue), and computer segmented nuclear membrane. (**C**) 3D Gaussian fit of 2 FISH signal in the same nucleus. Within the first box on the left, the left most column shows the original FISH signal in a projected image for one segmented nucleus. The second column is the 3D segmented cell. The last column is the 3D gauss fit of the dot and third column shows residual of fit. The box on the right is a similar box except in the 2nd FISH signal channel.

Variation in *omb* Transcriptional Activity is Relatively High

To quantify the fluctuation in *omb* transcriptional activity in the wing disc, I focused on the posterior half of the disc. While I cannot label any proteins to mark the AP boundary, due to limited availability of non-overlapping fluorescent markers, I decided to use *omb* expression as a guide to locate the AP boundary. Each image was divided along the AP direction into 50 equally spaced regions and mean integrated NTD intensities within each region were calculated to find where *omb* expression was highest. Using this location of high *omb* expression as a proxy for the AP boundary, I then quantified the integrated NTD intensities in the posterior half of the disc in 10 equal regions (Fig 3.3). Normalizing each individual profile to its maximum expression, and averaging over 4 individual discs; I found that the mean transcriptional activity of *omb* strongly resembles that of its protein counterpart (Shen et al., 2010). *Omb* transcriptional activity is high in the center and declines smoothly toward the posterior lateral margins of the pouch, in almost a linear fashion in both channels of NTDs (Fig 3.3).

The observed noise in *omb* transcriptional activity from each probe can be quantified in terms of the coefficient of variation (standard deviation over mean, CV). Here, the observed CV for *omb* transcriptional activity starts around 0.8-0.9 and gradually increases toward the posterior margins of the pouch, reaching over 1.3 for both probes. Using equation [7] above, I can calculate the true variance of the transcriptional activity for each probe within each region by accounting for the additional component of measurement noise, and re-calculate the CV to obtain an estimate of the

actual noise due to transcriptional activity. With the new calculation, I notice an overall drop in noise across the entire posterior region, decreasing CV by about 0.1 – 0.3 (Fig 3.3). Even with this corrected value, the amount of transcriptional noise, which goes from slightly above 0.7 to roughly 1.2, is much higher than what has been reported in the embryo, where the CV for *hb* transcriptional activity was found to be about 0.22 ± 0.03 (Little et al., 2013).

To be certain that extra variability is not being artifacturally introduced as a result of fluorescent attenuation along the z-axis of the image, I also quantified *omb* NTDs using only the "top layer" of nuclei present. Top layer nuclei refer to nuclei whose projection to the top of the stack is at least 80% free of overlap from any nuclei located above them, thus decreasing the effect of fluorescent attenuation due to amount of tissue emission light has to traverse (Rigaut and Vassy, 1991). Since the wing disc is not a completely flat epithelium and the mounting method I used preserves the curvature of the pouch around the lateral region, using this "top layered" approach permits signals that are deeper in absolute distance to the objective to be included. This way, I am able to capture all relevant nuclei and quantify their signals across the entire region of the disc. I observed similar mean transcriptional activity from these "top layer" nuclei, in which transcriptional activity decreases gradually from the medial to the posterior end of the disc in both channels (Fig 3.4A, B). Additionally, I found that transcriptional noise from both the observed and the corrected signal increased slightly compare to that of the entire z-stack. If the amount of noise measured from the entire z-stack was in part due to fluorescent attenuation, by using the top layer method, I would

expect to observe a decrease in the noise. The fact that I did not see a decrease, but rather an increase in noise shows that fluorescent attenuation is not a major contributor to the measured *omb* transcriptional noise. Furthermore, I tested to see if the slight increase in noise is unique to the top layer of the nuclei or just the result of decrease in sampling, in which each region had roughly 45-65 sample points. I randomly select 50 nuclei from each region and plotted the variation. Again, similar trends can be seen (Fig 3.4C), but the increase in noise from sampling just 50 nuclei is less than the additional noise observed when I quantified just the top layer.

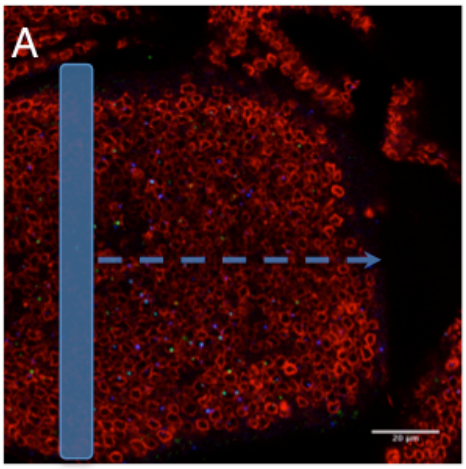
Considering that *omb* transcription activity steadily declines in the posterior half of the disc, even cells in the same rough region may not have the same exact positional values. If any of the regions into which cells were binned in these studies spanned a significant range of *omb* expression, the observed and corrected noise level could be higher than the true noise. I thus further subdivided the posterior half of the wing disc into smaller bins to confirm that observed variability in *omb* transcriptional activity was not affected. Repeating the quantification using 20 spatial regions, instead of 10 (Fig. 3.5) I find the observed and correct transcription signal to be roughly the same as previously, for similar positions on the disc. There are some regions with lowered noise from the first signal, but there is more region-to-region discontinuity. From the normalized mean *omb* transcriptional activity, I can see that the very last region closest to the lateral end of the disc has a slightly lowered value than previously. Interestingly, these results show that omb expression extends very close to the wing pouch border, i.e. *omb* is not completely suppressed by *brk*.

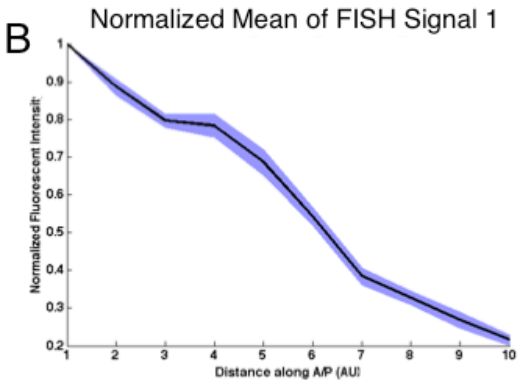
To understand what the observed level of noise represents, I must understand what factors go into creating each integrated NTD signal. Because each NTD represents a set of unfinished transcripts present at the same time on the *omb* gene, the expected signal intensity should depend upon the rate of transcriptional initiation, and the rate at which transcripts bearing the probed intron are spliced. Because transcripts are only detected once RNA polymerase reaches the sequence to which a probe binds, and become undetectable once they are spliced, the expected number of transcripts detected on the gene should equal the number of transcriptional initiations during a time interval equal to the time it takes for polymerase to travel from the probe sequence to the end of the intron, where co-translational splicing should occur (Brody et al., 2011; Pandya-Jones and Black, 2009).. In effect, the FISH procedure provides a time-average of a noisy signal, where the averaging time is dependent on the location of the probe-homologous sequence on the DNA. Probes located near the start (5'-end) of an intron reflect a longer averaging time, than probes located near the end (3'-end) (Zenklusen et al., 2008).

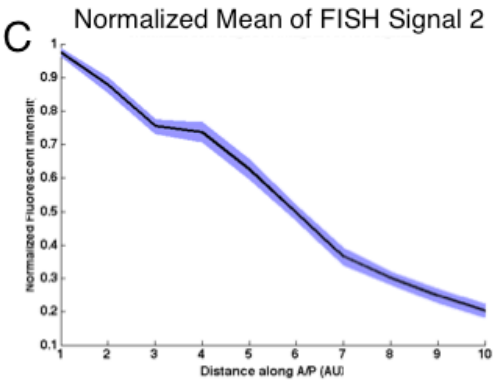
Since the rate of transcription elongation in *Drosophila* is ~1.1-1.5 kilobases/min (Ardehali and Lis, 2009; Garcia et al., 2013), and the probes designed for *omb* here are 27kb upstream of the next exon, I estimate that these two probes report transcriptional noise that is averaged over roughly 18-24 minutes. Based on this, I conclude that *omb* transcriptional noise in the wing disc is remarkably high throughout the disc. For example, near the Dpp production region, where Omb expression is highest due to very little repression from *brk*, I still measure a coefficient of variation of about 70%. In lateral

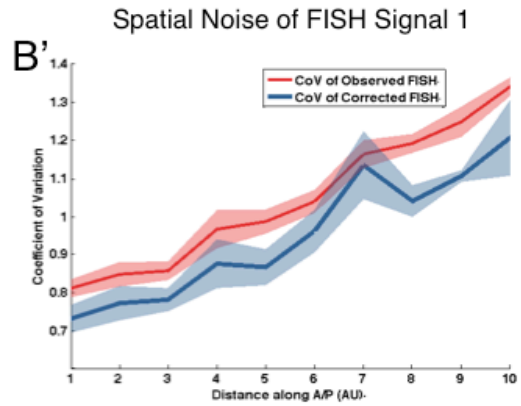
parts of the disc, where Omb is involved with induction of vein primordium genes when its expression drops to about 30% of maximum (Fig. 3.1E), the measured coefficient of variation is around 1-1.1 (this corresponds to region 8-9 in the FISH experiments; Fig. 3.3B'-C'). A coefficient of variation of 1 basically means that the standard deviation is as high as the mean, and would translate to a transition width that is twice that of the local length scale ($2cv \times \lambda$). What is surprising is to observe this level of noise after a transcriptional averaging time of ~18-24 minutes: In the *Drosophila* embryo, the transcriptional noise of *hb* (the gene for which is 2 kbp in length) is ~22% with an averaging time of only 2-3 minutes (Little et al., 2013). If initiation events were independently distributed in time (so that coefficients of variation should fall with the square root of the integration time), one could calculate the amount of noise that would be measured in these two systems using a common integration time. By that procedure, the noise in *omb* would exceed that of *hb* by a factor of >10.

Figure 3.3 Mean and Noise of *omb* transcriptional activity. (**A**) Single slice of *omb* FISH image; the posterior half of the disc is divided into 10 equal regions represented by the blue bar. The blue bar illustrates starting location and size of region that each point represents in (**B-C**). (**B**) Normalized mean transcriptional activity per nucleus as a function of position along the AP direction is given by the black line. Results from one of the FISH signal, blue shaded region is the standard error (**B'**) Mean transcriptional noise is quantified as integrated 3D Gaussians, plotted as a function of the coefficient of variation vs. the position along the AP direction. The red line and shaded region represent observed noise from the NTDs and its standard error (Fig. 3.2C). The blue line and shaded region represent the true transcriptional noise level after correction (Materials and methods). (**C**) and (**C'**) is the normalized mean and transcriptional noise results of FISH signal 2, plotted the same way as in (**B**).


A


B
Normalized Mean of FISH Signal 1
Normalized Fluorescent Intensity
Distance along A/P (AU)


C
Normalized Mean of FISH Signal 2
Normalized Fluorescent Intensity
Distance along A/P (AU)


B'
Spatial Noise of FISH Signal 1
CoV of Observed FISH
CoV of Corrected FISH
Coefficient of Variation
Distance along A/P (AU)

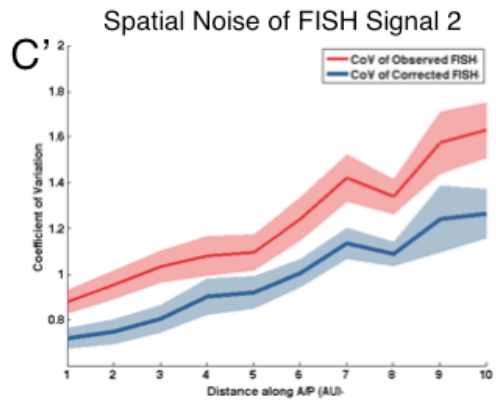
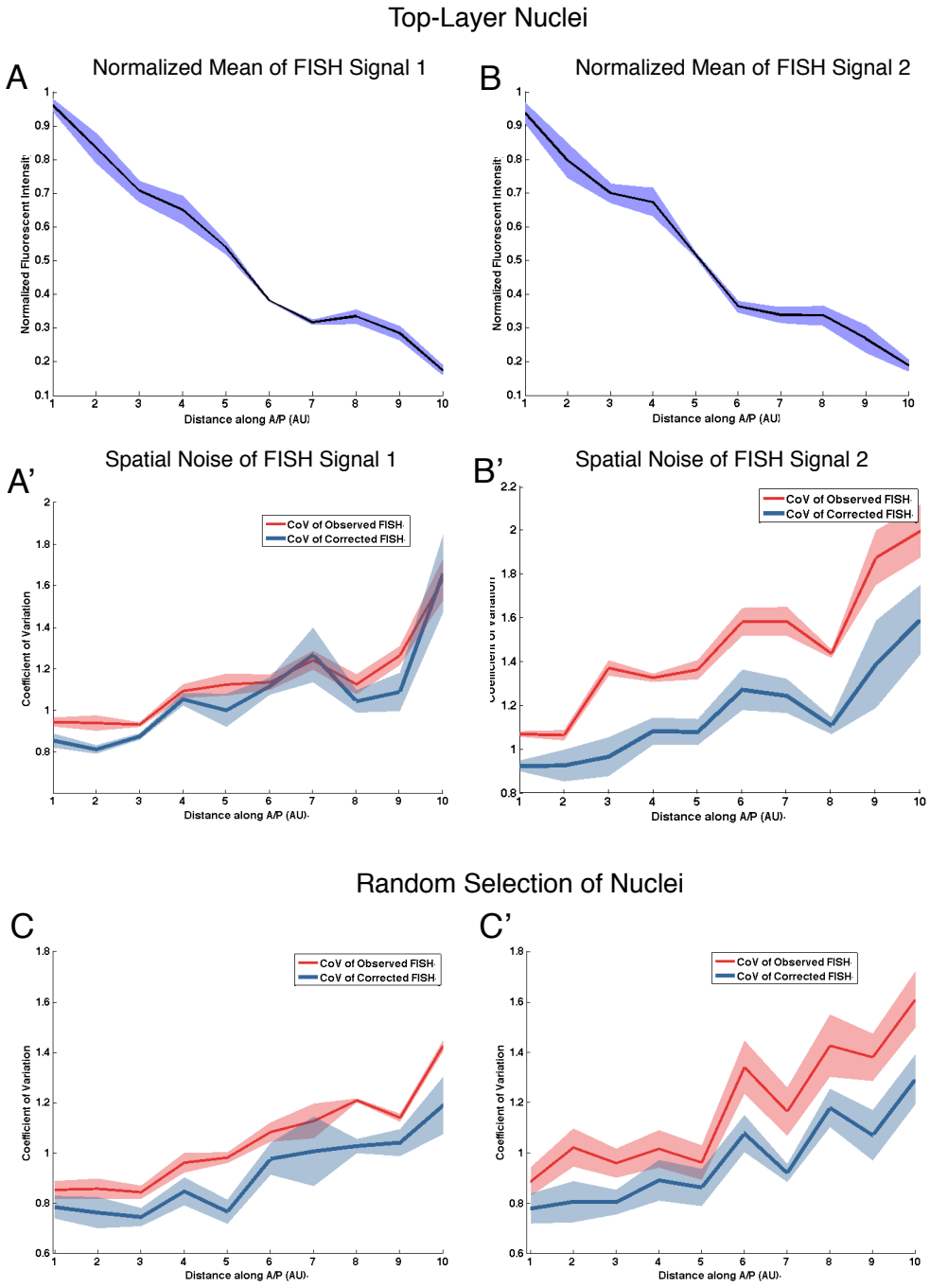

C'
Spatial Noise of FISH Signal 2
CoV of Observed FISH
CoV of Corrected FISH
Coefficient of Variation
Distance along A/P (AU)

Figure 3.4 Mean and Noise of *omb* transcriptional activity in "top layered" nuclei only. Results are calculated with the same method except only with "top layered" nuclei that are at least 80% free of overlap. (**A**) Normalized mean transcriptional activity per nucleus as a function of position along the AP direction from FISH signal 1. (**A'**) Mean transcriptional noise plotted as function of coefficient of variation as function of position along the AP direction. The red line and shaded region represent observed noise from the NTDs and its standard error (Fig. 3.2C). The blue line and shaded region represent the true transcriptional noise level after correction (Materials and methods). (**B**) and (**B'**) are the normalized mean and transcriptional noise results of FISH signal 2, plotted the same way as in (**A**). (**C**) Transcriptional noise of FISH signal 1 and (**C'**) FISH signal 2 from 50 randomly selected nuclei within each region along the AP direction.


Top-Layer Nuclei
A
Normalized Mean of FISH Signal 1
Normalized Fluorescent Intensit
Distance along A/P (AU)
B
Normalized Mean of FISH Signal 2
Normalized Fluorescent Intensit
Distance along A/P (AU)
A'
Spatial Noise of FISH Signal 1
CoV of Observed FISH
CoV of Corrected FISH
Coefficient of Variation
Distance along A/P (AU)
B'
Spatial Noise of FISH Signal 2
CoV of Observed FISH
CoV of Corrected FISH
Distance along A/P (AU)
Random Selection of Nuclei
C
CoV of Observed FISH
CoV of Corrected FISH
Coefficient of Variation
Distance along A/P (AU)
C'
CoV of Observed FISH
CoV of Corrected FISH
Coefficient of Variation
Distance along A/P (AU)

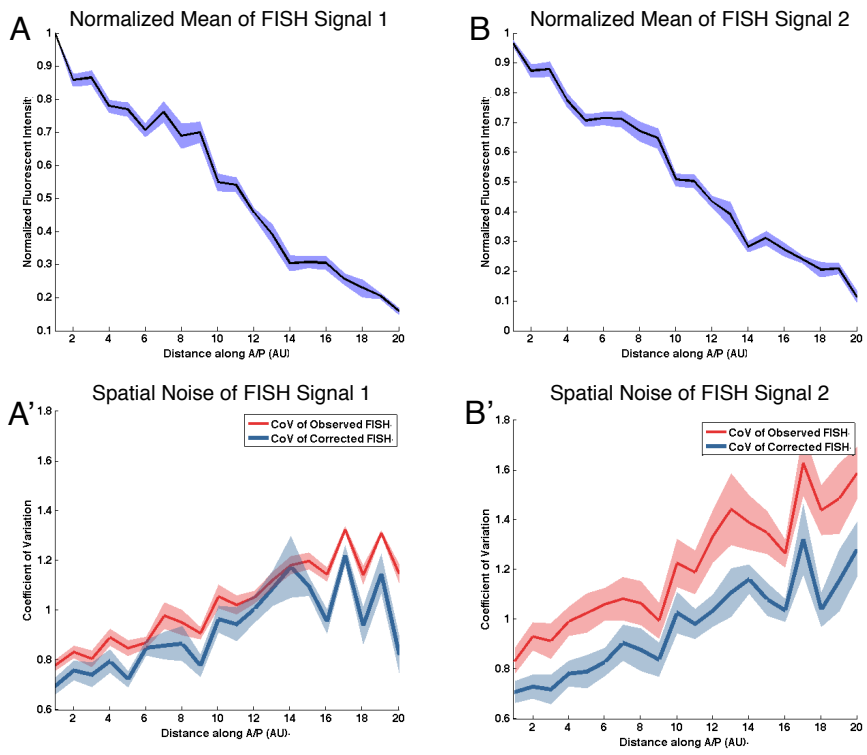


Figure 3.5 Mean and Noise of *omb* transcriptional activity with 20 equally spaced regions. Results are calculated with the same method as in Figure 3.3 except the disc is divided into 20 equal spaced regions instead of 10. Each region is now half the width of the region in Figure 3.3. (**A**) Normalized mean transcriptional activity per nucleus as a function of position along the AP direction from FISH signal 1. (**A'**) Mean transcriptional noise plotted as function of coefficient of variation as function of position along the AP direction. The red line and shaded region represent observed noise from the NTDs and its standard error (Fig. 3.2C). Blue line and shaded region represent the true transcriptional noise level after correction (Materials and methods). (**B**) and (**B'**) are the normalized mean and transcriptional noise results of FISH signal 2, plotted the same way as in (**A**).

Measuring Omb Protein Noise

Although the measured *omb* transcriptional noise is high, it is the Omb protein noise that will eventually affect its downstream targets. I quantitatively measured the protein noise of Omb by immunofluorescence, using an antibody specific to Omb (Shen et al., 2010), coupled with image segmentation (Fig. 3.6). Omb is a T-box transcription factor and I observed from confocal stacks that the detected signal is mostly inside the nucleus. To quantify Omb expression levels, I again started with a 3D segmentation of a nucleus from the Lamin stain, using the same active contour algorithm as in FISH experiments. This allowed me to focus on signals inside the nucleus and eliminate any nonspecific fluorescence in the cytoplasm. I assayed both the Omb concentration (average pixel intensity) and total Omb content (total pixel intensity) within each nucleus, and stored their values together with nucleus location. Again, just as in the FISH experiments, I divided the posterior half of the disc into 10 equal regions to examine distribution of Omb expression in space.

Normalizing each individual profile to its maximum expression and averaging over 6 individual wing discs: I found similar results using either Omb concentration or Omb content as a readout of Omb expression. In agreement with previous finding (Shen et al., 2010), Omb expression is graded, high in the center near Dpp production region and declining smoothly toward the lateral margins of the pouch (Fig. 3.6B,D). Overall, the observed noise for Omb protein expression was lower than observed for its mRNA; starting around 0.4 and staying relatively flat for half the posterior compartment, and gradually increasing after that to about 0.7 (Fig. 3.6C,E).

Similar to the noise components in NTD intensities, the observed Omb expression should also contain measurement noise. Ideally, I would use a similar method as in the FISH experiment, simultaneous detecting two Omb specific antibodies that target different parts of the Omb to subtract out experimental noise. However, since there is only one Omb-specific antibody available, I had to estimate the amount of experimental noise associated with immunofluorescence staining using a different approach. I assumed if the same immunostaining protocol was performed for a different protein, for which I did have access to two independent markers of expression level, and images were acquired with the same microscope settings, then the calculated measurement noise associated with detecting that protein would provide a reasonable estimate of the measurement noise associated with detected Omb.

I thus turned to a UAS-GFPN-LacZ *Drosophila* line, in which the transgene GFPN-LacZ encodes a nuclear localized GFP that is fused to bacterial β-galactosidase, that is expressed under the control of and Upstream Activating Sequence (UAS) (Shiga et al., 1996). Because GFP is fluorescent, and thus is no additional step involved in its detection and visualization, I may consider its intensity to be a “true” measure of the the amount of fusion protein. I used an Apterous-Gal4 driver to drive the expression of GFPN-LacZ in the dorsal half of wing discs, and then detected LacZ portion of the fusion protein using a specific anti β-galactosidase antibody, using the same procedures as used for Omb staining. To compare the detected (LacZ) signal and the true signal (GFP), I first segmented each nucleus in 3D with the same algorithm as above but using

DAPI as the boundary of each nucleus. Next, I assayed the GFP concentration and LacZ concentration for each individual nucleus and examined their relationship.

To estimate the experimental noise, I performed a simple linear regression of LacZ concentration against GFP concentration (Fig. 3.7B). Assuming correlation is linear, the regressed model should represent the true LacZ signal, $LacZ_{true} = \alpha + \beta GFP$. The observed LacZ fluorescent intensity can be defined as $LacZ_{obs} = LacZ_{true} + \Delta$, in which Δ is the experimental noise. Here, Δ can then be calculated as the residual of the observed LacZ value from the predicted value, and is equivalent of the difference between a random variable X from its mean. Therefore, the Var(X), or Var(Δ), is equal to SS_{res}/n (residual sum of squares divided by number of samples). If there is no experimental noise in the system, then LacZ concentration should lie perfectly with GFP concentration on a line and the variance of LacZ should be 0.

