UNIL | Université de Lausanne

Unicentre
CH-1015 Lausanne
http://serval.unil.ch

Year : 2011

Robustness of Drosophila early embryo and wing imaginal disc development

Neves da Silva Aitana

Neves da Silva Aitana, 2011, Robustness of Drosophila early embryo and wing imaginal disc development

Originally published at : Thesis, University of Lausanne

Posted at the University of Lausanne Open Archive.
http://serval.unil.ch

UNIL | Université de Lausanne
Faculté de biologie
et de médecine

DÉP. GÉNÉTIQUE MÉDICALE & INSTITUT SUISSE DE BIOINFORMATIQUE

Robustness of *Drosophila* early embryo and wing imaginal disc development

THÈSE DE DOCTORAT ÈS SCIENCES DE LA VIE (PhD)

PRÉSENTÉE À LA
FACTULTÉ DE BIOLOGIE ET MÉDECINE DE L'UNIVERSITÉ DE LAUSANNE
PAR

AITANA NEVES DA SILVA

PHYSICIENNE DIPLÔMÉE DE L'ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE, SUISSE

Jury
PROF. MEHDI TAFTI, PRÉSIDENT
PROF. SVEN BERGMANN, DIRECTEUR DE THÈSE
PROF. JACQUES BECKMANN, CO-DIRECTEUR DE THÈSE
PROF. RICHARD BENTON, EXPERT
PROF. ANDREAS WAGNER, EXPERT

LAUSANNE - 2011

Imprimatur

Vu le rapport présenté par le jury d'examen, composé de

Président	Monsieur Prof. Mehdi **Tafti**	
Directeur de thèse	Monsieur Prof. Sven **Bergmann**	
Co-directeur de thèse	Monsieur Prof. Jacques **Beckmann**	
Experts	Monsieur Prof. Andreas **Wagner**	
	Monsieur Prof. Richard **Benton**	

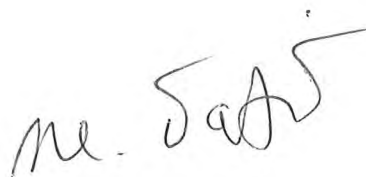le Conseil de Faculté autorise l'impression de la thèse de

Madame Aitana Neves Da Silva

Master of Sciences in Physics EPFL

intitulée

Robustness of *Drosophila* early embryo and wing imaginal disc development

Lausanne, le 7 octobre 2011

pour Le Doyen
de la Faculté de Biologie et de Médecine

Prof. Mehdi Tafti

Résumé grand public

Lorsqu'on se penche sur un organisme, on remarque qu'entre les individus de la même espèce, les proportions des différentes parties du corps sont généralement assez bien conservées. Bien qu'il existe des exceptions notables à cette règle quand on considère l'ensemble des individus d'une espèce, ces différences sont bien moindres lorsqu'on étudie une même lignée d'un organisme de laboratoire, comme la mouche *Drosophile*. Grâce à cette propriété d'*invariance d'échelle*, les veines sur une aile de mouche seront par exemple bien espacées entre elles, même sur une aile de plus grande taille, garantissant ainsi sa fonctionnalité.

Pendant mon doctorat, j'ai étudié l'invariance d'échelle dans le développement de la mouche *Drosophila melanogaster*. Afin de comprendre comment les veines sont positionnées correctement sur une aile en devenir, je me suis concentrée sur le stade larvaire du développement. C'est en effet à ce moment précis que l'aile de la mouche commence à se développer. Une question qui intrigue de nombreux biologistes est de savoir comment une cellule dans un tissu qui grandit sait où elle se trouve par rapport aux autres et peut donc décider en quel type cellulaire se différencier (par exemple, veine ou inter-veine). Ceci peut être obtenu par la production et la libération d'une molécule de signalisation, appelée morphogène, qui diffuse et forme un gradient de concentration: différents niveaux d'exposition au morphogène entraînent différentes spécifications cellulaires, grâce à l'activation de cascades de signalisation spécifiques. Par conséquent, en réponse au gradient de morphogène, les cellules commencent à exprimer certains gènes par endroits (appelés *domaines d'expression*) et forment dans le tissu des régions caractérisées par une combinaison unique des gènes qui y sont exprimés. En d'autres termes, la formation de ces domaines d'expression de gènes est le moyen par lequel les cellules déterminent leur sort. A mesure que le gradient de morphogène est décodé par les cellules dans le tissu, un motif de destinées cellulaires émerge.

Même si les morphogènes sont connus depuis quelques décennies, on comprend encore très mal comment cette spécification cellulaire, qui précède la différentiation, peut se faire de manière aussi précise et robuste. Au cours de mon doctorat, j'ai quantifié l'invariance d'échelle de ces domaines d'expression puis essayé de trouver des mécanismes pouvant rendre compte de ce phénomène, afin de garantir que l'aile adulte soit bien proportionnée. En règle générale, comprendre comment les cellules répondent à des signaux ou coordonnent leurs actions pourrait nous éclairer davantage sur des maladies comme le cancer et, en théorie, servir de base à l'élaboration de tissus artificiels.

Résumé

Dans un organisme en développement, les cellules doivent savoir où elles se trouvent afin de se différencier dans le type cellulaire adéquat. La formation de motifs d'expression de gènes est le processus par lequel les cellules acquièrent cette information spatiale et donc déterminent leur sort. Ceci peut être obtenu par la production et la libération d'une molécule de signalisation, appelée *morphogène*, qui diffuse et forme un gradient de concentration: différents niveaux d'exposition au morphogène entraînent différentes spécifications cellulaires, grâce à l'activation de cascades de signalisation spécifiques. Par conséquent, en réponse au gradient de morphogène, les cellules commencent à exprimer certains gènes par endroits et forment dans le tissu des domaines caractérisés par une combinaison unique de gènes qui y sont exprimés. A mesure que le gradient de morphogène est décodé par les cellules dans le tissu, on voit émerger un motif de destinées et spécifications cellulaires.

Même si les morphogènes sont connus depuis quelques décennies, on ne sait pas encore clairement comment ces gradients se forment et induisent des motifs d'expression de gènes robustes. Pendant ma thèse de doctorat, j'ai étudié les propriétés de Bicoid (Bcd) et de Decapentaplegic (Dpp), deux morphogènes qui participent à la spécification de l'axe antéro-postérieur de l'embryon et de l'aile de *Drosophile*, respectivement. En particulier, j'ai cherché à comprendre comment les *proportions* des motifs d'expression de gènes sont conservées parmi des embryons de tailles différentes ou dans un tissu en croissance. Cette propriété est appelée *invariance d'échelle* et permet de garantir que l'organisme (ou l'organe) final est bien proportionné et donc fonctionnel.

Afin de répondre à ces questions, j'ai analysé des images de fluorescence montrant les motifs d'expression de gènes dans l'embryo et le disque imaginal de l'aile. Après avoir caractérisé l'étendue de ces domaines de façon quantitative et systématique, j'ai introduit et appliqué une nouvelle *mesure d'invariance d'échelle* afin de déterminer à quel point les proportions sont effectivement maintenues. J'ai trouvé que l'invariance d'échelle est une propriété universelle à la fois chez l'embryon (du moins loin de la source de Bcd) et le disque imaginal de l'aile (à plusieurs stades du développement). Puisque nous voulions également comprendre les mécanismes sous-jacents à l'invariance d'échelle et comment elle est transmise du morphogène aux gènes cibles dans la cascade de signalisation, j'ai aussi quantifié l'invariance d'échelle dans des mouches mutantes où cette propriété pourrait être perturbée.

Alors que l'invariance d'échelle est largement conservée chez des embryons dont on a altéré les niveaux de *bcd*, ma modélisation suggère que la capture de Bcd par les noyaux ainsi que l'interpretation du gradient de morphogène à l'état de pré-équilibre sont essentiels pour assurer la robustesse de la cascade de signalisation de Bcd. Lors de la croissance du disque imaginal de l'aile, les résultats montrent que le gradient de Dpp s'adapte à la taille du tissu. Par ailleurs, l'invariance d'échelle n'est pas parfaite partout dans le tissu. En effet, l'invariance d'échelle des domaines d'expression cibles est meilleure là où ces gènes exercent une fonction. Par exemple, Spalt est davantage invariant à l'échelle dans le compartiment antérieur du disque, où il participe à la formation de l'une des veines de l'aile. L'analyse de mutants pour *pentagone*, qui est un gène cible de Dpp qui rétrocontrôle la cascade de signalisation, montre que Pentagone joue un rôle essentiel pour garantir l'invariance d'échelle du gradient d'activité de Dpp.

Summary

Within a developing organism, cells require information on where they are in order to differentiate into the correct cell-type. Pattern formation is the process by which cells acquire and process positional cues and thus determine their fate. This can be achieved by the production and release of a diffusible signaling molecule, called a *morphogen*, which forms a concentration gradient: exposure to different morphogen levels leads to the activation of specific signaling pathways. Thus, in response to the morphogen gradient, cells start to express different sets of genes, forming domains characterized by a unique combination of differentially expressed genes. As a result, a pattern of cell fates and specification emerges.

Though morphogens have been known for decades, it is not yet clear how these gradients form and are interpreted in order to yield highly robust patterns of gene expression. During my PhD thesis, I investigated the properties of Bicoid (Bcd) and Decapentaplegic (Dpp), two morphogens involved in the patterning of the anterior-posterior axis of *Drosophila* embryo and wing primordium, respectively. In particular, I have been interested in understanding how the pattern *proportions* are maintained across embryos of different sizes or within a growing tissue. This property is commonly referred to as *scaling* and is essential for yielding functional organs or organisms.

In order to tackle these questions, I analysed fluorescence images showing the pattern of gene expression domains in the early embryo and wing imaginal disc. After characterizing the extent of these domains in a quantitative and systematic manner, I introduced and applied a new *scaling measure* in order to assess how well proportions are maintained. I found that scaling emerged as a universal property both in early embryos (at least far away from the Bcd source) and in wing imaginal discs (across different developmental stages). Since we were also interested in understanding the mechanisms underlying scaling and how it is transmitted from the morphogen to the target genes down in the signaling cascade, I also quantified scaling in mutant flies where this property could be disrupted.

While scaling is largely conserved in embryos with altered *bcd* dosage, my modeling suggests that Bcd trapping by the nuclei as well as pre-steady state decoding of the morphogen gradient are essential to ensure precise and scaled patterning of the Bcd signaling cascade. In the wing imaginal disc, it appears that as the disc grows, the Dpp response expands and scales with the tissue size. Interestingly, scaling is not perfect at all positions in the field. The scaling of the target gene domains is best where they have a function; Spalt, for example, scales best at the position in the anterior compartment where it helps to form one of the anterior veins of the wing. Analysis of mutants for *pentagone*, a transcriptional target of Dpp that encodes a secreted feedback regulator of the pathway, indicates that Pentagone plays a key role in scaling the Dpp gradient activity.

Acknowledgement

4 years, 7 months, 7 days... to become a doctor. I am very happy now to take a few moments to thank all the people that participated in this journey into Science!

First of all, I would like to thank my PhD supervisor, Prof. Sven Bergmann, and my co-supervisor, Prof. Jacques Beckmann. Sven, I also remember very well when we first met for my interview. I had to think of a new problem, and I really enjoyed that you tried to follow my thoughts rather than just take me straight to the answer. I understood you would take further my reasoning and help me mature. In particular, I would like to thank you for your positive thinking and encouragements during all these years. Jacqui, tu as toujours su trouver les mots et les questions pour me faire avancer, merci pour toutes ces discussions! I also had the chance to be surrounded by very nice people in the lab and department, always willing to help and give me feedback. Thanks to Zoltán and Bastian who coached me at the beginning of my thesis. Diana, office-mate, room-mate, St Bernard, merci d'avoir été là à chaque moment important, avec tellement de générosité, tu comptes énormément pour moi. Et merci à tous les compagnons de midi, Valou, Gigi, Micha et les autres, pour tous les bons moments passés ensembles.

I would also like to thank the Swiss Institute of Bioinformatics PhD training network and all the students involved in it, I really had a great time participating in all these events and being part of the educational board! I am also very thankful to the SystemsX.ch WingX project. It was a great experience to meet scientists from all over Switzerland, with special thanks to Dr. Fisun Hamaratoglu and Prof. Markus Affolter with whom I closely collaborated. Fisun, I am really happy that I met you and that we started this collaboration together! I would also like to thank Thomas Schaffter, Dr. Sabine Schilling, Dr. Tinri Aegerter, Sophie Baranek, Simon Restrepo, Dr. Gerald Schwank, Dr. Oguz Kanca, Prof. Konrad Basler and Prof. Ernst Hafen for all the great discussions. I am also very grateful to Prof. Luc Pellerin and Dr. Robert Costalat for the enriching collaboration that we started.

Finally, I would like to thank my friends and family for their support all these years. Merci à ma famille, mes parents, mes frères, Lise, qui m'ont toujours encouragé dans mes études et ont su nourrir ma curiosité d'enfant. Merci David pour ta générosité, ton soutien et tes encouragements qui m'ont tellement fait grandir tout au long de ma thèse. Pour finir, un grand merci à toutes les mouches *Drosophiles* qui ont participé à l'étude, ainsi qu'au soleil qui me réveille tous les matins...

Abbreviations

After egg laying (AEL)

Anterior-posterior (A/P)

Bicoid (Bcd)

Brinker (Brk)

Dorsal-ventral (D/V)

Even-skipped (Eve)

Giant (Gt)

Hunchback (Hb)

Krüppel (Kr)

Length (L)

Length of the posterior compartment (L_p)

Optomoter blind (Omb)

Pentagone (Pent)

Phospho-Mad (P-mad)

Sal (Spalt)

Schnurri (Shn)

Time class (TC)

Contents

Prologue

. . . .

□ Developmental biology.   □ Morphogens and their interpretation.
□ Scaling and Precision.   □ Model organisms.   □ Thesis outlook and scope.

Developmental biology   "*To become an embryo, you had to build yourself from a single cell. You had to respire before you had lungs, digest before you had a gut, build bones when you were pulpy, and form orderly arrays of neurons before you knew how to think. One of the critical differences between you and a machine is that a machine is never required to funtion until after it is built. Every animal has to function as it builds itself.*" (Gilbert, 2006)

Developmental biology is the science of how animals evolve from a zygote to their adult stage, at the molecular/chemical, cellular/tissue, organ/organ-system, ecological and even evolutionary levels. Though experimental embryology has fascinated generations of biologists, it was only in the 1980s with the possiblity to investigate developmental processes at the molecular level that the field of *developmental biology* truely emerged. Indeed, with the advent of molecular biology, it became possible to alter the path of development by introducing or removing selected genes (Gilbert, 2006).

Developmental biology addresses essentially the fundamental question of how a fertilized egg gives rise to an adult organism. When the egg is fertilized, the embryo is made up of only one cell that eventually divides and gives rise to a large number of genetically identical cells. One may ask how cells containing the same DNA can give rise to an adult body containing different organs and limbs, as well as hundreds of different cell types. This is because development relies essentially on the *control of gene expression* (Levine and Davidson, 2005). However, since all cells are initially identical, the control of gene regulation needs to be started by an external factor, like the mother who breaks the initial symmetry very early in *Drosophila* development with the use of maternal cues (Driever and Nüsslein-Volhard, 1988a; Driever and Nüsslein-Volhard, 1988b). These cues, also called *morphogens*, allow to spatially *pattern* the embryo into compartments of differentially expressed genes that will eventually define the adult body plan (Wolpert, 1969; Morgan, 1901). Similarly, in a developing tissue, patterning is a prelude to *differentiation*, which is the process by which cells acquire a well-defined fate, contributing to the generation of cellular diversity. Because these differentiated cells are not randomly distributed, they can organize into intricate tissues and organs during a process called *morphogenesis* (Gilbert, 2006). Another important aspect of developmental biology is growth and size regulation. Indeed, it is intriguing how cell division can be so tightly regulated, allowing, e.g., an orgnanism to have both arms of the same size, meaning that a growing tissue actually "knows" when it has reached its final size.

Once the adult body has formed, comes the question of reproduction. The sperm and egg are indeed very specialized cells that are set apart to form the next generation (Gilbert, 2006). How this happens, and how this has happened throughout evolution to create new body forms under different constraints is another topic of great interest in the field. Finally, the developing organism often interacts with its environment,

like some butterflies that produce different wing colors based on the temperature or the amount of daylight experienced by the caterpillar before it undergoes metamorphosis, raising the question of how a developing organism integrates into its habitat and how it can potentially be altered by, e.g., the presence of some chemicals (Gilbert, 2006).

Since its beginning, developmental biology has integrated several fields of biology, like experimental embryology, developmental genetics and molecular biology (Slack, 2006). Interestingly too, physicists and mathematicians have also been working on these questions, seeking quantifiable laws of development in embryology (Salazar-Ciudad et al., 2000), tissue formation (Flenner et al., 2008) or patterning (Eldar et al., 2002; Turing, 1952; Ibañes and Belmonte, 2008) and growth (Aegerter-Wilmsen et al., 2007; Hufnagel et al., 2007), to give some examples. Following their wake, my PhD thesis focuses on developmental biology questions, undertaking a *systems biology* approach. This means that we describe the biological system with physical equations and use computational tools to solve them and to perform quantitative analyses.

Morphogens and their interpretation   In particular, I have focused on the *formation* and *interpretation* of morphogen gradients acting at different stages of *Drosophila* development. Morphogen gradient formation is shaped by production, degradation and transport. Production of a morphogen is generally assumed to be localized in a small region of a tissue and constant over time, although this domain can extend during growth (e.g. Dpp production region in *Drosophila* wing imaginal disc (Wartlick et al., 2011)). Other studies argue that a morphogen can be produced from a dynamic gradient of mRNA (e.g. *bcd* mRNA gradient in the *Drosophila* embryo (Spirov et al., 2009), *hoxd13* in the vertebrate developing limb bud (Ibañes et al., 2006) or *fgf8* in the vertebrate embryo (Dubrulle and Pourquié, 2004)). Degradation of the morphogen (or its interaction with some other molecules, e.g. receptors) can also help shape the gradient, in particular if it is controlled by the morphogen signal itself, where degradation was shown to increase robustness to gene dosage changes (Eldar et al., 2003). Lastly, morphogen transport (either active or passive) is essential to promote the formation of a long-range gradient from a localized source.

In 1970, Francis Crick proposed that gradients could form over a field of 50 cells within a few hours (Crick, 1970). Since then, diffusion has often been proposed as a mechanism for morphogen transport, in particular when morphogen movement is reported to be non-directional, which is in agreement with passive diffusion. However, since measured diffusion constants are much smaller than those measured in aqueous media, it is now common to talk about *effective* or *restricted* diffusion, when referring to non-directional passive transport in the crowded extracellular matrix (Strigini, 2005; Kicheva et al., 2007; Tabata and Takei, 2004). Apart from *restricted* extracellular diffusion (Lander et al., 2002), other forms of (non-directional) transport have been proposed for secreted molecules. These include vesicle-mediated transport in the extracellular media (Greco et al., 2001; Panáková et al., 2005), cell-to-cell transport by cycles of endocytosis and exocytosis (named *transcytosis*) (González-Gaitán, 2003; Bollenbach et al., 2007; Entchev et al., 2000), and propagation of intercellular signals through thin actin filaments called *cytonemes* (Latin for "cell thread") (Ramírez-Weber and Kornberg, 1999; Affolter and Basler, 2011). Although the study of morphogens has focused on secreted molecules that can form long-range gradients, there is increasing evidence for cellular-based transport, e.g. *via* gap junctions (Esser et al., 2006) or *via* cell growth, division and displacement which dilute and transport non-secreted molecules over space (Ibañes et al., 2006; Lecuit and
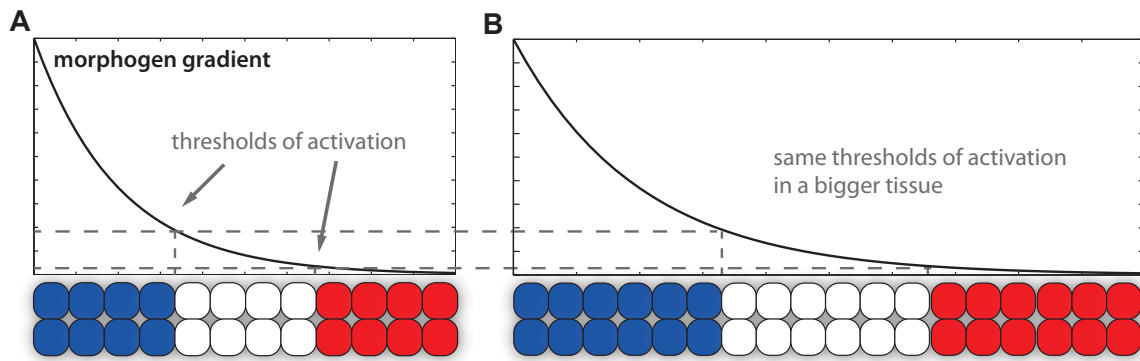
Figure 1: The *French flag* model for morphogen gradient interpretation and scaling *(Wolpert, 1969)*. *(A)* A morphogen gradient is established in the target field. The cells read out the morphogen concentrations and activate some target genes depending on the morphogen levels. Here, high concentrations activate a "blue" gene, lower concentrations induce a "white" gene, and at even lower concentrations the default "red" state is established. *(B)* In a bigger tissue where the morphogen gradient scales, the pattern will keep the same proportions (1/3 of blue, 1/3 of white and 1/3 of red cells, compare with (A)) and thus scale perfectly, since the target genes remain activated at the same threshold concentrations.

. . . .

Cohen, 1998; Teleman et al., 2001; Tabata, 2001; Dubrulle and Pourquié, 2004).

The morphogen gradient is meant to be interpreted by the cells in the patterning field. To be defined as such, the morphogen should form a gradient capable of activating or repressing target genes in a concentration-dependent manner. This means that the cells in the tissue read out the morphogen concentration and activate some target genes if the levels of morphogen are high enough. Because different genes can have different thresholds of activation, the cells in the tissue start to express different combinations of genes, depending on where they lie along the morphogen gradient (Figure 1). This model is known as the *French flag* (Wolpert, 1969), where the flag is used to demonstrate the effect of a morphogen on cell fate determination. High concentrations activate some gene *a*, lower concentrations induce another gene *b*, and at even lower concentrations the default gene *c* is transcribed. As a consequence, three cell fates emerge: "blue" for the cells expressing genes *a*, *b* and *c,* "white" for those expressing only genes *b* and *c*, and "red" for the remaining cells. These gene expression domains get more and more refined during development, eventually defining the segments of the adult body and specifying cell fates upon differentiation.

The cascade of activation or inhibition can be combined with several morphogens, further extended at the level of target genes which can, e.g., mutually repress each other (Jaeger et al., 2004), or even be feedback-driven (Eldar et al., 2002), hence extending the static and hierarchical French flag view of patterning (Jaeger and Reinitz, 2006). Some studies also suggest that pattern formation is a temporally dynamic process, coupled with oscillators (Dequéant and Pourquié, 2008), or raise questions about the influence of time on morphogen activity and desensitization to it (Kutejova et al., 2009; Dessaud et al., 2010). Finally, although many studies consider that morphogen gradients start to be decoded once they have reached steady state, increasing evidence suggests that in some cases pre-steady state decoding is important for robustness (Bergmann et al., 2007) and differential gene expression (Nahmad and Stathopoulos, 2009; Nahmad and Stathopoulos, 2010). All together, it appears that morphogen gradients are essential to organize and coordinate *temporal* and *spatial* development, assigning to each cell a role depending on its position within the embryo or tissue, in a dynamic and yet to be refined manner.

However, morphogen *gradients* as proposed by Thomas Hunt Morgan (Morgan, 1901) at the beginning of the last century are not the only possible mechanism to achieve patterning. In 1952, Turing proposed a model where two substances, an activator and an inhibitor, diffuse *and* interact, yielding complex periodic patterns from an *a priori* unperiodic prepattern (i.e. the morphogen sources, cf. Figure 2) (Turing, 1952; Meinhardt and Gierer, 2000). This process is self-regulating and can explain a variety of patterns present in Nature (see (Kondo and Miura, 2010) for a review). However, proving that these are actually Turing patterns requires investigating the dynamic properties of the pattern and not just reproducing it. For example, a recent study has shown that the zebrafish stripes are not stationary; following laser ablation of a pair of stripes, the pattern of stripes slowly evolved to a new stationnary state within few weeks, a behavior that was predicted by simulation using Turing's formalism (Yamaguchi et al., 2007). As a counter-example, the seven Even-skipped stripes in *Drosophila* embryo appeared like a good example of periodic Turing patterning, before experimental evidence showed the existence of several enhancers to control the positioning of the stripes (Small et al., 1992;

Figure 2: Schematic drawing showing the difference between the classical morphogen gradient and Turing models. *(A) A morphogen molecule produced at one end of an embryo forms a gradient by diffusion. Cells "know" their position from the concentration of the molecule. The gradient is totally dependent on the prepattern of the morphogen source (boundary condition). (B) Adding a second morphogen produces a relatively complex pattern; but with no interactions between the morphogens, the system is not self-regulating. (C) With addition of the interactions between the morphogens, the system becomes self-regulating and can form a variety of patterns independent of the prepattern. Art work by S. Miyazawa.* From (Kondo and Miura, 2010). Reprinted with permission from AAAS and S. Miyazawa.

Small et al., 1993; Small et al., 1996). On another level, Smith and co-workers have shown how a gene regulatory network can elaborate a complex spatio-temporal pattern of gene expression in the early sea urchin embryo without the use of a conventional morphogen (Smith et al., 2007).
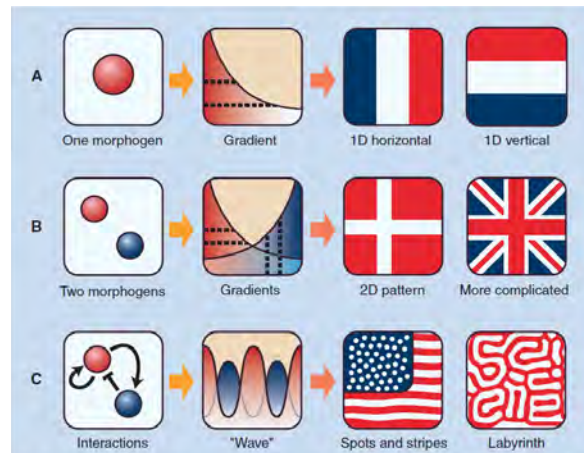
Scaling and Precision Despite noteworthy variation in the height and weight of various individuals from the same species (for example humans), the proportions of different body parts are much less variable. This scaling relationship between the size of body parts or organs with the overall body size is called allometry, a phenomenon that has long fascinated biologists (Stern and Emlen, 1999; Shingleton et al., 2007; Huxley, 1972; Thomson, 1917). Stern and Emlen described three kinds of allometry: ontogenetic, static and evolutionary. The growth trajectory of an organ relative to body size during the development of a single individual was coined ontogenetic allometry. Evolutionary allometry, by contrast, deals with the size relationship between organs across species. Lastly, static allometry refers to the scaling relationship between one body part and the total body size, when several individuals are compared at a single developmental stage. For example, the relative volumes of endoderm and ectoderm in a sea urchin embryo are constant in the face of an eight-fold size range of the embryos (Wolpert, 1969), and starved fly larvae and mice embryos
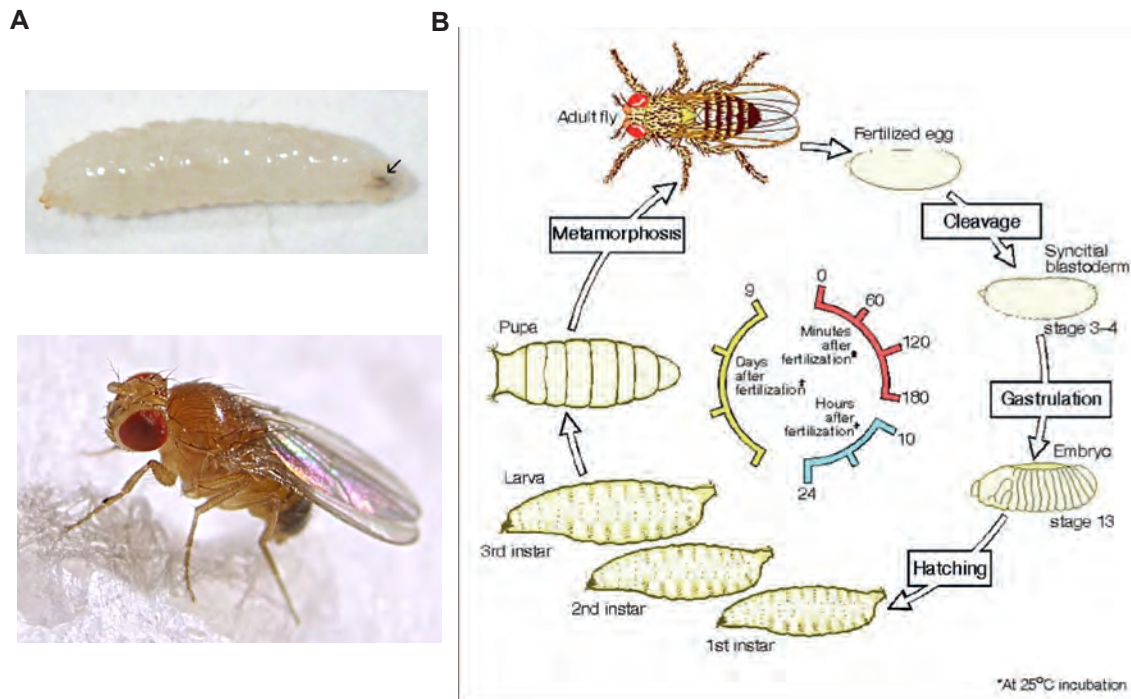
Figure 3: *Drosophila melanogaster.* **(A)** *(top) Second instar larva.* *Mouth parts are indicated with an arrow.* *Photo by J. Albert Vallunen, http://nl.wikipedia.org/wiki/Bestand:Drosophila_2nd_instar_larva.jpg (bottom) Adult fly.* *http://www.flickr.com/photos/max_westby/54275159/sizes/o/in/photostream/* **(B)** *Drosophila melanogaster life cycle.* *http://www.mun.ca/biology/desmid/brian/BIOL3530/DB_Ch02/fig2_29.jpg.*

. . . .

form smaller adults with proportionally smaller body parts (Edgar, 2006). With few exceptions (Nijhout and Grunert, 2010), the mechanisms underlying these allometries remain unknown.