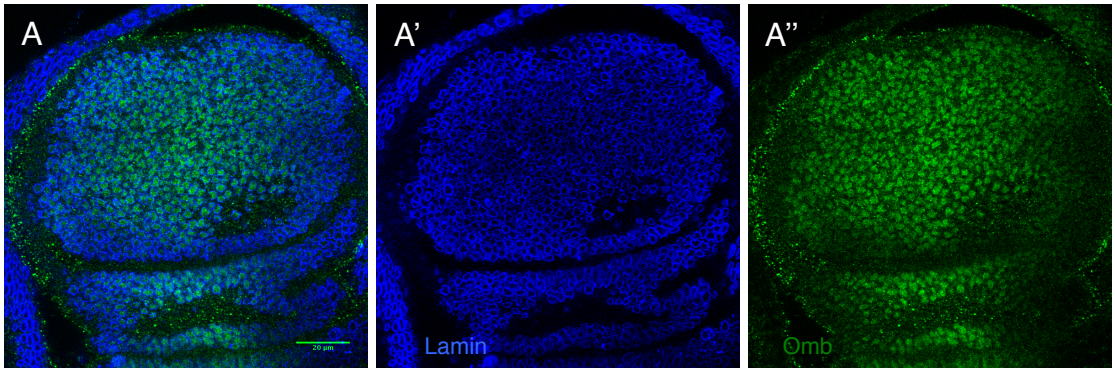
As there is no guarantee that the variance of Δ should remain constant for different levels of protein concentration. I divided the GFP-LacZ signal pairs based on their GFP value into 16 equal numbered parts and calculated variance of Δ for each sub-division. I found that the variance of experimental noise positively correlates with mean fluorescent intensity in fairly linear manner (Fig. 3.7C), as one might expect for a Poissonian process. To quantify this relationship between variance of experimental noise and mean fluorescent intensity, I fit the data with both a linear and a parabolic function. Using such functions, I can then interpolate the amount of variance for different levels of fluorescent intensity.

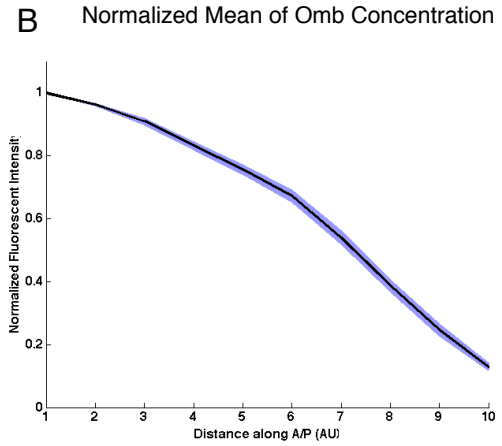
The observed Omb protein signal can be defined as $Omb_{obs} = Omb_{true} + \Delta$. Since I assumed variance of Δ (experimental noise) for similar levels of fluorescent intensity across different experiments to be the same, I can calculate the variance of Δ for different observed Omb intensity using the both (linear and parabolic) fitted relationships between Var(Δ) and mean fluorescent intensity from the GFP-LacZ experiment. Next, I calculated the variance of Omb_{true}, $Var(Omb_{true}) = Var(Omb_{obs}) - Var(\Delta)$, and plotted CV as a function of AP position to represent protein noise (Fig. 3.8).

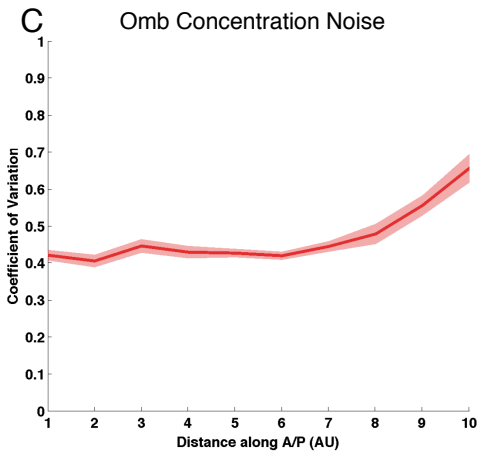
I found that the true Omb protein noise is associated with a CV about 0.2 lower than the observed Omb noise, throughout the wing disc. In region 8-9, where Omb protein plays an important role for induction of L5 organizing gene, I calculated that true Omb varies by about 35% (Fig. 3.8). The local length scale of Omb is the inverse of the relative derivative. It can also be numerically estimated from the location of interest as the by calculating the decay rate from local differences. Since the data here is in discrete units of regions, and C is the relative level at every region, $-\delta C/C$ is the relative rate of change per region. We can estimate the number of regions needed for C to decline by e^{-1} by dividing $-\delta C/C$ from e^{-1}, and then converting region to absolute distance. Therefore, $\lambda = \frac{-C}{\partial C} e^{-1} \Delta_x$ estimate the local length scale where Δ_x is the unit of distance per region, and λ represents the distance over which the gradient decreases by ~ 63% ($1 - e^{-1}$) (Lander et al., 2009). While the Omb protein gradient gradually declines almost linearly, the Omb length scale, λ, decreases in an exponential fashion. The calculated length scale in our region of interest is ~8μm. From these numbers, I find that the theoretical limit on how precise of a boundary the Omb gradient, with its

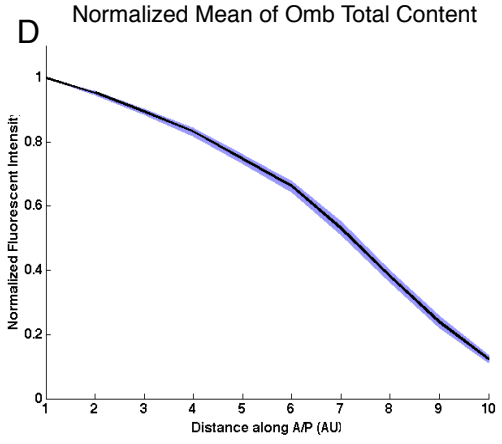
observed level of noise, can induce is around 5.6μm (2cv × λ). This transition width is just a bit smaller than the measured transition width of the Abrupt border at 7μm, and shows that the Omb protein gradient is precise enough for the purpose of pattern formation.

Figure 3.6 Mean and noise of Omb protein expression. (**A**) Immunostaining of Omb (green) and Lamin (blue). Posterior half of the disc is evenly divided into 10 regions in which the mean and noise of Omb protein expression is plotted in (**B-D**). (**B**) Normalized mean of Omb concentration (average fluorescent intensity per pixel) per nucleus as a function of position along the AP direction. Shaded blue represents the standard error around the mean. (**C**) Noise of Omb concentration quantified as CV, plotted as a function of position. Shaded red represents the standard error around the mean. (**D**) Normalized mean of total Omb content (sum of fluorescent intensity for all pixels) per nucleus as a function of position along the AP direction. (**E**) Noise of Omb content quantified as CV, plotted as a function of position along the AP direction.


A
A'
A''
Lamin
Omb


B
Normalized Mean of Omb Concentration
Normalized Fluorescent Intensit
Distance along A/P (AU)


C
Omb Concentration Noise
Coefficient of Variation
Distance along A/P (AU)


D
Normalized Mean of Omb Total Content
Normalized Fluorescent Intensit
Distance along A/P (AU)

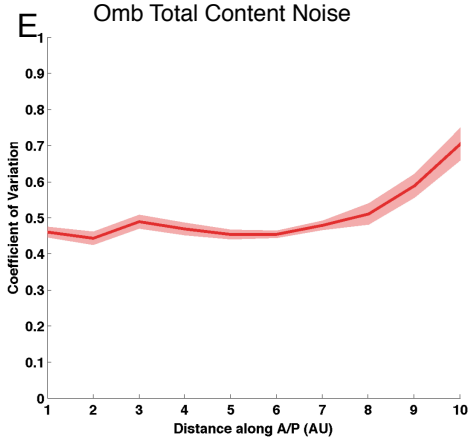
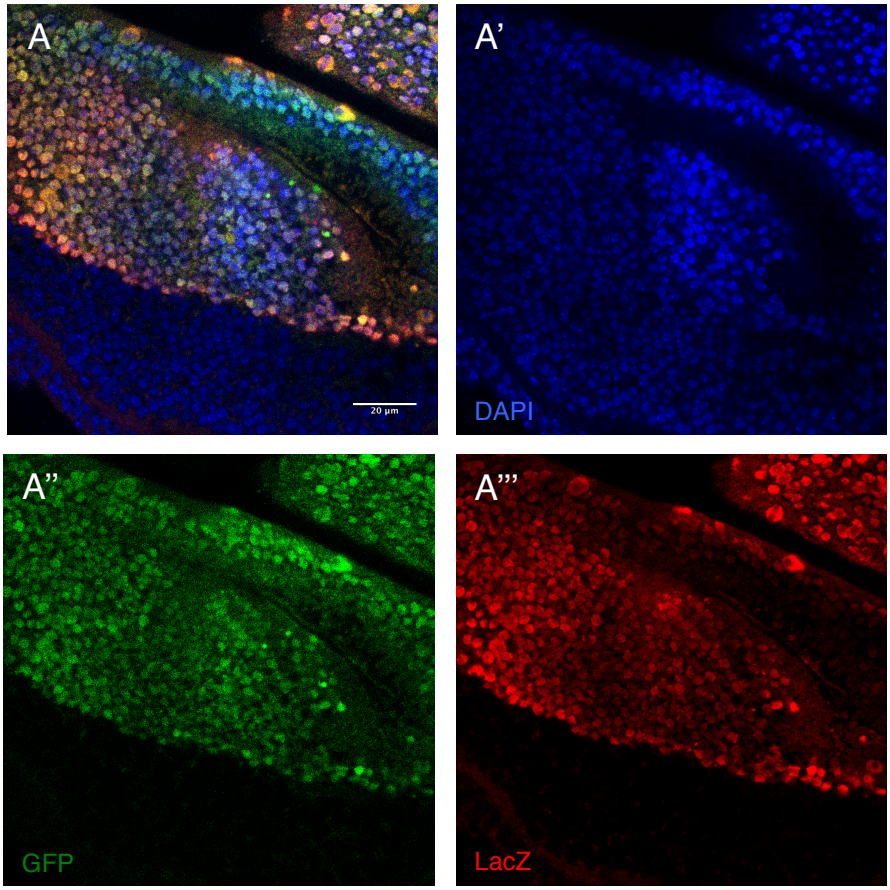
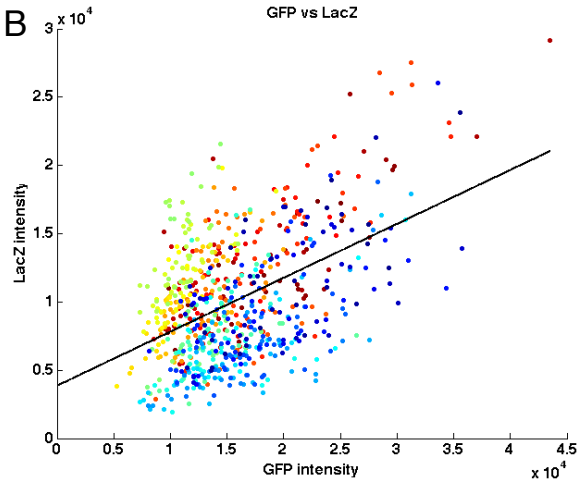
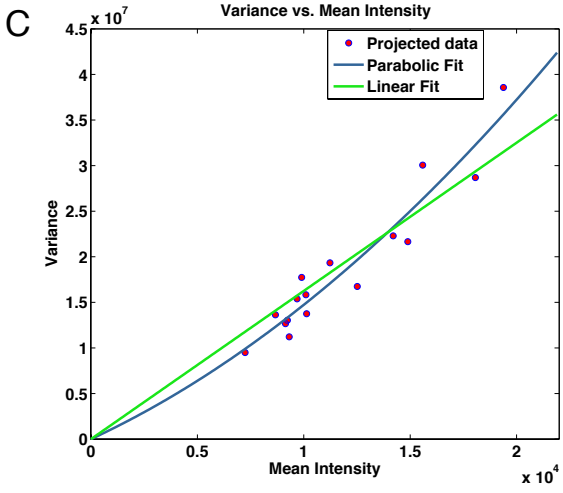

E
Omb Total Content Noise
Coefficient of Variation
Distance along A/P (AU)

Figure 3.7 Estimate of experimental noise introduced by immunostaining. (**A**) Late 3rd instar disc of *Ap-Gal4 > UAS-GFPN-LacZ*, imaged with GFP (green), DAPI (blue) and LacZ (red, immunostained with an antibody). (**B**) Scatter plot of GFP concentration (average fluorescent intensity per pixel) vs. corresponding LacZ concentration per nucleus. Each colored dot represent one pair of signals from the same nucleus. Intensities of GFP and LacZ positively correlated with a linear fit (black). (**C**) Variance of LacZ measurement error increases with level of GFP intensity, and their relationship is fitted with a linear function that goes through 0 (green), and a parabolic function that also goes through 0 (blue).


A
20 µm
A'
DAPI
A''
GFP
A'''
LacZ


B
GFP vs LacZ
x 10^4
LacZ intensity
GFP intensity
x 10^4
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5


C
Variance vs. Mean Intensity
x 10^7
Projected data
Parabolic Fit
Linear Fit
Variance
Mean Intensity
x 10^4
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5

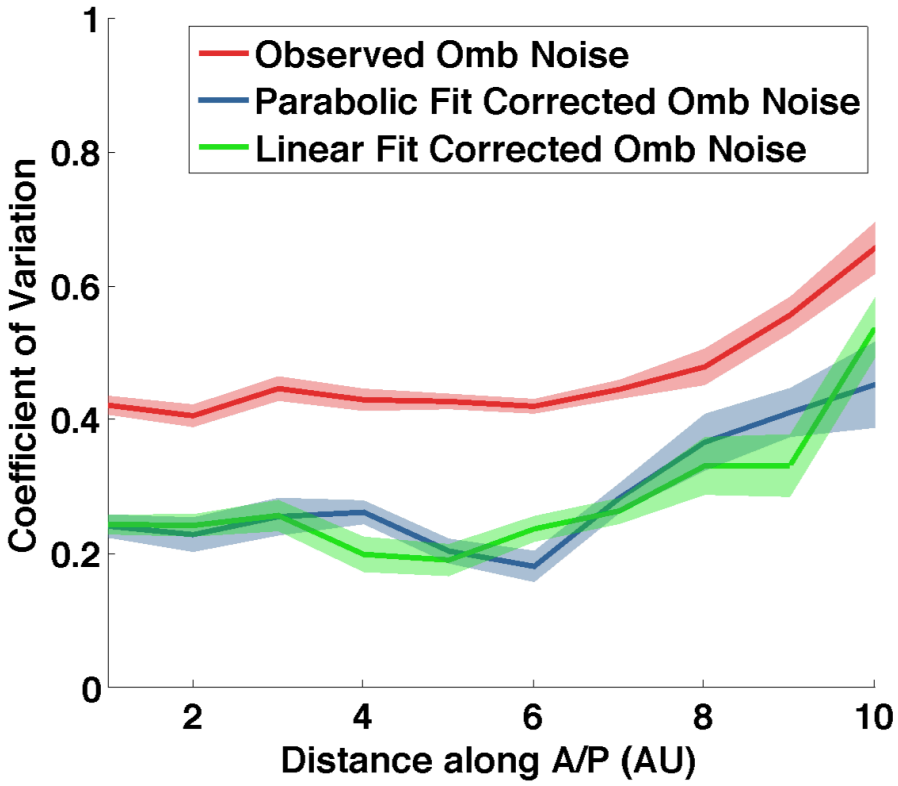


Figure 3.8 True Omb concentration noise after measurement error correction. Average of observed Omb concentration (average fluorescent intensity per pixel) noise as a function of position along the AP direction in the posterior compartment (red line), shaded region represents the standard error of the mean. True Omb concentration noise is corrected using two estimate of measurement noise, a linear and a parabolic relationship between noise and intensity (Figure 3.7). Blue line and its shaded region represent the mean and standard error of the true Omb concentration noise using correction from parabolic fit relationship between noise and intensity. Green line and its shaded region represent the mean and standard error of the true Omb concentration noise using correction from linear fit relationship between noise and intensity.

Omb Protein has a half-life of 6 hours

The above data indicate that, even though *omb* transcriptional noise is high in the wing disc, especially near the region where patterning takes place, the disc is able to reduce this noise by about 3 fold (1.1 / 0.35) in going from transcription to protein level. Since normal development of the wing disc takes roughly 3-4 days, and the cell cycle length is 8 hours or longer, this system might seem to provide a good environment for using simple temporal averaging as a way to bring down noise. Either a long half-life of the *omb* transcript or of Omb protein could allow substantial accumulation of Omb proteins in the nucleus, raising overall concentration level and effectively reducing noise. To see whether this might be occurring, I quantitatively measured the Omb protein half-life.

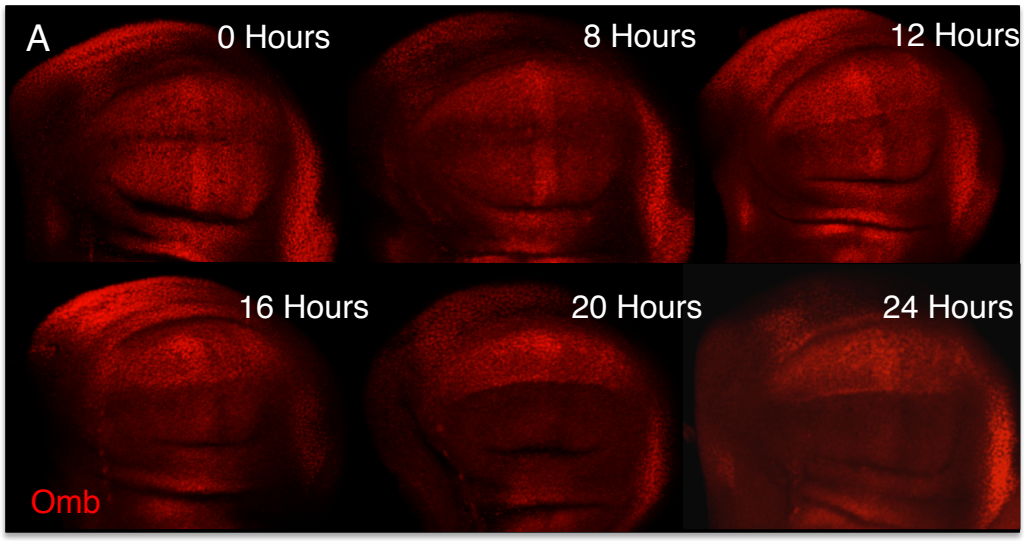
I setup an experiment using Ap-Gal4, tub-gal80^{ts} >UAS-Omb-RNAi to directly measure the rate of Omb protein degradation in otherwise normal discs. Apterous is expressed in the dorsal domain of the wing disc, thus leaving the ventral side to provide an un-perturbed internal control for comparison of Omb protein expression level. Because gal80^{ts} allows for temporal control of UAS-Omb-RNAi expression, it avoids any long-term effects that may be caused by long-term absence of Omb. Larvae were reared at 18°C for 3.5 to 4 days and then transferred to 29°C, where they remained until dissection. After the temperature switch, I took samples in 2-hour increments from 2 hours to 24 hours, and also at 48 hours. Discs were immediately dissected, fixed and immunostained with an Omb specific antibody. All images were then acquired on a laser scanning microscope with the same settings.

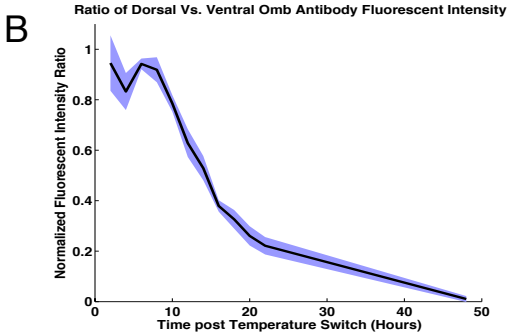
Omb antibody fluorescence in the dorsal compartment slowly started to decrease, compared to ventral compartment, approximately 6-8 hours after the initial temperature switch. By 22-24 hours, Omb signal had disappeared nearly completely (Fig. 3.9A). To quantify the rate of degradation, I took a background subtracted average of two rectangular regions, and computed the ratio of dorsal to ventral compartment fluorescent intensity. Next, I plotted the mean and standard error of the dorsal to ventral fluorescent intensity ratio as a function of time post-temperature switch (Fig. 3.9B). I observe from the plot that there is an initial delay of around 6-8 hours before the dorsal intensity starts to decline, after which dorsal intensity drops in an exponential manner.

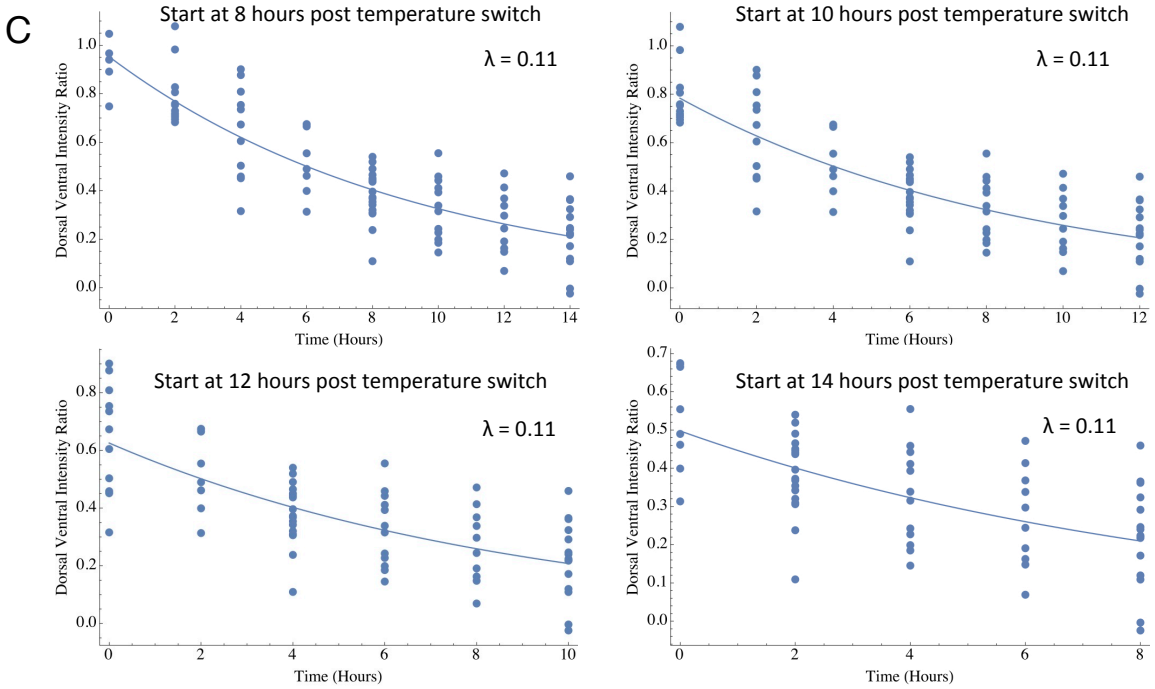
This final shape of the measured Omb degradation profile in Fig 3.9B is the result of several interacting parameters in the biological system: time delay from gene activation using Gal4>UAS, the rate of RNAi degradation and production, the rate of mRNA degradation from RNAi and its endogenous half-life, and protein production and half-life. To estimate the half-life of Omb from this experiment, I can build a mathematical model and fit our data to the model, or I can make some basic assumptions about how the system works. If I assume that Omb-RNAi effectively binds to the *omb* mRNA and half-life of the *omb* transcript is shorter than that of the protein, protein half-life will dominate the rate of degradation after Omb-RNAi reaches steady state and takes out all of the available transcript. Since the decay rate of an exponential function remains constant throughout, there should be regions in the later parts of the observed degradation profile in which the half-life of the protein becomes the dominant rate and remains constant. Accordingly, I fit the degradation profile from various starting

points to an exponential decay function ($\alpha e^{-\lambda t}$) and check if going forward from any point in time, the decay rate fits a single exponential. I found that this occurs about 8 hours after temperature switch mark. From this point on, the rate of disappearance of Omb, λ, remains stable around ~ 0.11, corresponding to a half-life ($\ln(2)/\lambda$) of a little over 6 hours (Fig. 3.9C).

Figure 3.9 Measurement of Omb protein half-life by turning off translation. (**A**) Time course of Omb protein expression in *Ap-Gal4,tub-gal80^{ts} > OmbRNAi* wing discs after temperature switch. Omb is visualized with an Omb specific antibody. Starting around 8 hours, dorsal expression of Omb starts to decrease, Omb protein expression slowly decreases over time and disappeared in most cells after 24 hours. (**B**) Mean of dorsal to ventral ratio in Omb expression as a function of time. 0 represents the time of temperature switch that turns on gal4>OmbRNAi, samples obtained every 2 hours until 22 hours and again at 48 hours. The blue shaded region is standard error of the mean. (**C**) Fitting different regions of Omb degradation profile as an exponential function. The profile starts from 8 hours after the temperature switch until 22 hours, 10 – 22 hours, 12 – 22 hours, and 14 – 22 hours.


A
0 Hours
8 Hours
12 Hours
16 Hours
20 Hours
24 Hours
Omb


B
Ratio of Dorsal Vs. Ventral Omb Antibody Fluorescent Intensity
Normalized Fluorescent Intensity Ratio
Time post Temperature Switch (Hours)


C
Start at 8 hours post temperature switch
λ = 0.11
Start at 10 hours post temperature switch
λ = 0.11
Start at 12 hours post temperature switch
λ = 0.11
Start at 14 hours post temperature switch
λ = 0.11
Dorsal Ventral Intensity Ratio
Time (Hours)

CONCLUSIONS:

Noise has always been a part of biology, but the study of noise in biological systems has not always been possible. With technological advances of the last few years, the study of noise has become one of the most important research topics. Single cell measurements of mRNA, protein and other cellular components have highlighted just how noisy biological processes can be. In a developing organism, morphogen-mediated pattern formation involves stochastic components in every step of the process, from morphogen binding and signal transduction to gene expression. These processes will create cell-to-cell variability in the system, but just how much noise is present, and how they propagate is not well studied (Lander, 2013). Evidence that has been previously published for developing organisms, has been mostly focused on the *Drosophila* embryo (Bollenbach et al., 2008; Gursky et al., 2011; Holloway et al., 2011; Little et al., 2013; McHale et al., 2011).

Extrinsic and intrinsic noise in *omb* transcription

I combined FISH, image segmentation and image analysis to present the first study of transcriptional noise in *Drosophila* wing disc. FISH with intronic probes has been previously used in the *Drosophila* embyo (McHale et al., 2011; Paré et al., 2009). Here, I adapted and modified the technique for the wing disc. An intronic probe allowed me to eliminate any cytoplasmic signals and focus on nascent transcriptional events. I picked *omb* as a gene of interest due to its long intronic region, thus allowing me to

have the flexibility of using intronic probes that report long transcriptional averaging times.

The finding that, with an averaging time of ~18-24 minutes, *omb* transcriptional activity varies with a CV of 0.7 to 1.1 means transcriptional activity is quite noisy. This measured variation (or spatial noise) in the morphogen gradient system presumably does not come from just the noise introduced from transcriptional kinetics as it can be contributed to by two potential components, intrinsic and extrinsic. Intrinsic noise comes from the stochasticity of transcription itself, as discussed above. It refers to the moment-to-moment variability in the transcriptional response to a signal. Extrinsic noise is the "input" noise that comes upstream of the transcriptional process; it is the accumulated noise from all the steps leading up to transcriptional regulation and the cell-to-cell variability in perceived morphogen signal. Some data exist for the input noise in pMad. It was estimated in the work of (Bollenbach et al., 2008), who reported data (their Fig. 5D) that imply a CV of around 0.15 near the Dpp source, which increases linearly to about 0.30 30μm away. Although their data are incomplete, since the posterior side of the disc is larger than 30μm, and were not corrected for measurement error, it is possible that at least some of the noise measured in omb transcriptional activity reflects input noise from pMad. It is not possible to calculate how much, however, because there are steps intervening between pMad and omb expression that could add or subtract noise. For example, *omb* is actually regulated by *brk,* and not directly by pMad.

The long half-life of Omb is sufficient to reduce *omb* transcriptional noise

Although I found *omb* transcription noise to be high, measured protein noise was substantially lower. Since I was not able to perform the same type of experiments as in FISH to calculate the experimental/measurement noise, I had a make some assumptions, and use an alternative method, in which I assumed that variation in signal for one antibody staining is comparable to any other antibody staining, as long as they are associated with the same measured fluorescent intensity. Consequently, it is possible that I over- or underestimated measurement error in calculating protein noise levels.

I independently measured the Omb protein half life to be about 6 hours. If translational events are independent and independently distributed in time (Poissonian) and the mean number of available mRNA remains constant, then on average, one should expect noise (coefficient of variation) to fall due to time averaging with the the square root of the fold change in integration time. This is because the coefficient of variation for a Poisson process is simply the reciprocal of the square root of the mean. In this case, the transcriptional integration time ranges between 18-24 minutes, so that a 6 hour protein half life allows for time averaging of 15-21 times the integration time, which would predict a potential reduction in CV of ~4 fold.

This 4-fold estimate represents a theoretical limit on how much temporal averaging can help. If translation is bursty, one may require longer averaging times to achieve a similar reduction in noise. In this case, a 4-fold decrease in CV from the level of transcriptional noise for *omb* would predict that Omb protein noise could be reduced by time averaging to a CV around 0.15 to 0.22, slightly better than measured protein

noise in the Omb protein gradient (suggesting that translation may be somewhat bursty). It is also possible that the half-life of the *omb* message plays some role in reducing the noise from *omb* transcription, although it would need to approach that of Omb protein in order to contribute significantly.

The Omb protein gradient is precise enough to generate vein primordium

Abrupt is the critical organizing gene for specifying L5, and I measured the transition width of its border at different developmental times. I found that the Ab border sharpens with increased disc size (which serves as a proxy for developmental time), suggesting that Ab expression starts out fuzzy but refines itself, possibly through the EGFR pathway.

The Omb protein gradient has a CV of approximately 0.35 in the region of the Ab border, and I measured the local length of Omb to be about ~8μm. The theoretical limit for the transition width produced by a gradient with said local length scale (8μm) and CV (0.35) is ($2cv \times \lambda$) = 5.6μm. This tells us that the Omb protein gradient is precise enough to provide the positional information the specifies the L5 pro-vein.

The cost of using time averaging, mediated via a long-lived protein, to overcome transcriptional noise in a morphogen gradient, is that the patterning events downstream of the morphogen are necessarily slower than they might otherwise be. Although the long duration of wing disc development makes this a feasible strategy, there are consequences that stem from the fact that the disc grows substantially on a time scale

similar to that over which Omb turns over. Some of these consequences will be discussed in the next chapter.

CHAPTER 4

Morphogen Gradient Interpretation and the Robustness of Wing Vein Primordium Formation in *Drosophila* Wing Disc

INTRODUCTION:

Multicellular organisms begin from a single cell, and through repeated division and differentiation, develop into a final form with patterns of various degrees of complexity. In order to generate pattern, developing cells must make decisions based on its temporal or spatial information. One source of spatial information is concentration gradients of secreted signaling molecules called morphogens. Morphogens are generally produced at a localized source, spread across a target tissue through diffusion (Muller et al., 2012; Schwank et al., 2011; Zhou et al., 2012) or active transport (Bollenbach et al., 2005; Erickson, 2011), and form a concentration gradient. Cells in the receiving region then respond to these gradients in a threshold-dependent manner and activate distinct patterns of gene expression (Wolpert, 2011). Many examples of morphogen gradient systems have been found in a broad spectrum of tissues and organisms: from BMPs in dorsoventral patterning of *Xenopus* and *Drosophila* (Dosch et al., 1997; Nellen et al., 1996), to HH in chick neural tube development (Briscoe et al., 2001), to Wnt in *Drosophila* wing development (Zecca et al., 1996).

The original model of morphogen theory assumes cells interpret their position based purely on concentration thresholds of the morphogen. While there has been some evidence of direct measurement of concentration through receptor occupancy in Activin regulated targets in the *Xenopus* embryo (Dyson and Gurdon, 1998; Smith,

2009); or differential binding affinity of target genes in Dorsal controlled D/V patterning in *Drosophila* (Jiang and Levine, 1993; Stathopoulos and Levine, 2004); other evidence has suggested that concentration thresholds are not the only source of information provided by a morphogen gradient (Nahmad and Lander, 2011). For example, in several morphogen systems, flattening or disrupting the morphogen gradient has been shown to have less than the predicted effect on downstream target expression (Alexandre et al., 2014; Lucchetta et al., 2008).

One of the key features of developmental patterning is robustness against fluctuation in environment or genetic makeup; in this regard, morphogen concentrations in steady-state diffusion gradients should not provide a particularly reliable source of information. Thus various feedback mechanisms (Eldar et al., 2002; Eldar et al., 2004; Perrimon and McMahon, 1999; Umulis et al., 2006), presteady-state interpretation (Bergmann et al., 2007), dynamic interpretation (Nahmad and Stathopoulos, 2009), or spatiotemporal interpretation (Dessaud et al., 2010) have all been proposed in different scenarios, to provide buffer against fluctuations.

Drosophila wing vein patterning is particularly robust to changes in levels of morphogen receptors, and other factors, but how downstream target genes interpret the morphogen gradient is only partly understood in this system. Decapentaplegic (Dpp) is one of the best known morphogens, and functions during several stages of *Drosophila* development. In the *Drosophila* wing disc, Dpp is involved in growth control (Schwank and Basler, 2010) and patterning of longitudinal veins along the anterior-posterior axis of the wing (Blair, 2007). Dpp is produced in a stripe of cells next to the

anterior/posterior border, and spreads to both sides of the disc through diffusion (Schwank et al., 2011; Zhou et al., 2012). The primary signaling event of Dpp starts with activation of its receptor Tkv, which leads to phosphorylation of the BMP specific Smad (Mad) to form pMad. PMad then forms a complex with Medea (Med) to directly activate downstream targets (Kim et al., 1997; Massague et al., 2005) or form Shn/Mad/Med complex to relieve repression of Brk (Gao et al., 2005; Winter and Campbell, 2004). Since pMad is involved in both mechanisms of gene control, it may be considered the primary readout of the BMP activity gradient.

There is also another BMP-family ligand in the wing disc that contributes to the BMP signaling activity, Glass bottom boat (Gbb, also known as 60A). It is expressed uniformly through the wing disc and can bind to the same receptors as Dpp, although with different affinities (Khalsa et al., 1998; Nellen et al., 1994; Ray and Wharton, 2001). It has been found that Dpp and Gbb can form heteromeric complexes, the absence of which has a strong effect on placement of the vein L5 (Bangi and Wharton, 2006a; Bangi and Wharton, 2006b).

Two downstream targets of Dpp are particularly important for the formation of longitudinal veins L2 and L5: *sal* and *omb.* Both *sal* and *omb* are transcription factors that are expressed in broad regions centered around the A/P boundary. L2 induction has been proposed to start with the production of an unknown short range signal “X” from *sal*-expressing cells, which then diffuses to neighboring cells and induce the L2-organizing gene *knirps* (Lunde et al., 1998). Similarly, it has been proposed that L5 is induced when *brk*-expressing cells release a similar short range signal “Y”, which

induces L5-organizing genes in regions with Omb expression (Cook et al., 2004). However, neither short range signals “X” & “Y” have yet been identified. Furthermore, a recent study actually shows Omb protein is expressed in a gradient that is high in the middle and gradually declines to the periphery of the disc, instead of in a broad region with a well-defined boundary.

Here, I explore how different downstream targets of Dpp respond to manipulations that change the shape of the Dpp gradient in the larval wing disc. I use both quantitative and qualitative experiments to look at different components of the signal transduction pathway and how they are affected, both at the end of larval development and at a variety of earlier developmental times.

Results:

Decreased expression of Dpp receptor Tkv lengthens the BMP activity gradient

As in many developmental patterning systems, wing vein patterning in adult *Drosophila* is exceptionally robust to environmental and genetic perturbations. A recent study analyzed the locations of seven adult wing vein landmarks across different natural genetic backgrounds and growth temperatures and found variations to lie within a single-cell width (Abouchar et al., 2014). Heterozygous mutations in a number of proteins involved in the morphogen gradient also result in nearly normal development of the veins. With a genetic perturbation of more than 50% to the morphogen receptor *tkv*, adult vein patterns in L2 and L5 are minimally affected (Fig. 4.1).

While the adult wing vein pattern is a result of morphogen-induced patterning of the imaginal disc, it is a multi-step process involving several levels of signaling and genetic control (Blair, 2007; De Celis, 2003). To examine just how much perturbations in receptor concentration affect the Dpp morphogen gradient in the wing disc, I started by looking at the most direct readout of Dpp, the BMP activity gradient. I performed immunofluorescence and visualized BMP activity gradient with an anti-p-Mad antibody (Fig. 4.2). Z-stack images were acquired and I extracted the pMad profile by first generating 20 orthogonal Z-slices along the A/P direction right below the D/V boundary (in the dorsal compartment). Next, I averaged those 20 orthogonal Z-slices and manually followed along the curvature of the disc proper cells in the disc, eliminating any signal that may be coming from the peripodial membrane. I focused primarily on pMad signal in the posterior half of the wing pouch, to avoid complication that comes from difference in modification of Dpp receptors Tkv by Hh (Funakoshi et al., 2001; Tanimoto et al., 2000).

To quantify pMad concentration profiles, I fit them to an exponential decay function $I_{pMad} = I_0 e^{-x/\lambda}$. Here, I_{pMad} is the intensity profile of pMad, I_0 is the intensity at x=0, or its maximum intensity, and λ is the characteristic length scale where the intensity profile falls to about e^{-1}, ~ 37% of I_0. The length scale parameter λ should be related to the processes of diffusion and decay, according to $\lambda = \sqrt{D/k}$, where D is the diffusion coefficient, and k is the effective rate of clearance, the latter related to the binding rate and receptor concentration for the morphogen. When I fit individual profiles separately, I found that mean length scale of pMad gradient in wild type animals is

~11.2 ± 0.5 μm (Fig 4.2C). It may be notted that my measurements of the pMad length scale is significantly shorter than a value published by another group, ~ 25.2 ± 4.5 μm (Bollenbach et al., 2008). The genotype used by that group was not actually wild type, but a rescue of a *dpp* deficient animal with DppGal4 driving a Dpp-GFP fusion protein. It's not hard to imagine a Dpp fusion protein having a different length scale from that of an endogenous protein, the change in protein size and structure may alter the diffusion coefficient as well as binding rate to receptors. In fact, when I used a genotype similar to the one in the Bollenbach et al, study, I found the pMad length scale to be ~ 20.6 ± 1.6 μm, match closer to the published value.

In $tkv^{7/+}$ animals, I found the mean length scale of the pMad gradient to be ~ 17.1 ± 1 μm (Fig. 4.2C). Tkv^{7} is an antimorphic allele, thus possibly reducing functional Tkv concentration by more than 50% (de Celis, 1997; Jazwinska et al., 1999). In the simple model of uniform morphogen decay described by the equation above, a 2 fold change in receptor concentration is expected to increase length scale by $\sqrt{2} \sim 1.4$. Thus the observed increase in length scale in heterozygous loss-of-function discs seems to match that of the theory, increasing length pMad length scale by ~ 1.5, slightly higher than expected for 2 fold decrease in receptor expression.

Additionally, I performed the same analysis with *pentagone* (*pent*, magu – FlyBase) heterozygous animals. *Pentagone* have been recently found to be a secreted feedback regulator of Dpp, and plays a central role as an expander of the Dpp gradient (Ben-Zvi et al., 2011; Hamaratoglu et al., 2011). Absence of *pent* significantly shrinks the pMad gradient and *pent* mutants fail to develop L5 vein in adults. Interestingly, *pent*

heterozygous animals are mostly normal, with no vein phenotypes. Nonetheless, I observed that *pent*$^{-/+}$ animals have a slightly reduced pMad length scale of ~8.8 ± 0.6μm (Fig 4.2C).

These data show that the shape of the BMP activity gradient can be altered in various genetic perturbations to components crucial to the formation of the Dpp morphogen gradient, yet, in the end, the positions of adult vein phenotypes are barely affected (Fig 4.1). This suggests that the interpretation of the morphogen gradient is somehow altered further downstream in the signaling pathway. I therefore decided to look at some of the downstream targets of BMP, to see how they respond to changes in the gradient.

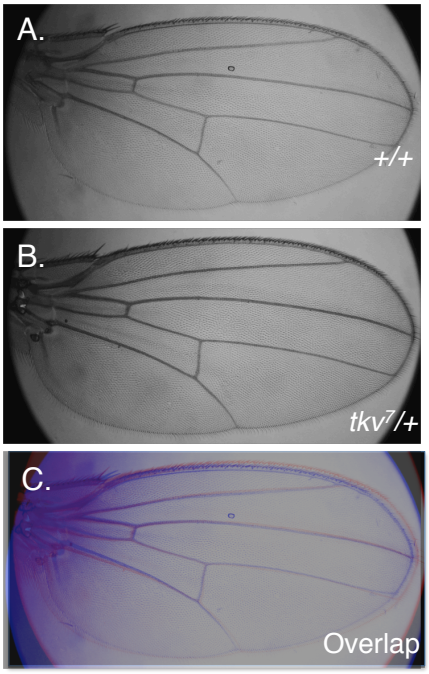
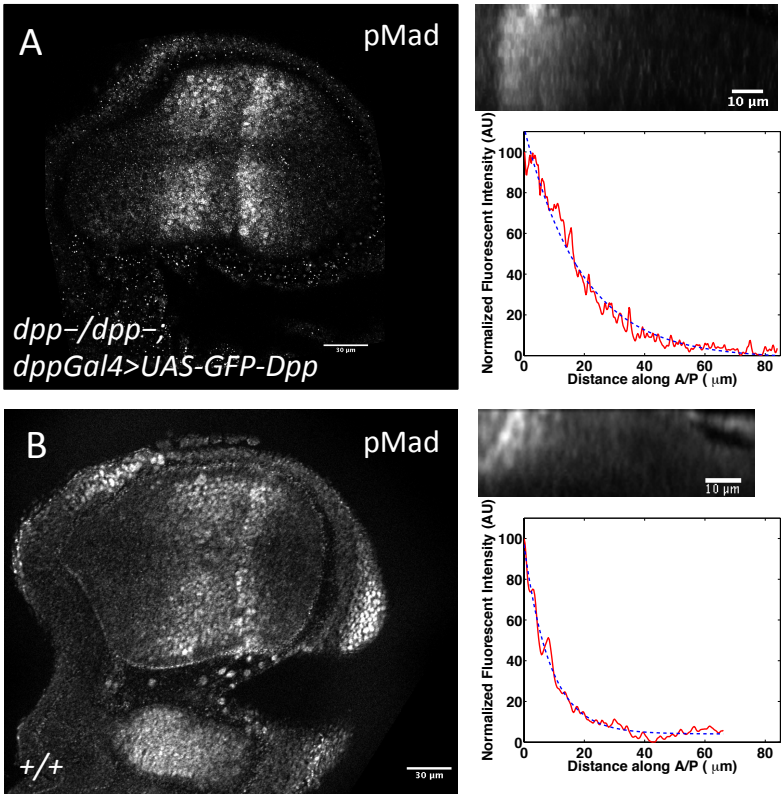
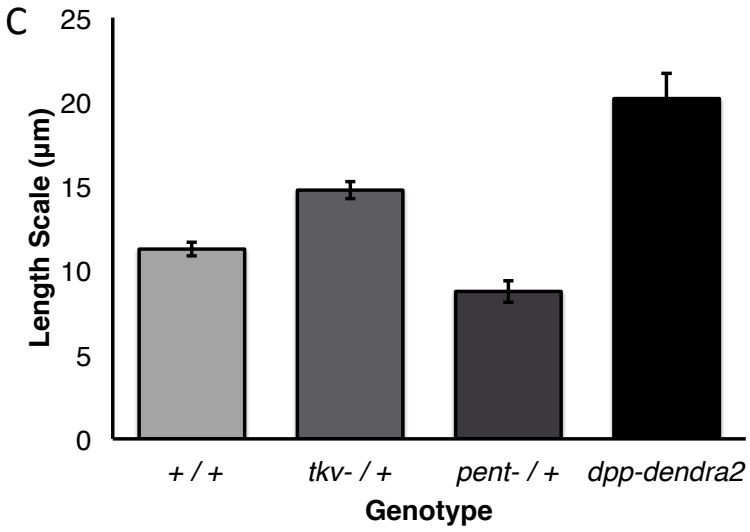


Figure 4.1 Adult wing vein positions are minimally affected by reduction of Dpp receptor. Comparison of adult *Drosophila* wing vein in wild type (**A**) and *tkv* antimorphic allele $tkv^{7/+}$ (**B**). Overlap (**C**) shows minimal difference in position of the wing veins.

Figure 4.2 Quantification of the BMP activity gradient length scale. (**A,B**) pMad immunostaining of d*pp⁻/dpp⁻; dppGal4>UAS-GFP-Dpp* and wild type respectively. Right panel shows orthogonal z-average of 20 slices in the dorsal posterior compartment, and corresponding normalized pMad intensity profile (red solid line). Blue dotted line, fit of the pMad profile to the exponential function of $I_{pMad} = I_0 e^{-x/\lambda}$. (**C**) Average length scale in μm of wild type, *tkv* antimorphic heterozygous, pent⁻ heterozygous, and d*pp⁻/dpp⁻; dppGal4>UAS-Dpp-Dendra2.*


A
pMad
dpp−/dpp−;
dppGal4>UAS-GFP-Dpp
30 µm
10 µm
Normalized Fluorescent Intensity (AU)
Distance along A/P (µm)
B
pMad
+/+
30 µm
10 µm
Normalized Fluorescent Intensity (AU)
Distance along A/P (µm)


C
Length Scale (µm)
+ / +
tkv- / +
pent- / +
dpp-dendra2
Genotype

Placement of vein primordia is unaffected by change in BMP activity gradient

The next logical target to focus on, downstream of pMad, is *brinker*. The interpretation of BMP activity in the wing disc is heavily dependent on the relief of gene repression mediated by *brk*. In the *Drosophila* wing disc, BMP signaling induces the formation of Shn/Mad/Med (SMM) complexes which repress *brk* transcription by binding to silencer elements of the *brk cis*-regulatory region (Arora et al., 1995; Gao et al., 2005; Muller et al., 2003; Pyrowolakis et al., 2004). To quantify how *brk* responds to changes in the shape of the pMad gradient, I used a lac-Z enhancer trap of *brk* to visualize *brk* expression. I extracted the brk-lacZ fluorescent intensity profiles just beneath the DV boundary in the dorsal compartment from the most anterior end of the pouch to the posterior end (Fig. 4.3 A-D). I also manually circled the wing pouch region to minimize effects from differences in wing disc size. Next, *brk* profiles were processed with a Savitzky-Golay filter to smooth the data and increase signal-to-noise ratio while preserving location of the peaks (Lieber and Mahadevan-Jansen, 2003; Luo and Bai, 2005). From the smoothed data, an automatic peak detection algorithm was used to detect the two peaks and find the distance between half-maximum points facing the central region of the disc. This distance is finally normalized to the square root of the wing pouch area. With the described method, I found that the normalized *brk* distance from anterior to posterior increases in the $tkv^{7/+}$ animals, going from 67 ± 1% to about 76 ± 2% (Fig. 4.3E). This increase, however, is not as much as one would have expected if *brk* were simply responding to threshold of pMad, where pMad length constant increased by ~1.5 times when receptor dosage was lowered by 2 fold. Thus, *brk*

expression seems to follow the concentration level of the BMP activity gradient, expanding when receptor levels are lowered, but not with the same magnitude as theory would have predicted.

Next, I examined how various organizing genes for vein primodium respond to changes in the morphogen gradient. I can visualize vein primoridia L5 and L2 in the wing discs with anti-Abrupt and anti-Knirps antibodies, respectively. Both Abrupt and Knirps act as organizing genes to induce vein fate (Cook et al., 2004; Lunde et al., 1998). I performed immunostaining in a DppGal4-UAS-GFP background, so that I could mark the Dpp production region (Fig. 4.4 A-D). Using similar methods as in the analysis of *brk* expression borders, I identified the anterior half maximum of Dpp, posterior half maximum of Dpp, peak expression of Ab, posterior half maximum of Kni and peak of Kni expression. Interestingly, I found no difference in the normalized distance from the Dpp source to the both provein borders or peak expressions between wild type and $tkv^{7/+}$ animals (Fig. 4.4E). The experiment was repeated using Delta staining to mark the L5 provein (Fig. 4.5 F-I), instead of Ab, and again there was no detectable difference in normalized distance from Dpp source to vein primordium.