During my thesis, I have considered gene expression domains as a molecular precursor for tissue patterning and asked whether their boundaries scale with embryo size (Part III) or with wing pouch size during growth (Part IV). For the wing imaginal disc, I also considered the morphogen activity gradient and tested whether it adapts to organ size. In the case of perfect scaling, the pattern of gene expression domains (e.g. the *French flag* in Figure 1A) should scale with tissue size. Assuming a *French flag* decoding model, this means that the morphogen gradient should spread further in bigger tissues, to ensure that the boundaries of the gene expression domains shift and therefore preserve the correct proportions of the pattern (see Figure 1B).

We define precision as the similarity in the localization of gene expression domains, when comparing several individuals at the same developmental stage. In case of imperfect scaling, the boundary position of gene expression domains across individuals may vary due to the fluctuations in tissue size. However, imperfect scaling is not the only source of patterning imprecision, since natural fluctuations in the environment also generally yield some dissimilarities in the final pattern. Interestingly, nature tends to show great robustness so that the resulting imprecision is usually lower than expected (e.g. having a fly embryo develop at two different temperatures yields nevertheless a very well patterned and viable fly (Lucchetta et al.,

2005)). This capacity to buffer fluctuations and correct for mistakes during development has fascinated and continues to intrigue generations of engineers seeking models and inspiration. This PhD thesis is an attempt to further elucidate some robustness mechanisms involved in *Drosophila* development, as a means to better understand the underlying biology.

Model organisms   Modern developmental biology has focused on a few animal species, which are commonly referred to as *model organisms*. We can also learn about humans by studying similar processes in these model organisms. Indeed, different species often share common proteins and signaling pathways, which are the key regulators of development. Thus, since all animals use similar mechanisms for their development, a model organism can serve as an example for how other animals develop (Slack, 2006).

During my thesis, I have focused on the fruit fly *Drosophila melanogaster*, which is one of the most widely studied model organisms (Figure 3A). For the past decades, molecular biologists have developed many tools for constructing stocks and carrying out mutagenesis screens for this model organism, allowing to discover the genes affecting its development (Slack, 2006).

The fruit fly has several practical advantages, since it is small and easy to feed, produces large numbers of eggs, develops within a few days and has a short life-cycle of two weeks. As can be seen in Figure 3B, *Drosophila* embryogenesis is very short, lasting about 24 hours. Hatching is followed by the larval stage which lasts about 5 days and ends with the metamorphosis into an adult fly with reproductive maturity. During my PhD thesis, I worked both on the embryonic (Part III) and the larval stages (Part IV) of *Drosophila* development.

Thesis outlook and scope   In summary, within a developing organism, cells need to "know" where they are in order to differentiate into the correct cell-type. Pattern formation is the process by which cells acquire their positional information and thus determine their fate. This can be achieved by the production and release of a diffusible signaling molecule, called a *morphogen*, which forms a concentration gradient: exposure to different morphogen levels leads to different cell fates. Though morphogens have been known for decades, it is not yet clear how these gradients form and yield such robust patterns. During my PhD thesis, I have investigated the properties of Bicoid and Decapentaplegic, two morphogens involved in the patterning of the anterior-posterior axis of the *Drosophila* embryo and the wing primordium, respectively. In particular, I have been interested in understanding how the pattern proportions are maintained across embryos of different sizes or within a growing tissue, which is essential to yield a correctly proportioned organism or organ. Ultimately, the general understanding of how cells respond to signals and coordinate their actions could bring new insights into some diseases where these processes are disregulated and provide the basis to make artificial tissues.

Part I

Setting the Stage: Reaction-Diffusion equations

. . . .

This chapter introduces the mathematical formalism of reaction-diffusion equations that I used during my thesis.

I.1   Derivatives and differential equations

Consider a car moving on a straight line (Figure I.1A). At time t, its position is given by $x(t)$. After some time Δt, its new position is $x(t + \Delta t)$. Thus, the average speed $v_{average}$ of the car is:

$$v_{average} = \frac{x(t + \Delta t) - x(t)}{\Delta t} \, .$$

In order to approximate as accurately as possible its instantaneous speed $v_{instant}$, the time window Δt should be as small as possible:

$$v_{instant} = \lim_{\Delta t \to 0} v_{average} = \lim_{\Delta t \to 0} \frac{x(t + \Delta t) - x(t)}{\Delta t} \, ,$$

where $\lim_{\Delta t \to 0}$ means that the time window should be infinitely small.

The derivative of x with respect to t, noted $\frac{dx}{dt}$, is defined as:

$$\frac{dx}{dt} \equiv \lim_{\Delta t \to 0} \frac{x(t + \Delta t) - x(t)}{\Delta t} \, .$$

Thus, the derivative of x with respect to t is a measure of how fast x is changing with respect to t, i.e. its instantaneous speed. It can be positive, meaning that the position x increases with time t, negative if the position decreases with time (e.g. the car goes backwards) or zero if the position does not change with time (e.g. the car stopped). In our example, x was the position of a car and t the time. However, the definition of the derivative is true no matter what x and t represent. For example, given a protein concentration P, dP/dt describes the temporal dynamics of the protein levels, i.e. how fast its concentration changes as time goes.

An ordinary differential equation (ODE) is simply an equation that contains derivatives of the variables, e.g.

$$\frac{dP}{dt} = -\alpha P + s \, .$$

Assuming that P is a protein concentration, the positive term s denotes a production term, while the negative term $-\alpha P$ is a degradation term. The fact that the degradation term is proportional to P ensures that there can only be degradation as long as there is some protein to degrade. Thus, given an initial concentration P_o, the derivative dP/dt will be positive if the production of protein is bigger than its degradation. In that

case, the protein levels will increase, until the point where the production of protein equals its degradation: $s = \alpha P$ and hence $dP/dt = 0$ (Figure I.1B, red curve). This is called the *steady state* (or *equilibrium*), since the protein levels will not change anymore (unless some external factor changes). Note that saying that the protein levels do not change does not mean that the protein stops being produced or degraded. It only means that the production and degradation compensate for one another, so that the protein levels remain the same. If we consider another intial concentration P'_o where degradation is higher than production, the derivative dP/dt will be negative and hence the protein levels will decrease, until reaching steady state as well (Figure I.1B, blue curve). From this example, we see that while the steady state solution of a differential equation can be unique (in this example, $P^{ss} = s/\alpha$), we need to have an *initial condition* in order to know the trajectory, $P(t)$. Note that a differential equation can have multiple steady states that can be reached from different initial conditions.

Let us now consider a protein whose concentration varies in time *and* space, $P(x,t)$. As we see, the protein levels now depend on two variables, namely the position x and the time t, meaning that in order to know the protein

Figure I.1: *(A) (top) The car drives with average speed $v_{average} = \Delta x/\Delta t$. The derivative is the instantaneous speed of the car, $v_{instant} = \frac{dx}{dt} = \lim_{\Delta t \to 0} \frac{\Delta x}{\Delta t}$. (bottom) We show the position of the car as a function of time. At the beginning, the car is stopped $(dx/dt = 0)$. Then it moves forward for some time $(dx/dt > 0)$, before going backwards $(dx/dt < 0)$. Hence, the derivative can also be thought of as the slope of the tangent to the trajectory. (B) Protein profile as a function of time for two different conditions, P_o and P'_o. The dashed line is the steady state solution, corresponding to $P^{ss} = s/\alpha$.*

levels, we need to provide both the position in the field and the time of measurement. We define the *partial derivative* of P with respect to t as:

$$\frac{\partial P}{\partial t} \equiv \lim_{\Delta t \to 0} \frac{P(x, t + \Delta t) - P(x, t)}{\Delta t} .$$

This partial derivative describes the temporal dynamics of the protein levels, at the *fixed* position x. Indeed, since the protein levels are not uniform in space, they can vary at different speeds in each position. Then, what does the partial derivative $\partial P/\partial x$ mean? $\partial P/\partial x$ describes how fast the protein levels vary as a function of position, for a given time t. Formally,

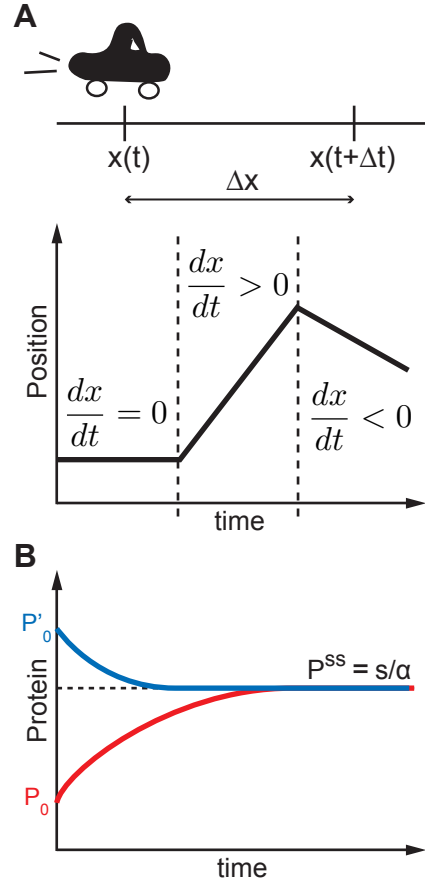$$\frac{\partial P}{\partial x} \equiv \lim_{\Delta x \to 0} \frac{P(x + \Delta x, t) - P(x, t)}{\Delta x} .$$

Thus, $\partial P/\partial x$ is a measure of the concentration gradient, at a *fixed* time t. Depending on its sign, the concentration gradient increases or decreases.

In the case where an equation contains partial derivatives (e.g. with respect to time and space), we call it a partial differential equation (PDE). In order to solve it, we also need to specify an initial condition. Importantly, partial differential equations also require a *boundary condition*, which describes how the solution behaves at the boundaries (e.g. the protein levels at the posterior boundary of the embryo should be zero).

I.2 Fick's laws of diffusion

We derive Fick's equations from the model of a random walk[1]. Consider particles moving in a 1d random walk from a "macroscopic" perspective. Particles can move either to the right or to the left. We are interested in estimating the flux J_x of particles moving to the right from some position x, i.e. the net number of particles moving from x to $x + \delta$, per unit area and time (Figure I.2A). We assume that it takes on average a time τ for the particles to move a distance δ. Thus, after some time τ, half the particles in x would have moved to the right at the position $(x + \delta)$ and half to the left at the position $(x - \delta)$. Assuming we know the number of particles at each position along the x axis, $N(x,t)$, the *net* number of particles crossing to the right at time t will be

$$\frac{1}{2}\left[N(x,t) - N(x + \delta,t)\right],$$

since on average, half the particles in x moved to the right, $\frac{N(x,t)}{2}$, and half the particles in $x + \delta$ moved to the left, $\frac{N(x+\delta,t)}{2}$. The net number of particles crossing to the right is therefore the difference between these two terms.

To obtain a flux J_x, i.e. the net number of particles per unit time and per unit area, we divide by the area normal to the x axis, A, and by the time interval τ (note that we have swapped the terms inside the brackets, hence the minus
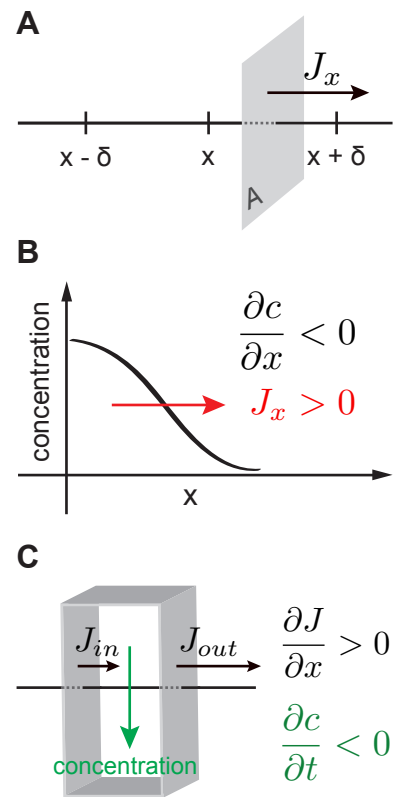
Figure I.2: *(A) Flux J_x of particles crossing the area A at position x. In this example, the flux goes to the right, hence $J_x > 0$. (B) Concentration of a protein as a function of position (e.g. in a tissue). Since the levels decrease, the concentration gradient is negative, $\partial c/\partial x < 0$. However, since there are more particles to the left, the net flux of particles will go to the right, hence $J_x > 0$. (C) The concentration in the volume changes if particles enter the box ($J_{in} \neq 0$) or if particles leave it ($J_{out} \neq 0$). In this example, $J_{out} > J_{in}$ and thus there are more particles leaving the box than entering it. The gradient of fluxes is therefore positive, $\partial J/\partial x > 0$. However, since the net number of particles in the box is decreasing, the concentration inside the box will decrease as a function of time, hence $\partial c/\partial t < 0$.*

[1]See *Random Walks in Biology, Howard C. Berg, Princeton University Press (1993)* for a nice introduction on the topic.

sign in front):

$$J_x = -\frac{1}{2}\left[N(x+\delta,t) - N(x,t)\right]/A\tau.$$

If we multiply by δ^2/δ^2, we obtain

$$J_x = -\frac{\delta^2}{2\tau}\frac{1}{\delta}\left[\frac{N(x+\delta,t)}{A\delta} - \frac{N(x,t)}{A\delta}\right].$$

The quantity $\delta^2/2\tau$ is defined as the diffusion coefficient, D, which has units of distance squared per time. $N(x)/A\delta$ is the concentration of particles at the point x, since $A\delta$ is a volume. Therefore, the flux becomes

$$J_x = -D\frac{1}{\delta}\left[C(x+\delta,t) - C(x,t)\right].$$

In the limit where δ is very small, we obtain Fick's first law:

$$J_x = -D\frac{1}{\delta}\left[C(x+\delta,t) - C(x,t)\right] \overset{\lim_{\delta\to 0}}{=} -D\frac{\partial C(x,t)}{\partial x}. \tag{I.1}$$

As we see, Fick's first equation states that the flux of particles J_x is proportional to minus the concentration gradient $\partial C/\partial x$. The proportionality factor is the diffusion coefficient. Indeed, if the concentration gradient between x and $x+\delta$ is e.g. *negative*, meaning that the concentration is higher at x than at $x+\delta$, there will be a net movement of particles from x to $x+\delta$, since there are more particles in x, hence a *positive* flux (cf. Figure I.2B).

Fick's second equation follows from the first, provided that the total number of particles is conserved. Consider a box where particles can only enter or exit through the sides (Figure I.2C). Thus, the concentration inside the box can only increase with time if the net flux of particles entering through the left, J_{in}, is bigger than the net flux of particles exiting through the right, J_{out}. In this case, the difference $J_{in} - J_{out}$ is positive and the gradient of fluxes, $\partial J/\partial x \equiv \lim_{\delta\to 0}\frac{J_{out}-J_{in}}{\delta}$, will be negative. In contrast, when the difference of fluxes $J_{out} - J_{in}$ is positive, the concentration inside the volume will decrease, hence $\partial C/\partial t < 0$. Finally, if $J_{in} = J_{out}$, the concentration inside the volume will remain unchanged, yielding $\partial C/\partial t = 0$. More generally, the continuity equation states that the concentration in a given volume can only change with time if there is a net flux of particles entering or exiting the volume:

$$\frac{\partial C}{\partial t} = -\frac{\partial J}{\partial x}. \tag{I.2}$$

Coupling Fick's first law with the continuity equation yields Fick's second law:

$$\frac{\partial C}{\partial t} = -\frac{\partial J_x}{\partial x} = -\frac{\partial}{\partial x}\left(-D\frac{\partial C}{\partial x}\right) = D\frac{\partial^2 C}{\partial x^2}. \tag{I.3}$$

Note that the last equality is only valid if the diffusion constant D is uniform in space. Fick's second law is also known as the diffusion equation. Indeed, given a diffusible protein with concentration $C(x,t)$, it states that the concentration at position x will change with time due to protein diffusion, which is proportional to the gradient of concentration fluxes.

I.3   The reaction-diffusion equation and the law of mass action

In the case where particles are created with rate s_0 at the boundary $x = 0$ and degraded with rate α, we obtain the continuous reaction-diffusion equation that I used to model morphogen gradient formation:

$$\frac{\partial C}{\partial t} = D\frac{\partial^2 C}{\partial x^2} - \alpha C\,, \tag{I.4}$$

with boundary condition $J(x = 0) = -D\left.\frac{\partial C}{\partial x}\right|_{x=0} = s_0$.

In order to model interactions between molecules or particles, e.g. Bcd trapping by the nuclei in the embryo or Dpp binding to the receptors on the cell surface of the wing imaginal disc, I also made use of the law of mass action. From a kinetics perspective, this law states that the rate of an elementary reaction (i.e. without transition states) is proportional to the concentration of the involved players. For example, a diffusing molecule that binds to a receptor has two forms: receptor-free, M_f, and receptor-bound, M_b. Thus, the binding of the molecule to the receptor (which has concentration R) occurs with a rate proportional to $M_f \cdot R$. Similarly, the unbinding of the molecule from the receptor is proportional to the concentration of the receptor-bound form, M_b. Assuming that the binding to the receptors is reversible, the equations for each form yield

$$\begin{cases} \frac{\partial M_f}{\partial t} = D\frac{\partial^2 M_f}{\partial x^2} - k_b M_f \cdot R + k_{-b} M_b \\ \frac{\partial M_b}{\partial t} = -k_{-b} M_b + k_b M_f \cdot R \end{cases}$$

where k_b and k_{-b} are the binding and unbinding affinities to the receptors, respectively. The term $k_b M_f \cdot R$ appears with a minus sign in the first equation because the concentration of free molecule will decrease as it binds to the receptor. On the other hand, the concentration of free molecule will increase as it unbinds from the receptor, hence the positive sign in front of $k_{-b} M_b$. In the second equation, these terms have opposite sign. Note that the diffusion term, $D\partial^2 M_f/\partial x^2$, only appears in the first equation because we assume that the molecule can no longer diffuse once it is receptor-bound.

Part II

Quantifying Scaling and Precision

....

This chapter introduces the quantification of scaling and precision that we apply in Parts III-IV.

II.1   Scaling

In developmental biology, understanding the robustness of gene expression patterns remains a great challenge. As we discussed in the *Prologue*, scaling ensures to keep a proportional body/organ plan despite the unavoidable fluctuations in embryo/tissue size. We define scaling as the relative response in a gene expression domain boundary x due to variations in tissue (or embryo) size L (de Lachapelle and Bergmann, 2010a):

$$S \equiv \frac{dx}{dL} \cdot \frac{L}{x} = \frac{dlog(x)}{dlog(L)} \tag{II.1}$$

Here, perfect scaling corresponds to $S = 1$. In this case, fluctuations in embryo length, dL/L, are exactly compensated by fluctuations in position, dx/x, implying perfectly conserved proportions (Figure II.1A). We use the terms hypo- and hyper-scaling to refer to $S < 1$ and $S > 1$, respectively. A position that hypo-scales does not compensate enough for a change in tissue size, meaning that in bigger tissues, the absolute domain boundary position is not shifted enough to keep the correct proportions (Figure II.1B). In contrast, hyper-scaling is the tendency to overcompensate for changes in tissue size (Figure II.1C).

In our analysis, we examined both morphogen gradients as well as several of their target genes, which are expressed in rather sharp domains where we assume that the boundary position matters more than the levels of the protein. In the case of sharp domains, we extracted the protein expression profiles (as deduced from fluorescent intensities) for each gene of interest in tissues/embryos of different sizes and asked whether the position of the boundary scales with tissue/embryo size (Figure II.2). The boundary of the domain was defined as the position with the steepest drop in protein concentration (Part III) or was obtained by fitting a Hill function to the protein profile (Part IV). We then estimated scaling by weighted linear regression :

$$w \cdot log(\frac{x}{\bar{x}}) = S \cdot log(\frac{L}{\bar{L}}). \tag{II.2}$$

where w are the normalized weights for each boundary position (coming from the fitting procedure) and $\bar{x}$, $\bar{L}$ are the average domain boundary and average tissue/embryo size, respectively (we divide by the means to center the data around zero in the plots, but this does not affect our definition of scaling since $dlog(x/\bar{x}) = dlog(x)$). The linear regression assumes that $log(x/\bar{x})$ and $log(L/\bar{L})$ are correlated.

Note that for small fluctuations in boundary position and embryo/tissue size (which is valid for Part III where fluctuations in embryo size are rather small), $log(x/\bar{x}) \simeq \frac{\Delta x}{\bar{x}}$ and $log(L/\bar{L}) \simeq \frac{\Delta L}{\bar{L}}$. Thus, we can
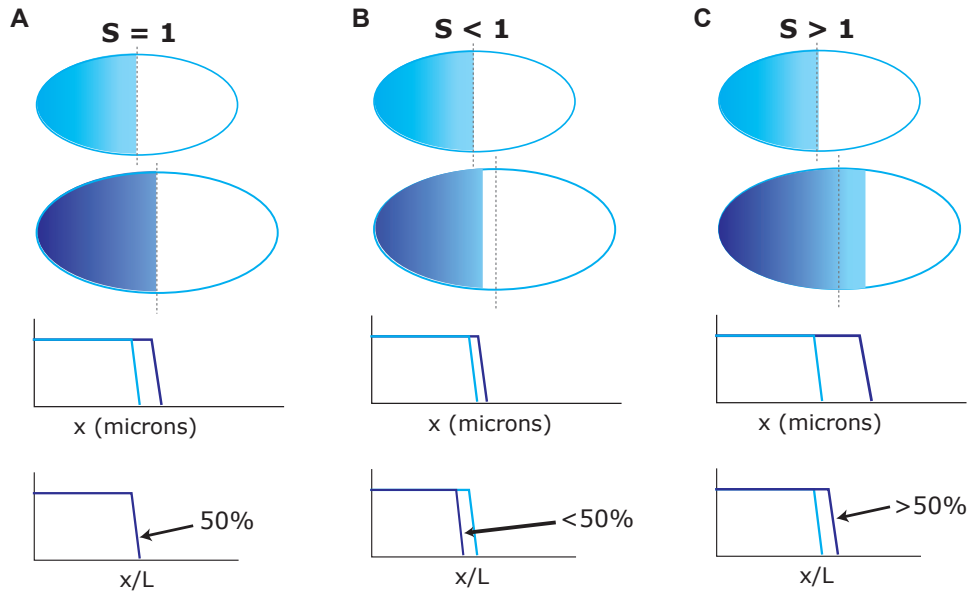
Figure II.1: Meaning of scaling. **(A)** Perfect scaling, $S = 1$. A domain that spans 50% of a small embryo (top) will also span 50% of a bigger embryo (below, dotted lines at 50% embryo size). The protein profiles collapse on relative units x/L, but not in absolute positions, x. Proportions are well preserved. **(B)** Hypo-scaling, $S < 1$. A domain that spans 50% of a small embryo does not expand enough in a bigger embryo, spanning less than 50% (dotted lines at 50% embryo size). **(C)** Hyper-scaling, $S > 1$. A domain that spans 50% in a small embryo expands too much in a bigger embryo, spanning more than 50% (dotted lines at 50% embryo size).

. . . .

estimate scaling by performing a linear regression of the domain positions x onto the embryo sizes L:

$$S \equiv \hat{\beta} \cdot \frac{\bar{L}}{\bar{x}} = \frac{cov(x,L)}{var(L)} \cdot \frac{\bar{L}}{\bar{x}} \tag{II.3}$$

where $\hat{\beta}$ is the *estimated* slope from a linear regression $x = x_0 + \beta L$ of the domain positions x_i onto their respective embryo sizes L_i.

In the case where we consider scaling across several developmental stages (like in Part IV where we investigate scaling throughout 40h of larval development), scaling might change with time. For example, the pattern could start hypo-scaling and then hyper-scaling, so that one should regress each time period separately to ensure that the data points are correlated, yielding a different scaling coefficient for each period.

For the morphogen gradients, we need to take into account the protein levels and not restrict ourselves to some artificial domain boundary (e.g. the decay length[2]), since protein levels are important for mophogen decoding and the gradient may scale differently across positions. As a first indication of scaling, one could look at the profiles in absolute and relative positions, x and x/L respectively. As we show in Figure II.3, scaling gradients from embryos/tissues of different sizes should collapse onto one single curve when plotted in relative units. In order to be more quantitative, we consider several thresholds of protein concentration read-out (*French flag* decoding (Wolpert, 1969)), though we could apply our scaling formalism to other decoding models (Figure II.4A). For a given threshold of activation, we plot the corresponding positions

---

[2]The decay length describes the position at which the concentration levels have decreased by a factor $e \simeq 2.7$, provided that we can fit a decaying exponential to the profile.
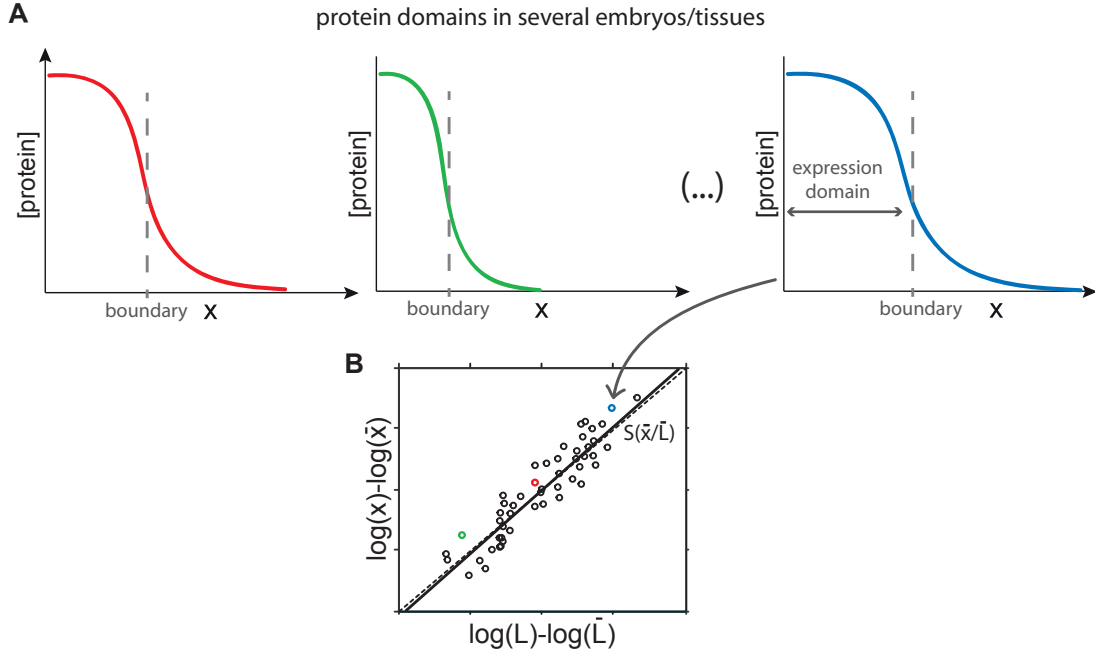
Figure II.2: Scaling of gene expression domains. *(A) For a collection of embryos (or tissues), we determine the boundary position x of some protein domain of interest as well as the embryo/tissue size L. (B) We plot the positions as a function of the sizes and determine scaling by linear regression (solid line; the dashed line is the diagonal which has a slope of one and thus corresponds to S = 1). In this example, scaling is close to perfect, with slight hyper-scaling. The colored circles refer to the protein profiles in A. Note that for small fluctuations, $log(x) - log(\bar{x}) \simeq \frac{\Delta x}{\bar{x}}$ and $log(L) - log(\bar{L}) \simeq \frac{\Delta L}{\bar{L}}$.*

. . . .

against the embryo/tissue sizes for our collection of embryos/tissues (Figure II.4B). Performing a linear regression (cf. Equation II.2) yields a scaling coefficient for that particular threshold, which we can then associate to the average relative position $\bar{x}/\bar{L}$. By repeating this procedure for several thresholds of activation, we obtain a scaling coefficient for several relative positions in the patterning field (Figure II.4C).

When modeling morphogen gradients, we can also determine scaling analytically, provided we can solve the equation for the morphogen gradient profile $M(x,t)$. Assuming again that the threshold concentration is fixed (implying $dM = (\partial M/\partial x)dx + (\partial M/\partial L)dL = 0$), it follows from Equation II.1 that the scaling coefficient is given by

$$S = -\left(\frac{\partial M}{\partial x}\right)^{-1} \cdot \frac{\partial M}{\partial L} \cdot \frac{L}{x} \tag{II.4}$$

Note that the above definition of scaling is generic and can be computed for any morphogen distribution $M(x,t)$ with explicit dependence on L.

Scaling and correlations

The scaling of gene expression domains has often been assessed as the correlation of the domain localization with embryo or tissue size (Aegerter-Wilmsen et al., 2005; Bollenbach et al., 2008; Houchmandzadeh et al., 2002; Holloway et al., 2006; Howard and ten Wolde, 2005). However, perfect correlation is in fact not
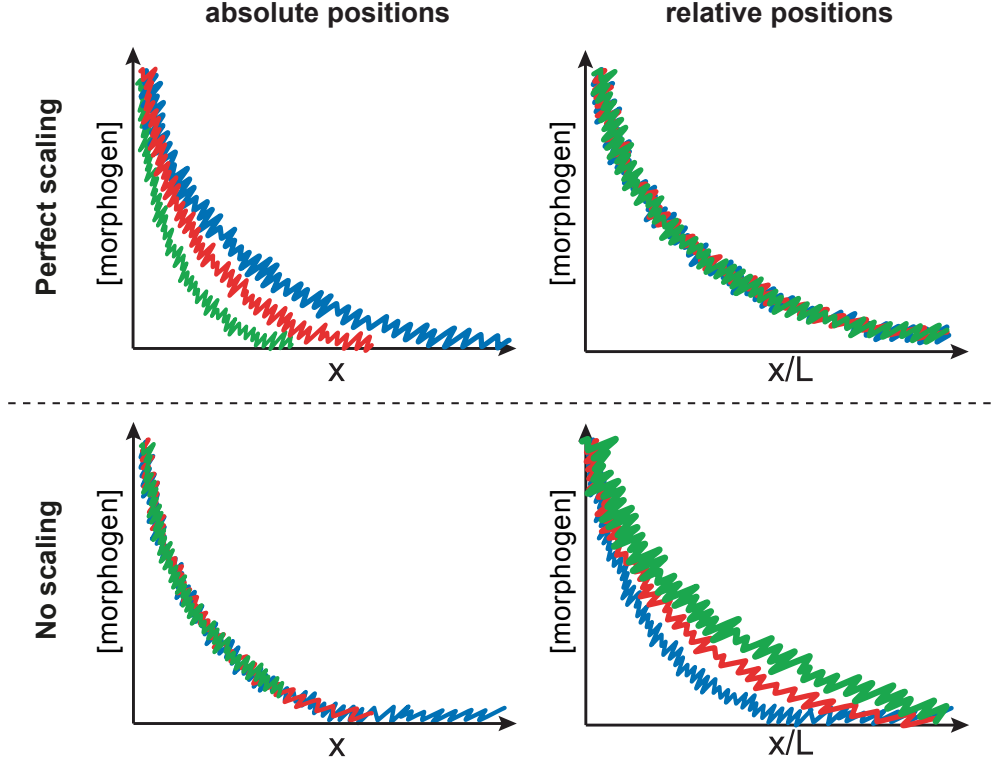
Figure II.3: Morphogen gradients in relative and absolute positions. *We consider three morphogen gradient profiles in a small (green), medium sized (red) and big (blue) embryo (or tissue). In the absence of scaling, the three profiles overlap in absolute positions x (top left). However, when plotting the profiles in relative positions x/L, they no longer collapse into one single curve (top right). In the case of perfect scaling, we observe the opposite scenario: profiles collapse in relative positions (bottom right) and not in absolute positions (bottom left).*

equivalent with perfect scaling, since it only guarantees a strictly linear response of the domain position with embryo/tissue size, but not the preservation of proportions. Thus, correlations are not informative on whether the shift is adequate, or if a specific domain position tends to hyper- or hypo-scale.

Importantly, the scaling measure that we introduced in Equation II.1 for continuous gradients and its equivalent for a discrete domain position in Equation II.3 not only provides a positional quantification of scaling, but also allows for proper disentanglement of scaling from precision, which is not the case when using correlations. This can be seen readily by writing the correlation in terms of our scaling measure in Equation II.1 as

$$corr(x, L) = S \cdot \frac{\sigma(L)/\bar{L}}{\sigma(x)/\bar{x}}. \tag{II.5}$$

Thus, the correlation depends explicitly on the fluctuations in the domain position $\sigma(x)$. However, fluctuations in domain position will generally also have a contribution that is not due to fluctuations in embryo/tissue size (e.g. internal noise). Consequently, correlations provide a scaling measure that does not properly disentangle such noise from the fluctuations induced by variations in length. Since fluctuations in parameters affecting gradient formation can give rise to position-dependent imprecision (cf. Part III), this is problematic. In contrast, our measure of scaling S only depends on the fluctuations in embryo length $\sigma(L)$ and the covariance $cov(x, L)$, since the slope was estimated by regressing x on L. Note that the correlation
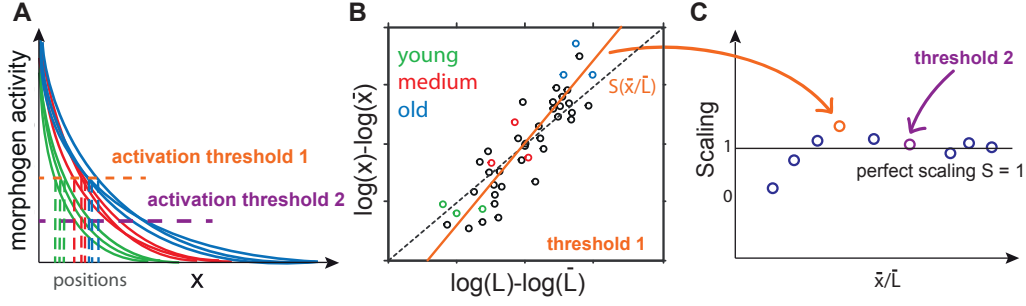
Figure II.4: Scaling of morphogen gradients. *We consider morphogen activity gradient profiles in small (green), medium sized (red) and big (blue) embryos (or tissues). (A) For a given morphogen activation threshold (e.g. threshold 1 in orange), we retrieve the corresponding positions (French flag decoding model (Wolpert, 1969)). (B) For that particular threshold, we plot the positions against the embryo (or tissue) sizes. Linear regression yields a scaling coefficient for that particular threshold, which in this example hyper-scales (solid line; the dashed line is the diagonal which has a slope of one and thus corresponds to $S = 1$). We relate this scaling coefficient to the average relative position $\bar{x}/\bar{L}$. Colored circles refer to the profiles in A. (C) We repeat this procedure for several activation thresholds and obtain a position-dependent picture of scaling for several relative positions in the patterning field. The black solid line indicates perfect scaling, $S = 1$. The arrows indicate the scaling coefficient obtained from the linear regression in B, where we had hyper-scaling (orange), as well as the scaling coefficient obtained for another concentration threshold (purple).*

. . . .

is equivalent to our scaling coefficient only if the mean fluctuation in domain position (relative to the mean domain position) is equal to that of the embryo/tissue size (relative to its mean size).

II.2   Precision

Here we use *precision* as a measure for the variability of domain position across individuals. For gene expression domains, we measure it as the variance of the domain boundary relative position, $Var(x/L)$.
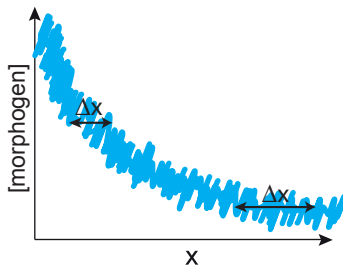

Figure II.5: Precision of morphogen gradients. *We show a morphogen profile as a function of position. Assuming a French flag decoding model (Wolpert, 1969), precision measures the variability in target gene domain position that would be induced by a particular threshold of activation (shown here for two different thresholds).*

. . . .

In the case of morphogen gradients, we want to assess the variation in position across individuals, Δx, which is induced by fluctuations of magnitude Δq in a parameter q affecting morphogen gradient formation. Δx can be estimated analytically to first order in Δq

$$\Delta x = \left| \frac{dx}{dq} \right| \Delta q = \left| \left(\frac{\partial M}{\partial x} \right)^{-1} \cdot \frac{\partial M}{\partial q} \right| \Delta q, \qquad \text{(II.6)}$$

where we used that the threshold concentration is fixed, yielding $dM = (\partial M/\partial x)dx + (\partial M/\partial q)dq = 0$ (note that Δx can be different at each position x). This estimate can be computed for any morphogen distribution $M(x,t)$ with explicit dependence on q.

Similarly to what we did for scaling, precision of morphogen gradients can also be assessed experimentally

at various thresholds of activation. In Figure II.5, we see a morphogen gradient with higher precision close to the source ($x = 0$), while further away, the imprecision Δx increases.

Variance induced by imperfect scaling

Imprecision, when measured as the standard deviation of the relative domain localizations $\sigma(x/L)$, is impacted both by "internal" noise as well as imperfect scaling. Therefore, in order to differentiate these two contributions, one should take into account that scaling can be position dependent (Figure II.4C). Assuming that the contributions to the total variance in domain position, $V_{tot} = \sigma_{tot}^2$, come from internal noise (V_{noise}) and imperfect scaling (V_{is}) and that these are independent, we have

$$V_{is}\left(\frac{x}{L}\right) = \bar{x}^2(1-S)^2 \cdot V\left(\frac{1}{L}\right). \tag{II.7}$$

In case of perfect scaling ($S = 1$), $V_{is}(x/L) = 0$ as expected.

Derivation of Equation II.7   Let us consider a set of embryos (or tissues) of different sizes having mean size $\bar{L}$. We assume that there are only fluctuations in embryo size and no other internal noise. If all the gradients scale perfectly with embryo size, they will look identical on a relative scale, so that the variance at a particular relative position will be zero, $V(x/L) = 0$. Our goal is to estimate the variance in the relative positions of expression domains due to imperfect scaling.

According to the definition of S in Equation II.1,

$$\frac{x - \bar{x}}{\bar{x}} = S \cdot \frac{L - \bar{L}}{\bar{L}},$$

in the case where fluctuations in boundary position and embryo size are small. Extracting x/L yields

$$\frac{x}{L} = S \cdot \frac{\bar{x}}{\bar{L}} + (1 - S) \cdot \frac{\bar{x}}{L}.$$

Assuming that $S \cdot \bar{x}/\bar{L}$ is constant when L fluctuates, we get Equation II.7

$$V_{is}\left(\frac{x}{L}\right) = \bar{x}^2(1-S)^2 \cdot V\left(\frac{1}{L}\right).$$

We see that for $S = 1$ (perfect scaling), there is no variance due to imperfect scaling (i.e. on a relative scale (x/L), all gradients coincide). However, when $S \neq 1$, the variance becomes larger for more posterior expression domains. This means that in the absence of perfect scaling, the imprecision of gradients measured in a relative scale increases towards the posterior pole, even without any internal noise, due to the fluctuations in embryo length.

Part III
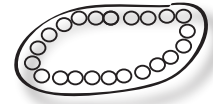
Drosophila early embryo development

. . . .

My first research project was on *Drosophila* early embryo development, with a focus on robustness. After spending a couple of months investigating theoretically the robustness of morphogen gradient formation and patterning to temperature changes (Lucchetta et al., 2005), I came across the work of (Gregor et al., 2005) and (Houchmandzadeh et al., 2002) on scaling of the Bicoid gradient and of its target gene Hunchback. Inspired by these papers and due to the available image data that had not yet been analyzed with regard to scaling, I made scaling the central focus of my thesis.

III.1   Overview of early *Drosophila* development

This work focused on the syncytial blastoderm period of embryogenesis (Figure III.1). In the following, I briefly present the different stages from oogenesis to embryogenesis until hatching into a larva. Larval stages will be described later, in Part IV. For a complete description of these stages, see (Gilbert, 2006; Slack, 2006) and references therein.

Even before fertilization, several important events occur during oogenesis. The initial germ cell divides four times, producing sixteen cells, one of which becomes the oocyte while the other fifteen become nurse cells. Interestingly, the position of the oocyte in the mother's ovary determines anterior-posterior (A/P) polarity while the position of the nucleus in the oocyte sets dorsal-ventral (D/V) polarity. Thus, when fertilization occurs, the egg is *already* activated and organizing itself as an embryo! Indeed, the mother, via the nurse cells, provides the oocyte with maternal mRNAs, like *bicoid* and *nanos*. These are translocated along polarized microtubules by motor proteins and end up in the anterior and posterior pole, respectively.

Following fertilization, the embryo undergoes a cleavage phase lasting about 2 hours at 26°C: nuclei divide synchronously eight times every 10 minutes and then start migrating to the periphery on concentric ellipsoids (cf. Figure III.1). Divisions continue, though at a slower rate, and at nuclear cycle 10, nuclei have reached the cortex where they keep dividing for another four times. During this period, the embryo is called a syncytial blastoderm, meaning that all the nuclei are contained within a common cytoplasm. Thus,
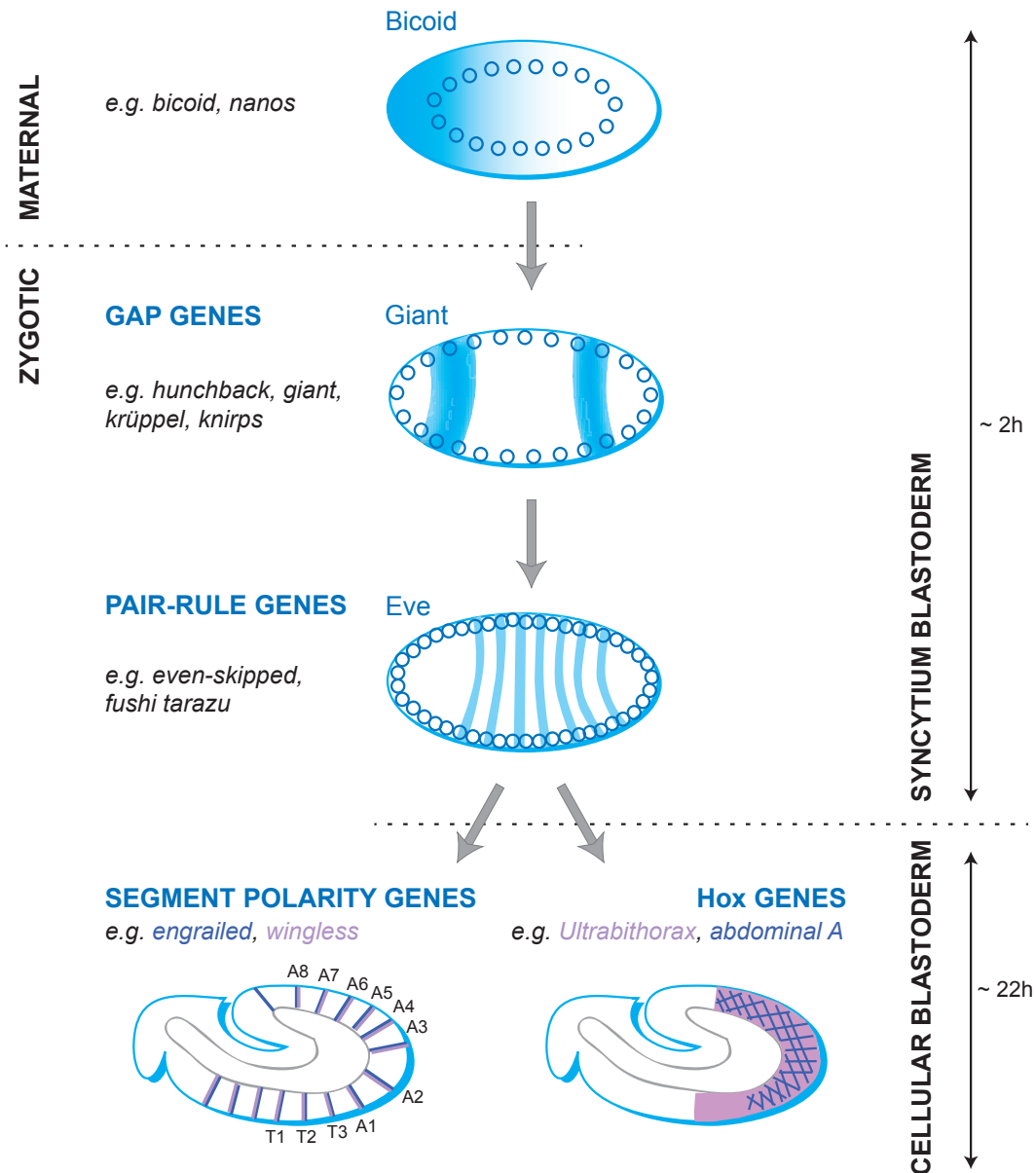
Figure III.1: Patterning along the A/P axis of the *Drosophila* embryo. *Several maternal genes are present in the syncytium embryo, yielding to the activation of target zygotic genes. The gap genes already subdivide the ectoderm into smaller domains. After activation of the pair-rule genes, the embryo cellularizes and undergoes gastrulation. The segment polarity genes and the Hox genes are mainly activated by the pair-rule genes, though a subset of gap genes also acts on the Hox genes. Together, segment polarity and Hox genes control the differentiation of the future larva segments.*

activation of genes does not require any sophisticated cell signaling. Instead, proteins acting at this stage are transcription factors that enter the nuclei where they can activate their target genes. At the top of the hierarchical cascade patterning the A/P axis, the maternal genes form protein gradients by diffusion and activate zygotic target genes, called gap genes. Gap genes specify the rough body plan along the head-to-tail axis and also interact with each other, but there is no evidence of any feedback from the zygotic back to the maternal genes. The signaling cascade continues with the activation of the so-called pair-rule genes. These genes control the formation of every second body segment.

Following division cycle 13, the oocyte membrane folds inwards and cellularizes the nuclei at the cortex, creating the cellular blastoderm. Patterning along the A/P axis continues while the D/V axis starts to be specified. The transition from specification (cell fate can still be altered in response to signals from other cells) to determination (cell fate becomes cell-intrinsic) is mediated by the segmentation genes (Gilbert, 2006). The segment polarity genes encode proteins from the Wingless and Hedgehog signal transduction pathways. They divide the embryo into a series of segmental primordia with A/P polarity, by activating *engrailed* in one half of each segment. After the segmental boundaries are set, the gap and pair-rule genes also interact to regulate the homeotic selector genes, which determine the identity of each segment. Thus, by the end of the blastoderm stage, each segment has acquired a unique identity through a combination of gap, pair-rule and homeotic gene products (Gilbert, 2006).

III.2   The Bicoid morphogen gradient

Mutagenesis screen experiments by Christiane Nüsslein-Volhard and Eric Wieschaus in the 1980s allowed to identify several genes involved in development, such as *bicoid*, granting them the Nobel Prize in 1995. The gradient of Bicoid (Bcd) was the first gradient to be detected experimentally (Driever and Nüsslein-Volhard, 1988b) and it was shown to determine position in the embryo in a concentration-dependent manner (Driever and Nüsslein-Volhard, 1988a). Indeed, altering *bcd* dosage shifted the embryo fate map posterior-wards in the case of an increase, anterior-wards in the case of a decrease. Thus, Bcd became the first known example of a morphogen gradient after Morgan and Wolpert had proposed the existence of diffusible substances capable of patterning developing tissues decades before (Morgan, 1901; Wolpert, 1969). Bcd is involved in the patterning of the anterior-posterior axis of the Drosophila embryo. In the absence of *bicoid*, embryos completely lack anterior structures like head and thorax, while in *exuperantia* and *swallow* mutants, which ensure proper localization of *bicoid* mRNA at the anterior pole in the wild-type (wt), the Bicoid gradient is less steep, extends farther into the posterior of the egg, and produces embryos lacking most anterior structures (Berleth et al., 1988).

Following this discovery, Driever and co-workers showed that Bcd is a transcription factor capable of activating several targets in a concentration-dependent fashion (Driever et al., 1989). In parallel, Struhl et al. showed that the *hunchback* gene contains multiple regulatory elements to respond to the Bcd protein (Struhl et al., 1989). Since then, many studies have revealed the existence of regulatory interactions between Bicoid and the gap and pair-rule genes (Driever et al., 1989; Driever and Nüsslein-Volhard, 1989; Struhl et al., 1989; Hoch et al., 1991; Stanojevic et al., 1991; Small et al., 1992; Small et al., 1993; Small et al., 1996; Ma et al., 1996; Burz et al., 1998; Gao and Finkelstein, 1998; Hoch et al., 1992).
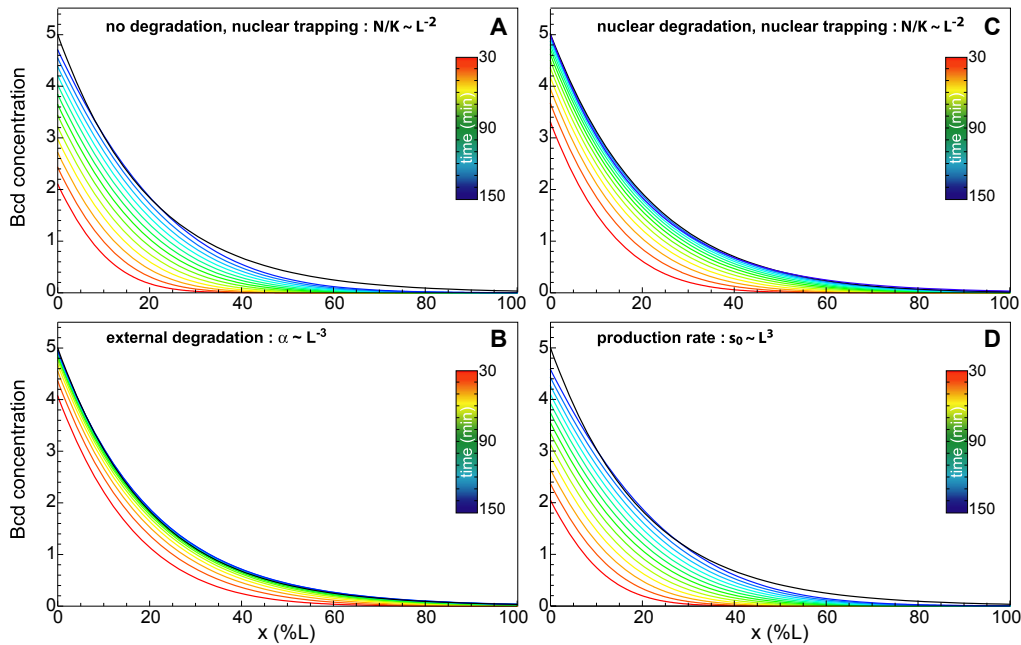
Figure III.2: Time evolution of the Bcd gradient for various models. *For comparison, a decaying exponential with length scale $\lambda = 0.2L$ is shown in black.* **(A)** *No degradation.* $N/K = 8 \cdot (L/\bar{L})^{-2}$. $D = 12\mu m^2/s$. **(B)** *Degradation is performed by a fixed number N_P of proteasomes, such that their density P scales as $k_\alpha P = 5 \cdot 10^{-4}\,\sec^{-1}(L/\bar{L})^{-3}$ (no nuclear trapping).* $D = 5.4\mu m^2/s$. **(C)** *Nuclear degradation, $\alpha_n = 3.3 \cdot 10^{-4}\,\sec^{-1}$, and nuclear trapping, $N/K = 2 \cdot (L/\bar{L})^{-2}$.* $D = 8.3\mu m^2/s$. **(D)** *The production rate is correlated with embryo size, $s_0 \propto L^3$ (no degradation nor nuclear trapping).* $D = 1.5\mu m^2/s$.

. . . .

III.3 Robustness of *Drosophila* early embryo development?

The 1980-1990s have unveiled many of the regulatory interactions between the genes involved in early embryo development. More recently, much attention has been drawn to how such regulatory networks ensure proper and robust patterning under varying conditions. Indeed, the positions of the gene expression domains downstream of Bcd are remarkably insensitive to fluctuations in the external environment (Crauk and Dostatni, 2005; Houchmandzadeh et al., 2002; Lucchetta et al., 2005) and their relative proportions are maintained across embryos of different sizes. The latter feature is referred to as scaling and occurs within a single species (Houchmandzadeh et al., 2002; Lott et al., 2007) and also across different species (Gregor et al., 2005; Gregor et al., 2008; Lott et al., 2007).

Recent experiments also showed that the Bcd gradient itself is rather precise and that its length scale correlates to the embryo size (Gregor et al., 2007a; He et al., 2008; He et al., 2010). These new findings suggested that the precision and scaling of Bcd target genes could, at least in part, be attributed to that of the morphogen gradient itself. Therefore, we focused on single morphogen models that aimed at explaining the precision and scaling of Bcd target genes at the level of the morphogen gradient. However, it is likely that other mechanisms (e.g. Bcd interactions with the *staufen* gene (Aegerter-Wilmsen et al., 2005), gap genes interactions (Jaeger et al., 2004; Jaeger et al., 2004; Jaeger et al., 2007; Jaeger and Reinitz, 2006; Manu et al., 2009a; Manu et al., 2009b), Bcd interactions with maternal Hunchback and the terminal system (Ochoa-Espinosa et al., 2009), or bistability (Lopes et al., 2008) also contribute to further increase

robustness in A/P patterning.

Recent data suggests that the total Bcd gradient does not stabilize before nuclear cycle 14 while nuclear concentrations appear to be quasi-steady (within 10%) from cycle 10 onwards (Gregor et al., 2007b). Indeed, recent work (Bergmann et al., 2008; Coppey et al., 2007) provided convincing arguments that steady nuclear concentrations can only be achieved by an external gradient which has not yet reached steady state. This is because the increase in external Bcd is compensated by the increasing number of nuclei of changing size, leaving the concentration in each nuclei rather constant. Yet, direct measurements of the tail of the Bcd profile were too noisy to demonstrate a super-exponential decay, which is a signature of a pre-steady state gradient (Bergmann et al., 2007). Thus, I decided to seek indirect evidence on the dynamics of the Bcd gradient in terms of its robustness properties. This included its sensitivity to changes in *bcd* dosage (Bergmann et al., 2007) as well as its precision and scaling that I investigated during my thesis.

I looked into these questions with an already published dataset of staining images for these target genes (Bergmann et al., 2007). I first investigated the embryo-to-embryo variability of the gene expression domain boundaries for various *bcd* dosages. I found that precision was position dependent more than gene dependent. Since increasing evidence suggested that Bcd precision was essential for the precision of the target genes, I modeled the Bcd gradient formation and investigated theoretically its precision and scaling at pre-steady and steady state. I used this model and my results on precision to predict the scaling of Bcd target genes as measured in this dataset.

III.4 Modeling Bicoid gradient formation

I considered a model where Bcd is produced at the anterior pole with production rate s_0, diffuses according to a uniform diffusion constant D, is degraded at uniform rate α and is trapped and released by the nuclei in the embryo at rates k_n, k_{-n}, respectively (Figure III.2; cf. Part I). The corresponding coupled partial differential equations for the free morphogen concentration and the nuclear concentration are (Bergmann et al., 2007; Coppey et al., 2007; Gregor et al., 2007b):

$$\begin{cases} \frac{\partial M}{\partial t} = D\frac{\partial^2 M}{\partial x^2} + s_0\delta(x) - \alpha M + k_{-n}M_n - k_nM \cdot N \\ \frac{\partial M_n}{\partial t} = -k_{-n}M_n + k_nM \cdot N \end{cases} \tag{III.1}$$

where N is the nuclei density and $s_0\delta(x)$ the source term localized at $x = 0$ (anterior pole). I considered zero-flux (Neumann) boundary conditions at the posterior pole and that there was no morphogen at the initial time ($t = 0$). Assuming that nucleo-cytoplasmic exchanges occur rapidly ($\partial M_n/\partial t|_M \approx 0$), as suggested by the experiments reported in (Gregor et al., 2007b), the following effective diffusion equation holds for the free morphogen (Bergmann et al., 2007):

$$\frac{\partial M(x,t)}{\partial t} = \tilde{D}\frac{\partial^2 M(x,t)}{\partial x^2} - \tilde{\alpha}M(x,t) + \tilde{s}_0\delta(x), \tag{III.2}$$

where $\tilde{D} = \frac{D}{1+N/K}$, $\tilde{\alpha} = \frac{\alpha}{1+N/K}$ and $\tilde{s}_0 = \frac{s_0}{1+N/K}$ (with $K = \frac{k_{-n}}{k_n}$). At time $t = 150$min (corresponding to nuclear cycle 14), the morphogen profile was numerically fitted to an exponentially decaying gradient with

length scale $\lambda = 0.2L$. Thus, given a degradation rate, the diffusion constant was adjusted accordingly (see figure legends for numerical values). I accounted for the presence of the nuclei by setting $N/K = 1$ (i.e. to be specific, I assumed that the probability that external Bcd is trapped equals the probability that nuclear Bcd is released; yet the qualitative results for precision were robust with respect to the exact choice of N/K). According to the nuclear trapping model (Equation III.1), the nuclear concentration c_n, which determines the precision and scaling of Bcd target genes, is given by

$$c_n \equiv \frac{M_n}{N \cdot v_n} = \frac{M}{K \cdot v_n}, \qquad \text{(III.3)}$$

where v_n is the nuclear volume. Importantly, according to Equation III.3 the fluctuations in the nuclear concentrations c_n are proportional to those in the external gradient, provided we can neglect variability in v_n and K.

III.5 Imaging Bicoid target genes

I used an interface developed in MATLAB® (Figure III.3) in order to extract and analyze expression profiles from the staining images of different Bcd target genes, yielding quantitative information on the positions of each protein domain (Bergmann et al., 2007). I further improved this interface in order to measure precision and scaling in a total of 154 staining images using a semi-automated analysis tool where the position of the anterior and posterior poles were marked 50 times for each embryo. Based on this input, the software extracted a rectangular region from the image from which it generated protein concentration profiles and automatically determined the positional information of the gap and pair-rule gene expression domains. Thus, for each embryo, its total length and the domain localizations were characterized by a mean value and a standard deviation computed from the 50 markings. The right boundary of the anterior Hb expression domain was defined as the position at which the decline in the Hb concentration was the steepest. Gt and Kr domains were described similarly by their left and right boundaries, while the position with maximal staining intensity characterized their center. Since the Eve stripes were not so broad, we only extracted the position at which the Eve intensity was maximal. (All the staining images are available for download at http://www.unil.ch/cbg/morphogen/Images_paper.zip.)

With the help of Dr. Zoltan Kutalik in our lab, I also developed a second, almost fully automated analysis tool that only required human input for selecting which of the two poles was the anterior one (while their position was detected automatically). This tool was used to analyze the expression domain positions of a set of 119 embryos, which included 78 (51%) from the first analysis. Precision and scaling results turned out to be similar for the two analysis approaches, suggesting that measurement errors and inter-embryo variation (e.g. because of different embryo orientations) could be safely neglected (Figure III.15). Note however that part of the observed fluctuations in embryo size may be induced by the fixation procedure. Yet, computing precision and scaling by considering embryos on different slides separately yielded similar results (Figure III.16), suggesting that fixation or other batch effects have no positional bias towards either of the embryo poles.
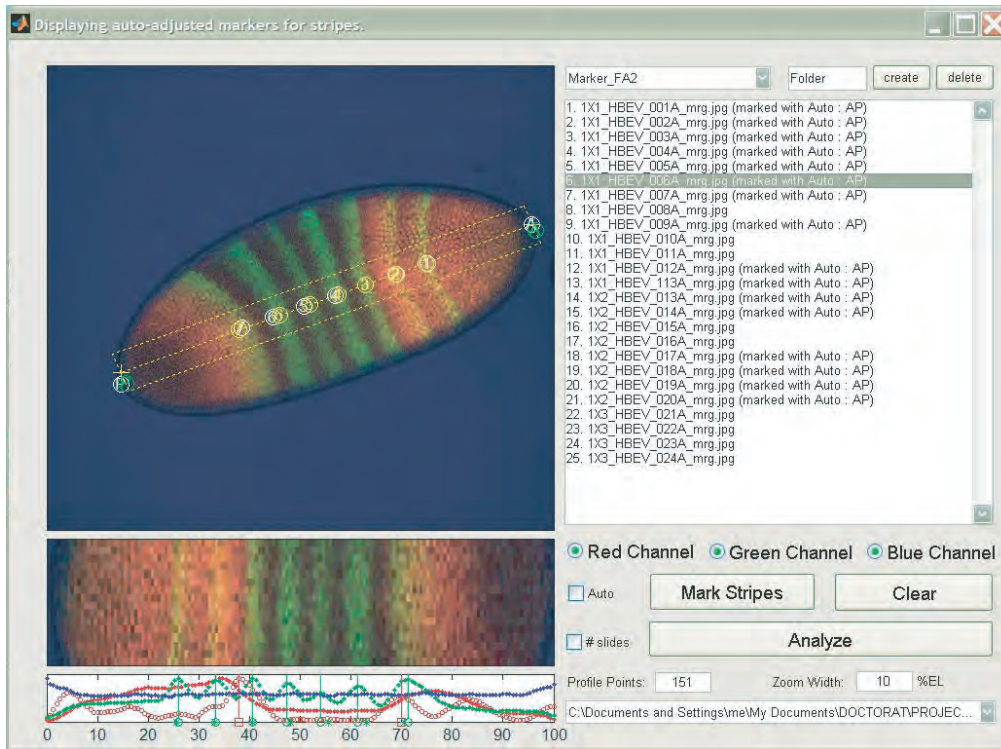
Figure III.3: Graphical user-interface for analysis of the gap and even-skipped gene expression domains. *The three color profiles (green: Eve stripes, red: Hb, Gt or Kr, blue: background) are extracted for every embryo and shown below the main embryo image. Once all the embryos have been marked, the data can be analyzed, returning precision and scaling features over all the embryos and for a wide range of positions in the embryo.*

. . . .