In addition to using *tkv* heterozygotes to change the shape of the Dpp gradient, I also performed knock down experiments using Tkv-RNAi to control levels of *tkv* expression. Utilizing the spatial temporal control offered by the Gal4/$Gal80^{ts}$ system, I limited spatial expression of Tkv-RNAi to sub-regions of the wing disc by using different Gal4 drivers. The temperature-sensitive mutant $Gal80^{ts}$ is active at 18°C but does not repress Gal4 at 29°C. I found that between 25°C and 29°C, I could use the system to

continuously vary the strength of Gal4-induced gene expression. To observe how Dpp downstream targets respond to differing levels of Tkv, while maintaining early BMP signaling, I reared Apterous-Gal4; tub-Gal80^{ts} > UAS-Tkv-RNAi animals at 18°C until early first instar, then transferred them to 25.5°C until dissection. Apterous is only expressed in the dorsal compartment of the disc, which sets up ventral compartment as an internal control. Since pMad expression also affects growth of the disc, by temporally controlling for *tkv* expression level, I can limit the effect of size difference between the compartments. In this setup, if location of L5 is dependent on pMad level, genes marking L5 would be expected to be expressed in a jagged manner between ventral and dorsal compartments.

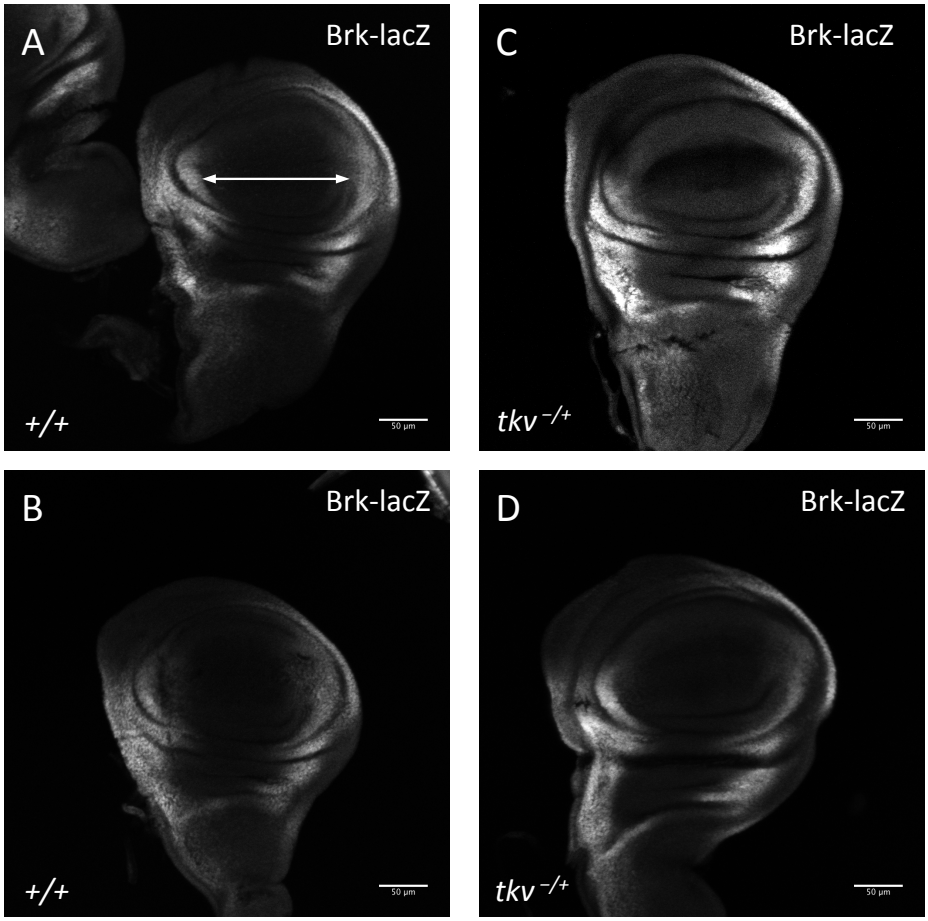
Strikingly, I observed highly elevated and posterior-shifted pMad expression in the dorsal (Gal4-expressing) compartment compared with the ventral (control) (Fig. 4.5 B'), with Delta (marking L5 position) expression that was essentially unchanged in position compared to the anterior side (Fig. 4.5 B''). PMad expression is extremely sensitive to titration of Gal4/Gal80^{ts} levels and varies even in tightly controlled temperatures. In some cases, pMad expression remains at a high level near the Dpp production region beforere dropping off, showing signs of receptor saturation. Delta expression also seemed to be slightly elevated in some discs. Similar to L5 or Delta position, Omb visualized by antibody staining showed no change in dorsal compartment, it is still expressed in a gradient extending to near the end of the wing disc pouch. On the other hand, levels of *Brk* expression, visualized with brk-lacZ, is

elevated in regions of low pMad expression, and seems to respond to concentration levels of pMad (Fig 4.5 D'').

The proteoglycan Dally, a co-receptor of Dpp, also regulates the distribution of and cellular responses to the Dpp morphogen. Dally antagonizes the effect of Tkv in taking up Dpp (Akiyama et al., 2008) and has also been linked to *pent* function (Ben-Zvi et al., 2011) in assisting in the spread of the Dpp morphogen gradient. I over-expressed Dally in the dorsal compartment of the wing disc, and again looked at the effect on pMad and Delta expression (Fig. 4.7). As in the experiments in which I downregulated Tkv, the pMad gradient expanded further into the posterior compartment, but Delta expression remains unchanged, aligning with the Delta stripe in the ventral (control) half of the disc (Fig. 4.7 C).

In summary, variety of perturbations that change the BMP activity gradient can affect the downstream target *brk*. However, it appears that Omb, Spalt, and other downstream targets are robust to changing the thresholds of pMad. Even though *brk* has been linked to the induction of L5 primordium (along with Omb; (Cook et al., 2004), its expression border seems to have little effect on the position of L5 primordium.

Figure 4.3 Quantification of normalized *Brinker* border-to-border distance. (**A, B, C, D**) Imunostaining of Brk-LacZ in 2 wild type animals and 2 $tkv^{7/+}$ animals, respectively. White arrow shows distance measured between anterior border or Brinker to posterior border of Brinker. (**E**) Brinker border-to-border distance measured below the DV boundary in the dorsal compartment, quantified as half-maximum to half-maximum of florescent intensity, normalized to squared root of the wing pouch area, for wild type and $tkv^{7/+}$ animals. Brinker border to border distance increased in $tkv^{7/+}$.


A
Brk-lacZ
+/+
C
Brk-lacZ
tkv −/+
B
Brk-lacZ
+/+
D
Brk-lacZ
tkv −/+

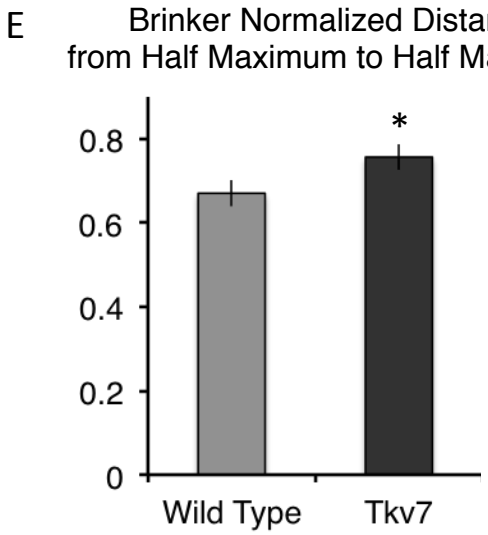
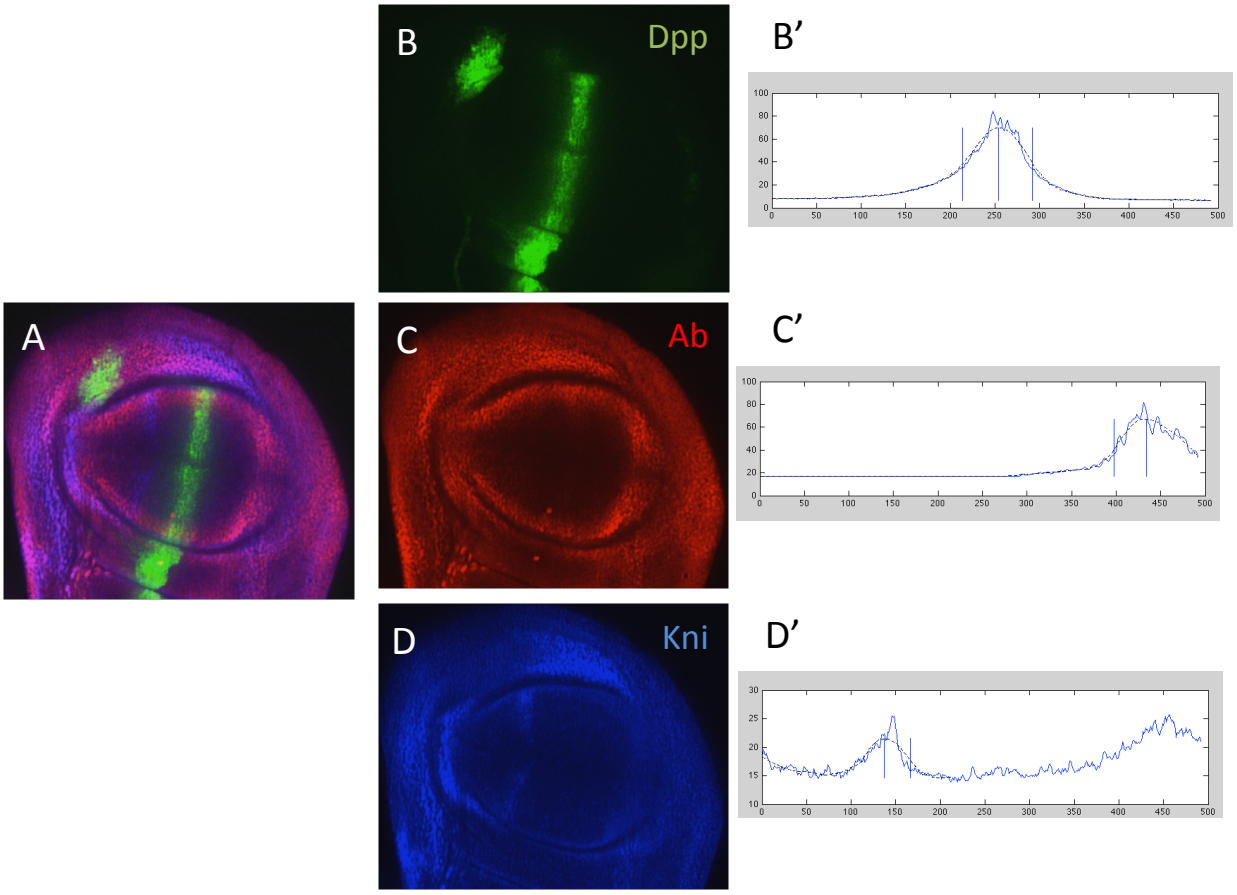
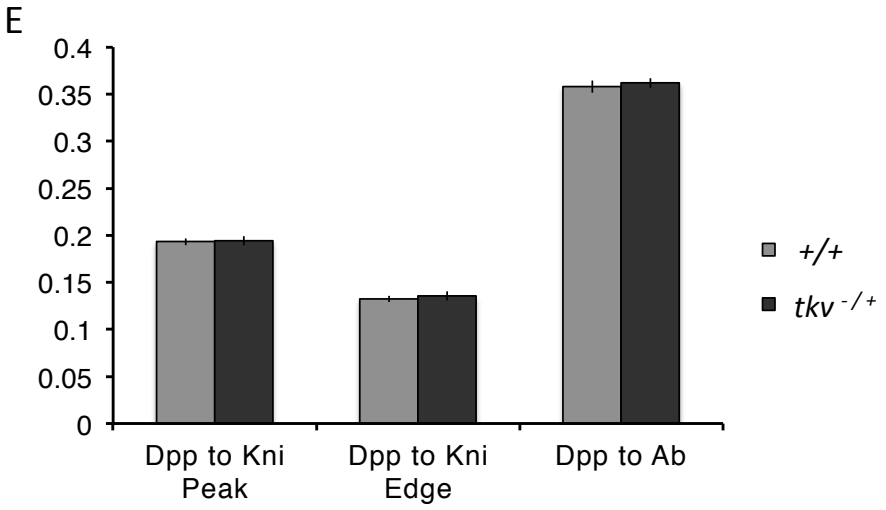

E
Brinker Normalized Distance
from Half Maximum to Half Maximum
*
0.8
0.6
0.4
0.2
0
Wild Type
Tkv7

Figure 4.4 Normalized distance from Dpp source to L2 and L5 vein primordia. (**A-D**) Wild type immunostaining of DppGal4-UAS-GFP, Abrupt, and Knirps, expression profiles of Dpp, Ab and Kni are extracted from average z-stacks of Apotome images in (**B'-D'**). Dotted lines are post Savitzky-Golay filtered profiles and vertical lines denotes automated detection of peaks and half-maximum positions. (**E**) Normalized distances of anterior Dpp to Kni peak, anterior Dpp to Kni half maximum facing Dpp source, and posterior Dpp to Abrupt maximum for wild type and $tkv^{7/+}$ animals. No difference in normalized distances were detected. (**F-I**) Alternative immunostaining were performed with Delta antibody to mark L3-L5 vein primordium expressions.


A
B
Dpp
B'
C
Ab
C'
D
Kni
D'


E
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
Dpp to Kni Peak
Dpp to Kni Edge
Dpp to Ab
+/+
tkv -/+

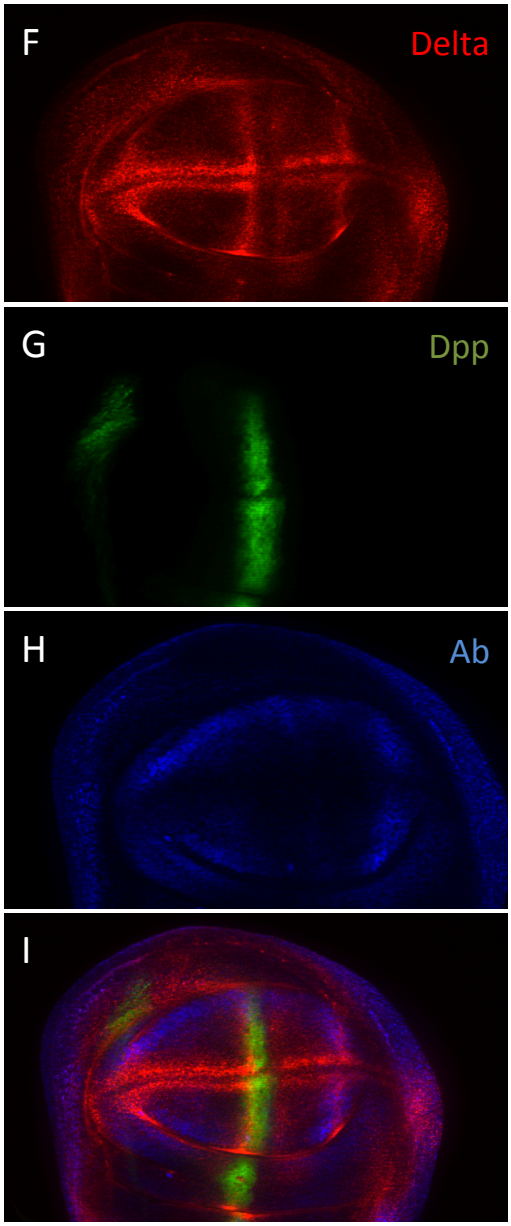

F
Delta
G
Dpp
H
Ab
I

Figure 4.5 Effect of Tkv-RNAi expressed in dorsal compartment of wing disc on Dpp downstream target gene expressions. Wild type control images in group (**A**) and (**C**), ApGal4;tubGal80^{ts}>UAS-Tkv-RNAi reared in 25.5 °C starting from 1st instar in group (**B**) and (**D**). (**B'**) PMad expression is clearly expanded in the dorsal compartment of the disc, yet (**B''**) shows delta staining inline with that of the ventral compartment. Delta staining also seem to be strongly expressed in comparison to ventral compartment. (**D'**) shows pMad expression in dorsal compartment expanded and shifted. Regions of extreme low pMad corresponds with elevated levels *brk-lacZ* expression, even in the medial region of the disc (**D''**). Omb expression detected with anti-Omb antibody seems to be unchanged in (**D'''**) compare to the ventral side.

A
+ / +
A'
pMad
A''
delta
B
apGal4>UAS-TkvRNAi
B'
pMad
B''
delta
B'''
C
+ / +
C'
pMad
C''
brk
C'''
omb
D
apGal4>UAS-TkvRNAi
D'
pMad
D''
brk
D'''
omb

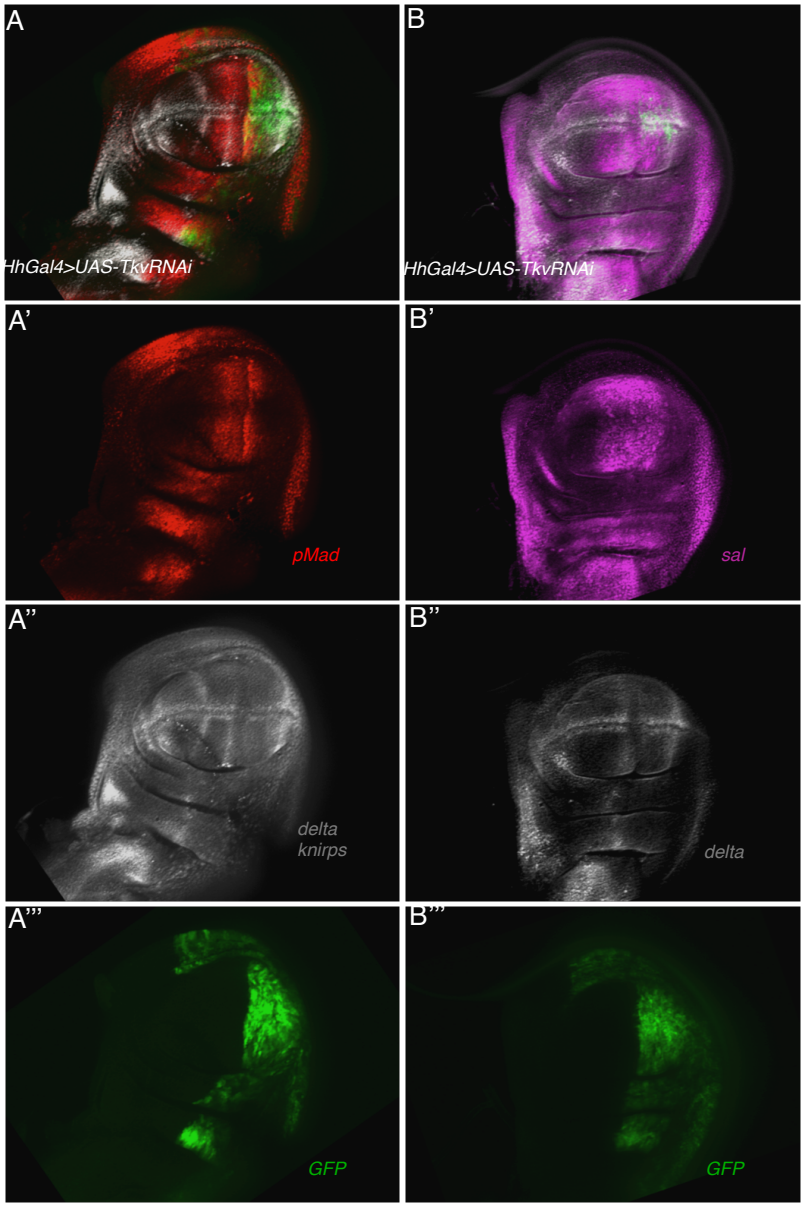


Figure 4.6 Effect of Tkv-RNAi expressed in posterior compartment of wing disc on Dpp downstream target gene expressions. HhGal4,tubGal80^{ts}>UAS-Tkv-RNAi reared in 26.2 °C starting from 1st instar for both (**A-B**). pMad is slightly expanded in the posterior direction (**A'**), while Delta expression seems to be unchanged in (**A'', B''**).

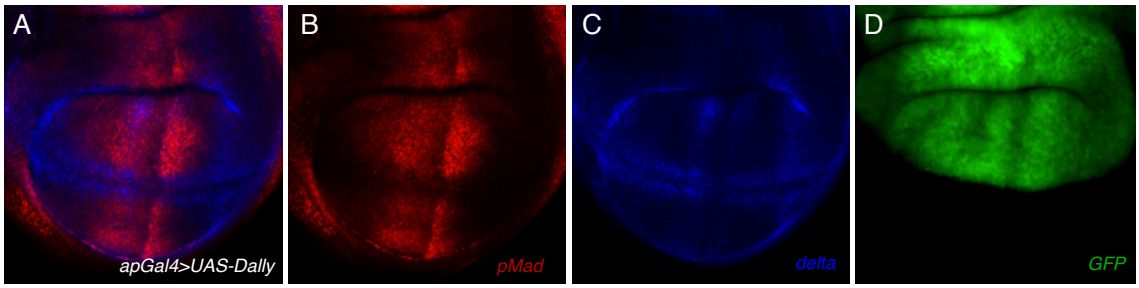


Figure 4.7 Increased Dally expression in dorsal compartment shows no effect on L5 position. *ApGal4>UAS-Dally* (**A**) Composite image of pMad, Delta and UAS-GFP expression. (**B**) pMad expression in the dorsal compartment is expanded in the posterior compartment. (**C**) Delta expression, especially expression of L5 is unchanged and remained inline with Delta expression in the ventral compartment.

L2 knirps faithfully follows Spalt border

Currently, theories on L2 and L5 formation share some similarities. Two theories exist on formation of L2 at the moment, one involves a short range signal X produced by Sal expressing cells that induce *knirps* expression in neighboring cells, and is repressed by *brk* expression (Lunde et al., 1998). Another theory proposes that *knirps* is induced in low levels of Sal expression, typically at the limit of Salm/Salr domain (de Celis and Barrio, 2000). Similar to L5, L2 (marked by Knirps) position is also robust to *tkv* dosage. In Tkv knock-down experiments, pMad expression expands in the anterior compartment, but L2 position never moves outward and away from the Dpp source (Fig. 4.8 C). In some instances, L2 may even move inwards toward the Dpp source (Fig. 4.8 B). Both results show that L2 does not simply rely on the level of pMad to define its location.

However, there are some differences between L2 and L5 formation in response to changes in the pMad gradient. While L5 expression is almost never affected by the genetic perturbations I tested, ectopic expression of Knirps does occur when Tkv levels are low. Within the same experiment of spatiotemporally controlled Tkv-RNAi, Sal expression detected by an antibody is absent in regions lacking BMP activity. This is in agreement with previous descriptions of Sal, that its activation and transcriptional rate comes from BMP signaling, while its boundary is regulated by Brk (Campbell and Tomlinson, 1999; Marty et al., 2000). On the contrary, Omb expression is only regulated by Brk repression and does not require activation by BMP signaling (Winter and Campbell, 2004). Omb expression in Tkv-RNAi experiments is only downregulated in

instances where BMP signaling is completely absent and Brk expression is expressed in the entire region.

Interestingly, gaps of missing Sal expression in medial region of the disc can induce ectopic expression of Knirps (Fig. 4.8 B'', C''). It is likely that those regions contain high expression of *brk* as seen in other discs where *tkv* levels are knocked down. Knirps expression also seems to closely follow borders of Sal in the anterior compartment; any curves or slants of Sal will induce curves in Kni expression (Fig. 4.8 B''). To confirm that these results of ectopic Kni expression is not due to over-expression of Gal4, I stained for Kni in ApGal4>UAS-GFP animals and found normal expression pattern of knirps (Fig. 4.8 D). While it is hard to tell whether Sal expression expands where pMad gradient expands, these results show that Kni expression follows closely that of its proposed upstream signal. This results argues against the theory that brk represses knirps expression and may favor the theory that *knirps* is induced by low levels of Sal expression.

A separate observation I made is when Tkv-RNAi is so highly expressed that it knocks down all Tkv, and thereby prevents all BMP signaling. *Sal* is paradoxically highly expressed everywhere. It is possible that cells under these conditions are re-capitulating *sal* expression that normally occurs in the periphery of the disc, which is know to be independent of BMP regulation (Campbell and Tomlinson, 1999; de Celis et al., 1999).

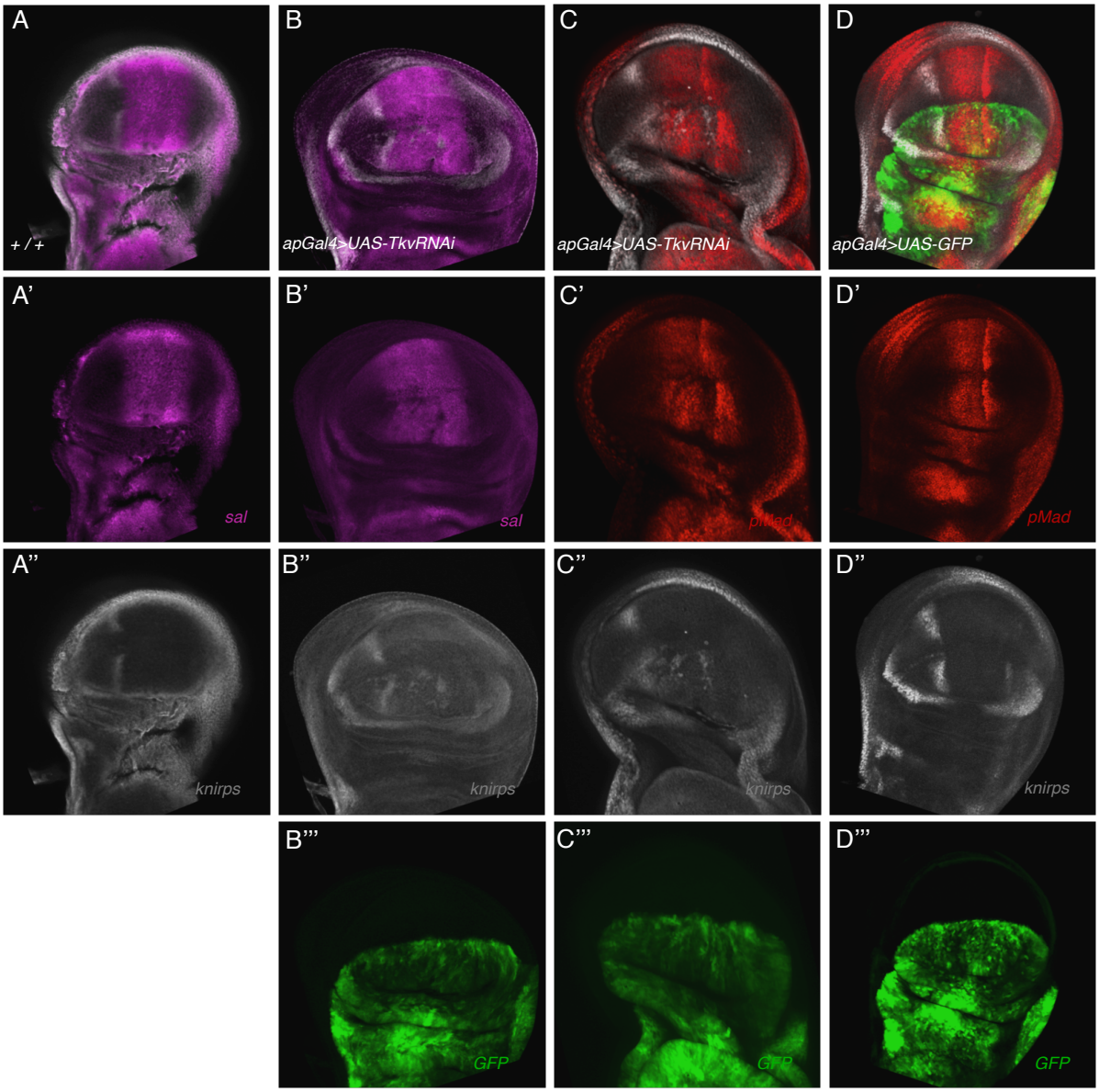


Figure 4.8 Knirps expression closely follows Spalt border. (**A**) Wild type control of Sal and Knirps expression. (**B-C**) *ApGal4; tubGal80^{ts}>UAS-Tkv-RNAi* animals reared at 25.5°C starting from 1st instar. Knirps expression follows Sal expression closely when Sal expression is slanted (**B''**). In regions of low Sal expression, even near central region of the disc, Knirps is abnormally expressed (**B'',C''**). Sal requires pMad expression to be activated. In regions of low pMad expression, Sal expression is downregulated. (**D**) Negative control of Gal4 induced knirps expression, Kni is not expressed abnormally in regions of Gal4 expression.

Value and Slope Hypothesis

So far, numerous pieces of evidence point to a disconnect, of sorts, between Omb and L5 expression, from the level of Dpp signaling. This suggests that the current model of L5 specification needs to be revised. Other reasons to consider revising that model are that the Omb expression domain does not actually have a sharp posterior border, declining gradually in transcriptional rate (Chapter 3) and protein expression (Shen et al., 2010), and the fact that the short range signal proteins proposed to be involved in the induction of vein primordial have yet to be found. I therefore considered whether alternative mechanisms of morphogen interpretation could explain the existing data.

A first hypothesis was inspired by a study of growth regulation in the *Drosophila* wing disc. Dpp regulates growth by two distinct mechanisms: one depends on the concentration level of Dpp, and the other depends on the slope of the gradient (Rogulja and Irvine, 2005). It has been proposed that these two distinct mechanism act in conjunction to create a relatively even distribution of cell proliferation, in which the gradient response mechanism promotes proliferation proportional to the slope of the gradient. Apparently, the slope of the Dpp gradient influences growth by modulating the activity of the Fat signaling pathway. Both known regulators of Fat, Four-jointed (Fj) and Dachsous (Ds, are expressed in gradients in the developing imaginal discs, and both are regulated by Dpp signaling (Rogulja et al., 2008).

Evidence from genetic experiments show that Fat pathway genes may also affect wing vein patterns and not just growth. Strong allelic combinations of a downstream

target of Fat, dachs, result in a fusion of veins; the knockdown of Dachs by RNAi or ubiquitous over-expression of Dachs lead to the appearance of a segment of extra vein parallel to L2 (Mao et al., 2006). Uniform expression of Ds also results in extra vein segments parallel to L2 (Rogulja et al., 2008).

Given that the Fat signaling pathway appears to have effects on vein specification, I proposed an alternative mechanism of morphogen gradient interpretation that includes input from both the value and slope of the morphogen gradient, the latter measured via the Fat pathway as in Rogulja et al. 2008b. In this case, *brk* expression would depend on the value of the BMP gradient and receive little or no contribution from the slope; whereas *omb* would combine information from both the value and slope, still other targets could potentially even get the majority of their information from the slope (Fig. 4.9).

Since cells get local slope information through the Fat signaling pathway, which is dependent on the gradient of dachsous expression, I wanted to directly test what would happen if the ability to measure slope information was disabled in regions near vein primordium. Using a system of drug inducible Gal4 flp-out clones (See Methods), I generated clones of ubiquitous *ds* and *fj* cells, effectively removing the cells' ability to measure slope within the clone, but creating a dramatic slope at the border of the clones (Fig 4.10 and 4.11).

In clones of ubiquitous *ds* expression, I first showed that pMad is unaffected as expected, so value information from the morphogen gradient should stay the same (Fig. 4.10 B). Next, I examined the effect of clones on three different provein markers, Delta,

Ab, and Kni. Clones of UAS-ds generated cell-autonomous ecotopic expression of Delta, regardless of their location (Fig 4.10 C). Ab expression in UAS-ds clones also exhibited ectopic expression both cell autonomously and near the border on the anterior side. In the posterior compartment, UAS-ds clone pushed expression of Ab toward the Dpp source mostly within the clone (Fig 4.10 D). Kni expression in UAS-ds clones near the normal Kni-expression region were unaffected and showed no change in expression (Fig. 4.10 E).

In parallel experiments making clones of UAS-fj, I also observed no effect on pMad signaling (Fig 4.11 A). Delta expression in clones of UAS-fj animals was elevated only within clones and did not seem to respond to changes in slope information at the boundary of the clones (Fig 4.11 B-C).

These experiments showed that Delta and Ab expression can be affected locally by Fat pathway signals, but they did not support the value-and-slope hypothesis, at least in its simplest form. According to that hypothesis, I would have expected to see see vein genes bend away from the clone of ubiquitous Ds and Fj clones laterally. This is because although these clones removed the cell's ability to measure slope, they created an artificial dramatic slope measurement at the border of the clones. I would have expected a lot non-cell-autonomous effects from these clones. Clones genes sensitive to slope measurement would respond to this artificial slope by moving away further from the morphogen source. Since they would now have a big input from slope, and they would achieve the same positional information at a lower value of concentration

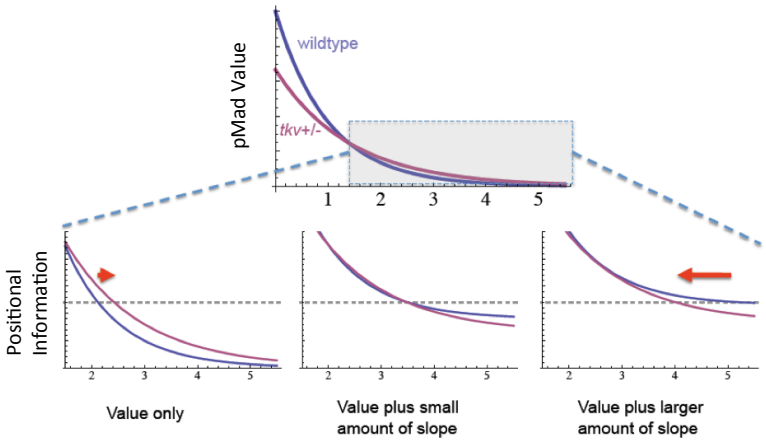
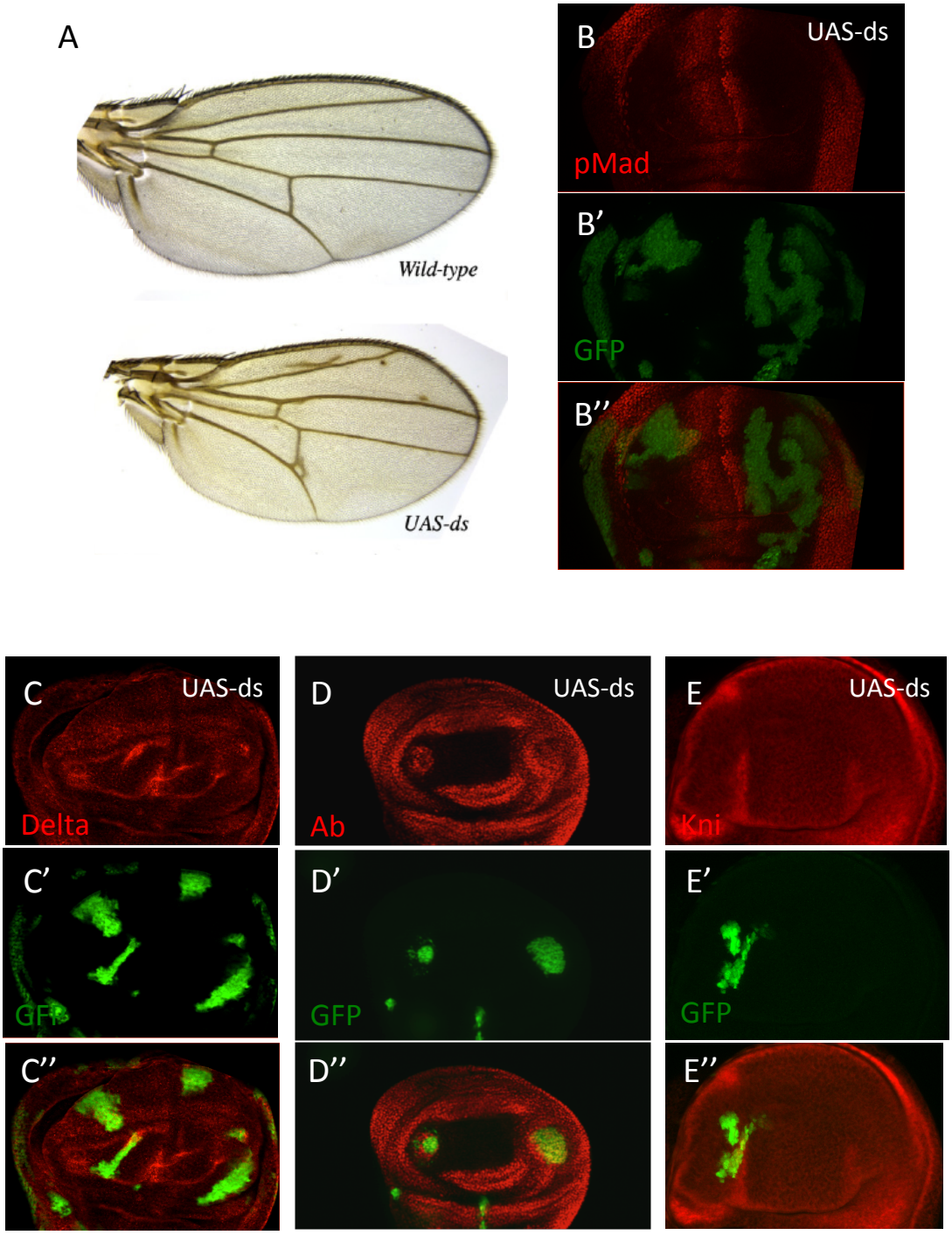


Figure 4.9 Slope and Value hypothesis. Top panels shows the expansion of the pMad gradient as seen in *tkv*$^{+/-}$ mutants. Lower panels show three possible scenarios of combining value and slope using the curve from upper panel. The dashed line represents the hypothetical positional signal. When only value is interpreted, the position moves to the right (as seen in *brk*); when value plus small amount of slope is used, the positions stays put (as seen in *omb*); when value plus larger amount of slope is interpreted, the position moves to the left.

Figure 4.10 Clones of ubiquitous ds expression do not provide evidence that gradient slope controls patterning. (**A**) Wild type and UAS-ds adult wings, modified from (Rogulja et al., 2008). (**B**) Clones of UAS-ds does not affect pMad gradient. (**C**) Delta expression is upregulated mostly cell-autonomously in clones of UAS-ds. (**D**) Abrupt expression is pushed toward the Dpp source within the clone of UAs-ds cells. Anterior ectopic expression of Abrupt is also mostly within the clone. (**E**) Kni expression is unaffected by UAS-ds clones.


A
Wild-type
UAS-ds
B
UAS-ds
pMad
B'
GFP
B''
C
UAS-ds
Delta
C'
GFP
C''
D
UAS-ds
Ab
D'
GFP
D''
E
UAS-ds
Kni
E'
GFP
E''

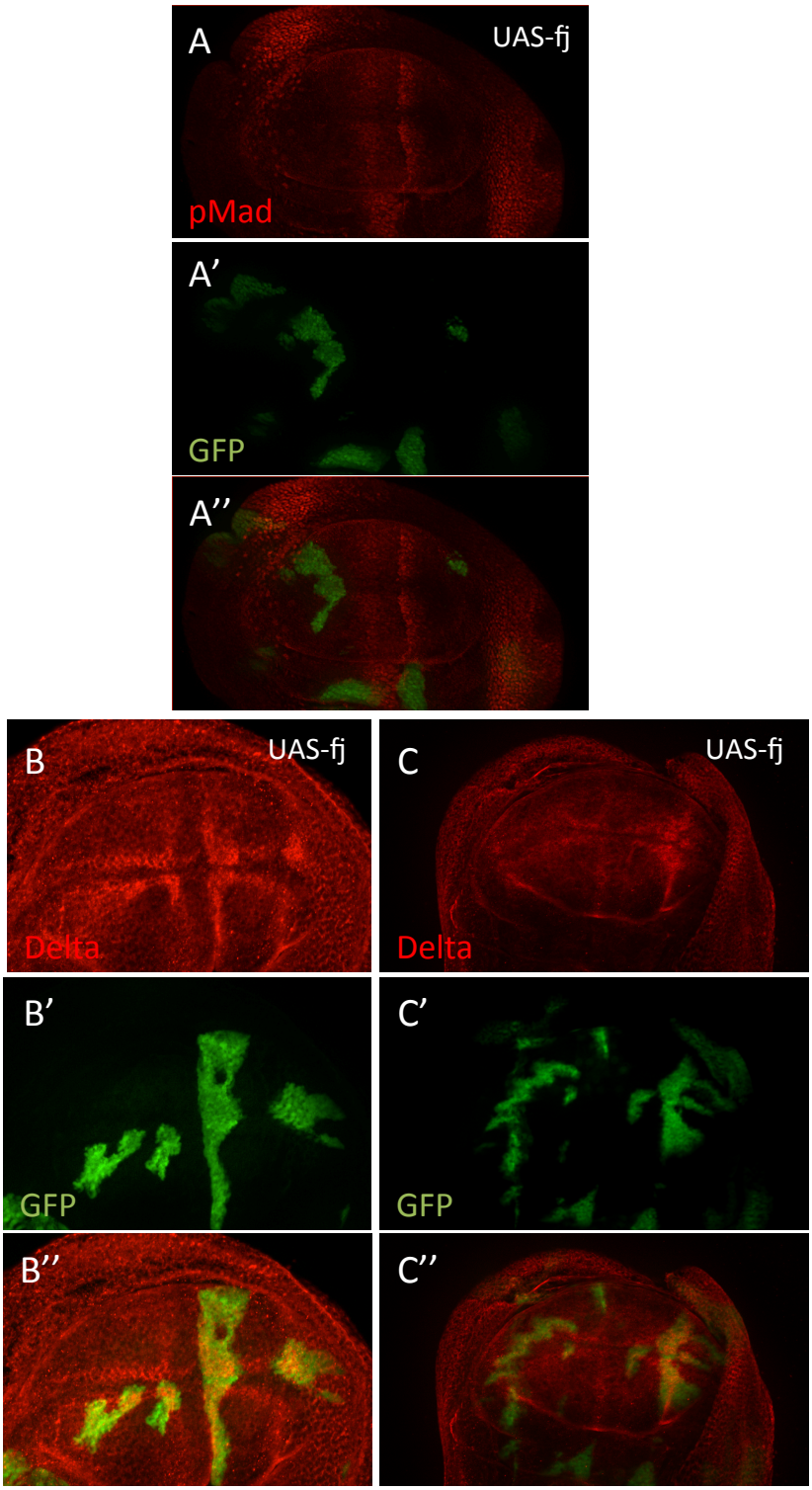


Figure 4.11 Clones of ubiquitous fj expression show autonomously increased Delta expression. (**A**) Clones of UAS-fj does not affect pMad gradient. (**B-C**) Delta expression (red) is upregulated cell-autonomously in clones of UAS-fj.

pMad gradient dynamics during growth

Drosophila imaginal discs are widely thought to exemplify a morphogen system at steady state. Discs grow over a period of several days, and cells in the disc divide with an average cell-cycle length of ~8 hours (Milan et al., 1996), much longer than the times associated with morphogen diffusion or decay. Most studies in the wing disc focus on the last stretch of development, in the late 3rd instar larvae, assuming that, since the morphogen gradient is at steady state, the relevant patterning information will be apparent then.

Recently, some studies have investigated how Dpp and pMad change quantitatively over time. One study, which used a Dpp-GFP fusion protein expressed under Dpp regulating elements, found that both the amplitude and decay length of the Dpp gradient increases significantly during growth (Wartlick et al., 2011). Another study quantified pMad gradients from various developmental times, and found that pMad profiles also expand as the disc grows (Hamaratoglu et al., 2011).

If pMad expands, and its amplitude increases over time, I wondered how downstream targets of Dpp might respond to a changing BMP activity gradient, and how that response would be related to the absolute or relative concentration of morphogen. I therefore used an Omb-LacZ enhancer trap line, and co-stained β-gal and pMad, in wild type and *tkv*$^{7/+}$ wing discs ranging from early 3rd instar to late 3rd instar larvae. As in the first part of this chapter, I extracted pMad concentration profiles from the posterior compartment using averages of orthogonal Z-slices, to avoid contamination from signals in the peripodial membrane, and to follow curvature of the disc proper cells. I then fit the

profile to an exponential decay function $I_{pMad} = I_0 e^{-x/\lambda}$ to obtain a characteristic length scale λ (Fig. 4.12). I quantified that pMad concentration profiles in early 3rd instar discs (Fig. 4.12A) and confirmed that they expand as the disc grows (Fig. 4.12B). Two unexpected observations were made: first, I found that very early gradients tended to be much more linear (i.e. less exponential) in shape than at later stages. Second, I found that the decay lengths of pMad gradients do not continue to increase throughout larval development, but actually decrease during mid-3rd instar, and stabilize by late 3rd instar (Fig. 4.14). Similar analyses were performed with *tkv*$^{7/+}$ animals, and again I observed a similar trend of pMad gradient contraction and stabilization (Fig 4.13A-D). Over the entire developmental time, *tkv*$^{7/+}$ animals maintain a slightly higher pMad length scale compared to the wild type of similar disc sizes (Fig. 4.14).

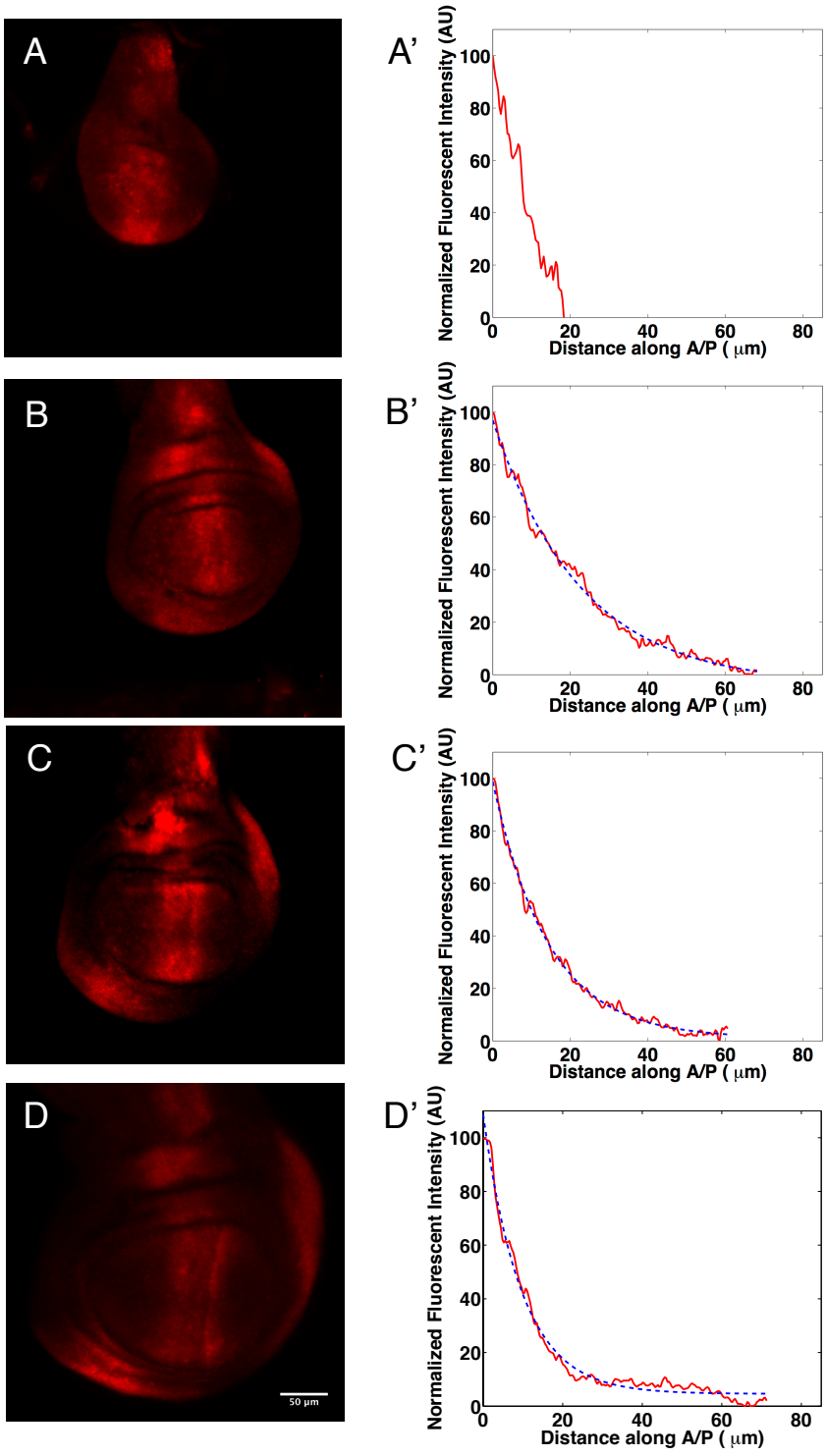


Figure 4.12 Wild type pMad profiles through developmental time. Representative wild type wing imaginal discs with pMad immunostaining from early to late 3rd instar (**A-D**), using wing disc pouch as a proxy for developmental time. (**A'-D'**) Extracted pMad profiles in the posterior compartment in red line and exponential fit in dotted blue line.