III.6   Precision of Bcd and its targets

III.6.1   The gap and pair-rule gene expression domains are precise

I studied staining images of 154 *Drosophila melanogaster* embryos at cleavage cycle 14 (see Methods in (Bergmann et al., 2007)). Using image processing tools described in Section III.5, I measured the relative positions x/L of the protein domains of Gt, Hb, Kr and Eve. I screened both wild type (wt) embryos with two copies of *bcd* and mutant strains with one or four *bcd* copies resulting in shifted expression domains. I observed that in mid-embryo, the standard deviation of the relative domain localizations, $\sigma(x/L)$, was between 1 and 2% and it tended to be higher $(2-4\%)$ toward the anterior and posterior poles (Figure III.4).

These results were in good agreement with previous results on the Hb domain (Houchmandzadeh et al., 2002). Importantly, they showed that the precision of the target genes had the same magnitude and positional trend as that of Bcd according to Gregor et al. (Gregor et al., 2007a). Moreover, this experimental analysis showed that precision was more *position* dependent than *gene* dependent. Indeed, when expression domains were shifted due to different *bcd* mRNA dosages, the precision of the domain seemed to change according to the new domain position. Overall, all the investigated domains followed similar precision trends. This suggested that a major contribution to the precision of the Bcd target genes could be attributed to the morphogen gradient itself. Note that position dependent precision may also arise as an experimental artifact
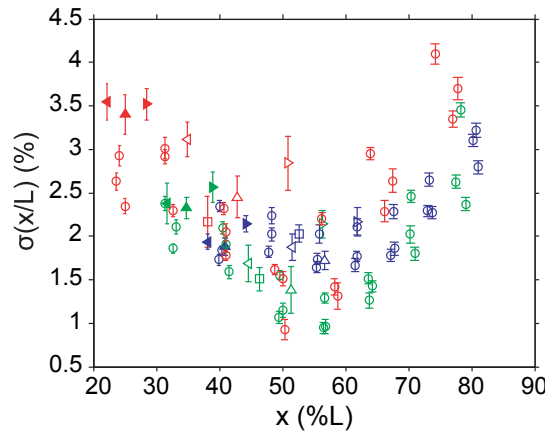
Figure III.4: Measuring precision of Bcd target genes. *Measure of imprecision $\sigma(x/L)$ of the gap and pair-rule gene expression domains as a function of position x along the A/P axis. Errors (bars) were estimated by computing the standard deviation across 50 independent (semi-automated) markings of expression domain boundaries. For Gt (filled triangles) and Kr (empty triangles), we show results for their left ($\triangleleft$) and right ($\triangleright$) boundaries, as well as for the center (Δ) of their expression domain. The Hb ($\square$) domain is characterized by the boundary where its concentration drops and the Eve stripes ($\circ$) by the position at which their intensity is maximal. Color code: 1xbcd (red), 2xbcd = wild type (green), 4xbcd (blue).*

. . . .

when analyzing embryos with different orientations and varying ages (times classes T5-T8 of nuclear cycle 14) or because of imperfect scaling (cf. Equation II.7). However, estimating the maximal contributions of such effects, I still found that the higher precision at mid-embryo was statistically significant (see below).

Statistical significance of the precision results

The analysis of the gap and pair-rule gene expression domain boundaries showed that precision is highest at mid-embryo (cf. Figure III.4). However, position dependent precision may also arise as an experimental artifact when analyzing embryos with different orientations (Surkova et al., 2008). At mid-embryo, this effect is the smallest, which could explain at least part of the observed dip in the precision profile (Figure III.4). In the analysis, I measured expression profiles close to the line connecting the anterior and posterior pole, where the impact of the embryo orientation was the smallest. I performed a control experiment where expression profiles were extracted 50 times moving uniformly from the top of the embryo (above the A/P line) to its bottom (below the A/P line). I found that the positional variations induced by this (extreme) procedure were indeed smallest at mid-embryo (Figure III.5). Nevertheless, for all the seven Eve stripes, I found that the median intra-embryo variation (across these 50 extractions) was below the measured inter-embryo variation. Thus, while I cannot exclude that part of the observed positional dependence of precision is due to our experimental procedure, it cannot explain it entirely. In particular, it can hardly account for the relatively strong increase of positional fluctuations towards the anterior pole.

A second potential systematic error could arise from the fact that the posterior gap gene expression domains exhibit significant drift towards the anterior pole in cleavage cycle 14 (see Figure 1 in (Jaeger et al., 2004)). Apparently, such a drift is also transmitted to the Eve expression domains, where it is strongest (~5%L) for stripe 7 (see Table 1 in (Surkova et al., 2008)). This is relevant because the staining images were taken

in time classes T5-T8 of cycle 14 (based on the fact that the Eve expression domains have been clearly separated) and part of the observed positional fluctuations $\sigma(x/L)$ may be a result of combining data from different time classes. Yet even for stripe 7, this effect may at best contribute half of the observed variance, while it cannot account for the increased imprecision observed for the more anterior expression domains (see Figure III.6).
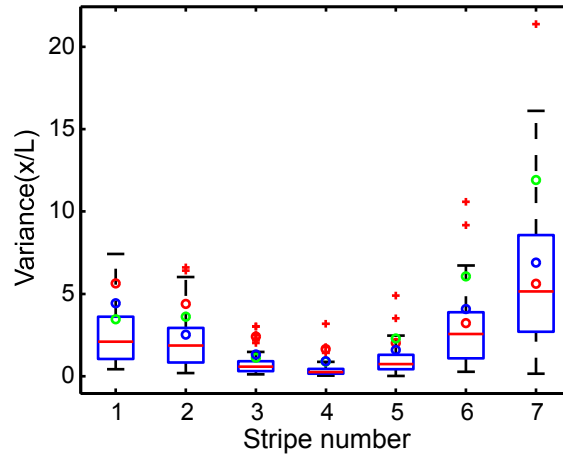

Figure III.5: Contribution of embryo orientation to the precision of Eve stripes. *Measured variance of the Eve stripes (○) co-stained with Hb (red), Kr (green) and Gt (blue), for wild-type bcd mRNA dosage (cf. Figure III.4). In the control experiment, expression profiles in a given embryo were extracted 50 times moving uniformly from the top of the embryo (above the A/P line) to its bottom (below the A/P line), giving an estimate of the maximal positional variance induced by embryo orientation. This was repeated for all embryos in our dataset with wild-type bcd mRNA dosage. Thus, the boxplot represents the distribution of maximal positional variabilities due to embryo orientation.*

. . . .
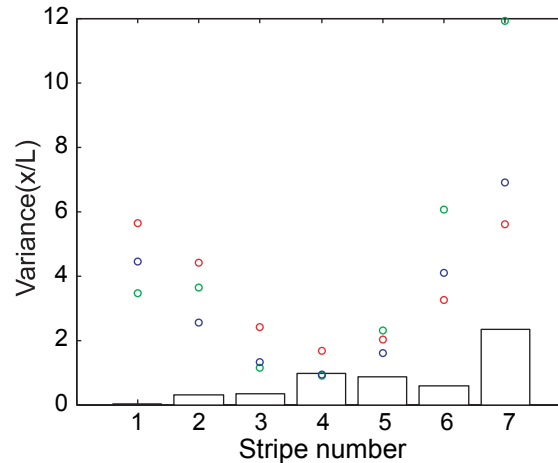
Figure III.6: Contribution of domain drift to the variance of the Eve stripes. *Measured variance of the Eve stripes (○) co-stained with Hb (red), Kr (green) and Gt (blue), for wild-type bcd mRNA dosage (cf. Figure III.4), together with the estimated maximal variances (bars) that could arise due to the drift of these domains. These estimates were computed as $d_i^2/12$, i.e. the variance of a uniform distribution between 0 for images taken early (T3 for stripes i=1,2,3,4,7 and T4 for stripes i=5,6), and the maximal drift d_i for images of the last time class T8 (cf. Table 1 of (Surkova et al., 2008)). These are upper bounds since our images were likely taken in T5-T8 based on the differentiation of the Eve domains.*

. . . .

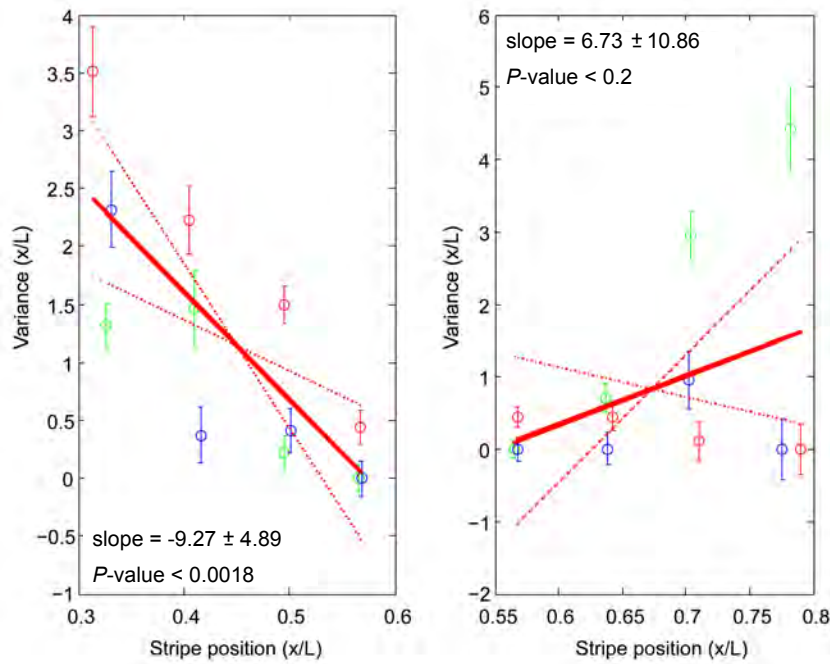**Figure III.7:** Statistical significance of the increase in precision from the anterior pole to mid-embryo. *Measured variance of the Eve stripes (○) co-stained with Hb (red), Kr (green) and Gt (blue), for wild-type bcd mRNA dosage (cf. Figure III.4), corrected for domain drift and measurement errors. Errors (bars) were estimated by computing the standard deviation across 50 independent (semi-automated) markings of expression domain boundaries. The two linear regressions (red lines) were perfomed to assess statistically if mid-embryo domains are significantly more precise. On the left, the linear regression (performed with data from stripes 1 to 4) shows that the increase in precision towards the mid-embryo is indeed highly significant (P-value < 0.0018). On the right, we also observed a trend of increasing precison towards the mid-embryo using data from stripes 4 to 7), however this effect is not significant (P-value < 0.2) when using the corrected variance.*

· ·

In Figure III.7, I show the residual variance when subtracting the variance due to changes in embryo orientation *and* domain drifts (negative values are set to zero). In order to test for the statistical significance of position dependence in this residual variance, I performed two linear regressions: one with the data of stripes 1 to 4, and a second with the data of stripes 4 to 7. I found that the increase in precision from the anterior pole towards the mid-embryo was still highly significant with a P-value < 0.0018. In contrast, while I also observed a trend of increasing precison towards the mid-embryo from the posterior pole, this effect was not very significant (P-value < 0.2) when using the corrected variance. Thus within this very conservative estimate of the residual positional variance, I cannot rule out that the larger contributions towards the posterior pole mainly stem from averaging signals of embryos with different orientation and domain drifts.

III.6.2 Modeling precision in a single morphogen framework

Given the aforementioned pieces of evidence for achieving precision at the level of the gradient formation, I considered a French-flag model (Wolpert, 1969), where a single morphogen gradient induces expression domain boundaries in a concentration dependent manner (see Section III.4). During the first stages of embryonic development, thirteen nuclear divisions occur about every 10 minutes. Gregor et al. (Gregor

et al., 2007b) have shown that nuclear concentrations stabilize at cycle 10. From a robustness perspective, this means that Bcd read-out is based on a similar number of molecules and so the associated stochastic noise does not change significantly from cycle 10 on. Therefore, I focused on the noise propagated from the *external* Bcd gradient formation and investigated the fluctuations in domain localization when perturbing the various parameters characterizing gradient formation (i.e. production, degradation, diffusion and nuclear trapping rates).

Within my modeling framework (cf. Section III.4), I investigated the precision of morphogen gradient formation upon fluctuations in the production, diffusion, degradation and nuclear trapping rates (cf. Equation II.6). In Figure III.8, I plot the imprecision measure Δx for small fluctuations (5%) in each of these parameters (Figure III.8A-D) as well as a combination of fluctuations in the production and diffusion rates (Figure III.8E). I choose $t_i = 100$ min as the time when patterning was initiated, but qualitatively the results only depended on the relation between t_i and the Bcd decay time $\tau = 1/\alpha$ (for $t_i \ll \tau$ the gradient has not yet reached steady state). I found that fluctuations in s_0 (Figure III.8A) gave rise to decreasing Δx toward the posterior pole if the pre-steady state gradient was read out (i.e. for small α). In contrast, for fluctuations in D or α, Δx increased toward the posterior pole (Figure III.8B-C). In the case of Bcd pre-steady state decoding (bluish curves) and fluctuations both in the production and degradation or diffusion rates (Figure III.8E), maximal precision of the Bcd gradient around mid-embryo can arise naturally, as was indeed observed directly by Gregor et al. (Gregor et al., 2007a) and indirectly by my analysis (cf. Figure III.4).

Within this modeling framework, the positional increase of precision toward the anterior pole observed in Figure III.7 by itself favored the pre-steady state decoding of Bcd. This is because such an increase could only be driven by noise in the production rate (see Figure III.8A) or nuclear binding (Figure III.8D) at early times, before the gradient had reached steady state. Thus, my modeling approach provided a proof of principle that maximal precision at mid-embryo can arise if the gradient is decoded before steady state. Yet, I cannot rule out that other mechanisms may yield such a pattern of precision, even for a steady-state Bcd gradient (e.g. in the wing disc, it was shown that cell-to-cell variability in the production, diffusion and degradation rates can yield higher precision around mid-field (Bollenbach et al., 2008)).

In this analysis, I assumed the classical French-flag model (Wolpert, 1969), where domain boundaries are determined from critical morphogen concentrations. This was a simplification and recent work by the Reinitz group (Jaeger et al., 2004; Jaeger et al., 2004; Jaeger and Reinitz, 2006; Jaeger et al., 2007; Surkova et al., 2008; Manu et al., 2009a; Manu et al., 2009b) showed that the gap gene expression domains are established by a highly dynamic process characterized by drifting domain boundaries. In my model, I neglected the gap gene dynamics and rather focused on how precision can be achieved already at the level of the Bcd gradient, as suggested by our data (as well as those reported in (Gregor et al., 2007a; He et al., 2008)). Thus, I assumed that the positional information was transmitted relatively early to the gap genes, around cycle 10 (as suggested in (Lucchetta et al., 2008)). The precision and proper scaling of this read-out could then be maintained (and possibly refined) by the gap genes independently of Bcd, due to their cross- and auto-regulation (Bergmann et al., 2007). Meanwhile, the Bcd gradient continued to evolve, eventually decaying at the onset of gastrulation (e.g. by activation of its PEST domain (Niessing et al., 1999)).

I want to emphasize that my search for models was guided by simplicity, but also took into account the results from (Gregor et al., 2007b) who showed that nuclear Bicoid concentrations remain constant from
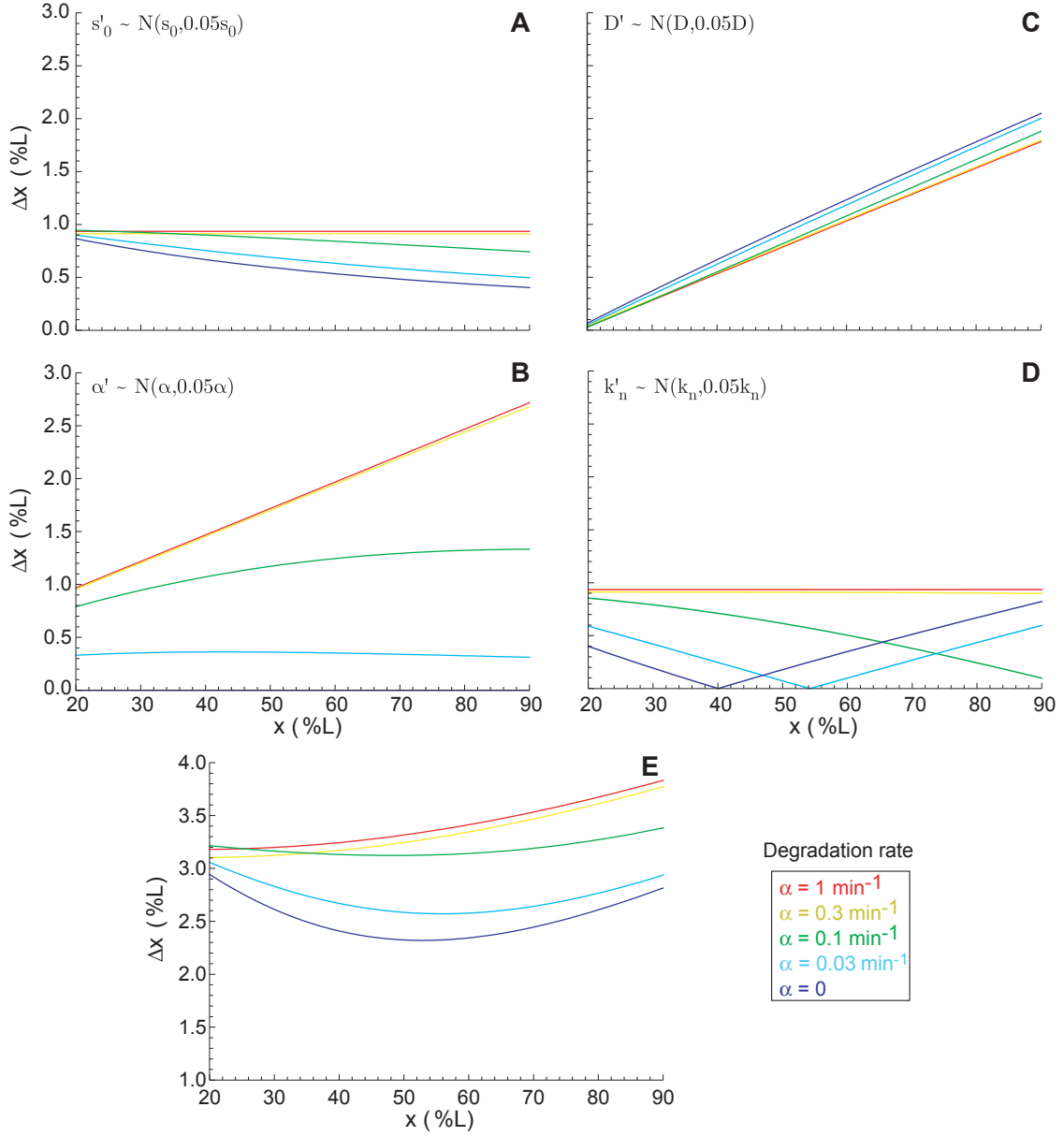
Figure III.8: Modeling precision of Bcd. *Predicted positional variability at $t = 100\,$min for various Bcd degradation rates $\alpha = \{1, 0.3, 0.1, 0.03, 0\}\,$min^{-1} and their corresponding diffusion constants $D = \{146, 46, 17, 9, 3.5\}\,\mu m^2/s$, ensuring $\lambda = 0.2L$ at $t = 150\,$min (numerical fit); 5% noise was added to the Bcd profile at the level of (A) the production rate, (B) the diffusion constant, (C) the degradation rate and (D) the nuclear trapping rate. In (E), a combination of noise was added to the Bcd profile (17% on the production- and 6% on the diffusion rate; all contributions are assumed to be independent). (L = 1 for all embryos.)*

cycle 10 on. Yet, as was argued previously (Bergmann et al., 2008), this is fully consistent with, and actually requires an evolving external gradient. Thus, nuclear trapping slows down the evolution of the external gradient which remains longer at pre-steady state, while the divisions and changing sizes of the nuclei effectively stabilize the nuclear concentrations. Interestingly, this intriguing combination might actually yield an increase in overall pattern robustness, since the external super-exponentially decaying profile is more robust to fluctuations in the production, degradation and nuclear trapping rates (Bergmann et al., 2007; de Lachapelle and Bergmann, 2010a), while stabilized nuclear concentrations allow for longer averaging times, thus reducing read-out stochastic noise (Saunders and Howard, 2009). Coexistence of an external pre-steady state gradient and *stabilized* nuclear concentrations might therefore be the start to robust patterning in *Drosophila* early embryo development (de Lachapelle and Bergmann, 2010b).

Figure III.9: Modeling scaling of Bcd. *Parameters are chosen such that the profile is closest to an exponential decay with length scale $\lambda = 0.2L$ at $t = 150$min (see Figure III.2 for temporal evolution).* **(A)** *No degradation, scaling is brought about by nuclear trapping, with $N/K = 8 \cdot (L/\bar{L})^{-2}$. $D = 12 \mu m^2/s$.* **(B)** *Degradation is performed by a fixed number N_P of proteasomes, such that their density P scales as $k_\alpha P = 5 \cdot 10^{-4} \sec^{-1} (L/\bar{L})^{-3}$ (no nuclear trapping). $D = 5.4 \mu m^2/s$.* **(C)** *Nuclear degradation, $\alpha_n = 3.3 \cdot 10^{-4} \sec^{-1}$, and nuclear trapping, $N/K = 2 \cdot (L/\bar{L})^{-2}$. $D = 8.3 \mu m^2/s$.* **(D)** *No degradation nor nuclear trapping, scaling appears because the production rate is correlated with embryo size, $s_0 \propto L^3$. $D = 1.5 \mu m^2/s$.*

Figure III.10: Modeling scaling of Bcd with nuclear trapping. *Parameters are chosen such that the profile is similar to an exponential decay with length scale $\lambda = 0.2L$ at $t = 100$min. (A) No degradation, scaling is induced by nuclear trapping, with $N/K = 8 \cdot (L/\bar{L})^{-1}$. (A') As in (A), but with $N/K = 100 \cdot (L/\bar{L})^{-1}$. With $n = 1$, scaling is position dependent at late times, but too small except for the most anterior region. (B) No degradation, scaling is induced by nuclear trapping, with $N/K = 8 \cdot (L/\bar{L})^{-3}$. (B') As in (B), but with $N/K = 2 \cdot (L/\bar{L})^{-3}$. Later times exhibit stronger hyper-scaling in the anterior region.*

. . . .

III.7 Scaling of Bcd and its targets

III.7.1 Modeling scaling

I considered four models that imposed a functional relationship between the embryo length and model parameters. For each model, I investigated how the scaling coefficient $S(x,t)$ changed as a function of position and time (cf. Section II.1).

In the first model (Figure III.9A), there was no morphogen degradation at all, such that the external profile never reached steady state. This extreme case, also considered by (Coppey et al., 2007), assumed that there was no significant morphogen degradation until the decoding time, though it might be triggered latter on. Indeed, there was no evidence of Bcd degradation until a very recent publication this year which estimated Bcd half life to be < 15min (Liu and Ma, 2011). I assumed that at the decoding time, all embryos had the same number of nuclei independent of their size (which was in good agreement with the deterministic doubling of nuclei at each cycle). Thus the nuclei density N depended on the embryo size like $N \propto L^{-n}$, where $n \in [1; 3]$ (here $n = 3$ corresponds to a uniform distribution of nuclei, while $n = 2$ is true if nuclei are distributed on a shell with a fixed width (Gregor et al., 2007b)). I found that in general scaling was time and position dependent (in Figure III.9A I set $N/K = 8 \cdot (L/\bar{L})^{-2}$, see Figure III.10 for dependence
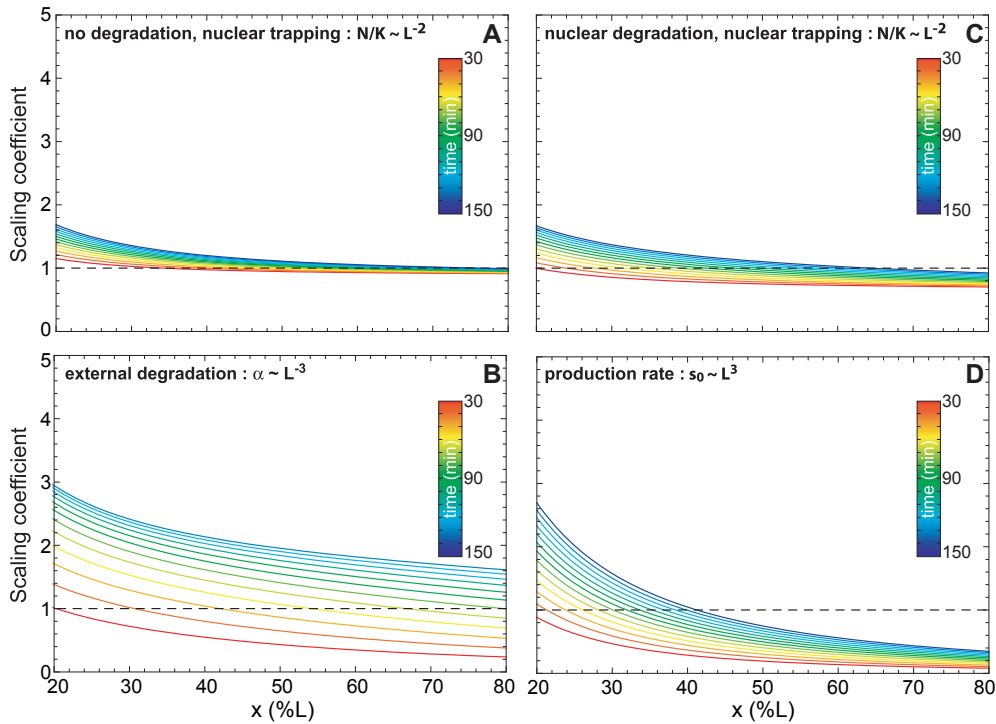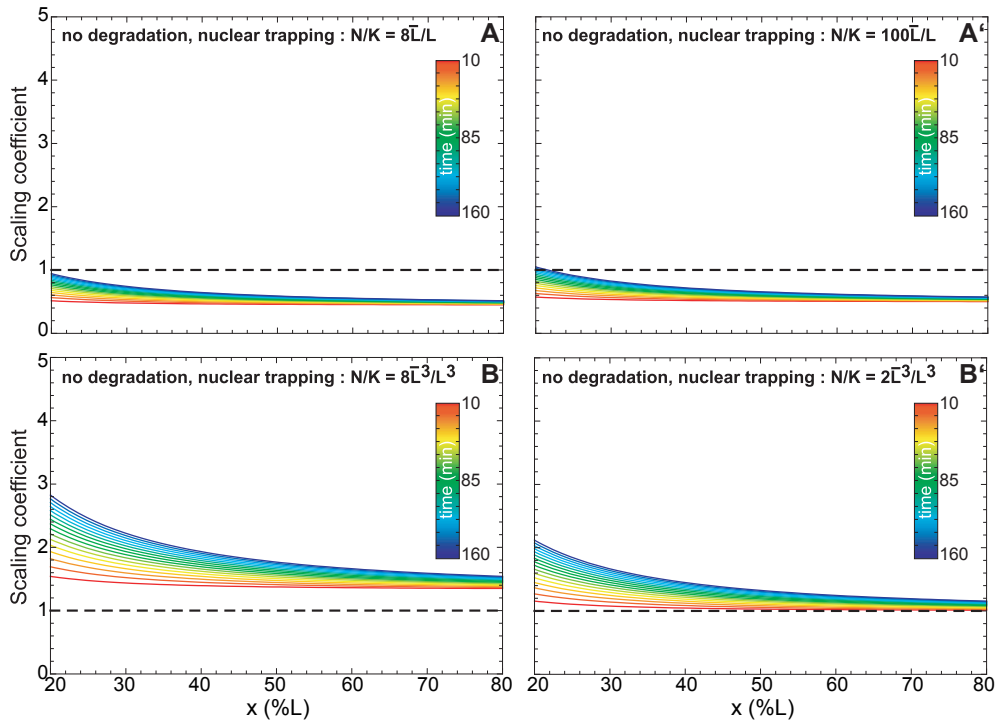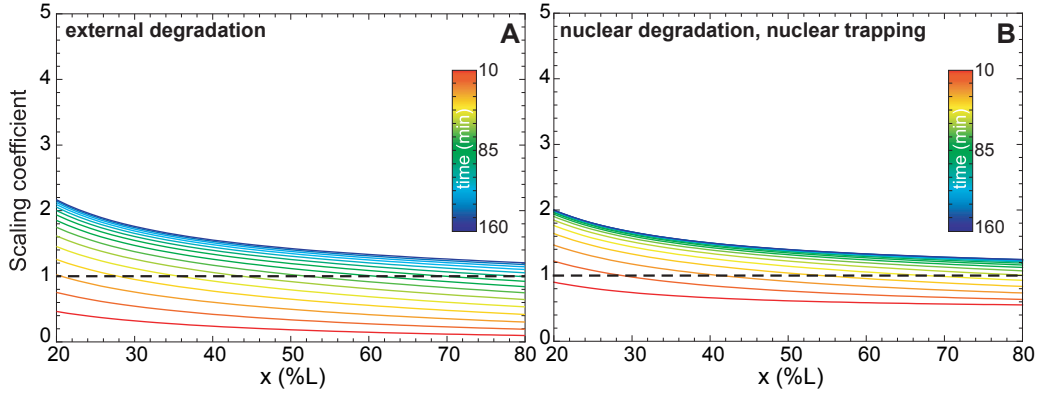
Figure III.11: Modeling scaling of Bcd with degradation. *Parameters are chosen such that the profile is closest to an exponential decay with length scale $\lambda = 0.2L$ at $t = 100$min. (A) Degradation is performed by a fixed number N_P of proteasomes, with $k_\alpha P = 5 \cdot 10^{-4} (L/\bar{L})^{-2} \sec^{-1}$ (no nuclear trapping). (B) Nuclear degradation, $\alpha_n = 0.002 \sec^{-1}$, and nuclear trapping, $N/K = 1 \cdot (L/\bar{L})^{-2}$.*

of scaling on N/K and n). Specifically, anterior domain boundaries hyper-scaled, in particular if decoding occurred relatively late, while posterior domains showed very good scaling for a wide range of decoding times. Note that if there were also cytoplasmic degradation, S would actually decrease towards zero as converging to the steady state, since the length scale is given by $\lambda = \sqrt{\tilde{D}/\tilde{\alpha}} = \sqrt{D/\alpha}$, which is independent of L, cf. Equation III.2. Importantly, the qualitative behavior of scaling did not depend on the particular value of n (see Figure III.10). Therefore, even if the density of nuclei at the embryo surface were to correlate more weakly with embryo size (Fowlkes et al., 2008), effectively yielding $1 < n < 2$, scaling could still be achieved. Note also that nuclei progressively migrate to the cortex of the embryo, arranged on concentric shells (Foe and Alberts, 1983). At the cortex, nuclei continue dividing, but their radii become smaller (Gregor et al., 2007b). Thus, at each nuclear cycle, it is likely that the ratio N/K does not increase by a factor 2, but much less. For simplicity I therefore considered this ratio to be constant in my simulations and analytical results.