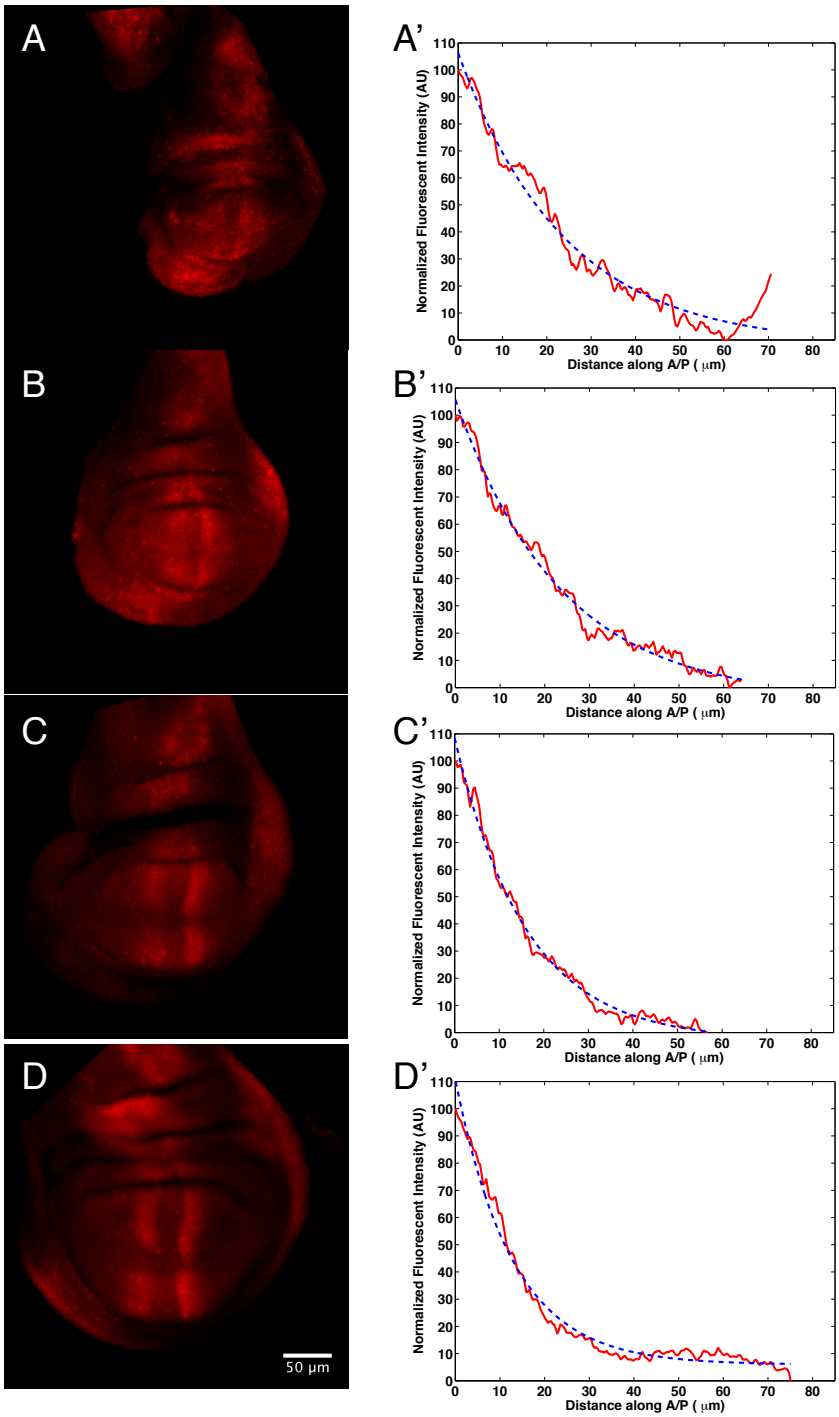


Figure 4.13 *Tkv*$^{7/+}$ pMad profiles through developmental time. Representative *Tkv*$^{7/+}$ wing imaginal discs with pMad immunostaining from early to late 3rd instar (**A-D**), using wing disc pouch as a proxy for developmental time. (**A'-D'**) Extracted pMad profiles in the posterior compartment in red line and exponential fit in dotted blue line.

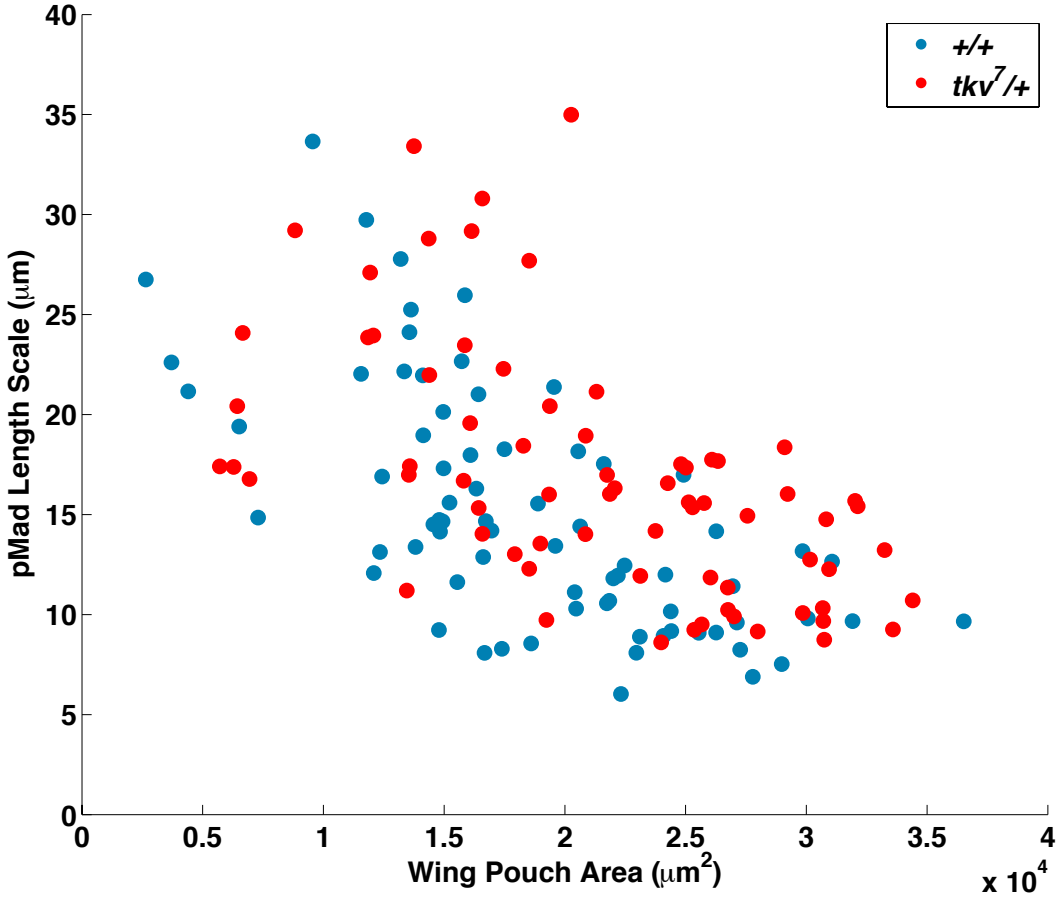


Figure 4.14 pMad length scale contracts and stabilizes as disc size increases. Individual pMad profiles from wild type (blue) and *tkv*$^{7/+}$ (red) animals were fitted with an exponential function to extract its pMad decay length scale, which is then plotted against manually extracted wing pouch area. Wing pouch area can be viewed as a proxy for developmental time as the size of the disc increases exponentially with time. Value for pouch area < 1000 μm^2 are less accurate because they were obtained by fitting an exponential function to gradient that are not very exponential in shape.

Omb response to pMad

The observation that the pMad gradient length scale of the wing disc falls, from about mid-3rd larval instar onward, was unexpected (although, in retrospect, hints of such behavior can be found in data reported by others, e.g. Fig. 1A of (Wartlick et al., 2011)). To maintain the location at which any threshold of pMad is crossed at a constant absolute position, it would be necessary for the amplitude of the pMad gradient to rise considerably; to fix such a threshold to a constant relative position (as the disc grows), the amplitude would need to rise even more. For example, for a pMad length scale that declines from 20 μm to 10 μm as the wing pouch grows from 1.7 μm^2 to 2.7 μm^2 (as in Fig. 4.14), to keep pMad constant at a relative position corresponding to 50 μm from the Dpp source at the end of growth, it would be necessary to increase the amplitude of the gradient by nearly 150-fold. Although it has been reported that Dpp signaling rises over time during disc development, such increases are at most a few fold.

Accordingly, I would expect that, as pMad gradients contract, so should domains of target genes that act as a direct readout of Dpp signaling. To test this, I co-stained for pMad and *omb-LacZ* (a transcriptional reporter for *omb*) in growing discs. Although *omb-LacZ* is not a perfect readout for *omb*, (it displays a somewhat sharper boundary than the omb protein profile (Shen et al., 2010)), it still provides a good indication of how far away Omb is expressed from the Dpp source. I extracted the *omb-LacZ* profile as a function of distance to the Dpp source, using the same cut-off of maximum expression

of pMad to mark the end of Dpp source (Fig. 4. 14D). I then used fitting to a sigmoidal function $s(x) = s_0(\tanh((d-x)/w)+1)/2$ to find the half-maximal values of each profile.

Interestingly, I observed that the position of the *Omb* boundary increases approximately linearly with the area of wing pouch (Fig. 4.15E), in both wild type and *tkv*$^{7/+}$ animals. Because this occurs at the same time that the pMad gradient length scale is decreasing, it results in an inverse correlation between omb border position and pMad length scale (Fig. 4.15F). Thus, not only are the final positions of vein primordia relatively insensitive to manipulations that alter the final form of the pMad gradient (Fig. 4.5-4.7), the time-evolution of the *omb* expression pattern, which determines the position of the L5 vein primordium, is very different from that of the pMad gradient.

One possible explanation for this observation is that the gradient of positional information specified by Dpp is not truly at steady state. On a slowly growing domain, such as a wing disc, fast diffusion and uptake of Dpp only ensure that the gradient of *free* Dpp may be considered to be at a steady state. If bound Dpp, or signals downstream of Dpp, turn over at a rate that is sufficiently slow compared with the displacement of cells due to growth, the shapes of their gradients will not be at steady state, and could be very different from that of free Dpp. For example, for a molecule such as Omb, which turns over with an apparent half-life of approximately 6 hr (Chapter 3), and assuming a cell cycle time of ~10 hr, we may expect that the omb within a single cell 50 μm from the center of the disc will be advected by 7.5 μm within a single half-life.

This issue was recently explored by Iber, using mathematical modeling (Fried and Iber, 2014), in which the standard reaction-diffusion equation that is commonly used

to model morphogen gradients on stationary domains was modified to account for both advection (the movement of molecules due to the growth-driven movement of the cells) and dilution (decrease in concentration due to net growth).

By extending this analysis to account explicitly for successive levels of morphogen signaling (e.g. bound Dpp, pMad, downstream target genes), it is possible to explore the predicted spatiotemporal differences in the time-evolution at different stages in the morphogen signaling pathway. Fig. 4.16 displays typical results in which the concentrations of free morphogen (A,D), morphogen bound to receptors that turn over quickly (or, alternatively, quickly-decaying signals that are produced by morphogen-bound receptors)(B,E), and a long-lived signal, with an observed half-life on the order of that of Omb (C,F) are displayed as kymographs, using heatmaps to show the positions of thresholds of concentration. Notice that the length scales of the three different components evolve differently, with faster-decaying components failing to spread out as effectively as slower-decaying ones (in simulations that include the regulation of receptor expression by morphogens, the decay lengths of faster-decaying signals may even retract).

Not only do such simulations provide a possible explanation for the lack of concordance between the evolution of pMad and Omb profiles (Fig. 4.15), they also recapitulate the different responses of pMad and downstream signaling molecules to manipulation of tkv expression. As shown in Fig. 4.16E-F, under conditions in which pMad (or other short term signals) become expanded by changes in receptor expression, the profile of Omb (or other long term signals) can remain effectively

unchanged. The reason for such robustness is not just that the dynamics of long-term signals are slower, but that, in the presence of reduced receptor levels, gradients progress less far along their dynamical trajectories.

Figure 4.15 Omb-lacZ response to pMad. (**A-B**) Wild type wing disc immunostainings for pMad (**A**, green) and omb-lacZ (**B**, red). (**C**) Normalized fluorescent intensity of pMad in red is fitted with an exponential function in blue dotted line. (**D**) Normalized fluorescent intensity of Omb-Lacz, used to determine the Omb distance from Dpp, "d", by fitting the sigmoidal function $s(x) = s_0(\tanh((d-x)/w)+1)/2$, blue dotted line. (**E**) Omb distance d from Dpp plotted against wing pouch area, wild type in blue dots and $tkv^{7/+}$ in red dots. (**F**) Omb distance d from Dpp plotted against the pMad length scale from the same imaginal disc, wild type in blue dots and $tkv^{7/+}$ in red dots.

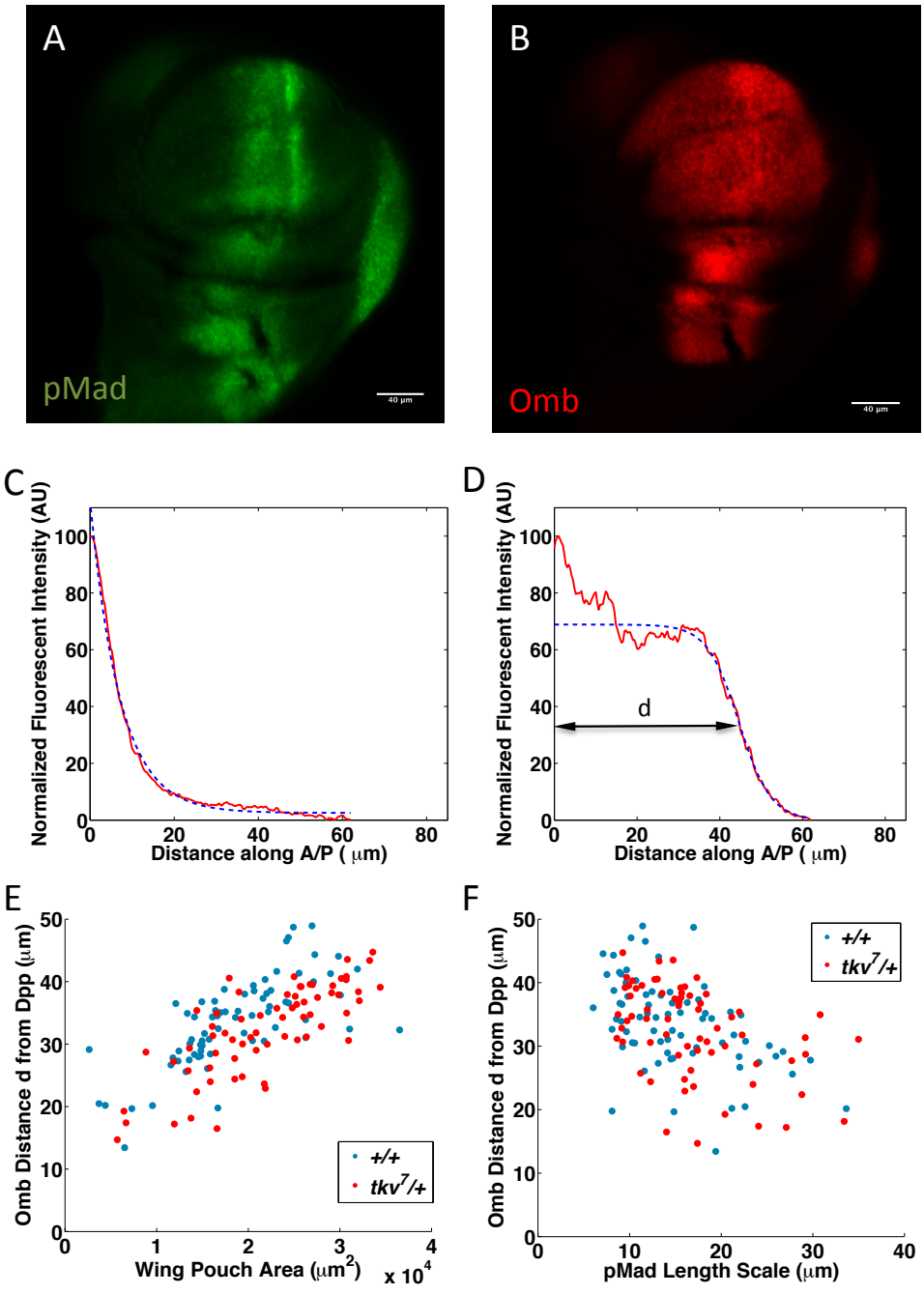
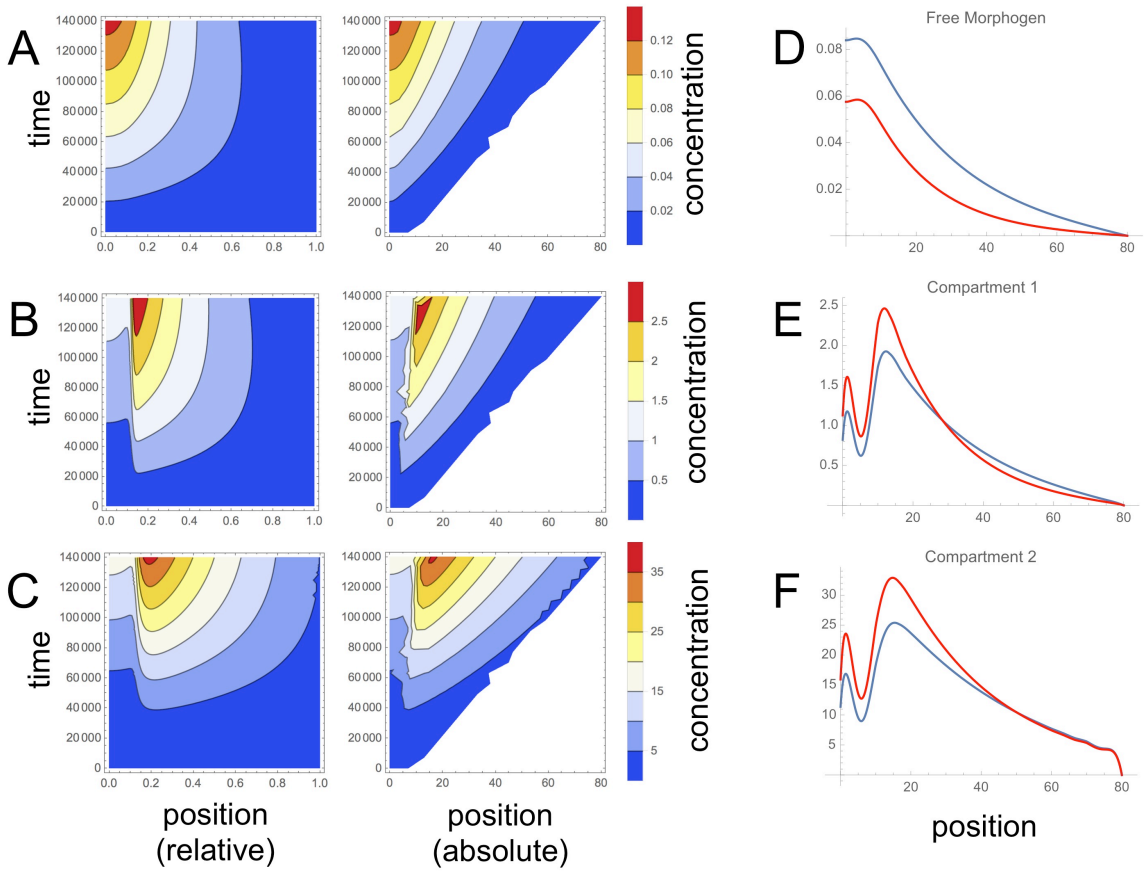

A
pMad
40 µm
B
Omb
40 µm
C
Normalized Fluorescent Intensity (AU)
Distance along A/P (µm)
D
Normalized Fluorescent Intensity (AU)
d
Distance along A/P (µm)
E
Omb Distance d from Dpp (µm)
+/+
tkv7/+
Wing Pouch Area (µm2)
x 10^4
F
Omb Distance d from Dpp (µm)
+/+
tkv7/+
pMad Length Scale (µm)

Figure 4.16. Modeling morphogen gradients and downstream signals on a growing domain. The evolution of a morphogen gradient on a linearly growing domain was modeled as described (Fried and Iber, 2014), except that an explicit production region was incorporated (representing the leftmost 12% of the total domain) that produced morphogen at a constant rate, and receptors were explicitly represented as immobile, non-saturable sites (and which were present at reduced levels in the morphogen production region). The total time simulated was approximately 39 hours. **(A-C)** depict kymographs of gradient evolution, in which time (in seconds) runs from bottom to top, position from left to right, and concentrations are given by heat maps at left. In the left panels positions are scaled relative to the gradient field; in the right, true positions are given (in μm). Panel A shows results free morphogen, panel B for bound morphogen; this plot would also represent the profile for any signaling molecule that turns over quickly relative to receptors (the receptor half-life was taken to be 23 minutes). Panel C shows the time evolution of a species that is produced in response to receptor-bound morphogen and has a half-life of 9.6 hours. Note that, due to effects of dilution, the observed half-life of such a species would be about 6 hours, similar to that measured for Omb in Chapter 3. (**D-F**)**.** Final positions of the gradients in A-C are shown for two receptor concentrations. Blue curves are simulations in which the receptor concentration was halved compared to the red curves. Notice that the gradient of bound morphogen (and short-term signaling) in panel F is expanded by halving receptor concentration, as observed for pMad in $tkv^{7/+}$ wing discs. In contrast, the gradient of the long-lived signal (panel F) is barely altered. A position specified by this signal would be robust to changes in receptor expression.


A
B
C
time
position (relative)
position (absolute)
concentration
D
E
F
Free Morphogen
Compartment 1
Compartment 2
position

CHAPTER 5

Discussion

Precision of morphogen interpretation

Morphogen gradient orchestrated pattern formation has been a topic of research for several decades, and one of its most serious objections has to do with how reliable it is. It would seem that, with the amount of variability coming from every angle, developing organisms are all faced with the tall task of discerning noise from signal. Yet, somehow, most systems emerge from a single embryo with remarkably complex patterns all in the right place.

Little is known about the actual levels of noise in morphogen gradient responses, how the noise propagates, and what mechanisms are available to control it. To investigate this, I quantified one downstream target of Dpp in the *Drosophila* wing disc. Omb is a transcription factor that plays an important role in specifying the location of the vein primordium L5. I combined FISH with image segmentation and image analysis to quantitatively measure the rate of transcription in *omb*. I found that the mean transcriptional rate decreased in a spatial gradient similar to that of the Omb protein, high in the medial region of the disc near Dpp production decreasing towards both posterior and anterior ends. While *Drosophila* embryo has reported an instantaneous transcriptional noise of about 50%, I found that with an averaging time of 18-24 minutes, *omb* transcriptional noise starts around 70% in regions of high expression, and increases all the way to around 110% in regions where Omb was expected to induce L5. In the embryo where *hb* noise was measured, the authors saw a noise level

characterized by CV of 50% for each transcription loci. *Hb* is a short gene of 3.2kb, and this would mean that their measured noise has an averaging time of roughly 2 minutes. If we assume that transcription were Poissonian and not bursty, we would estimate the CV of omb with an averaging time of 2 minutes to be 300%, rough 6 folds higher than that detected in embryo.