In a second model (Figure III.9B), I assumed that the effective degradation rate scaled inversely with the size of the embryo, i.e. $\alpha = k_\alpha P \propto L^{-p}$, with $p \in [1; 3]$. For example, one may imagine that degradation is performed by a fixed number of proteasomes N_P originally deposited in the embryos such that the proteasome density P varies between embryos as a function of their length: $P \propto N_P/L^p$. Interestingly, this model ensured some scaling even at steady state (we set N/K to zero, to disentangle this effect from scaling due to nuclear trapping). From Figure III.9B, it is apparent that early decoding times show large posterior hypo-scaling, while later decoding times exhibit hyper-scaling even in the posterior-most domains. Specifically, good scaling was only achieved in a small region, which varied as a function of the decoding time. Note that if I were to consider $p = 2$, the steady state length scale would correlate perfectly with embryo size, $\lambda \propto L$. However, this would not be enough to ensure a scaling coefficient $S = 1$ everywhere in the embryo (Figure III.11A). This is because the amplitude of the profile, $s_0/\sqrt{\alpha D}$, also depends through α on the embryo size ($\propto L^{p/2}$) and thus, domain boundaries very close to $x = 0$ would move much more than expected because of the change in amplitude at the source, yielding hyper-scaling.

A third model (Figure III.9C) assumed that degradation occurred in the nuclei with rate α_n, as suggested
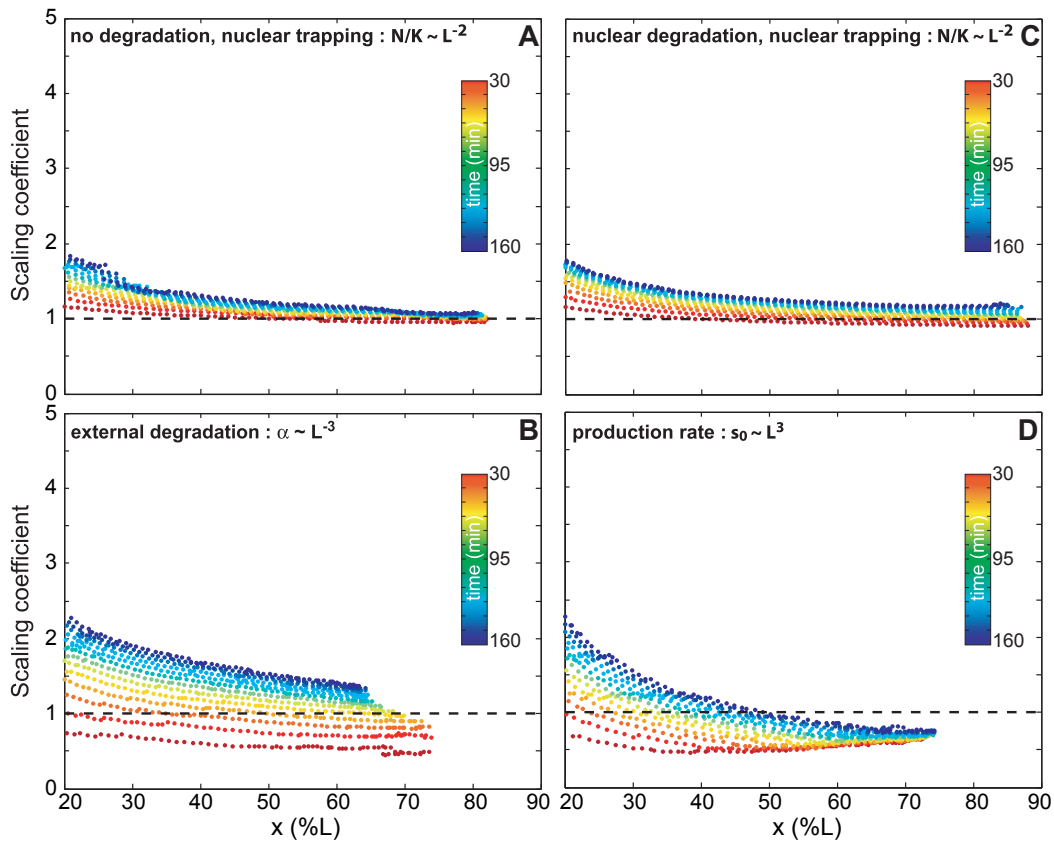
Figure III.12: Numerical results for Bcd scaling in the presence of noise. *Parameters are chosen such that the profile is closest to an exponential decay with length scale $\lambda = 0.2L$ at $t = 100$ min. For many Bcd read-out thresholds, we plot the fluctuations in domain position as a function of the fluctuations in embryo size. The scaling coefficient at each threshold is obtained using Equation II.3. 25% Gaussian noise is added to all parameters to test for robustness. (A) No degradation, scaling is brought about by nuclear trapping, with $N/K = 8 \cdot (L/\bar{L})^{-2}$. (B) Degradation is performed by a fixed number N_P of proteasomes, such that their density P scales as $k_\alpha P = 5 \cdot 10^{-4} \sec^{-1} (L/\bar{L})^{-3}$ (no nuclear trapping). (C) Nuclear degradation, $\alpha_n = 5 \cdot 10^{-4} \sec^{-1}$, and nuclear trapping, $N/K = 8 \cdot (L/\bar{L})^{-2}$. (D) No degradation nor nuclear trapping, scaling appears because the production rate is correlated with embryo size, $s_0 \propto L^3$.*

in (Gregor et al., 2007b), yielding an effective degradation rate $\tilde{\alpha} = \frac{\alpha_n \cdot N/K}{1 + N/K}$. In this model, the morphogen profile converged to a steady state which still exhibited some degree of scaling. For simplicity, I assumed that there was no external degradation ($\alpha = 0$) and that the early embryo only used its nuclear machinery to degrade Bcd. Compared to the simple nuclear trapping model (Figure III.9A), it seemed that nuclear degradation yielded more hypo-scaling in posterior domains, as well as more variability in time. Note that for $N/K \gg 1$, scaling results were very similar to those obtained only with nuclear trapping. In this case, increasing nuclear degradation speeded up the convergence to steady state, but did not really change the scaling behavior at *decoding* time.

Finally, in a fourth model (Figure III.9D), the number of ribosomes or the amount of *bcd* mRNA deposited by the mother was assumed to be correlated with embryo size, such that the production rate scaled as $s_0 \propto L^s$ (as proposed by (He et al., 2008)). In this model, the amplitude of the profile was bigger in larger embryos, giving rise to hyper-scaling in the regions close to the source (in Figure III.9D we set $s = 3$ and $N/K = 0$, to avoid scaling due to nuclear trapping). However, this model was less flexible to decribe scaling, since the

exponent *s* was its only tunable parameter. Within our parameter constraints to ensure the correct gradient length scale at nuclear cycle 14, it appeared that this model yielded strong hypo-scaling in the posterior domains, in particular at early decoding times.

I also investigated the robustness of scaling when adding noise to the parameters involved in gradient formation (Figure III.12). I introduced noise by drawing these parameters from a normal distribution centered at their mean μ and with standard deviation $\sigma = 0.25 \cdot \mu$. Scaling was robust in all scenarios. However, the scenarios including nuclear trapping were less sensitive to fluctuations in Bcd decoding time, since the scaling behavior had a similar trend over a large window of times (Figure III.9A/C). Moreover, these models showed good scaling in most of the embryo, avoiding strong hyper- and hypo-scaling in the anterior and posterior regions, respectively.

In order to fully validate the nuclear trapping models, more experiments should be performed. According to these models, Bcd should not scale at all in the absence of nuclei. This prediction could in principle be tested in unfertilized eggs using our expression for scaling. Yet it would require direct measurements of continuous Bcd concentrations along the gradient (e.g. using a *bcd*-GFP fusion) for many embryos, which is more difficult than quantification by proxy using its target genes (as I did in this study). Note that if Bcd were not to scale in unfertilized eggs, considering the profiles in relative units (x/L) could be misleading, since the profiles would fluctuate due to the variability in embryo size, in particular in the posterior compartment.

III.7.2   The gap and pair-rule gene expression domains scale with embryo size

The nuclear trapping model investigated above predicted that scaling should be *position* rather than *gene* dependent and close to perfect except for anterior most domains. In order to test these two predictions experimentally on our collection of staining images, I adapted the continuous definition of the scaling coefficient $S(x,t)$ in Equation II.1 for discrete measurements (cf. Equation II.3 and Figure II.2). In our data, the relative embryo size fluctuations were of the order of 10-15%. The results of my scaling analysis for Gt, Hb, Kr and Eve in embryos with single, double and quadruple *bcd* dosage are presented in Figure III.13. I observed that anterior domains indeed tended to hyper-scale, while mid-embryo and posterior domains showed good scaling (P-value < 0.00021 in Figure III.14). Moreover the magnitude of the scaling coefficient depended mainly on the position of the respective domain (boundary) rather than the associated gene. This was in good agreement with the model discussed above, where scaling was a long-range feature achieved *via* nuclear trapping of external Bcd and was then transmitted to the nuclear concentrations (c.f. Equation III.3). Indeed, although other models produced anterior hyper-scaling, only those considering nuclear trapping with (Gregor et al, 2007b) or without nuclear degradation avoided significant hypo-scaling in posterior domains and were robust to fluctuations in decoding time. Still, these *ad hoc* models merely served to provide a proof of principle that the observed scaling pattern could arise naturally, while a more realistic scenario may combine several of their features, as well as other mechanisms that channel information on the embryo size to the morphogen gradient.

The precision analysis of the gap and pair-rule gene expression domains revealed that precision was maximal at mid-embryo, in good agreement with Bcd precision (Gregor et al., 2007a). Thus, considering that
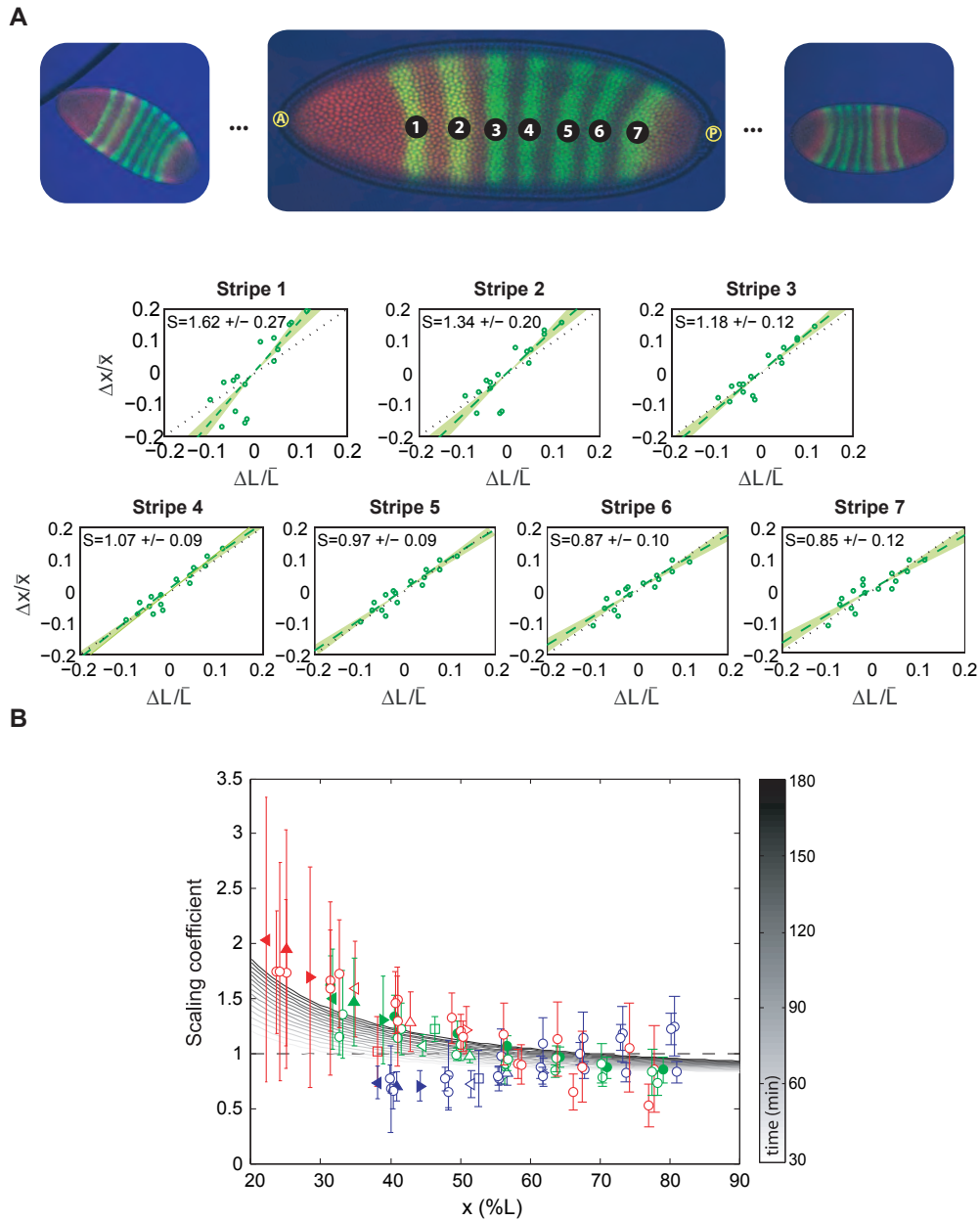
Figure III.13: Measuring scaling of Bcd target genes. *(A) Scaling of the Eve stripes for wild-type bcd mRNA dosage, co-stained with Hb. For each stripe in each embryo (images on top), fluctuations in the domain position $\Delta x/\bar{x}$ are plotted against fluctuations in embryo size $\Delta L/\bar{L}$. Scaling coefficients are then estimated by linear regression. Errors show 68% confidence intervals from the regression analysis. (B) Measured scaling coefficients of the gap and pair-rule gene expression domains as a function of position x along the A/P axis. Errors (bars) represent 68% confidence intervals from the linear regression. For Gt (filled triangles) and Kr (empty triangles), we show results for their left (◁) and right (▷) boundaries, as well as for the center (△) of their expression domain. The Hb (□) domain is characterized by the boundary where its concentration drops and the Eve stripes (○) by the position at which their intensity is maximal (filled circles represent Eve co-stained with Hb, as in (A)). Color code: 1xbcd (red), 2xbcd = wild type (green), 4xbcd (blue). We also show scaling as predicted by our model (grayscale). Parameters are chosen such that the profile is closest to an exponential decay with length scale $\lambda = 0.2L$ at $t = 150$ min (see Figure III.2A for Bcd temporal evolution).*

Bcd target genes exhibited good scaling in general (Figure III.13B), the standard deviation of the relative positions $\sigma(x/L)$ presented in Figure III.4 essentially measured internal noise, since $V_{is}(x/L) \approx 0$ (cf. Equation II.7, except for the anterior most region of the embryo).

Previous works employed the correlation of the decay length of a (measured) morphogen profile with embryo size as a proxy for scaling (Bollenbach et al., 2008; Gregor et al., 2005). Focusing on the decay length has however several disadvantages: (i) It ignores a potentially position dependent signal for scaling, (ii) it is only applicable for profiles that are well-described with a single fixed length scale (like an exponential decay), and (iii) it does not capture the impact of the morphogen gradient amplitude on scaling. In this work, I have argued that scaling should in general be viewed as a position dependent feature (cf. Figure III.13B). In this sense, the measure of scaling introduced in Equations II.1-II.3 can be used for gradients with an arbitrary functional dependence and accounts for the gradient amplitude. Importantly, it provides a unified quantification of scaling for both gradient profiles (cf. Equation II.1 that I used for the modeling approach) and expression domains (cf. Equation II.3 which is the numerical equivalent that I applied to the experimental study). Note that Equation II.4 introduced an explicit model for the morphogen gradient interpretation (i.e. the French-flag model (Wolpert, 1969)). Yet, in principle dx/dL could be computed also for other, more intricate mechanisms that transmit positional information from the morphogen to its target genes, either numerically or in some cases analytically.

Statistical significance of the scaling results

An important conclusion drawn from the image analysis was the hyper-scaling of anterior-most domains. Given the large error bars for these domains, I performed a nested model analysis in order to investigate
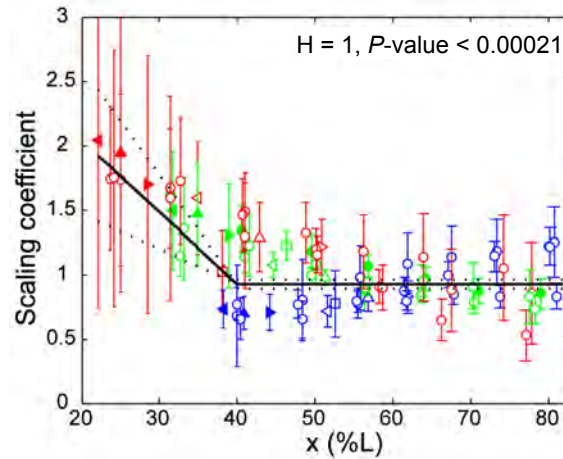
Figure III.14: Statistical significance of anterior hyper-scaling. *Measured scaling coefficients of the gap and pair-rule gene expression domains as a function of position x along the A/P axis. Errors (bars) represent 68% confidence intervals on the linear regression. For Gt (filled triangles) and Kr (empty triangles), we show results for their left (◁) and right (▷) boundaries, as well as for the center (△) of their expression domain. The Hb (□) domain is characterized by the boundary where its concentration drops and the Eve stripes (○) by the position at which their intensity is maximal (filled circles represent Eve co-stained with Hb, as in III.13A). Color code: 1xbcd (red), 2xbcd = wild type (green), 4xbcd (blue). The linear regression (black line) was performed to assess the significance of anterior hyper-scaling, against a model where scaling is uniform and equal to the weighted mean of the measured scaling coefficients. The P-value was given by an F test (see Subsection III.7.2).*

Figure III.15: Scaling and precision with the semi-automated (50x marking) vs automated extration tool. *(A) Scaling coefficients obtained using embryos marked 50 times manually or once automatically. (B) Precision results obtained using embryos marked 50 times manually or once automatically. For Gt (filled triangles) and Kr (empty triangles), we show results for their left (◁) and right (▷) boundaries, as well as for the center (Δ) of their expression domain. The Hb (□) domain is characterized by the boundary where its concentration drops and the Eve stripes (○) by the position at which their intensity is maximal. Color code: 1xbcd (red), 2xbcd = wild type (green), 4xbcd (blue).*

. . . .

the statistical significance of this conclusion (Figure III.14). My first model was a simple regression $S = w_{mean}$, where $w_{mean} = \frac{\Sigma_i \left(\frac{1}{\sigma_i}\right)^2 \cdot S_i}{\Sigma_i \left(\frac{1}{\sigma_i}\right)^2}$ is the weighted mean of the scaling coefficients (with σ_i the standard error of the scaling measurement). This model was nested within my second model $S = w_{mean} + \beta(x - x_0)(x < x_0)$ which accounted for hyper-scaling of domains localized in $[0; x_0]$ (note that the estimate of the slope, β, was obtained taking into account the standard errors of the scaling measurements). Because the second model had more parameters than the first one, it always fitted the data better. Thus, in order to determine if the second model gave a *significantly* better fit to the data, I made use of an **F** test. The **F** statistic is given by $F = \frac{\frac{RSS_1 - RSS_2}{p_2 - p_1}}{\frac{RSS_2}{n - p_2}}$, where RSS_i is the weighted residual sum of squares of model i, p_i the number of parameters of model i and n the number of data points used to estimate the parameters of both models. Under the null hypothesis, the second model does not provide a signicantly better fit than the first model and **F** will have an **F** distribution with $(p_2 - p_1; n - p_2)$ degrees of freedom. In Figure III.14, we see that the null hypothesis was rejected with a P-value < 0.00021. Therefore, despite the big error bars in our measurements, the fact that all the anterior measurements ($x < x_0$ with $x_0 = 40\% L$) showed hyper-scaling was still a significant result.

III.8 Bicoid-related controversies

This work has been published in January 2010 (de Lachapelle and Bergmann, 2010a). It also introduced the measure for scaling quantification under the approximation of small fluctuations in embryo size (cf. Part II). In the following, I discuss the debate raised by our publication as well as by very recent findings in the

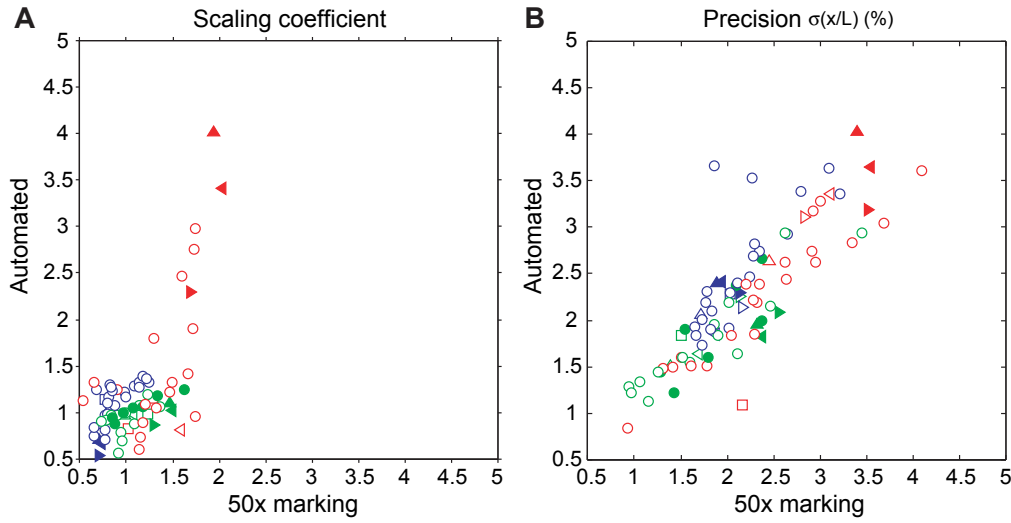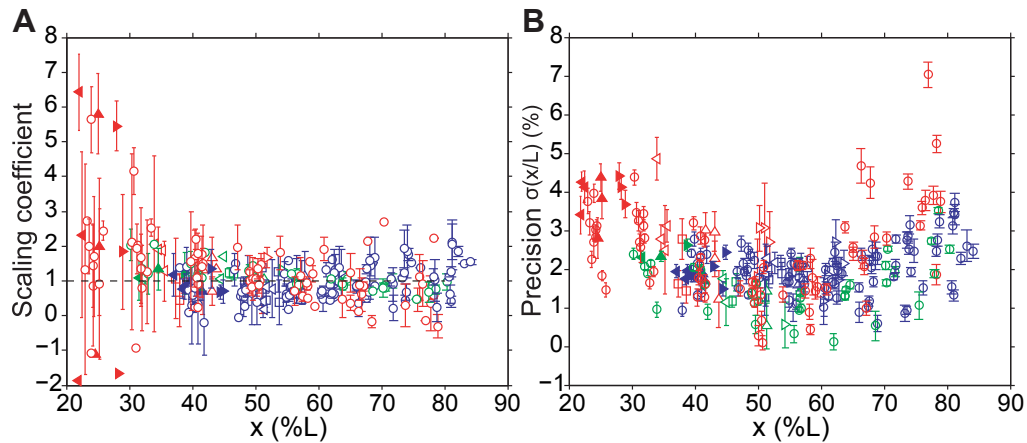**Figure III.16:** Scaling and precision of Bcd target genes on different slides. *(A) Measured scaling coefficients of the gap and pair-rule gene expression domains as a function of position x along the A/P axis, for embryos on the same slide. Errors (bars) represent 68% confidence interval on the linear regression. (B) Measure of precision σ(x/L) of the gap and pair-rule gene expression domains as a function of position x along the A/P axis, for embryos on the same slide. Measurement errors (bars) were estimated by computing the standard deviation across 50 independent (semi-automated) markings of expression domain boundaries. For Gt (filled triangles) and Kr (empty triangles), we show results for their left (◁) and right (▷) boundaries, as well as for the center (△) of their expression domain. The Hb (□) domain is characterized by the boundary where its concentration drops and the Eve stripes (○) by the position at which their intensity is maximal. Color code: 1xbcd (red), 2xbcd = wild type (green), 4xbcd (blue).*

. . . .

field. Indeed, for the past 10 years, research on Bicoid has been very prolific and has raised controversy on various questions.

First, some studies have argued that the precision and scaling of the Bicoid gradient cannot explain that of its target genes (Houchmandzadeh et al., 2002). In this sense, many models have been proposed to explain the increased precision and scaling of the boundaries of the target gene expression domains, e.g. considering a second gradient (Howard and ten Wolde, 2005) or the interactions among the target genes (Jaeger et al., 2004). My work, however, suggested that the Bicoid morphogen gradient is essential for the precision and scaling of the target gene domains, though mechanisms certainly exist to further refine their boundaries. Interestingly, as my investigations were under progress, two studies reported that the Bicoid gradient was actually precise and scaled with embryo size to some extent (He et al., 2008; Gregor et al., 2007a). However, suggesting that Bicoid precision and scaling are necessary for target genes robustness was strongly debated during the reviewing process of our manuscript and once it was published (see (Jaeger, 2010) and our reply (de Lachapelle and Bergmann, 2010b)). My conclusions were based on two observations: (i) Precision and scaling seemed more position dependent than gene dependent and (ii) the observed imprecision of the target genes was in good agreement with the measured Bcd imprecision. However, one of the main criticisms of (Jaeger, 2010) was that the embryos were not always oriented the same way during imaging, suggesting that my results were actually artifactual. This issue was also raised by one of the reviewers of our manuscript and I analyzed this effect very carefully, as explained in Section III.6. The debate on morphogen robustness is still going on and for the past three years, there has been increasing evidence that the Bicoid gradient scales with embryo size (Cheung et al., 2011)[3] and that Bcd high precision is essential for patterning precision (He et al., 2010), in agreement with a simple morphogen gradient model (He et al.,

---

[3]This work however suggests that Bcd scaling is achieved by a volume-dependent production rate.

2008; de Lachapelle and Bergmann, 2010b).

Second, since morphogen gradients usually act on large time scales, it is generally assumed that they have reached steady state (or quasi-steady state) when decoding by the target genes occurs. *Drosophila melanogaster* embryogenesis is however peculiar because of the syncytium blastoderm phase that lasts for about 2h only (see Figure III.1). A study by my PhD thesis supervisor and co-workers suggested that the Bicoid gradient might be decoded prior to steady state, using indirect experimental observations and theoretical analyses (Bergmann et al., 2007). Interestingly, my results on the precision of the target genes supported a model where Bcd is decoded at pre-steady state as well. However, this was also debated by (Jaeger, 2010) and our reviewers (cf. Subsection III.6.1) and at first sight, these results might indeed seem contradictory with the recent observation that Bicoid nuclear concentrations are stable from nuclear cycle 10 on (Gregor et al., 2007b). However, it has been argued that stable nuclear concentrations are fully consistent with an external pre-steady state gradient (Bergmann et al., 2008), and that this combination can actually increase the robustness of the boundaries of the target gene domains (de Lachapelle and Bergmann, 2010b). I believe that the currently available data are fully consistent with the idea that evolution has found this combination of an external pre-steady state gradient and stabilized nuclear concentrations as the best compromise for initiating robust patterning of the anterior-posterior axis of *Drosophila* embryos.

Third, it is generally assumed that morphogen gradients are produced from a localized source. As mentioned in the *Prologue* however, Markus Noll and co-workers have argued that they observe a gradient of *bcd* mRNA in the early embryo (Spirov et al., 2009). Following this work, I re-examined with Dr. Sascha Dalessi part of their images, which was not conclusive, and he investigated theoretically the consequences of such a model on Bicoid degradation and diffusion. The main conclusion was that the diffusion rate of the Bicoid protein could be significantly smaller since *bcd* mRNA also participated in transport, which agreed better with the experiments (Gregor et al., 2007b). However, since we were not fully convinced by how the image data had been processed and in light of a recent paper that minimized the importance of the source delocalization for the protein gradient formation (Little et al., 2011), we did not investigate this question any further. More generally, Sascha also investigated the effects of various more realistic sources on morphogen gradient formation[4].

III.9   Perspectives

Precision and scaling are important to achieve robust pattern formation. My new position dependent measures provide a unified quantification of scaling for both functional descriptions of gradient profiles derived from models (Equation II.1) and experimentally determined expression domains (Equation II.3). These measures will be useful to address scaling also in different systems like the wing imaginal disc. Importantly, these measures clearly identify perfect scaling ($S = 1$) corresponding to the correct preservation of proportions, which is not necessarily equivalent to perfect correlation. As I have shown, gradient formation of a single morphogen can naturally ensure maximal robustness only in a limited patterning domain. For Bcd this appears to be mid-embryo, where most of its targets are expressed. Outside of this domain, other

---

[4] *Modeling morphogen gradient formation from arbitrary realistic structured sources. S Dalessi, **AMdL** and S Bergmann, submitted to Journal of Theoretical Biology.*

systems and mechanisms may have evolved to cooperate in maintaining and potentially increasing robustness. Considering precision and scaling as position- (and possibly time-) dependent features will allow researchers to better characterize and develop models for these systems. In this sense, a recent publication describes a method to simultaneously detect Bcd protein by antibody staining and nascent transcripts of target genes by fluorescence *in situ* hybridization (He et al., 2011). Combining these measurements at various nuclear cycles could give very useful insights into the decoding time of target genes, the sustained role of Bcd and the dynamics of precision and scaling.

Part IV

Drosophila wing imaginal discs development

. . . .

This project was part of the WingX SystemsX.ch consortium, which started in 2008, and aimed towards a better understanding of wing imaginal disc development, undertaking a systems biology approach. Following our work in the embryo, and considering the debate and controversies on the topic of scaling in the wing imaginal disc as well, we decided to establish a collaboration with experimentalists to investigate this question in detail. Since then, I have been working very closely with Dr Fisun Hamaratoglu (lab of Prof. Markus Affolter, Uni. Basel), who performed all the experiments and imaging. I performed the image analysis and all the quantifications. Thomas Schaffter (lab of Prof. Dario Floreano, EPFL) helped me with the image analysis at the beginning. Importantly, I have always updated and discussed the methods and results with Fisun, and we have planned together new experiments several times following the results of the analyses. These fruitful exchanges have given rise to a careful and robust analysis, which I will detail in the following.