However, it important to note that the measure noise is a component of both input noise and extrinsic noise, so it makes sense that the transcriptional noise in wing disc is much higher than that of the embryo. It most likely has a much bigger input noise. The first three hours of embryo development is crucial to its survival and tightly controlled transcriptional events is understandable for two reasons. First, input noise in the measured hb transcription is likely low, as previous studies have found Bcd expression to be very precise, with Bcd concentration varying by less than 10% in comparable regions (Dubuis et al., 2013; Gregor et al., 2007). Second, *Drosophila* embryogenesis takes places with a built-in spatial noise filter in the form of a syncytium embryo (Sauer et al., 1996). The process of diffusion, receptor binding and activation, and the relief of repression through brk cis-regulatory element, all contributes to the input noise of *omb.*

Next, I measured the protein noise in Omb and found it to be much lower. Just as previously reported, Omb protein is expressed in a gradient with a noise slightly over 20% in regions where it is highly expressed, increasing only near the end of the gradient to around 35% in regions of Abrupt expression. While the reduction in noise from transcription to protein expression is high, I found Omb half-life to be quite long, around

6 hours. If I assume translation is Poissonian, this can mean that the long half-life of Omb can provide a noise reduction of 4 fold.

I measured the Omb local length scale to be 8μm, and with a noise level of 35%, I believe it is precise enough to place a stripe with a transition width of around 6μm. It turns out that the transition width of Abrupt, the L5 organizing gene, is around 7μm in late 3rd instar, which suggests that Omb has the potential to set a stripe of roughly the same precision.

I have found that in the *Drosophila* wing disc, even though transcriptional noise is very high, the stability of the protein is able to provide enough temporal noise integration to bring the noise level down in protein expression. The level of protein precision in the disc is sufficient to lay down a stripe in the same order of magnitude as the width of the vein primordial.

Furthermore, our measured length scale of pMad in wild type animals is quite shorter than previously described, at slightly over 11μm. This makes the interpretation of positional information farther away from the source difficult, as the absolute level of pMad becomes very low while the noise increases relative to the signal. One recent study points to an increase in amplitude of both Dpp and Daughters against Dpp (Dad) during wing development (Wartlick et al., 2011), which can provide some help with signal to noise ratio, but may cause problems in interpreting the gradient using a purely local threshold value theory, since the local thresholds for Dpp downstream target genes would also need to increase over time to maintain proportional patterning. Although concentrations of actual BMP activity gradient of pMad have not been carefully

quantified, several studies show that pMad does not increase significantly over time (Hamaratoglu et al., 2014; Hamaratoglu et al., 2011). This difference between Dpp and pMad dynamics might be explained by the known negative feedback loop formed by the Dpp target Dad (Tsuneizumi et al., 1997). An increase in Dpp signals translates to an increase of Dad signal and consequently, the stronger inhibition of pMad (Ogiso et al., 2011).

Robustness of morphogen interpretation

In the 2nd part of the thesis, I found that the downstream target of Dpp responds differently to changes in the morphogen gradient. Using quantitative measurements from immunostained markers, I found that *brk* responded to the level of morphogen concentration, moving in the same direction as pMad, but L2 and L5 primordium are robust to pMad changes. Using a qualitative approach, I was able to show the robustness of L5 position is by altering Dpp receptor levels in half of the disc, dramatically affecting the BMP activity gradient.

Furthermore, the pMad gradient in the 3rd instar wing disc seemed to change over time. Through development of the 3rd instar, pMad first expands, then contracts, and finally stabilizes to a decay length of around 11μm. This was not fully explored before, although some evidence existed in previous studies (Wartlick et al., 2011). In figure 1E of Wartlick et al. 2011, for example, the decay length of Dpp gradient rose for the first 120 hours after egg lay, then showed signs of decrease between 120-140 hours after egg lay. My own measured used area of the wing pouch as a proxy for time, and

could go up to 140+ hours. One possible explanation of why pMad gradient actually contracts could be due to the slowing growth in the disc as it prepares to enter into the pupal stage. If Dpp production slows down, pMad level will decrease and thus shrink the gradient. Dad may also be affecting pMad, especially if it also happens to be slow varying. As Dad catches up to pMad, it represses pMad signaling and shrinks it toward the production region.

The fact that we see that the Omb border increases linearly with the size of the disc provides even more evidence that *omb* does not simply mimic the morphogen signal, and confirms that Omb indeed scales with the size of the disc, as other studies have reported (Hamaratoglu et al., 2011). This antagonistic movement of pMad and Omb border during mid to late 3rd instar larvae actually creates an inverse relationship between them (Fig. 4.15F). While I propose this is due to the combined effect of advection, dilution and a long half-life, it is also possible that other mechanisms of feedback regulation in the wing disc plays a more important role in Omb regulation.

Stability in Omb provides precision and robustness

Having a very stable protein in Omb proves to be crucial for pattern formation in two ways. First, being a very stable protein in cells with very long cell cycles provides significant amounts of noise-filtering capabilities. Because wing disc develops over a period of several days, it is inevitably exposed to environmental fluctuations in temperature and food source, and a global strategy for organisms to dampen the effect of noise, is to provide temporal averaging through stable proteins.

Second, because Omb has a long half-life, its border expansion during growth is slower than that of the pMad, and is essentially reporting a time lagged version of the pMad gradient. The protein itself is also being diluted and advected so that there is some level of spatial averaging of the signal. This can help to reduce the effect from receptor concentration, even though having less receptors eventually leads to a more expanded pMad gradient, the time lagged version of this gradient is much closer in concentration than the wild type.

Future Directions

Is the transcriptional dynamics of omb bursty?

Stochastic transcription seems to be a universal process. It has been found in both single cell organism grown in cultures (Raj et al., 2010; So et al., 2011; Stewart-Ornstein et al., 2012) as well as in different developmental settings (Little et al., 2013; Paré et al., 2009; Saffer et al., 2011).

Early studies of transcription often described the production and degradation of mRNA according to the statistics of a Poisson process. That is, the probability of a transcript being produced within any given time period is constant. However, experimental studies have found evidence that mRNA production does not occur with a constant probability in time, but rather with increased probability at certain time periods over others, often referred to as transcriptional bursts. Although I was able to conclude that transcriptional noise in *omb* is quite high, I did not have enough information to determine whether or not *omb* transcription is bursty. Experimentally distinguishing

bursty transcription from Poissonian transcription requires transcriptional rates to be measured in molecular units of mRNA per cell. This is because variability scales as a function of the mean mRNA number. In the Poissonian model, as the mean increases, the relative variability decreases with the mean. In transcriptional bursting, there can be high variability even with high mean expression levels. When measurements are made in arbitrary fluorescent units, it is hard to differentiate between Poissonian models of transcription at low levels from bursty transcription with high means. Fano factor, defined as the ratio of variance to the mean, is usually used to determine Poissonian models from bursty models. When measurements are in molecular units, a Fano factor of 1 describes a Poisson distribution, and any Fano factor higher than 1 can be described as bursty (Paulsson, 2004; Raj and van Oudenaarden, 2009). Unless I am able to confer number of mRNA molecule per range of fluorescent units, I would not be able to make arguments of *omb* transcriptional dynamic.

Self-organized pattern in wing vein

Several mechanisms of a combined system, involving both the self-organized reaction diffusion-style pattern formation and the morphogen mediated positional information, have been proposed. One way that reaction diffusion can contribute is by acting upstream of the morphogen, where Turing pattern only has enough space to create a single stripe. In the case of left-right patterning, where Nodal and Left2 molecules constitute a Turing system that has only enough space to make half a

wavelength, it creates a perfect gradient that theoretically could be used for further interpretation of positional information (Green and Sharpe, 2015).

Alternatively, reaction diffusion can also act in concert with a morphogen gradient, in that cells integrate information from both systems simultaneously. This can be seen in current models of digit specification in the vertebrate limb. In the limb bud, a reaction diffusion system produces periodic patterning along the anterior posterior axis that can be used as information to distinguish digit from inter-digit. At the same time, a morphogen gradient is also present in the system that provides positional information to each digit so that they can adopt different identities (Raspopovic et al., 2014; Sheth et al., 2012).

A third scenario is one where the reaction diffusion acts downstream of the morphogen system. Such systems can use the information provided by the morphogen gradient to create different wavelengths of the reaction diffusion pattern.

A hint that wing vein primordium might also be involved in self organized pattern formation is the fact that the borders of Ab, the L5 organizing gene, sharpens over time, while the upstream regulator of L5, *omb*, does not seem to respond to threshold levels in the morphogen gradient system. It is possible that during early or mid 3rd instar stages, borders of *omb* expression, or regions of low *omb* expression, first induces a broad expression of Ab, that refines itself overtime. In order to truly figure out whether self-organization is part of the vein induction mechanism, direct evidence of gene interaction is necessary. A good starting point would be EGF related proteins expressed

in the wing pouch, as several experiments have shown vein phenotypes when EGF proteins were altered.

Some observed phenotypes of extra vein segments in adult wings could also be a result of having a broad pro-vein region. Our work on the slope sensing potential of the wing disc showed that UAS-ds clones induce an elevated level of vein genes. It's possible that in UAS-ds animals, most vein genes are now expressed in slightly broader domains, and the vein-intervein refinement process in the pupal now has enough physical space to make two period waves instead of one, thus resulting in segments of extra vein nearby.

Mechanical forces in the disc

Mechanical forces could play a role in morphogenesis. Even in his seminal paper on morphogens, Turing defined systems undergoing morphogenesis as consisting of "two parts, the mechanical and the chemical." Mechanical stress has been measured in the *Drosophila* wing disc (Nienhaus et al., 2009) and implicated in growth control. Another example of long-range mechanical forces playing a crucial role in patterning is the development of the *Drosophila* wing during the pupa stage. Contraction at the hinge region of wing orients the wing epithelial cells in order to elongate and divide the cells in the same direction, causing the planar polarity of the cells to realign and elongate with the long axis of the developing wing (Aigouy et al., 2010).

Active stress differences in cells can lead to the movement of materials such as the actin cortex (Engler et al., 2006; Mayer et al., 2010), and morphogens embedded in

these materials could potentially be moved along with them. Any movement of morphogens by directed transport with motor proteins, or by division of cells that then spreads morphogen signals, can be considered as a type of “advection.” The effective length scale of combined advection and diffusion depends on the product of the two. When advection moves materials in the opposite direction of diffusion, the new length scale is $\lambda=D/v$, where D is the diffusion coefficient and v is the speed of the advective transport (Howard et al., 2011). Interestingly, Omb is also involved in cell adhesive properties of the disc and organization of epithelial cytoskeleton. Disruption of Omb gradient long term can lead to cells retracting toward the basal membrane and pulling out of the disc proper cells (Shen et al., 2010). These properties of Omb can provide a source of input from mechanical properties of the disc that could potentially be incorporated into cell fate decisions.

Perspectives

The *Drosophila* wing has become an important model system for the study of growth and pattern formation in developmental biology. The remarkable robustness and precision of developing systems suggest that a considerable amount of cellular machinery is devoted to achieve such goal. Multiples genes (i.e. *dpp, fat/dachsous, omb*) and pathways (BMP, EGFR) are often involved in the regulation of several developmental processes, affecting both pattern formation and growth during larval and pupal development, implying that these processes are interconnected. Although considerable efforts have been made in recent years to understand and identify

mechanisms of morphogen interpretation for the encoding and detecting of positional information, or to achieve reliability, there is much we still do not know. There are layers of complexity in the transcriptional and feedback networks, the coupling of growth and pattern, and the regulation through mechanical properties, which could all contribute to the ultimate goal of robustness and precision in development. Tackling these questions would require the exploration of new approaches and methods to understand not just "how," but also "why" developing organisms choose certain specific set of mechanisms for its ultimate objectives.

References

Abouchar, L., Petkova, M. D., Steinhardt, C. R. and Gregor, T. (2014). Fly wing vein patterns have spatial reproducibility of a single cell. *J R Soc Interface* **11**, 20140443.

Aigouy, B., Farhadifar, R., Staple, D. B., Sagner, A., Roper, J. C., Julicher, F. and Eaton, S. (2010). Cell flow reorients the axis of planar polarity in the wing epithelium of Drosophila. *Cell* **142**, 773-786.

Akiyama, T., Kamimura, K., Firkus, C., Takeo, S., Shimmi, O. and Nakato, H. (2008). Dally regulates Dpp morphogen gradient formation by stabilizing Dpp on the cell surface. *Dev Biol* **313**, 408-419.

Alexandre, C., Baena-Lopez, A. and Vincent, J.-P. (2014). Patterning and growth control by membrane-tethered Wingless. *Nature* **505**, 180-185.

Apionishev, S., Katanayeva, N. M., Marks, S. A., Kalderon, D. and Tomlinson, A. (2005). Drosophila Smoothened phosphorylation sites essential for Hedgehog signal transduction. *Nat Cell Biol* **7**, 86-92.

Ardehali, M. B. and Lis, J. T. (2009). Tracking rates of transcription and splicing in vivo. *Nat Struct Mol Biol* **16**, 1123-1124.

Arora, K., Dai, H., Kazuko, S. G., Jamal, J., O'Connor, M. B., Letsou, A. and Warrior, R. (1995). The Drosophila schnurri gene acts in the Dpp/TGF beta signaling pathway and encodes a transcription factor homologous to the human MBP family. *Cell* **81**, 781-790.

Bangi, E. and Wharton, K. (2006a). Dpp and Gbb exhibit different effective ranges in the establishment of the BMP activity gradient critical for Drosophila wing patterning. *Developmental biology* **295**, 178-193.

Bangi, E. and Wharton, K. (2006b). Dual function of the Drosophila Alk1/Alk2 ortholog Saxophone shapes the Bmp activity gradient in the wing imaginal disc. *Development* **133**, 3295-3303.

Bar-Even, A., Paulsson, J., Maheshri, N., Carmi, M., O'Shea, E., Pilpel, Y. and Barkai, N. (2006). Noise in protein expression scales with natural protein abundance. *Nat Genet* **38**, 636-643.

Ben-Zvi, D., Pyrowolakis, G., Barkai, N. and Shilo, B.-Z. (2011). Expansion-repression mechanism for scaling the dpp activation gradient in Drosophila wing imaginal discs. *Curr Biol* **21**, 1391-1396.

Bergmann, S., Sandler, O., Sberro, H., Shnider, S., Schejter, E., Shilo, B. Z. and Barkai, N. (2007). Pre-steady-state decoding of the Bicoid morphogen gradient. *PLoS Biol* **5**, e46.

Blair, S. S. (2007). Wing vein patterning in Drosophila and the analysis of intercellular signaling. *Annu Rev Cell Dev Biol* **23**, 293-319.

Blake, W. J., Balazsi, G., Kohanski, M. A., Isaacs, F. J., Murphy, K. F., Kuang, Y., Cantor, C. R., Walt, D. R. and Collins, J. J. (2006). Phenotypic consequences of promoter-mediated transcriptional noise. *Mol Cell* **24**, 853-865.

Bland, J. M. and Altman, D. G. (1999). Measuring agreement in method comparison studies. *Stat Methods Med Res* **8**, 135-160.
---- (2007). Agreement between methods of measurement with multiple observations per individual. *J Biopharm Stat* **17**, 571-582.
Boireau, S., Maiuri, P., Basyuk, E., de la Mata, M., Knezevich, A., Pradet-Balade, B., Backer, V., Kornblihtt, A., Marcello, A. and Bertrand, E. (2007). The transcriptional cycle of HIV-1 in real-time and live cells. *J Cell Biol* **179**, 291-304.
Bollenbach, T., Kruse, K., Pantazis, P., Gonzalez-Gaitan, M. and Julicher, F. (2005). Robust formation of morphogen gradients. *Phys Rev Lett* **94**, 018103.
Bollenbach, T., Pantazis, P., Kicheva, A., Bokel, C., Gonzalez-Gaitan, M. and Julicher, F. (2008). Precision of the Dpp gradient. *Development* **135**, 1137-1146.
Briscoe, J., Chen, Y., Jessell, T. M. and Struhl, G. (2001). A hedgehog-insensitive form of patched provides evidence for direct long-range morphogen activity of sonic hedgehog in the neural tube. *Mol Cell* **7**, 1279-1291.
Brody, Y., Neufeld, N., Bieberstein, N., Causse, S. Z., Bohnlein, E. M., Neugebauer, K. M., Darzacq, X. and Shav-Tal, Y. (2011). The in vivo kinetics of RNA polymerase II elongation during co-transcriptional splicing. *PLoS Biol* **9**, e1000573.
Cai, A. Q., Radtke, K., Linville, A., Lander, A. D., Nie, Q. and Schilling, T. F. (2012). Cellular retinoic acid-binding proteins are essential for hindbrain patterning and signal robustness in zebrafish. *Development* **139**, 2150-2155.
Campbell, G. and Tomlinson, A. (1999). Transducing the Dpp morphogen gradient in the wing of Drosophila: regulation of Dpp targets by brinker. *Cell* **96**, 553-562.
Carey, L. B., van Dijk, D., Sloot, P. M., Kaandorp, J. A. and Segal, E. (2013). Promoter sequence determines the relationship between expression level and noise. *PLoS Biol* **11**, e1001528.
Chen, H., Xu, Z., Mei, C., Yu, D. and Small, S. (2012). A system of repressor gradients spatially organizes the boundaries of Bicoid-dependent target genes. *Cell* **149**, 618-629.
Cohen-Saidon, C., Cohen, A. A., Sigal, A., Liron, Y. and Alon, U. (2009). Dynamics and variability of ERK2 response to EGF in individual living cells. *Mol Cell* **36**, 885-893.
Cook, O., Biehs, B. and Bier, E. (2004). brinker and optomotor-blind act coordinately to initiate development of the L5 wing vein primordium in Drosophila. *Development (Cambridge, England)* **131**, 2113-2124.
Coulon, A., Chow, C. C., Singer, R. H. and Larson, D. R. (2013). Eukaryotic transcriptional dynamics: from single molecules to cell populations. *Nat Rev Genet* **14**, 572-584.
Cowden, J. and Levine, M. (2003). Ventral dominance governs sequential patterns of gene expression across the dorsal-ventral axis of the neuroectoderm in the Drosophila embryo. *Dev Biol* **262**, 335-349.
Crick, F. (1970). Diffusion in embryogenesis. *Nature* **225**, 420-422.
Dar, R. D., Razooky, B. S., Singh, A., Trimeloni, T. V., McCollum, J. M., Cox, C. D., Simpson, M. L. and Weinberger, L. S. (2012). Transcriptional burst frequency

and burst size are equally modulated across the human genome. *Proc Natl Acad Sci U S A* **109**, 17454-17459.
de Celis, J. F. (1997). Expression and function of decapentaplegic and thick veins during the differentiation of the veins in the Drosophila wing. *Development* **124**, 1007-1018.
De Celis, J. F. (2003). Pattern formation in the Drosophila wing: The development of the veins. *BioEssays : news and reviews in molecular, cellular and developmental biology* **25**, 443-451.
de Celis, J. F. and Barrio, R. (2000). Function of the spalt/spalt-related gene complex in positioning the veins in the Drosophila wing. *Mech Dev* **91**, 31-41.
de Celis, J. F., Barrio, R. and Kafatos, F. C. (1999). Regulation of the spalt/spalt-related gene complex and its function during sensory organ development in the Drosophila thorax. *Development* **126**, 2653-2662.
De Solorzano, C. O., Malladi, R., Lelievre, S. A. and Lockett, S. J. (2001). Segmentation of nuclei and cells using membrane related protein markers. *J Microsc* **201**, 404-415.
Denef, N., Neubuser, D., Perez, L. and Cohen, S. M. (2000). Hedgehog induces opposite changes in turnover and subcellular localization of patched and smoothened. *Cell* **102**, 521-531.
Dessaud, E., Ribes, V., Balaskas, N., Yang, L. L., Pierani, A., Kicheva, A., Novitch, B. G., Briscoe, J. and Sasai, N. (2010). Dynamic assignment and maintenance of positional identity in the ventral neural tube by the morphogen sonic hedgehog. *PLoS Biol* **8**, e1000382.
Dessaud, E., Yang, L. L., Hill, K., Cox, B., Ulloa, F., Ribeiro, A., Mynett, A., Novitch, B. G. and Briscoe, J. (2007). Interpretation of the sonic hedgehog morphogen gradient by a temporal adaptation mechanism. *Nature* **450**, 717-720.
Dosch, R., Gawantka, V., Delius, H., Blumenstock, C. and Niehrs, C. (1997). Bmp-4 acts as a morphogen in dorsoventral mesoderm patterning in Xenopus. *Development* **124**, 2325-2334.
Driever, W. and Nusslein-Volhard, C. (1988a). The bicoid protein determines position in the Drosophila embryo in a concentration-dependent manner. *Cell* **54**, 95-104.
---- (1988b). A gradient of bicoid protein in Drosophila embryos. *Cell* **54**, 83-93.
Dubuis, J. O., Samanta, R. and Gregor, T. (2013). Accurate measurements of dynamics and reproducibility in small genetic networks. *Mol Syst Biol* **9**, 639.
Duester, G. (2008). Retinoic acid synthesis and signaling during early organogenesis. *Cell* **134**, 921-931.
Dyson, S. and Gurdon, J. B. (1998). The interpretation of position in a morphogen gradient as revealed by occupancy of activin receptors. *Cell* **93**, 557-568.
Eldar, A., Dorfman, R., Weiss, D., Ashe, H., Shilo, B. Z. and Barkai, N. (2002). Robustness of the BMP morphogen gradient in Drosophila embryonic patterning. *Nature* **419**, 304-308.
Eldar, A., Rosin, D., Shilo, B. Z. and Barkai, N. (2003). Self-enhanced ligand degradation underlies robustness of morphogen gradients. *Dev Cell* **5**, 635-646.

Eldar, A., Shilo, B. Z. and Barkai, N. (2004). Elucidating mechanisms underlying robustness of morphogen gradients. *Curr Opin Genet Dev* **14**, 435-439.
Engler, A. J., Sen, S., Sweeney, H. L. and Discher, D. E. (2006). Matrix elasticity directs stem cell lineage specification. *Cell* **126**, 677-689.
Entchev, E. V., Schwabedissen, A. and Gonzalez-Gaitan, M. (2000). Gradient formation of the TGF-beta homolog Dpp. *Cell* **103**, 981-991.
Erickson, J. L. (2011). Formation and maintenance of morphogen gradients: an essential role for the endomembrane system in Drosophila melanogaster wing development. *Fly (Austin)* **5**, 266-271.
Femino, a. M., Fay, F. S., Fogarty, K. and Singer, R. H. (1998). Visualization of single RNA transcripts in situ. *Science (New York, N.Y.)* **280**, 585-590.
Ferguson, E. L. and Anderson, K. V. (1992). Decapentaplegic acts as a morphogen to organize dorsal-ventral pattern in the Drosophila embryo. *Cell* **71**, 451-461.
Field, Y., Kaplan, N., Fondufe-Mittendorf, Y., Moore, I. K., Sharon, E., Lubling, Y., Widom, J. and Segal, E. (2008). Distinct modes of regulation by chromatin encoded through nucleosome positioning signals. *PLoS Comput Biol* **4**, e1000216.
Fried, P. and Iber, D. (2014). Dynamic scaling of morphogen gradients on growing domains. *Nat Commun* **5**, 5077.
Funakoshi, Y., Minami, M. and Tabata, T. (2001). mtv shapes the activity gradient of the Dpp morphogen through regulation of thickveins. *Development* **128**, 67-74.
Gao, S., Steffen, J. and Laughon, A. (2005). Dpp-responsive silencers are bound by a trimeric Mad-Medea complex. *J Biol Chem* **280**, 36158-36164.
Garcia, H. G., Tikhonov, M., Lin, A. and Gregor, T. (2013). Quantitative imaging of transcription in living Drosophila embryos links polymerase activity to patterning. *Current biology : CB* **23**, 2140-2145.
Goentoro, L. and Kirschner, M. W. (2009). Evidence that fold-change, and not absolute level, of beta-catenin dictates Wnt signaling. *Mol Cell* **36**, 872-884.
Goentoro, L., Shoval, O., Kirschner, M. W. and Alon, U. (2009). The incoherent feedforward loop can provide fold-change detection in gene regulation. *Mol Cell* **36**, 894-899.
Golding, I., Paulsson, J., Zawilski, S. M. and Cox, E. C. (2005). Real-time kinetics of gene activity in individual bacteria. *Cell* **123**, 1025-1036.
Gould, A., Itasaki, N. and Krumlauf, R. (1998). Initiation of rhombomeric Hoxb4 expression requires induction by somites and a retinoid pathway. *Neuron* **21**, 39-51.
Gould, A., Morrison, A., Sproat, G., White, R. A. and Krumlauf, R. (1997). Positive cross-regulation and enhancer sharing: two mechanisms for specifying overlapping Hox expression patterns. *Genes Dev* **11**, 900-913.
Green, J. B., New, H. V. and Smith, J. C. (1992). Responses of embryonic Xenopus cells to activin and FGF are separated by multiple dose thresholds and correspond to distinct axes of the mesoderm. *Cell* **71**, 731-739.
Green, J. B. and Sharpe, J. (2015). Positional information and reaction-diffusion: two big ideas in developmental biology combine. *Development* **142**, 1203-1211.

Green, J. B. and Smith, J. C. (1990). Graded changes in dose of a Xenopus activin A homologue elicit stepwise transitions in embryonic cell fate. *Nature* **347**, 391-394.
Gregor, T., Tank, D. W., Wieschaus, E. F. and Bialek, W. (2007). Probing the limits to positional information. *Cell* **130**, 153-164.
Guptasarma, P. (1996). Cooperative relaxation of supercoils and periodic transcriptional initiation within polymerase batteries. *Bioessays* **18**, 325-332.
Gurdon, J. B., Harger, P., Mitchell, A. and Lemaire, P. (1994). Activin signalling and response to a morphogen gradient. *Nature* **371**, 487-492.
Gursky, V. V., Panok, L., Myasnikova, E. M., Manu, Samsonova, M. G., Reinitz, J. and Samsonov, A. M. (2011). Mechanisms of gap gene expression canalization in the Drosophila blastoderm. *BMC Syst Biol* **5**, 118.
Hamaratoglu, F., Affolter, M. and Pyrowolakis, G. (2014). Dpp/BMP signaling in flies: From molecules to biology. *Seminars in Cell and Developmental Biology* **32**, 128-136.
Hamaratoglu, F., de Lachapelle, A. M., Pyrowolakis, G., Bergmann, S. and Affolter, M. (2011). Dpp signaling activity requires Pentagone to scale with tissue size in the growing Drosophila wing imaginal disc. *PLoS Biol* **9**, e1001182.
Hamburger, V. (1969). Hans Spemann and the organizer concept. *Experientia* **25**, 1121-1125.
He, F., Wen, Y., Deng, J., Lin, X., Lu, L. J., Jiao, R. and Ma, J. (2008). Probing intrinsic properties of a robust morphogen gradient in Drosophila. *Dev Cell* **15**, 558-567.
Holloway, D. M., Lopes, F. J., da Fontoura Costa, L., Travencolo, B. A., Golyandina, N., Usevich, K. and Spirov, A. V. (2011). Gene expression noise in spatial patterning: hunchback promoter structure affects noise amplitude and distribution in Drosophila segmentation. *PLoS Comput Biol* **7**, e1001069.
Hornung, G., Bar-Ziv, R., Rosin, D., Tokuriki, N., Tawfik, D. S., Oren, M. and Barkai, N. (2012). Noise-mean relationship in mutated promoters. *Genome Res* **22**, 2409-2417.
Howard, J., Grill, S. W. and Bois, J. S. (2011). Turing's next steps: the mechanochemical basis of morphogenesis. *Nat Rev Mol Cell Biol* **12**, 392-398.
Huppert, S. S., Jacobsen, T. L. and Muskavitch, M. A. (1997). Feedback regulation is central to Delta-Notch signalling required for Drosophila wing vein morphogenesis. *Development* **124**, 3283-3291.
Itzkovitz, S., Lyubimova, A., Blat, I. C., Maynard, M., van Es, J., Lees, J., Jacks, T., Clevers, H. and van Oudenaarden, A. (2012). Single-molecule transcript counting of stem-cell markers in the mouse intestine. *Nat Cell Biol* **14**, 106-114.
Jazwinska, A., Rushlow, C. and Roth, S. (1999). The role of brinker in mediating the graded response to Dpp in early Drosophila embryos. *Development* **126**, 3323-3334.
Jia, J., Tong, C., Wang, B., Luo, L. and Jiang, J. (2004). Hedgehog signalling activity of Smoothened requires phosphorylation by protein kinase A and casein kinase I. *Nature* **432**, 1045-1050.

Jiang, J. and Levine, M. (1993). Binding affinities and cooperative interactions with bHLH activators delimit threshold responses to the dorsal gradient morphogen. *Cell* **72**, 741-752.

Jortikka, L., Laitinen, M., Lindholm, T. S. and Marttinen, A. (1997). Internalization and intracellular processing of bone morphogenetic protein (BMP) in rat skeletal muscle myoblasts (L6). *Cell Signal* **9**, 47-51.

Khalsa, O., Yoon, J. W., Torres-Schumann, S. and Wharton, K. a. (1998). TGF-beta/BMP superfamily members, Gbb-60A and Dpp, cooperate to provide pattern information and establish cell identity in the Drosophila wing. *Development (Cambridge, England)* **125**, 2723-2734.

Kim, J., Johnson, K., Chen, H. J., Carroll, S. and Laughon, A. (1997). Drosophila Mad binds to DNA and directly mediates activation of vestigial by Decapentaplegic. *Nature* **388**, 304-308.

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

Kunnapuu, J., Bjorkgren, I. and Shimmi, O. (2009). The Drosophila DPP signal is produced by cleavage of its proprotein at evolutionary diversified furin-recognition sites. *Proc Natl Acad Sci U S A* **106**, 8501-8506.

Lander, A. D. (2013). How cells know where they are. *Science (New York, N.Y.)* **339**, 923-927.

Lander, A. D., Lo, W.-C., Nie, Q. and Wan, F. Y. M. (2009). The Measure of Success: Constraints, Objectives, and Tradeoffs in Morphogen-mediated Patterning. *Cold Spring Harbor Perspect Biol* **1**, a002022.

Larson, D. R., Zenklusen, D., Wu, B., Chao, J. A. and Singer, R. H. (2011). Real-time observation of transcription initiation and elongation on an endogenous yeast gene. *Science* **332**, 475-478.

Lawrence, P. A. (2001). Morphogens: how big is the big picture? *Nat Cell Biol* **3**, E151-154.

Lecuit, T. and Cohen, S. M. (1998). Dpp receptor levels contribute to shaping the Dpp morphogen gradient in the Drosophila wing imaginal disc. *Development* **125**, 4901-4907.

Lecuyer, E., Parthasarathy, N. and Krause, H. M. (2008). Fluorescent in situ hybridization protocols in Drosophila embryos and tissues. *Methods Mol Biol* **420**, 289-302.

Lek, M., Dias, J. M., Marklund, U., Uhde, C. W., Kurdija, S., Lei, Q., Sussel, L., Rubenstein, J. L., Matise, M. P., Arnold, H. H., et al. (2010). A homeodomain feedback circuit underlies step-function interpretation of a Shh morphogen gradient during ventral neural patterning. *Development* **137**, 4051-4060.

Lieber, C. A. and Mahadevan-Jansen, A. (2003). Automated method for subtraction of fluorescence from biological Raman spectra. *Appl Spectrosc* **57**, 1363-1367.

Little, S. C., Tikhonov, M. and Gregor, T. (2013). Precise developmental gene expression arises from globally stochastic transcriptional activity. *Cell* **154**, 789-800.

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

Lucchetta, E. M., Vincent, M. E. and Ismagilov, R. F. (2008). A precise Bicoid gradient is nonessential during cycles 11-13 for precise patterning in the Drosophila blastoderm. *PLoS One* **3**, e3651.

Lunde, K., Biehs, B., Nauber, U. and Bier, E. (1998). The knirps and knirps-related genes organize development of the second wing vein in Drosophila. *Development* **125**, 4145-4154.

Luo, J. W. and Bai, J. (2005). Savitzky-Golay smoothing and differentiation filter of even length: A gram polynomial approach. *Spectroscopy* **20**, 27-+.

Manu, Surkova, S., Spirov, A. V., Gursky, V. V., Janssens, H., Kim, A. R., Radulescu, O., Vanario-Alonso, C. E., Sharp, D. H., Samsonova, M., et al. (2009). Canalization of gene expression in the Drosophila blastoderm by gap gene cross regulation. *PLoS Biol* **7**, e1000049.

Mao, Y., Rauskolb, C., Cho, E., Hu, W. L., Hayter, H., Minihan, G., Katz, F. N. and Irvine, K. D. (2006). Dachs: an unconventional myosin that functions downstream of Fat to regulate growth, affinity and gene expression in Drosophila. *Development* **133**, 2539-2551.

Marty, T., Muller, B., Basler, K. and Affolter, M. (2000). Schnurri mediates Dpp-dependent repression of brinker transcription. *Nat Cell Biol* **2**, 745-749.

Massague, J., Seoane, J. and Wotton, D. (2005). Smad transcription factors. *Genes Dev* **19**, 2783-2810.

Mayer, M., Depken, M., Bois, J. S., Julicher, F. and Grill, S. W. (2010). Anisotropies in cortical tension reveal the physical basis of polarizing cortical flows. *Nature* **467**, 617-621.

McHale, P., Mizutani, C. M., Kosman, D., MacKay, D. L., Belu, M., Hermann, A., McGinnis, W., Bier, E. and Hwa, T. (2011). Gene length may contribute to graded transcriptional responses in the Drosophila embryo. *Dev Biol* **360**, 230-240.

Methot, N. and Basler, K. (1999). Hedgehog controls limb development by regulating the activities of distinct transcriptional activator and repressor forms of Cubitus interruptus. *Cell* **96**, 819-831.

Milan, M., Campuzano, S. and Garcia-Bellido, A. (1996). Cell cycling and patterned cell proliferation in the Drosophila wing during metamorphosis. *Proc Natl Acad Sci U S A* **93**, 11687-11692.

Mueller, F., Senecal, A., Tantale, K., Marie-Nelly, H., Ly, N., Collin, O., Basyuk, E., Bertrand, E., Darzacq, X. and Zimmer, C. (2013). FISH-quant: automatic counting of transcripts in 3D FISH images. *Nature methods* **10**, 277-278.

Muller, B., Hartmann, B., Pyrowolakis, G., Affolter, M. and Basler, K. (2003). Conversion of an extracellular Dpp/BMP morphogen gradient into an inverse transcriptional gradient. *Cell* **113**, 221-233.

Muller, P., Rogers, K. W., Jordan, B. M., Lee, J. S., Robson, D., Ramanathan, S. and Schier, A. F. (2012). Differential diffusivity of Nodal and Lefty underlies a reaction-diffusion patterning system. *Science* **336**, 721-724.

Muramoto, T., Cannon, D., Gierlinski, M., Corrigan, a., Barton, G. J. and Chubb, J. R. (2012). Live imaging of nascent RNA dynamics reveals distinct types of transcriptional pulse regulation. *Proceedings of the National Academy of Sciences* **109**, 7350-7355.

Nahmad, M. and Lander, A. D. (2011). Spatiotemporal mechanisms of morphogen gradient interpretation. *Curr Opin Genet Dev* **21**, 726-731.

Nahmad, M. and Stathopoulos, A. (2009). Dynamic interpretation of hedgehog signaling in the Drosophila wing disc. *PLoS Biol* **7**, e1000202.

Nellen, D., Affolter, M. and Basler, K. (1994). Receptor serine/threonine kinases implicated in the control of Drosophila body pattern by decapentaplegic. *Cell* **78**, 225-237.

Nellen, D., Burke, R., Struhl, G. and Basler, K. (1996). Direct and long-range action of a DPP morphogen gradient. *Cell* **85**, 357-368.

Neumann, C. J. and Cohen, S. M. (1997). Long-range action of Wingless organizes the dorsal-ventral axis of the Drosophila wing. *Development* **124**, 871-880.

Newman, J. R., Ghaemmaghami, S., Ihmels, J., Breslow, D. K., Noble, M., DeRisi, J. L. and Weissman, J. S. (2006). Single-cell proteomic analysis of S. cerevisiae reveals the architecture of biological noise. *Nature* **441**, 840-846.

Nienhaus, U., Aegerter-Wilmsen, T. and Aegerter, C. M. (2009). Determination of mechanical stress distribution in Drosophila wing discs using photoelasticity. *Mech Dev* **126**, 942-949.

Ochoa-Espinosa, A., Yu, D., Tsirigos, A., Struffi, P. and Small, S. (2009). Anterior-posterior positional information in the absence of a strong Bicoid gradient. *Proc Natl Acad Sci U S A* **106**, 3823-3828.

Ochoa-Espinosa, A., Yucel, G., Kaplan, L., Pare, A., Pura, N., Oberstein, A., Papatsenko, D. and Small, S. (2005). The role of binding site cluster strength in Bicoid-dependent patterning in Drosophila. *Proc Natl Acad Sci U S A* **102**, 4960-4965.

Ogiso, Y., Tsuneizumi, K., Masuda, N., Sato, M. and Tabata, T. (2011). Robustness of the Dpp morphogen activity gradient depends on negative feedback regulation by the inhibitory Smad, Dad. *Dev Growth Differ* **53**, 668-678.

Padgett, R. W., St Johnston, R. D. and Gelbart, W. M. (1987). A transcript from a Drosophila pattern gene predicts a protein homologous to the transforming growth factor-beta family. *Nature* **325**, 81-84.

Pandya-Jones, A. and Black, D. L. (2009). Co-transcriptional splicing of constitutive and alternative exons. *RNA* **15**, 1896-1908.

Paré, A., Lemons, D., Kosman, D., Beaver, W., Freund, Y. and McGinnis, W. (2009). Visualization of individual Scr mRNAs during Drosophila embryogenesis yields evidence for transcriptional bursting. *Current biology : CB* **19**, 2037-2042.

Paulsson, J. (2004). Summing up the noise in gene networks. *Nature* **427**, 415-418.

Perrimon, N. and McMahon, A. P. (1999). Negative feedback mechanisms and their roles during pattern formation. *Cell* **97**, 13-16.
Pyrowolakis, G., Hartmann, B., Muller, B., Basler, K. and Affolter, M. (2004). A simple molecular complex mediates widespread BMP-induced repression during Drosophila development. *Dev Cell* **7**, 229-240.
Raj, A., Peskin, C. S., Tranchina, D., Vargas, D. Y. and Tyagi, S. (2006). Stochastic mRNA synthesis in mammalian cells. *PLoS Biol* **4**, e309.
Raj, A., Rifkin, S. a., Andersen, E. and van Oudenaarden, A. (2010). Variability in gene expression underlies incomplete penetrance. *Nature* **463**, 913-918.
Raj, A., van den Bogaard, P., Rifkin, S. A., van Oudenaarden, A. and Tyagi, S. (2008). Imaging individual mRNA molecules using multiple singly labeled probes. *Nat Methods* **5**, 877-879.
Raj, A. and van Oudenaarden, A. (2009). Single-molecule approaches to stochastic gene expression. *Annu Rev Biophys* **38**, 255-270.
Raser, J. M. and O'Shea, E. K. (2005). Noise in gene expression: origins, consequences, and control. *Science* **309**, 2010-2013.
Raspopovic, J., Marcon, L., Russo, L. and Sharpe, J. (2014). Modeling digits. Digit patterning is controlled by a Bmp-Sox9-Wnt Turing network modulated by morphogen gradients. *Science* **345**, 566-570.
Ray, R. P. and Wharton, K. A. (2001). Context-dependent relationships between the BMPs gbb and dpp during development of the Drosophila wing imaginal disk. *Development* **128**, 3913-3925.
Riemer, D., Stuurman, N., Berrios, M., Hunter, C., Fisher, P. A. and Weber, K. (1995). Expression of Drosophila lamin C is developmentally regulated: analogies with vertebrate A-type lamins. *J Cell Sci* **108 (Pt 10)**, 3189-3198.
Rigaut, J. P. and Vassy, J. (1991). High-resolution three-dimensional images from confocal scanning laser microscopy. Quantitative study and mathematical correction of the effects from bleaching and fluorescence attenuation in depth. *Anal Quant Cytol Histol* **13**, 223-232.
Rogulja, D. and Irvine, K. D. (2005). Regulation of cell proliferation by a morphogen gradient. *Cell* **123**, 449-461.
Rogulja, D., Rauskolb, C. and Irvine, K. D. (2008). Morphogen control of wing growth through the Fat signaling pathway. *Dev Cell* **15**, 309-321.
Rusakov, D. A. and Kullmann, D. M. (1998). Geometric and viscous components of the tortuosity of the extracellular space in the brain. *Proc Natl Acad Sci U S A* **95**, 8975-8980.
Saffer, A. M., Kim, D. H., van Oudenaarden, A. and Horvitz, H. R. (2011). The Caenorhabditis elegans synthetic multivulva genes prevent ras pathway activation by tightly repressing global ectopic expression of lin-3 EGF. *PLoS Genet* **7**, e1002418.
Sanchez, A., Choubey, S. and Kondev, J. (2013). Regulation of noise in gene expression. *Annu Rev Biophys* **42**, 469-491.
Sanchez, A. and Golding, I. (2013). Genetic determinants and cellular constraints in noisy gene expression. *Science* **342**, 1188-1193.

Sauer, F., Rivera-Pomar, R., Hoch, M. and Jackle, H. (1996). Gene regulation in the Drosophila embryo. *Philos Trans R Soc Lond B Biol Sci* **351**, 579-587.
Schwank, G. and Basler, K. (2010). Regulation of organ growth by morphogen gradients. *Cold Spring Harbor perspectives in biology* **2**, a001669.
Schwank, G., Dalessi, S., Yang, S. F., Yagi, R., de Lachapelle, A. M., Affolter, M., Bergmann, S. and Basler, K. (2011). Formation of the long range Dpp morphogen gradient. *PLoS Biol* **9**, e1001111.
Shen, J., Dahmann, C. and Pflugfelder, G. O. (2010). Spatial discontinuity of optomotor-blind expression in the Drosophila wing imaginal disc disrupts epithelial architecture and promotes cell sorting. *BMC Dev Biol* **10**, 23.
Sheth, R., Marcon, L., Bastida, M. F., Junco, M., Quintana, L., Dahn, R., Kmita, M., Sharpe, J. and Ros, M. A. (2012). Hox genes regulate digit patterning by controlling the wavelength of a Turing-type mechanism. *Science* **338**, 1476-1480.
Shifrin, S. and Steers, E., Jr. (1967). The effect of urea on subunit interactions of beta-galactosidase from Escherichia coli K12. *Biochim Biophys Acta* **133**, 463-471.
Shiga, Y., Tanaka-Matakatsu, M. and Hayashi, S. (1996). A nuclear GFP/ß-galactoside fusion protein as a marker for morphogenesis in living Drosophila. In *Develop. Growth Differ.*, pp. 99-106.
Shimmi, O., Umulis, D., Othmer, H. and O'Connor, M. B. (2005). Facilitated transport of a Dpp/Scw heterodimer by Sog/Tsg leads to robust patterning of the Drosophila blastoderm embryo. *Cell* **120**, 873-886.
Sivasankaran, R., Vigano, M. A., Muller, B., Affolter, M. and Basler, K. (2000). Direct transcriptional control of the Dpp target omb by the DNA binding protein Brinker. *EMBO J* **19**, 6162-6172.
Smith, J. C. (2009). Forming and interpreting gradients in the early Xenopus embryo. *Cold Spring Harb Perspect Biol* **1**, a002477.
So, L. H., Ghosh, A., Zong, C., Sepulveda, L. A., Segev, R. and Golding, I. (2011). General properties of transcriptional time series in Escherichia coli. *Nat Genet* **43**, 554-560.
Sopory, S., Kwon, S., Wehrli, M. and Christian, J. L. (2010). Regulation of Dpp activity by tissue-specific cleavage of an upstream site within the prodomain. *Dev Biol* **346**, 102-112.
Stathopoulos, A. and Levine, M. (2002). Dorsal gradient networks in the Drosophila embryo. *Dev Biol* **246**, 57-67.
---- (2004). Whole-genome analysis of Drosophila gastrulation. *Curr Opin Genet Dev* **14**, 477-484.
---- (2005). Genomic regulatory networks and animal development. *Dev Cell* **9**, 449-462.
Stewart-Ornstein, J., Weissman, J. S. and El-Samad, H. (2012). Cellular noise regulons underlie fluctuations in Saccharomyces cerevisiae. *Mol Cell* **45**, 483-493.
Tabata, T. (2001). Genetics of morphogen gradients. *Nat Rev Genet* **2**, 620-630.
Tabata, T. and Takei, Y. (2004). Morphogens, their identification and regulation. *Development* **131**, 703-712.

Tanimoto, H., Itoh, S., ten Dijke, P. and Tabata, T. (2000). Hedgehog creates a gradient of DPP activity in Drosophila wing imaginal discs. *Mol Cell* **5**, 59-71.
Teleman, A. A. and Cohen, S. M. (2000a). Dpp gradient formation in the Drosophila wing imaginal disc. *Cell* **103**, 971-980.
Teleman, A. A. and Cohen, S. M. (2000b). Dpp gradient formation in the Drosophila wing imaginal disc. In *Cell*, pp. 971-980.
Tirosh, I. and Barkai, N. (2008). Two strategies for gene regulation by promoter nucleosomes. *Genome Res* **18**, 1084-1091.
Tsuneizumi, K., Nakayama, T., Kamoshida, Y., Kornberg, T. B., Christian, J. L. and Tabata, T. (1997). Daughters against dpp modulates dpp organizing activity in Drosophila wing development. *Nature* **389**, 627-631.
Turing, A. M. (1952). The chemical basis of morphogenesis. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* **237**, 37-72.
Umulis, D. M., Serpe, M., O'Connor, M. B. and Othmer, H. G. (2006). Robust, bistable patterning of the dorsal surface of the Drosophila embryo. *Proc Natl Acad Sci U S A* **103**, 11613-11618.
Wada, Y., Ohta, Y., Xu, M., Tsutsumi, S., Minami, T., Inoue, K., Komura, D., Kitakami, J. i., Oshida, N., Papantonis, A., et al. (2009). A wave of nascent transcription on activated human genes. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 18357-18361.
Wartlick, O., Mumcu, P., Kicheva, a., Bittig, T., Seum, C., Jülicher, F. and González-Gaitán, M. (2011). Dynamics of Dpp signaling and proliferation control. *Science (New York, N.Y.)* **331**, 1154-1159.
White, R. J., Nie, Q., Lander, A. D. and Schilling, T. F. (2007). Complex regulation of cyp26a1 creates a robust retinoic acid gradient in the zebrafish embryo. *PLoS Biol* **5**, e304.
White, R. J. and Schilling, T. F. (2008). How degrading: Cyp26s in hindbrain development. *Dev Dyn* **237**, 2775-2790.
Wilkie, G. S., Shermoen, A. W., O'Farrell, P. H. and Davis, I. (1999). Transcribed genes are localized according to chromosomal position within polarized Drosophila embryonic nuclei. *Curr Biol* **9**, 1263-1266.
Winter, S. E. and Campbell, G. (2004). Repression of Dpp targets in the Drosophila wing by Brinker. *Development* **131**, 6071-6081.
Wolpert, L. (1969). Positional information and the spatial pattern of cellular differentiation. *J Theor Biol* **25**, 1-47.
---- (2011). Positional information and patterning revisited. *J Theor Biol* **269**, 359-365.
Xu, H., Sepulveda, L. A., Figard, L., Sokac, A. M. and Golding, I. (2015). Combining protein and mRNA quantification to decipher transcriptional regulation. *Nat Methods* **12**, 739-742.
Yan, D. and Lin, X. (2009). Shaping morphogen gradients by proteoglycans. *Cold Spring Harbor perspectives in biology* **1**, a002493.
Yunger, S., Rosenfeld, L., Garini, Y. and Shav-Tal, Y. (2010). Single-allele analysis of transcription kinetics in living mammalian cells. *Nat Methods* **7**, 631-633.

Zecca, M., Basler, K. and Struhl, G. (1996). Direct and long-range action of a wingless morphogen gradient. *Cell* **87**, 833-844.

Zenklusen, D., Larson, D. R. and Singer, R. H. (2008). Single-RNA counting reveals alternative modes of gene expression in yeast. *Nat Struct Mol Biol* **15**, 1263-1271.

Zhang, H., Levine, M. and Ashe, H. L. (2001). Brinker is a sequence-specific transcriptional repressor in the Drosophila embryo. *Genes Dev* **15**, 261-266.

Zhou, S., Lo, W.-C., Suhalim, J. L., Digman, M. a., Gratton, E., Nie, Q. and Lander, A. D. (2012). Free extracellular diffusion creates the Dpp morphogen gradient of the Drosophila wing disc. *Current biology : CB* **22**, 668-675.

Zopf, C. J., Quinn, K., Zeidman, J. and Maheshri, N. (2013). Cell-cycle dependence of transcription dominates noise in gene expression. *PLoS Comput Biol* **9**, e1003161.