IV.1 Overview of *Drosophila* larval stages

The two wing imaginal discs, which are the wing primordia, develop and get patterned during *Drosophila* larval stages that I briefly describe here (see also (Gilbert, 2006; Slack, 2006)).

After hatching, the larva passes through three *instar* phases. The first and second instar stages last together about three days. The third instar phase is the longest stage, lasting three days and characterised by wandering larvae. During these stages, the larva feeds substantially and grows a lot. Importantly, clusters of undifferentiated cells called *imaginal discs*[5] get patterned during this time. Thus, larvae are made up of both larval cells and imaginal cells, which await the signal to differentiate (Gilbert, 2006). Imaginal discs divide rapidly and form the cuticular structures of the adult, including the wings, legs, antennae, eye, head, thorax and genitalia (the specification of the general cell fates occurs in the embryo, see Section III.1) (Gilbert, 2006). The transformation into the adult occurs in the pupa, where most of the larval body is destroyed, while the adult organs develop from 19 undifferentiated (yet *patterned*) imaginal discs (Figure IV.1).

---

[5]i.e. relating to the *imago*, which is the fully developed adult stage of an insect.

Wing imaginal disc development

The wing primordium arises from a small cluster of cells at the border between parasegments 4 and 5 in the embryo (selector genes of the Antennapedia and bithorax complexes). *Drosophila* wing imaginal discs give rise to the adult second thoracic segment, which includes the mesonotum (thorax) and wing blade. The wing imaginal disc consists of one layer of columnar cells covered by another layer of squamous epithelia (Figure IV.2A). The wing disc contains about 30 cells at the beginning of the first instar stage and reaches at metamorphosis about 50000 cells (Milan et al., 1996). During pupation, the wing disc eversion produces the adult wing whose cells neither divide nor further grow (Affolter and Basler, 2007).

In the first instar larva, the anterior-posterior axis of the wing imaginal disc is characterized by the expression of *engrailed* (*en*) in the posterior compartment. Engrailed, a transcription factor, activates the *hedgehog* (*hh*) gene, which acts as a short range morphogen and activates *dpp* at the A/P border. To restrict Dpp expression to a narrow stripe at the A/P boundary, En also represses *cubitus interruptus* (*ci*) and *patched* (*ptc*), which are therefore only present in the A compartment (Aza-Blanc and Kornberg, 1999). Ptc is a membrane protein which negatively regulates Hh signaling and restricts its range. Ci is a complex transcription factor that, in the absence of Hh signaling (i.e. in most of the A compartment) represses *hh* and *decapentaplegic* (*dpp*). However, it also responds to Hh by upregulating *ptc*, thereby helping to restrict the range of the Hh signal, and by inducing *dpp* (Figure IV.2B). In summary, patterning starts with the presence of two compartments characterized by the presence or absence of a selector gene. The compartment in which the selector gene is active directs the expression of a short range signal. The selector gene also makes these same cells refractory to the signal. However, in the adjacent compartment where the selector gene is inactive, cells are sensitive to the signal, within a few rows of cells. These cells respond by becoming a line source of a long range morphogen which diffuses away in all directions (Lawrence and Struhl, 1996).

The dorsal-ventral axis forms during second-instar stage and is characterized by the expression of the *apterous* gene in the dorsal compartment. Apterous induces expression of *fringe*, which results in the activation
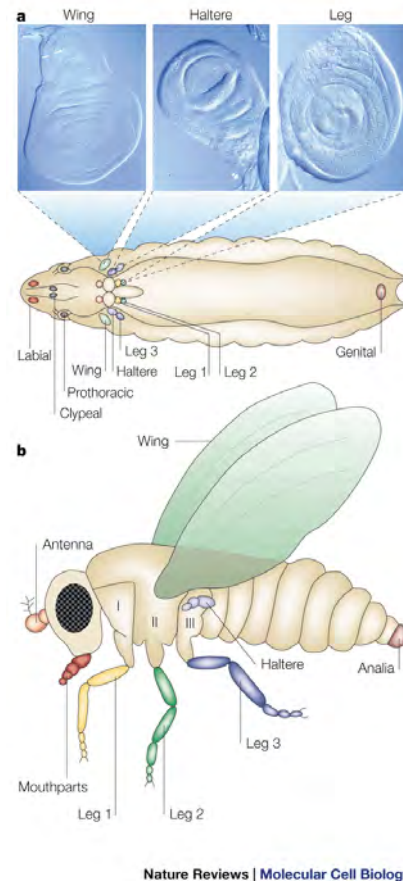
Figure IV.1: Locations and developmental fates of the imaginal discs and tissues in the third instar larva. *(A) A third-instar larva showing the position of all imaginal discs (colour coded according to the appendages that they will develop into). Photographs of wing, haltere and leg discs are shown. In the leg discs the presumptive regions corresponding to trunk and appendage are indicated. (B) The body of the adult Drosophila, indicating the trunk and the different appendages.* Reprinted by permission from Macmillan Publishers Ltd: Nature Reviews Molecular Cell Biology (Morata, 2001), copyright 2001.

Figure IV.2: Wing imaginal disc development. *(A) Vertical cross-section : large peripodial cells cover the surface of the imaginal disc, which consists of tall columnar cells in a tight epithelium. The lumen lies between these two types of cells. Reprinted from (Reeves et al., 2006), copyright 2006, with permission from Elsevier. (B) Cartoon of a Drosophila wing imaginal disc at third larval instar depicting the sites of synthesis of the morphogens discussed in the text. During metamorphosis, the wing pouch will extend out of the plane of the page, with the dorsal (D) an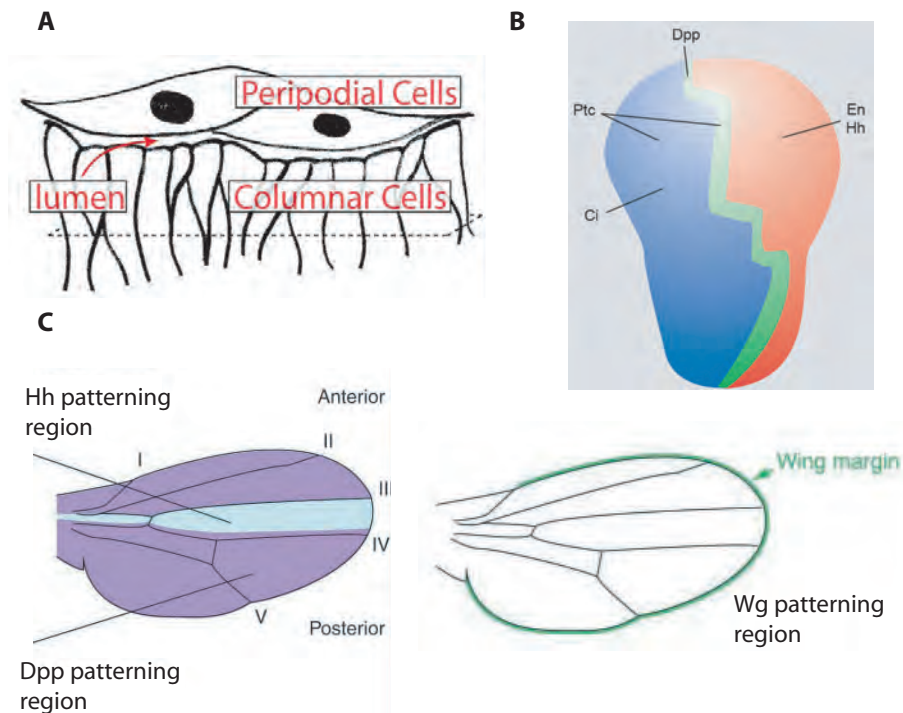d ventral (V) sides adhering together to form the wing blade. Reprinted from (Aza-Blanc and Kornberg, 1999), copyright 1999, with permission from Elsevier. (C) Patterning domains of Hh and Dpp (left), and of Wg (right). Reproduced / adapted with permission from Development (Tabata and Takei, 2004).*

. . . .

of the Notch receptor pathway at the D/V border. Activated Notch induces Wingless (Wg) synthesis at the D/V border, where it functions as a morphogen to organize wing patterning (Tabata and Takei, 2004) (Figure IV.2B).

Thus, in the wing imaginal disc, three proteins act as morphogens: Dpp (a TGF-β family member), Wg and Hh. Hh signals from the posterior compartment to control gene expression in the posterior-most part of the anterior compartment (patterning the central domain of the wing), Dpp specifies cell fates along the anterior-posterior axis (patterning beyond the central domain) and Wg along the dorsal-ventral axis (the D/V border develops into the wing margin) (Figure IV.2C).

IV.2   Dpp signaling

The Dpp signal transduction pathway is highly conserved and relatively simple (Figure IV.3B). Ligand-mediated receptor activation induces phosphorylation of Mothers-against-Dpp (Mad, P-Mad in its phosphorylated and active form) and nuclear translocation of a heteromeric complex of P-Mad and the co-Smad Medea. This complex directly regulates expression of a large number of target genes and has the ability to
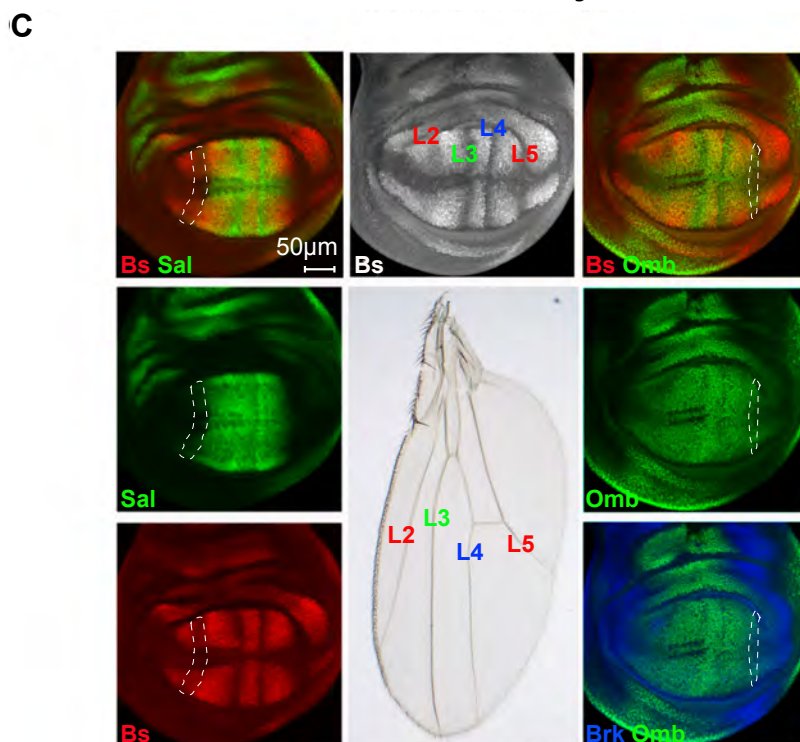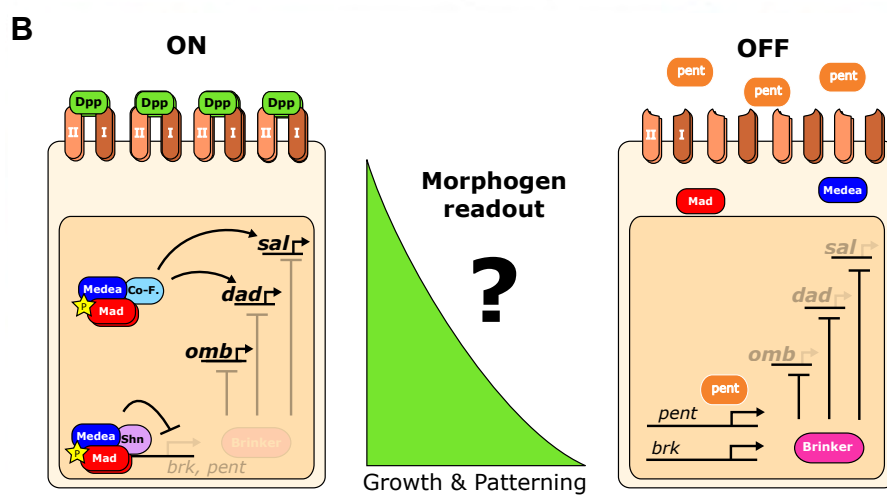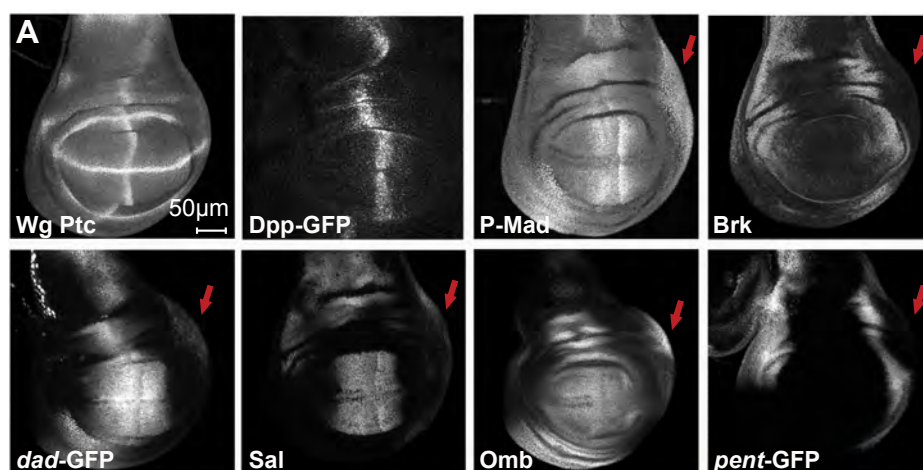
Figure IV.3: Dpp Signal Transduction and Vein Formation. *(A) Representative images of expression patterns referred to in this study. Red arrows point to the effect of posterior Dpp source on various target gene expression patterns. (B) In the medial cells (left), type I (Thickveins-Tkv) and type II (Punt) receptors –both Ser/Thr kinases- from heterodimers upon Dpp binding allowing consitutively active Punt to phosphorylate and activate Tkv, which in turn phosphorylates the Mad proteins (Receptor-Smad). P-Mad molecules form complexes with Medea (co-Smad) and translocate into the nucleus where they can both activate as well as inhibit transcription of target genes with the help of co-factors. P-Mad/Medea/Schnurri complex represses transcription of brk and pent via binding to the so-called silencer elements (SEs), hence brk and pent can only be expressed in the lateral cells. Transcription of sal and dad are positively regulated by the P-Mad/Medea complexes. Brk is the default repressor of Dpp target genes, and its removal results in derepression of omb transcription. In the absence of Dpp signal (right), Brk and Pent are expressed at high levels, Brk keeps sal, dad and omb off. Pent is secreted and helps movement of Dpp laterally via binding to the HSPG Dally. In between these two extremes, cells read both P-Mad and Brk gradients and the sensitivity of enhancers to these two factors as well as others determine their response. Modified from (Affolter and Basler, 2007), figure template courtesy of Alex Weiss. (C) Third instar wing imaginal discs stained for Blistered (Bs) (red) and Sal (green) on the left and for Bs (red), Omb (green), and Brk (blue) on the right. Bs expression is suppressed in the future veins. The middle panel shows the vein positions in a third instar disc and an adult wing. L2, marked with white tracing on the left, is formed within the anterior edge of the Sal/Salr expression domain, overlapping with very low Sal/Salr levels (de Celis and Bray, 2000). The L5 primordium, marked with white tracing on the right, forms within the posterior edge of the Omb domain adjacent to cells expressing high levels of Brk (Cook et al., 2004).*

. . . .

activate as well as suppress transcription (Affolter and Basler, 2007). One of the most critical functions of Dpp signaling is to suppress *brinker* (*brk*) transcription because Brk acts as a transcriptional repressor of many Dpp target genes (Campbell and Tomlinson, 1999; Jazwinska et al., 1999; Minami et al., 1999) (Figure IV.3). Repression of *brk* is achieved via short 'silencer elements' (SEs) in the *brk* enhancer; the *Drosophila* Smad proteins P-Mad and Medea bind as a trimer (two Mad, one Medea) to the SEs and recruit the co-repressor Schnurri (Shn) (Pyrowolakis et al., 2004; Gao et al., 2005; Yao et al., 2006). Consequently, the extracellular Dpp gradient and the resulting intracellular P-Mad gradient are translated into an inverse nuclear gradient of Brk (Muller et al., 2003). In the lateral regions of the wing disc, where Dpp signaling is not active, Brk gradient delimits the expression domains of the Dpp target genes *daughters-against-dpp* (*dad*), *spalt* (*sal*) and *optomotor blind* (*omb*) (Figure IV.3). In patches of marked cells where *brk* function is deleted (i.e. *brk* loss of function clone), *dad*, *sal* and *omb* are derepressed (Affolter and Basler, 2007; Campbell and Tomlinson, 1999; Minami et al., 1999). P-Mad/Medea complex can also directly bind to and activate transcription of *dad* and *sal* (Weiss et al., 2010; Barrio and de Celis, 2004). Hence *dad* and *sal* enhancers read both P-Mad and Brk levels, and their sensitivity to these two factors as well as others determines their expression domains.

While the role of Dad is less well studied, Sal and Omb are crucial for the determination of the positioning of veins L2 and L5 of the adult wing, respectively (Sturtevant et al., 1997; Han et al., 2005; Blair, 2007) (Figure IV.3C). How are the positions of these veins determined? The pouch section of the wing imaginal disc, which will give rise to the adult wing, is subdivided into alternating vein and intervein territories during the larval stages (Figure IV.3C). The combined activity of the Epidermal Growth Factor, Notch, Hedgehog and Dpp pathways culminate in the restricted expression patterns of transcription factors that identify different veins. For example, the zinc-finger proteins Knirps and Abrupt are expressed and required specifically in L2 and L5 respectively. Loss of function mutations of these genes result in the loss of the corresponding veins (de Celis and Bray, 2000; Biehs et al., 1998). Knirps is expressed within the anterior edge of the Sal expression domain, while the L5 primordium forms within the posterior edge of the Omb

domain adjacent to cells expressing high levels of Brk (Figure IV.3C) (de Celis and Bray, 2000; Cook et al., 2004). Hence, Sal, Omb and Brk play instrumental roles in setting the positions of L2 and L5.

Recently, *pentagone* (*pent*) emerged as an important target gene of Dpp signaling, playing essential roles for both growth and patterning functions of the pathway. Pent is a secreted protein that directly interacts with the heparan sulfate proteoglycan Dally to promote long-range distribution of Dpp. Absence of *pent* causes a severe contraction of the Dpp activity gradient resulting in patterning and growth defects of the adult wing. *pent* transcription is repressed by Dpp signaling via SEs and acts as an inbuilt feedback loop with a crucial role in shaping and fine-tuning the Dpp morphogen gradient read-out (Vuilleumier et al., 2010; Vuilleumier et al., 2011).

IV.3   Scaling of Dpp and its targets?

The fruit fly *Drosophila melanogaster,* with its exquisite genetic tool kit, represents an excellent model system for quantitative analysis. The position of the veins in the adult wing scale rather precisely with the final wing size, ensuring functionality (Birdsall et al., 2000; Bollenbach et al., 2008). This observation indicates that there is an active mechanism that coordinates the patterning and growth of the wing. The easiest imaginable way of coordinating growth and patterning is by having the same players control both processes. *Drosophila* Decapentaplegic (Dpp), a TGF-β superfamily member, is essential for the formation of all imaginal discs. Dpp signaling has been extensively studied in the wing imaginal disc. In this tissue, Dpp is produced in a stripe of cells anterior to and abutting the A/P compartment boundary, spreads into both compartments to form a gradient and patterns the growing tissue (Figure IV.3A). Dpp is a morphogen with the capability to specify distinct target gene expression boundaries at different distances from its source; these expression boundaries are instrumental in setting the positions of veins during subsequent development of the wing imaginal disc (Han et al., 2005; Affolter and Basler, 2007; Blair, 2007) (Figure IV.3C). Furthermore, Dpp is required for normal growth of the wing, and able to induce excess growth when ectopically activated (Zecca et al., 1995; Burke and Basler, 1996; Martin-Castellanos and Edgar, 2002). Dpp is therefore an attractive candidate to investigate scaling.

Indeed, several contradicting studies have debated this topic recently (Teleman and Cohen, 2000; Bollenbach et al., 2008; Hufnagel et al., 2007; Wartlick et al., 2011) (discussed in Section IV.10). Here, we made use of the wealth of information available with regard to the molecular read-out of the Dpp signaling activity in the wing imaginal disc and investigated whether the Dpp activity gradients (P-Mad and Brk) and the downstream domain boundaries (Dad, Sal and Omb) scale and thus adapt to the size of the growing tissue. After establishing a protocol to reliably quantify the spatial and temporal changes in the Dpp activity gradients, we found that both P-Mad and Brk scale rather well with the tissue size.

Next, we tried to uncover the molecular mechanisms that ensure proper scaling of these activity gradients. A recent mathematical model termed expansion-repression integral feedback control suggested that scaling emerges as a natural consequence of a feedback loop which is based on two diffusible components: a morphogen and a hypothetical molecule termed *expander* (Ben-Zvi and Barkai, 2010). This molecule facilitates the spread of the morphogen and in turn is repressed by it. As a consequence, the gradient expands as long as the expander molecule can be produced. The gradient stops expanding once the morphogen

levels are high enough to inhibit expander production in the whole field. Because the expander molecule is assumed to be very stable and diffusible, the morphogen gradient remains expanded, even when no more expander is produced. In the context of a slowly growing tissue, more expander could be produced in the lateral regions as the tissue grows. The morphogen gradient would therefore expand again, until it can inhibit expander production once again everywhere. Since Dpp signaling controls the expression of Pent, which itself regulates the Dpp activity gradient, we tested whether Pent acts as the expander of the Dpp gradient during disc growth.

Dpp activity gradients are read out by several target genes, such as *dad*, *sal* and *omb*, domains of which we found scale with tissue size. How is scaling transmitted from the activity gradients to the target gene domains? Inspired by the French flag model for pattern formation (Wolpert, 1969), we tested whether the target genes *dad*, *sal* and *omb* respond to similar concentration thresholds of P-Mad and Brk activity during growth. In this case, provided that the activity gradients scale, the boundaries characterized by these constant thresholds would shift as the gradients expand, ensuring perfect scaling of the target gene domains with tissue size (cf. Figure 1 in the *Prologue*). Interestingly, our results do not support such a model, but rather suggest that P-Mad and Brk activity gradients are combined in a gene-specific fashion to ensure proper scaling of the targets. Finally, we took advantage of our datasets to test the validity of a recently proposed uniform growth model in the wing imaginal disc (Wartlick et al., 2011).

IV.4   Generating a dataset for quantitative analysis

Before we could ask questions regarding the scaling behavior of Dpp signaling read-out during wing imaginal discs growth, it was necessary to establish methods to acquire images in a quantifiable manner. We concentrated our analysis on the pouch of the wing imaginal disc that gives rise to the future wing. To extract the pouch and determine the A/P and D/V compartment boundaries, we co-stained all discs with Wingless (Wg) and Patched (Ptc) antibodies. Ptc is induced at very early stages and is restricted to the anterior side of the A/P boundary and hence marks the A/P boundary at all the examined stages. Wg expression gets refined later; by 65-70h into development, it outlines the pouch and the D/V boundary (Figures IV.4A). Since we wanted to measure parameters exclusively in the pouch, the discs from the 70h old larvae were the youngest we included in our analysis and we spanned the third instar larval stage with 10h intervals. In this manner, we defined five time classes (TC) and color-coded them as following: TC1 in purple: 65−75h hours after egg laying (AEL), TC2 in green: 75−85h AEL, TC3 in orange: 85−95h AEL, TC4 in blue: 95−105h AEL, and TC5 in red: 105−120h AEL (Figure 2A). (Figure IV.6A). Note that timing here refers to hours after egg laying (AEL) and not to hours after hatching; a larvae of 70h AEL would be 46h old after hatching. To minimize variation within a single time class (TC), we collected eggs for 1h only, and dissected only male larvae. To minimize errors due to sample handling, we processed all samples in parallel in identical solutions and mounted discs from all TCs on the same slide and imaged them under identical settings. Further details on sample collection and processing are provided below.

To account for the amplitudes of the protein gradients in our analysis and to test for the hypothesis that downstream target gene domains are defined at constant thresholds of P-Mad and Brk, we treated fluorescence intensities as a measure of protein concentrations. To ascertain that changes in fluorescent intensities
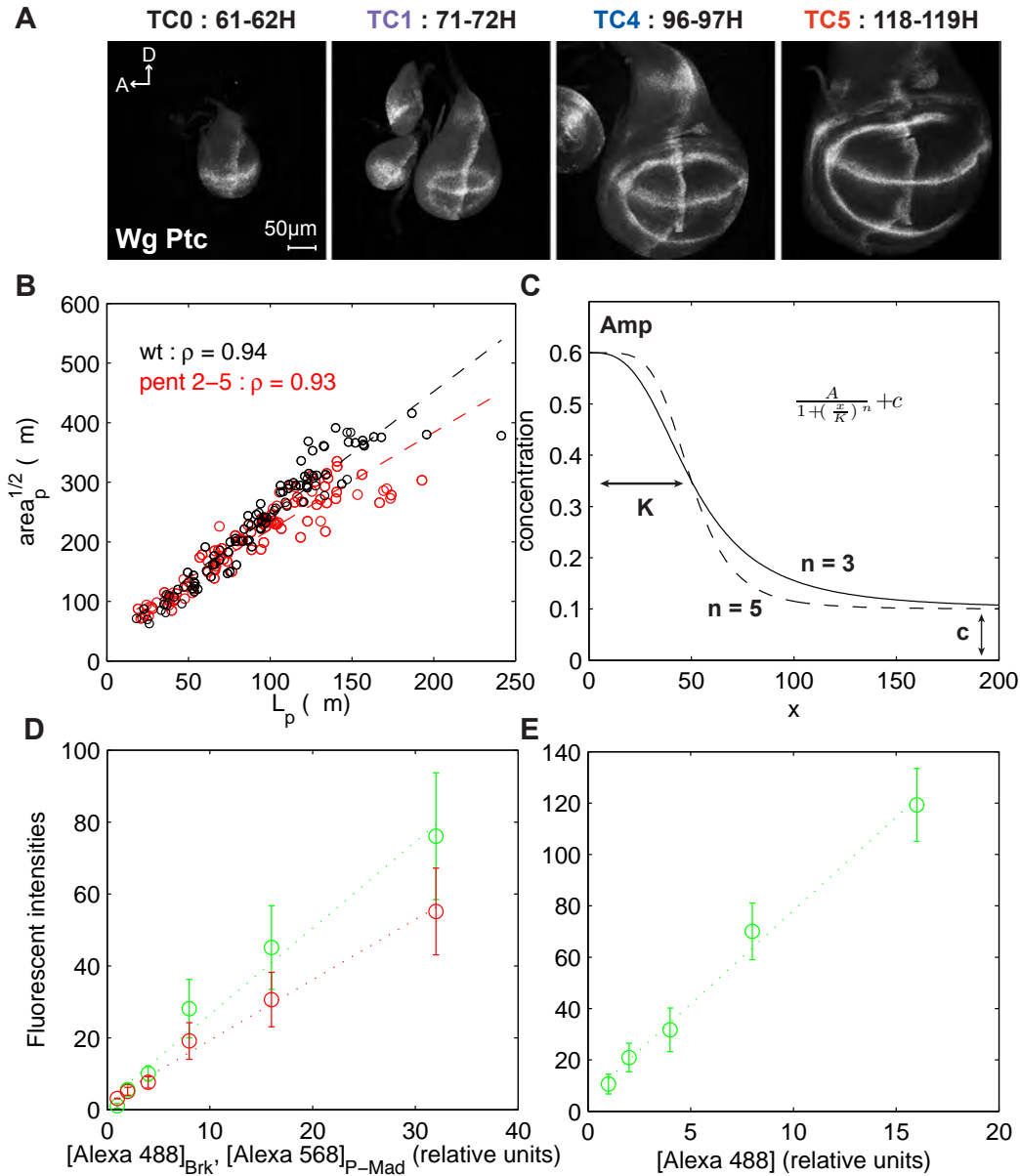
Figure IV.4: Methods. *(A) Discs of varying ages stained with Wg and Ptc antibodies. Wg staining gets refined by 71-72h AEL. (B) The posterior compartment length measured along the D/V axis (L_p) correlates well with the square root of the posterior compartment area ($area_p$) both in wt (black) and pent 2-5 mutant discs (red). Each dot represents a disc. (C) The Hill function used to fit the gene expression domains returns four parameters : the amplitude Amp, the transition point of the domain K, the sharpness of the domain boundary n and a constant offset c. (D-E) Linear range imaging for P-Mad/Brk dataset 1 (D) and Omb/Brk dataset 2 (E). Several dilutions of the secondary antibodies Alexa 488 (green) and Alexa 568 (red) yield fluorescent intensities that are proportional to their concentrations under our imaging conditions. Mean intensities in the whole field and the standard deviations were obtained using the Histogram function in Image J. We measured background by imaging an empty slide and subtracted this value. Linear regressions are indicated with dotted lines.*

reflected changes in protein concentrations in a linear manner, we imaged fluorescent dyes of known concentrations at the same settings we used for our images and determined the linear range for our imaging conditions. We found that the intensities obtained in our experimental recordings indeed fell into the linear range of our imaging conditions (Figure IV.4D-E).

To measure scaling, we modified and applied an approach we devised for evaluating the scaling properties of the Bicoid gradient in the embryo (de Lachapelle and Bergmann, 2010a) (cf. Part II). We defined scaling as the conservation of proportions across various sizes where perfect scaling requires changes in domain size to be proportional to those in tissue size. While we measured the scaling of P-Mad and Brk activity gradients at various positions (cf. Figure II.4), the scaling of downstream gene expression domains (Dad, Sal, Omb) were measured only at a single position -their domain boundary- by fitting a Hill function to the profile, which returned the position of the domain boundary; note that the amplitudes were not informative for determining this boundary position (cf. Figure II.2 for assessing scaling and Figure IV.4C for the Hill fit).

IV.4.1 Sample collection

Flies were constantly kept in a 26°C incubator and the eggs were collected on grape juice plates. It is known that the females can keep the fertilized eggs for up to 8h, so a freshly laid egg can be anywhere between minutes to 8h old. We circumvented this problem by treating flies with CO_2 prior to collection, which is thought to relax the muscles and facilitate the deposition of old eggs. This first collection was discarded and the flies were transferred to a clean collection chamber. Additionally, as sexual dimorphism exerts itself early on, only male larvae were included in our analysis where possible. Male larvae were positively selected for by the presence of a clear, oval genital disc which is clearly visible starting from 80h AEL. Therefore, our 70h collections had both male and female larvae. We observed that 70h AEL corresponds to the beginning of the third instar stage at 26°C as hatching larvae were frequently encountered. Dissected larvae were fixed immediately, washed and stored at 4°C. Once all time classes were obtained (usually within 2 days), all samples were processed for antibody staining in parallel using identical solutions.

IV.4.2 Image acquisition

Larvae of different time classes (TC1: 65−75h AEL, TC2: 75−85h AEL, TC3: 85−95h AEL, TC4: 95−105h AEL, TC5: 105−120h AEL) were transferred into cold fixative (4% pfa in PBS, pH=7) and fixed for 25 min at room temperature on a rotator. Following extensive washes in PBT (PBS+ 0.03% TritonX), the discs were blocked in PBTN (PBT + 2% Normal Donkey Serum, Jackson Immuno Research Laboratories) for 1h at 4°C on a rotator, and incubated with primary antibodies overnight at 4°C. The discs were washed several times with cold PBT and incubated in secondary antibodies for 2h at room temperature on a rotator. After another round of washes with PBT, the excess fluid was removed and replaced with Vectashield mounting media (Vector Labs). All discs from a dataset (i.e. all 5 TCs) were mounted on the same slide to reduce potential variation in thickness between the slide and the coverslip across different samples. Brain discs were used as spacers. All discs from a dataset were imaged under identical microscopy settings using a Leica SP5 confocal microscope (1μm thick sections).
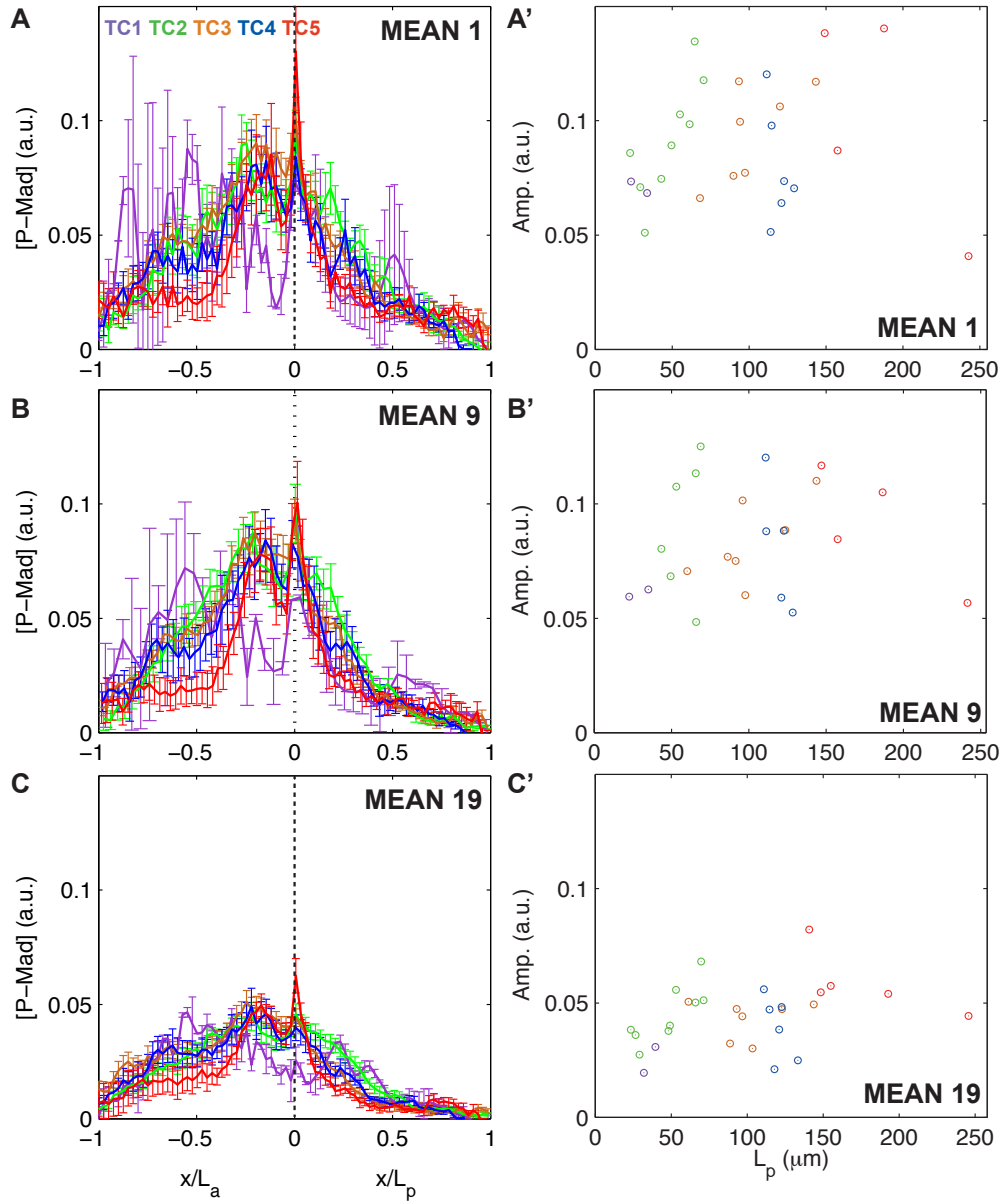
Figure IV.5: Mean projection. *(A-A')* Average P-Mad profiles per TC obtained from the brightest section (A) and their corresponding amplitudes (i.e. the concentration at $x = 0$ of the individual profiles, A'). Positions in the posterior compartment are normalized relative to the posterior compartment length L_p, while positions in the anterior compartment are normalized relative to the anterior compartment length L_a. *(B-B')* Same as (A-A'), but the profiles are obtained from 9 sections from which we performed a mean projection. *(C-C')* Same as (B-B'), but with 19 sections.

IV.4.3 Image processing

Generally, the confocal images contained about 25 to 40 sections depending on the age of the disc. After the image acquisition, we manually selected four consecutive slices above and below the brightest slice from each stack and performed a mean projection of these 9 slices. Using a reduced number of slices and performing the mean projection allowed to reduce the noise as well as avoid the signal from the peripodial membrane. Still, we made sure that we still had enough signal from the columnar cells of the pouch and that we were not losing signal due to "empty" sections. We show in Figure IV.5 the P-Mad profiles and corresponding amplitudes (i.e. the concentration at $x = 0$ of the individual profiles) obtained by mean projection of 1, 9 and 19 sections. We see that the levels with 1 or 9 sections are similar, while they are significantly lower with 19 sections. We also performed a maximum projection (after removing some sections from the peripodial membrane), which yielded similar, but somewhat more noisy profiles, in particular far from the A/P boundary, where some remaining signal from the peripodial membrane affected the profiles. Thus, we finally settled on the mean projection of nine sections.

We then manually contoured the inner pouch boundary as well as the anterior-posterior and dorsal-ventral boundaries, as marked by the Wg and Ptc stainings, in ImageJ. The remaining analyses were automated in MATLAB® and applied solely to the pouch. All the pouchs were rotated to have anterior to the left and dorsal upwards orientation. We extracted the profiles along the D/V boundary, since it is a natural coordinate in the wing pouch, or parallel to it with a small offset of 5% of the height of the pouch into the dorsal compartment to avoid potential interference with Wg, which is expressed at the D/V boundary (cf. Figure IV.7A). We repeated our analysis also with 15% and 25% offsets into the ventral compartment. Note that since the D/V boundary is not a thin line but a stripe, we applied mean filtering with a rectangular sliding window of fixed size [20 x 3] pixels (height x width) to smoothen the images of [1024 x 1024] pixels before further analyses. Also, because we used nuclear markers, the 1d extracted profiles looked very rugged and we therefore applied Gaussian filtering before quantifying scaling.

IV.5 P-Mad gradient expands as the disc grows

The first event downstream of receptor activation is the phosphorylation of the signal transducer Mad, which we visualized and quantified with P-Mad antibodies. We analyzed P-Mad gradients in wing imaginal discs from different TCs (Figure IV.6A). We extracted the P-Mad profiles exactly along the D/V boundary and with different offsets into the dorsal and ventral compartments within the wing pouch (Figure IV.7A). This approach yielded a global view of the dynamics of the P-Mad gradients during development (Figure IV.6B, see Figures IV.7-IV.8 for P-Mad profiles in relative and absolute distances extracted at several offsets). We observed that the P-Mad levels were significantly suppressed at the D/V boundary in the last TC (red). Along the D/V axis, the average amplitudes were 25-30% lower than in the other TCs and the profiles became steeper at this last stage (Figures IV.7B). However, this effect diminished gradually away from the D/V boundary, suggesting that it was caused by a factor centered at the D/V boundary (Figure IV.7). Therefore, in order to minimize the impact of this effect and of the influence of a secondary Dpp source expressed in the dorsal posterior compartment (red arrows in Figure IV.3A) (Foronda et al., 2009), we per-
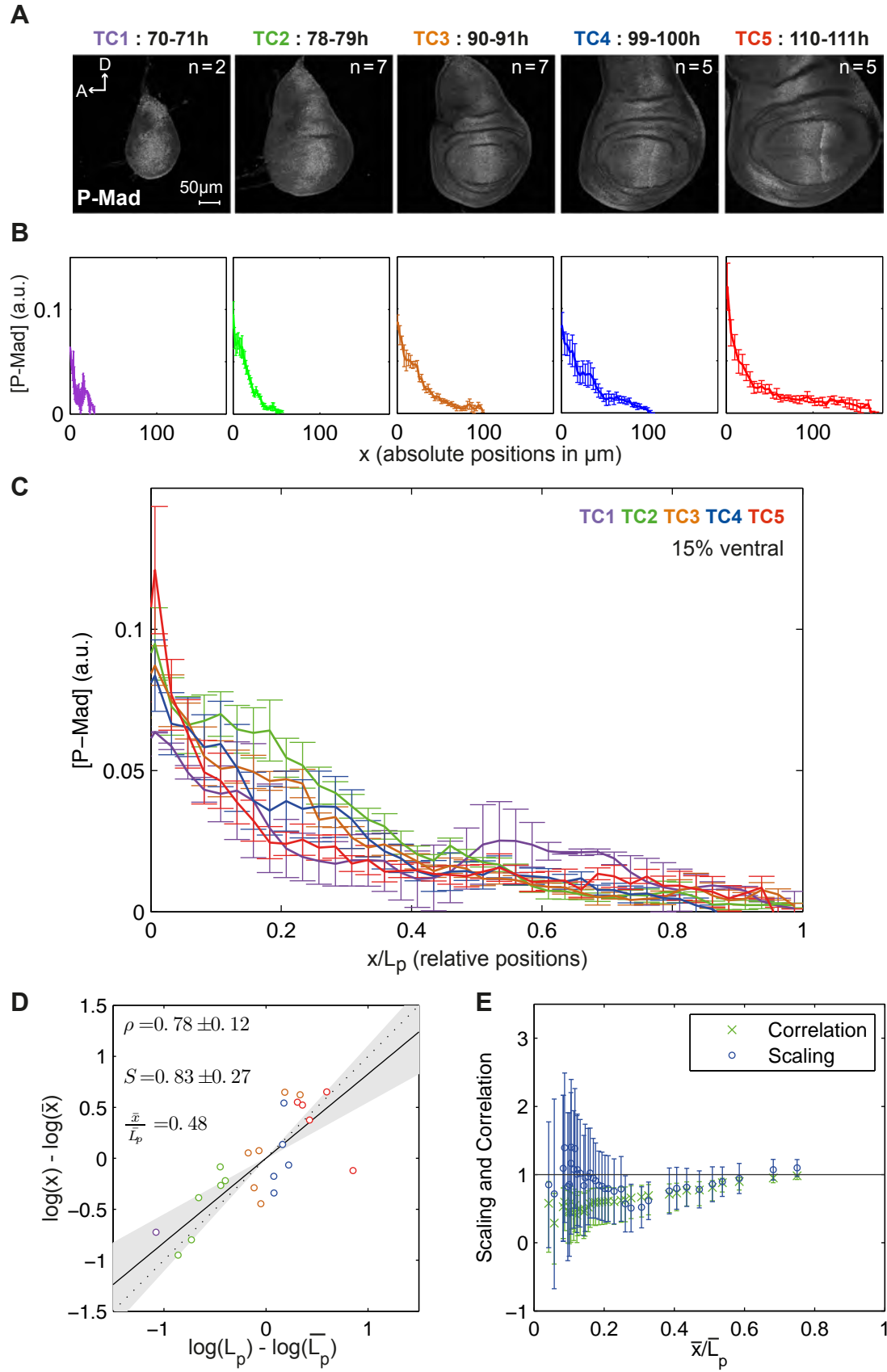
Figure IV.6: P-Mad gradients expand as the tissue grows. *(A) Representative wing imaginal discs from each time class (TC) stained for P-Mad antibodies. (B) Average P-Mad profiles at 15% ventral offset. Profiles are shown separately for each TC and in absolute positions. Error bars show the standard error here and in the similar panels of the following figures. (C) Profiles in (B) shown in relative positions. (D) P-Mad scaling corresponding to the average relative position $\bar{x}/\bar{L} = 0.48$. For each disc, log-deviations in the position, $log(x/\bar{x}) = log(x) - log(\bar{x})$, are plotted against log-deviations in the posterior compartment length, $log(L_p/\bar{L}_p) = log(L_p) - log(\bar{L}_p)$, for a given P-Mad threshold concentration. The scaling coefficient S is obtained by linear regression (95% confidence interval on the slope, i.e. on the scaling coefficient, is shown in gray). The dotted diagonal line represents perfect scaling (S = 1) here and in subsequent figures. Correlation ρ with its 95% confidence interval is shown. (E) P-Mad scaling (o) and correlation (x) for several threshold concentrations. Error bars represent the 95% confidence intervals.*
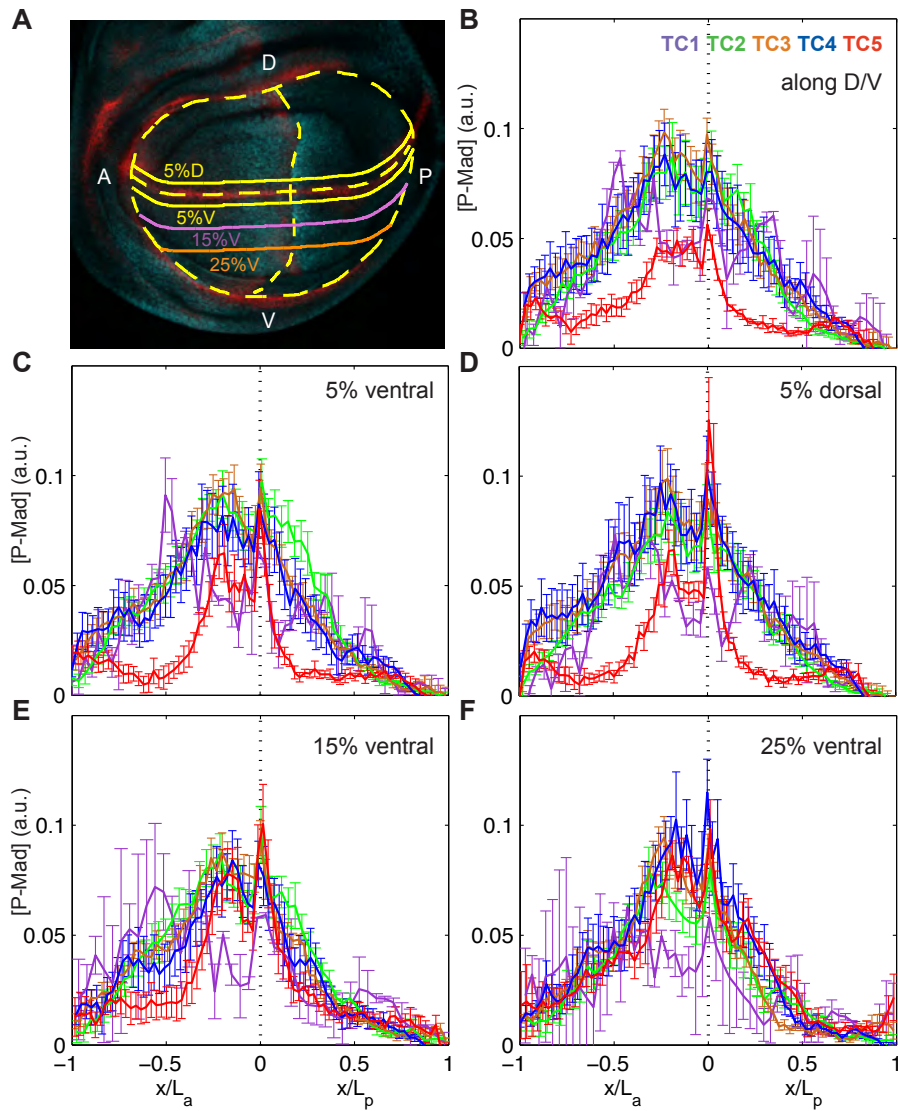
Figure IV.7: P-Mad is repressed along the D/V boundary at the end of third instar. *(A) The dashed yellow lines outline the pouch as well as the A/P and the D/V compartment boundaries, as defined by Wg and Ptc stainings in red. P-Mad (blue) profiles were extracted along the D/V and with 5% (yellow), 15% (purple), 25% (orange) offsets from it. (B-F) P-Mad profiles averaged per TC in relative positions along the D/V (C), with 5% offset into the dorsal compartment (E), and with 5% (D) 15% (F) and 25% (G) offsets into the ventral compartment. Positions in the posterior compartment are normalized relative to the posterior compartment length L_p, while positions in the anterior compartment are normalized relative to the anterior compartment length L_a.*
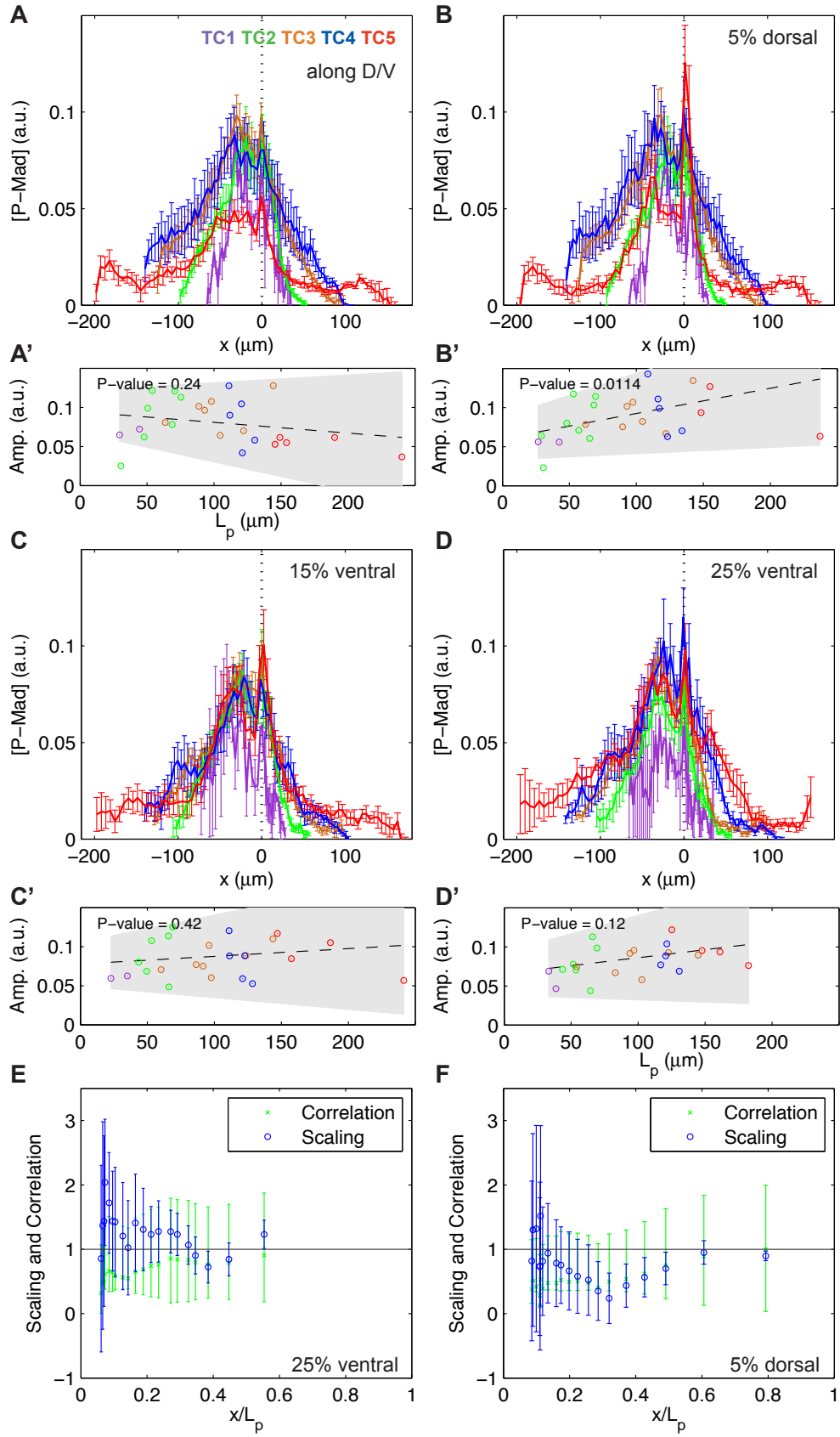
Figure IV.8: P-Mad profiles and amplitudes at various positions. *(A-D) P-Mad profiles averaged per TC along the D/V (A), and with 5% dorsal (B), 15% ventral (C), 25% ventral (D) offsets. (A'-D') The amplitude of the P-Mad profile (i.e. the concentration at A/P compartment boundary, $x = 0$) plotted versus the posterior compartment length. Each dot represents a disc and is color-coded according to its age. The linear regression with 95% confidence interval (gray area) and its t-test P-value under the null hypothesis that the slope is equal to zero are shown. (E-F) P-Mad scaling (o) and correlation (x) for several threshold concentrations using the P-Mad profiles that were extracted with 25% ventral offset (E) and 5% dorsal offset (F). Error bars represent the 95% confidence intervals, obtained from the linear regressions in the case of scaling.*

. . . .

formed our scaling analysis for protein profiles in the ventral compartment with 15% offsets (light purple demarcation in Figure IV.7A).

For further analyses, we concentrated on the posterior half as to exclude Dpp secreting cells from the analysis as well as the modifications of Dpp receptor levels via Hh, which is active only in the anterior compartment (Tanimoto et al., 2000; Funakoshi et al., 2001). Qualitatively, from TC2 to TC4 the P-Mad gradient expanded and adjusted to the disc size, displaying this trend regardless of where it was measured (Figures IV.6B-C and IV.7). Quantitatively, P-Mad showed close to perfect scaling with $S = 0.83 \pm 0.27$ at a threshold concentration corresponding approximately to the mid-posterior compartment ($\bar{x}/\bar{L}_p = 0.48$) (Figure IV.6D, the individual discs are represented with color-coded circles according to their age).

In order to obtain a position dependent picture of scaling, we considered several other protein concentration thresholds and calculated a scaling coefficient at each threshold (Figure IV.6E). Scaling coefficients (blue circles) and correlation coefficients (green crosses) -informative for the goodness of fit- were plotted as a function of the relative positions, with $\bar{x}/\bar{L}_p = 0$ corresponding to the intercept with the A/P axis and $\bar{x}/\bar{L}_p = 1$ to the end of the pouch for each disc. We found that the P-Mad gradient showed overall very good scaling for $x > 0.4L_p$ (Figures IV.6E and IV.8E-F; closer to the A/P boundary, the error bars were too large for a meaningful conclusion). Accordingly, the scaling of a target gene domain that strictly depends on P-Mad levels should follow the same trend. We therefore analyzed known P-Mad targets with this question in mind.

IV.6 Brk adapts to disc size with increasing amplitudes

brk is a direct target of Dpp signaling. Its transcription is repressed in cells with high levels of Dpp activity, such as the medial cells, and is expressed at varying levels in response to the Dpp activity gradient in the lateral parts of the wing disc (Figure IV.3) (Campbell and Tomlinson, 1999; Jazwinska et al., 1999; Minami et al., 1999). We generated two independent datasets for Brk that we present in two separate figures (Figures IV.9 and IV.10). Importantly, they yielded very similar results demonstrating the reproducibility of our protocol. We examined whether the P-Mad dynamics were mirrored in Brk protein levels. We found that, as the disc grows, the region of maximum Brk expression was found at further absolute distance from the A/P boundary, consistent with the fact that P-Mad gradients now reach further out and suppress *brk* transcription (Figure IV.9A-B). Cells expressing highest levels of Brk were found around 80% L_p throughout development, suggesting that behind this relative position, P-Mad levels were too low to suppress *brk* transcription (Figures IV.9C, IV.10D).
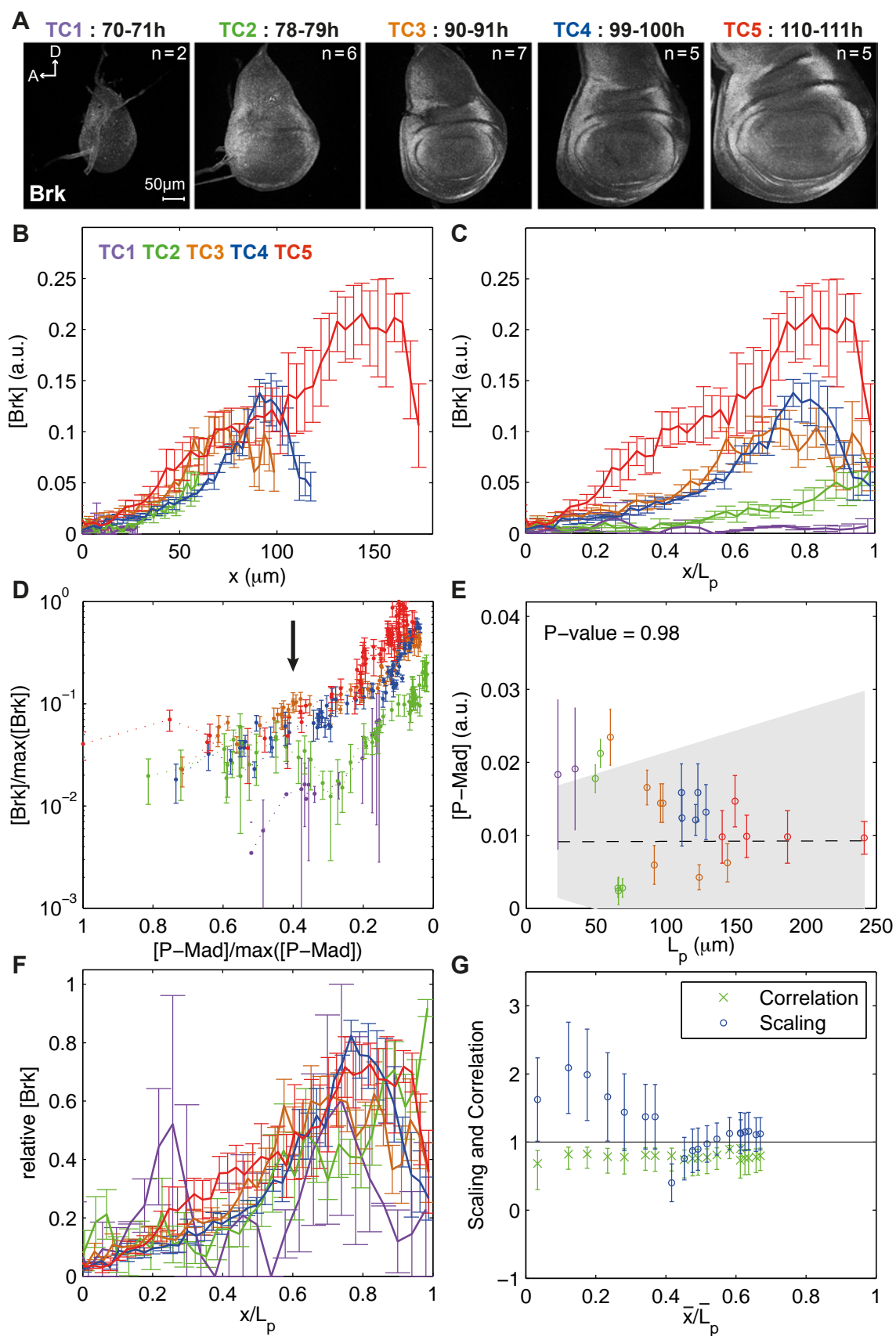
Figure IV.9: The range of Brk scales with tissue size. *(A) Representative images of wing discs for each TC stained with Brk antibodies. (B) Average Brk profiles per TC, measured at 15% ventral offset. (C) Profiles in (B) normalized for L_p. Note that the Brk concentration at a relative position increases with time. (D) Brk versus P-Mad concentrations for every position in the posterior compartment at 15% ventral offset (P-Mad and Brk from dataset 1 were co-stained). P-Mad and Brk concentrations were normalized to their maximum over all TCs: $max([P-Mad]) = 0.12$, $max([Brk]) = 0.22$. Arrow points to 40% of $max([P-Mad])$. (E) Brk profiles were fitted with a decaying exponential function, yielding a decay length λ_{Brk}. For each disc, the average P-Mad concentration at the position $x = \lambda_{Brk}$ was calculated. The weighted linear regression with 95% confidence interval (gray area) and its t-test P-value under the null hypothesis that the slope is equal to zero are shown. (F) Average Brk profiles in relative positions, after normalizing the intensities of the individual profiles. (G) Brk scaling (o) and correlation(x) for several threshold concentrations. Error bars represent the 95% confidence intervals. Note that scaling was computed using the individual Brk profiles with normalized amplitudes.*

. . . .

As a result of P-Mad gradients scaling and keeping their amplitudes roughly constant, cells at the same relative position had very similar P-Mad levels across development. Interestingly, the same cells were subject to increasing levels of Brk: while the magnitude of the increase tended to be smaller away from the D/V boundary, we detected 10 to 20 fold increase in the average amplitudes in the 40h interval we studied (Figure IV.10A/B). Our collaborators observed a similar trend in the expression levels of Brk, when they used a *brk*-GFP reporter line in which GFP expression was driven by the wing enhancer of *brk* (Muller et al., 2003) (not shown). Hence, the changes in Brk protein levels are unlikely to be due to post-transcriptional events.

How are the constant P-Mad levels translated into increasing Brk levels? Since discs were co-stained for Brk and P-Mad, we investigated how Brk concentrations related to P-Mad concentrations within each disc. Figure IV.9D shows the average measured response of P-Mad and Brk within each TC (after normalizing their maximum concentrations over all TCs to one). At P-mad concentrations above 40% of maximum levels, Brk responded similarly to P-Mad at all TCs. However, below this threshold (to the right of the arrow), it appears that Brk could accumulate (Figure IV.9D). Finally, when we fitted decaying exponentials to the Brk expression profiles, the resulting decay lengths corresponded roughly to mid-posterior compartment throughout development, a position where P-Mad scaled very precisely (Figure IV.9E). Hence, the decay length of the Brk expression profile corresponded to very similar P-Mad levels across TCs (Figures IV.9E and IV.10G). Therefore, the expression pattern of Brk depended on P-Mad while the increase in protein levels could not be explained with the P-Mad dynamics alone.

Finally, we studied the scaling properties of the Brk profiles. Apart from being non-quantitative, looking at the collapse of the profiles adjusted to compartment size was a good indicator of the level of scaling (cf. Figure II.3). Brk profiles showed good scaling only between TC3 and TC4 (Figure IV.9C, compare orange and blue profiles). However, after we normalized the individual Brk gradients to their maximum, all the profiles collapsed rather well (Figures IV.9F and IV.10E), and we measured nearly perfect scaling for $x > 0.4L_p$ (while they seemed to hyper-scale more medially), in agreement with the measured scaling for P-Mad (Figures IV.9G and IV.10F, compare to Figure IV.6E). Overall, we concluded that the range but not the levels of the Brk gradients scaled with the tissue size.
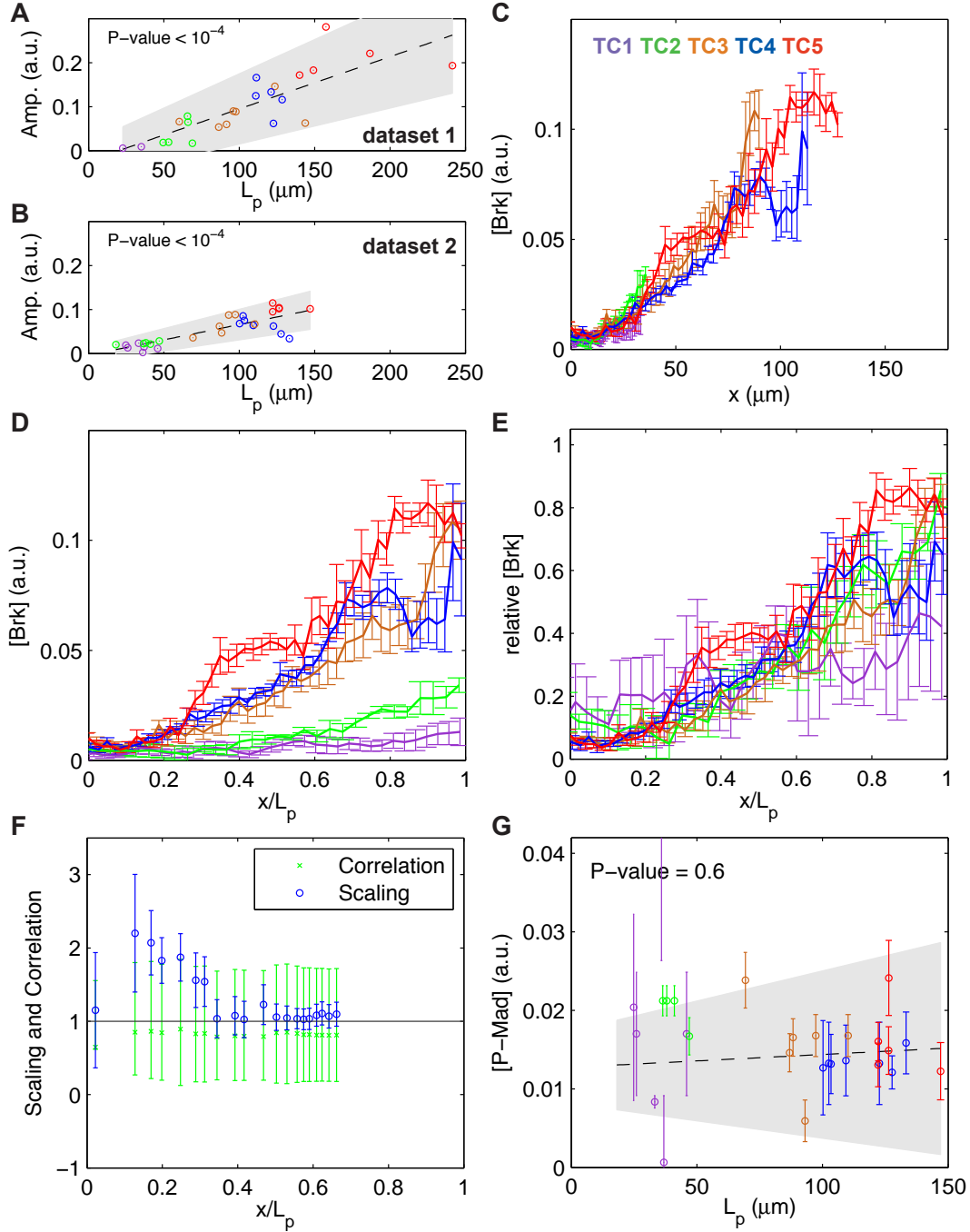
Figure IV.10: A second dataset for Brk. *(A-B) The amplitudes of the Brk profiles (i.e. the peak concentration in the lateral region) at 15% ventral offset versus the posterior compartment length for dataset 1 (A) and dataset 2 (B). Taking the extremes, the ratio between the extreme values of the Brk amplitudes (Amp) are $Amp^{max}/Amp^{min} = 48.7$ for dataset 1 and $Amp^{max}/Amp^{min} = 48.3$ for dataset 2, respectively. (C) Brk profiles averaged per TC with 15% ventral offset (dataset 2). (D) Profiles in (C) in relative positions. (E) Profiles in (D) with normalized amplitudes. (F) Brk scaling for several threshold concentrations. Error bars represent the 95% confidence intervals. To compute scaling at each position, the Brk profiles with normalized amplitudes were used (dataset 2). (G) Brk profiles were fitted with a decaying exponential function (dataset 2) to obtain the decay length λ_{Brk} of the profile. For each disc, the average P-Mad concentration at the position $x = \lambda_{Brk}$, were plotted against the L_p of the disc. The weighted linear regression with 95% confidence interval (gray area) and its t-test P-value under the null hypothesis that the slope is equal to zero are shown.*
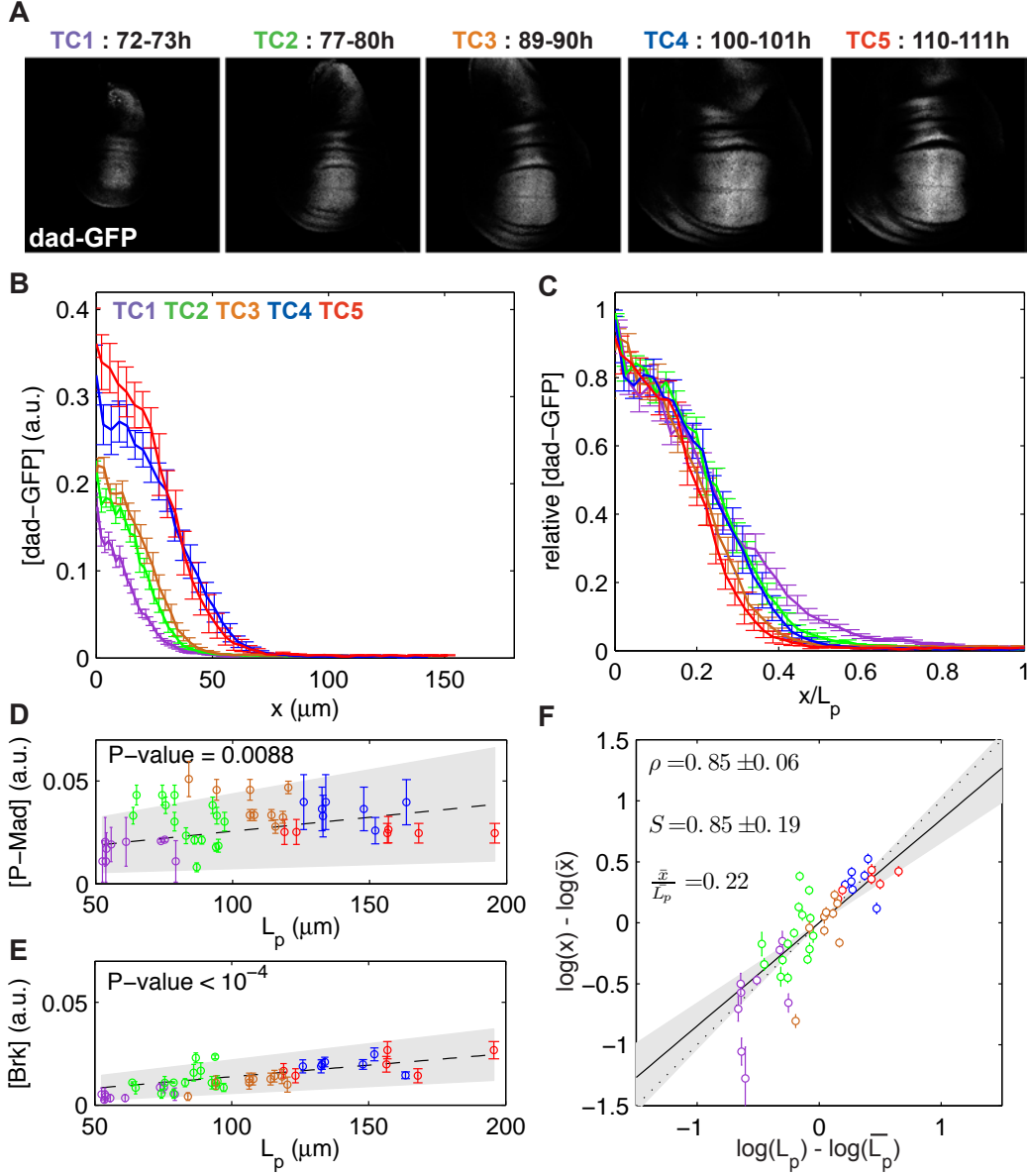
Figure IV.11: *dad-GFP displays position-dependent scaling. (A) Representative images of wing discs from larvae carrying the dad-GFP transgenes. (B) Average dad-GFP profiles per TC at 15% ventral offset. (C) Profiles in (B), normalized for L_p and for their amplitudes. (D) dad-GFP profiles were fitted with a Hill function describing the transition point $K_{dad-GFP}$ of the domain (see Figure IV.4C). We plot the average P-Mad concentration at the position $x = K_{dad-GFP}$ against L_p for each disc and show the weighted linear regression with 95% confidence interval (gray area) and its t-test P-value under the null hypothesis that the slope is equal to zero. (E) Average Brk concentration at the position $x = K_{dad-GFP}$ plotted against L_p for each disc. (F) For each disc, the log-deviations in dad-GFP domain boundary position (as defined by the Hill fit; error bars represent the 95% confidence interval of this fit) were plotted as a function of the log-deviations in L_p. The scaling coefficient S is obtained by weighted linear regression (95% confidence interval in gray). We show the correlation ρ of the data with 95% confidence interval.*
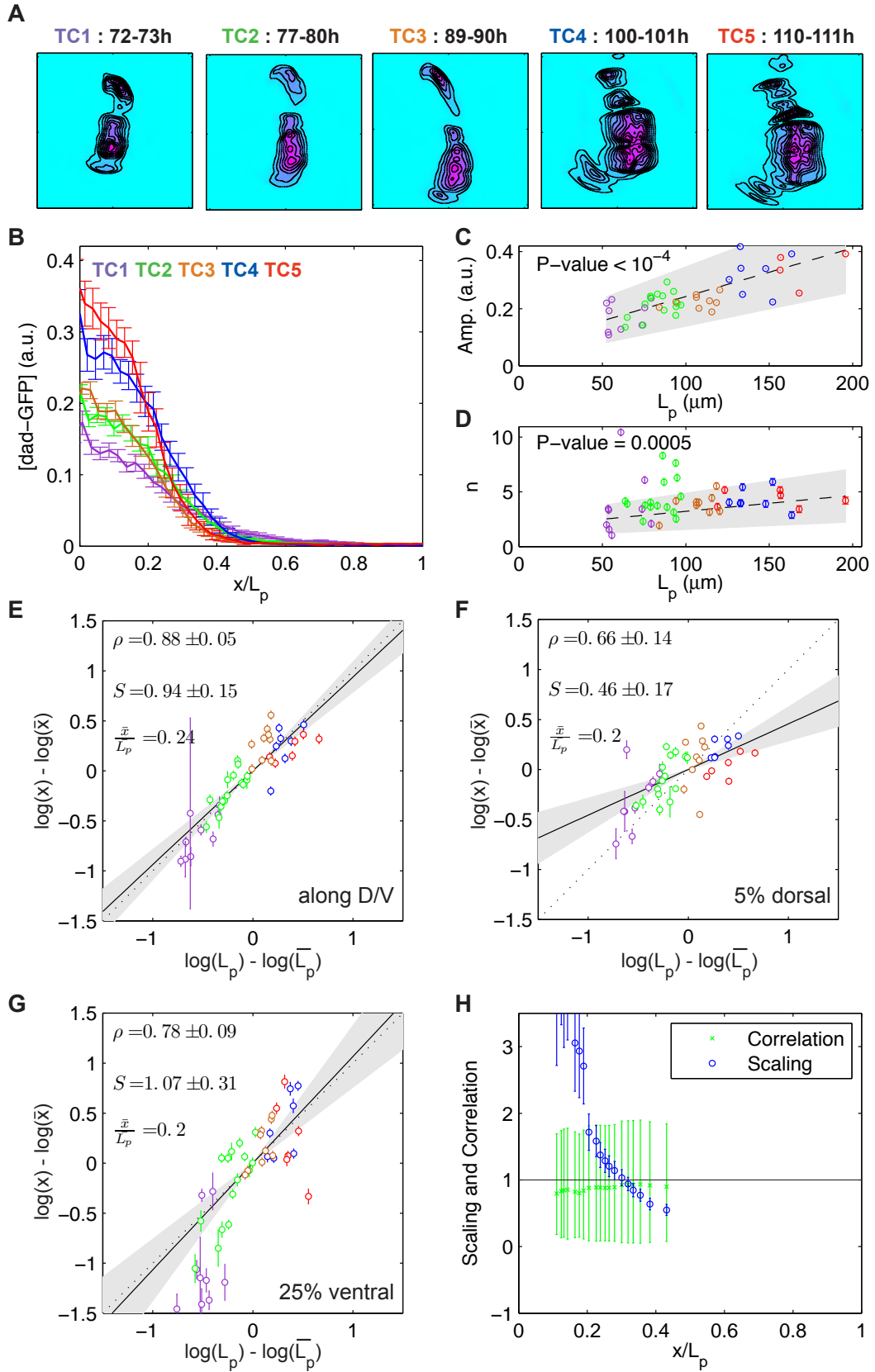
Figure IV.12: *dad-GFP shows position-dependent scaling. **(A)** For each time class, a representative dad-GFP contour plot is shown. Lower concentrations in light blue, higher concentrations in pink. **(B)** dad-GFP profiles averaged per TC in relative positions, at 15% ventral offset. **(C)** The amplitudes of the dad-GFP profiles (i.e. the concentration at A/P compartment boundary, $x = 0$) are plotted against the L_p for each disc, at 15% ventral offset. **(D)** dad-GFP domain boundary at 15% ventral offset gets slightly sharper across growth. Error bars represent 95% confidence intervals from the Hill fits (see also Figure IV.4C). The weighted linear regression with 95% confidence interval (gray area) and its t-test P-value under the null hypothesis that the slope is equal to zero are shown. **(E-G)** Scaling and correlation of dad-GFP domain boundary along the D/V (E), and with 5% dorsal (F), 25% ventral (G) offsets. **(H)** Scaling of dad-GFP for several threshold concentrations when treated as a gradient. Error bars represent the 95% confidence intervals, obtained from the linear regressions in the case of scaling. Consistent with the profiles in (B), positions anterior to $0.4L_p$ show hyper-scaling due to the increasing dad-GFP levels, while positions posterior to it show hypo-scaling.*

. . . .

IV.7   Dad expression incorporates P-Mad and Brk activities

Since there were no antibodies available for Dad, we visualized changes in *dad* expression over time with a *dad-GFP* transgene where the GFP expression was controlled by a 2kb *dad* enhancer fragment (Figure IV.3A). Note that this enhancer fragment incorporates positive input from P-Mad as well as negative input from Brk (Weiss et al., 2010). Hence, *dad-GFP* represented a good tool to monitor combined activity of P-Mad and Brk. While *dad-GFP* formed a gradient reminiscent of P-Mad, it did not tail off as far, presumably because it was sharpened by Brk (Figure IV.11B). As a result, while *dad-GFP* could be treated like a gradient (Figure IV.12H), a Hill function also yielded a good fit. The boundary of the *dad-GFP* domain was obtained by fitting a Hill function to the individual profiles and corresponded to $x = K_{dad-GFP}$ (cf. Figure IV.4C).

We investigated scaling of the *dad* expression domain along the D/V boundary and with several offsets. Because the *dad*-GFP domain boundary was not straight but contracted at the D/V boundary, especially during the last TCs (Figures IV.11A and IV.12A), scaling was quite different depending on where it was measured: the *dad*-GFP domain boundary showed good scaling along the D/V axis ($S = 0.947 \pm 0.15$), at 25% ($S = 1.07 \pm 0.31$) and 15% ($S = 0.85 \pm 0.19$) ventral offsets, while it hypo-scaled when measured close to the D/V boundary ($S = 0.46 \pm 0.17$ at 5% dorsal offset where the expression domain was narrowest) (Figures IV.11F and IV.12E-G). Therefore, while the *dad* expression domain definitely expanded with the growing tissue, it scaled to varying degrees at different distances from the D/V boundary. Similar to Brk levels, *dad*-GFP levels increased with time (Figures IV.11A-B and IV.12C). Since *dad* transcription is controlled by both P-Mad and Brk, we asked whether the *dad* domain was specified at constant concentration thresholds of these gradients. We found that levels of both P-Mad and Brk corresponding to the *dad* domain boundary position $x = K_{dad-GFP}$ increased over time (Fig. IV.11D-E). Hence, the Dad domain was not defined at constant P-Mad and Brk concentration thresholds.

IV.8   Sal and Omb domains scale well where they define vein positions

Proper positioning of the veins in the developing wing requires Dpp signaling and is important to ensure adult wing functionality (Birdsall et al., 2000; Blair, 2007; Bollenbach et al., 2008). We asked whether the Sal and Omb domains, informative for positioning veins L2 and L5, respectively, already scaled with tissue
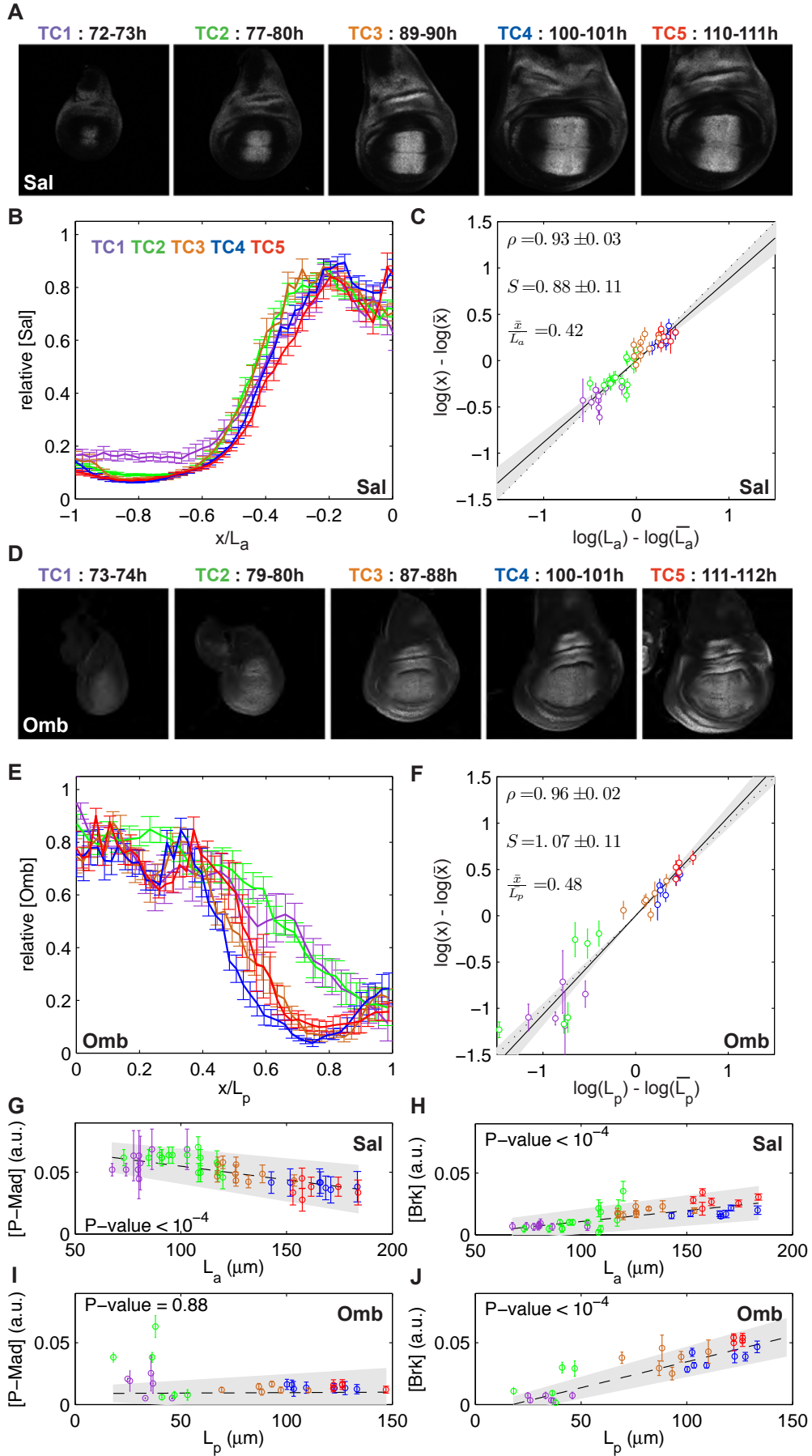
Figure IV.13: L2 and L5 vein positions already scale with tissue size in the larval stages. *(A) Representative images of Sal antibody stainings at each TC. **(B)** Average Sal profiles obtained from the individual profiles normalized to their maximum in the anterior compartment at 15% ventral offset, in relative positions. **(C)** For each disc, the log-deviations in Sal anterior domain boundary position are plotted as a function of the log-deviations in the anterior compartment length L_a. We show the scaling coefficient S and the correlation ρ of the data with 95% confidence intervals. **(D)** Representative images of Omb antibody stainings at each TC. **(E)** Average Omb profiles at 15% ventral offset, obtained from the individual profiles normalized to their maximum. **(F)** Like (C), but for Omb in the posterior compartment. **(G-H)** Average P-Mad (G) and Brk (H) concentrations at the position $x = K_{Sal}$ plotted against L_a for each disc. **(I-J)** Average P-Mad (I) and Brk (J) concentration at the position $x = K_{Omb}$ plotted against L_p for each disc.*

. . . .

size during the larval stages. We found that Sal started to be expressed in the pouch only from the beginning of the third instar stage while Omb was already induced earlier (Figures IV.13A/D). The Sal domain in the anterior compartment spanned 40-45% of the pouch, while it was much narrower in the posterior compartment reaching only up to 15% (Figures IV.13A-B and IV.14A-B). Since the vein L2 is located in the anterior compartment, we investigated scaling properties of Sal in both compartments. Sal profiles expanded with the growing disc with increasing amplitudes of roughly 4-folds, and most of this amplitude increase took place within the first 20 hours of the third instar stage (Figures IV.13A and IV.14B/C/E). It was previously reported that the Sal domain boundary position correlates with the disc size at the end of development (Bollenbach et al., 2008). Consistent with this, we found that the Sal domain boundary actually correlated with the lengths of both the anterior and posterior compartments throughout development. However, a good correlation is not sufficient to ensure scaling, and indeed we observed that the Sal domain boundary exhibited significant hyper-scaling in the posterior compartment ($S = 1.44 \pm 0.3$, Figure IV.14G), while it scaled in the anterior compartment ($S = 0.88 \pm 0.11$, Figure IV.13C). This finding suggested that there might be additional factors at work in the anterior compartment to ensure proper Sal scaling.

The *omb* gene was expressed in a domain larger than Sal, spanning about half the pouch in the posterior compartment (Figure IV.13D-E), and its domain boundary sharpened during development (Figure IV.15D). The position of the boundary was well correlated with the length of the posterior compartment and Omb exhibited close to perfect scaling ($S = 1.07 \pm 0.11$ at 15% ventral offset, see Figure IV.13F).

Finally, we investigated whether the boundaries of Sal and Omb were defined at constant P-Mad and Brk levels. We found that the anterior Sal domain boundary corresponded to decreasing P-Mad levels and increasing Brk levels over time (Figures IV.13G-H and IV.14H-I). In the case of Omb, clonal analysis with *brk* loss of function alleles suggested that positioning of the Omb domain boundary strictly depends on Brk activity (Campbell and Tomlinson, 1999; Sivasankaran et al., 2000). In fact, Omb has no direct input from Mad/Medea complexes and is only indirectly activated by Dpp via repression of Brk (Affolter and Basler, 2007). It was surprising in this light that the Omb domain boundary corresponded to similar levels of P-Mad and increasing amounts of Brk during growth (Figure IV.13I-J). This finding suggested that the *dad*, *sal* and *omb* enhancers desensitized to Brk over time.

Overall, we have shown here that P-Mad and its targets scaled rather well with the growing tissue. Additionally, Teleman et al. found that when the posterior compartment was enlarged or reduced in size via modifications of Insulin signaling activity, the size of the Sal domain adjusted accordingly (Teleman and Cohen, 2000). These observations suggest that there might be an active mechanism in place to ensure scaling of Dpp activity with tissue size, and raise the question of the identity of the involved players.
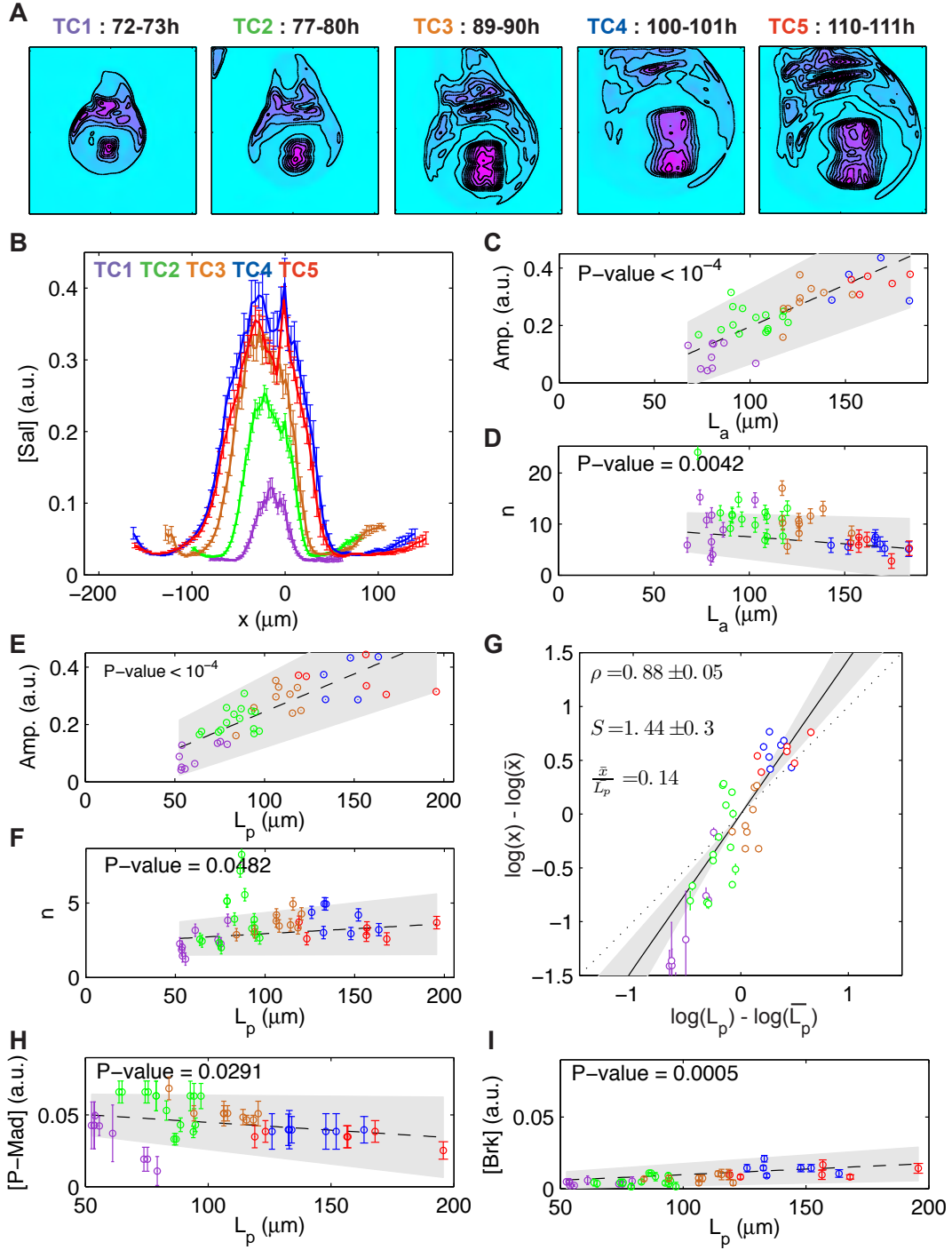
Figure IV.14: Sal amplitudes increase over time and the posterior Sal domain hyper-scales. *(A) Representative Sal contour plots for each TC. Lower concentrations in light blue, higher concentrations in pink. (B) Sal profiles averaged per TC at 15% ventral offset. (C, E) Sal amplitudes (i.e. the concentration at A/P compartment boundary, $x = 0$) for each disc at 15% ventral offset versus the anterior (C) and posterior (E) compartment lengths. (D, F) Sharpness of the Sal domain boundary (n) in the anterior (D) and the posterior (F) compartments plotted against tissue size. (G) Scaling and correlation of the Sal domain boundary in the posterior compartment at 15% ventral offset. (H-I) The average P-Mad (H) and Brk (I) concentrations at the position $x = K_{Sal_P}$ (Sal boundary in the posterior compartment) were plotted against the L_p for each disc.*

Figure IV.15: The Omb domain scales with tissue size. *(A) Representative Omb contour plots for each TC. Lower concentrations in light blue, higher concentrations in pink. (B) Omb profiles averaged per TC at 15% ventral offset. (C) Omb amplitudes (i.e. the concentration at A/P compartment boundary, $x = 0$) versus the posterior compartment length for each disc at 15% ventral offset. (D) Sharpness of the Omb domain boundary.*

. . . .
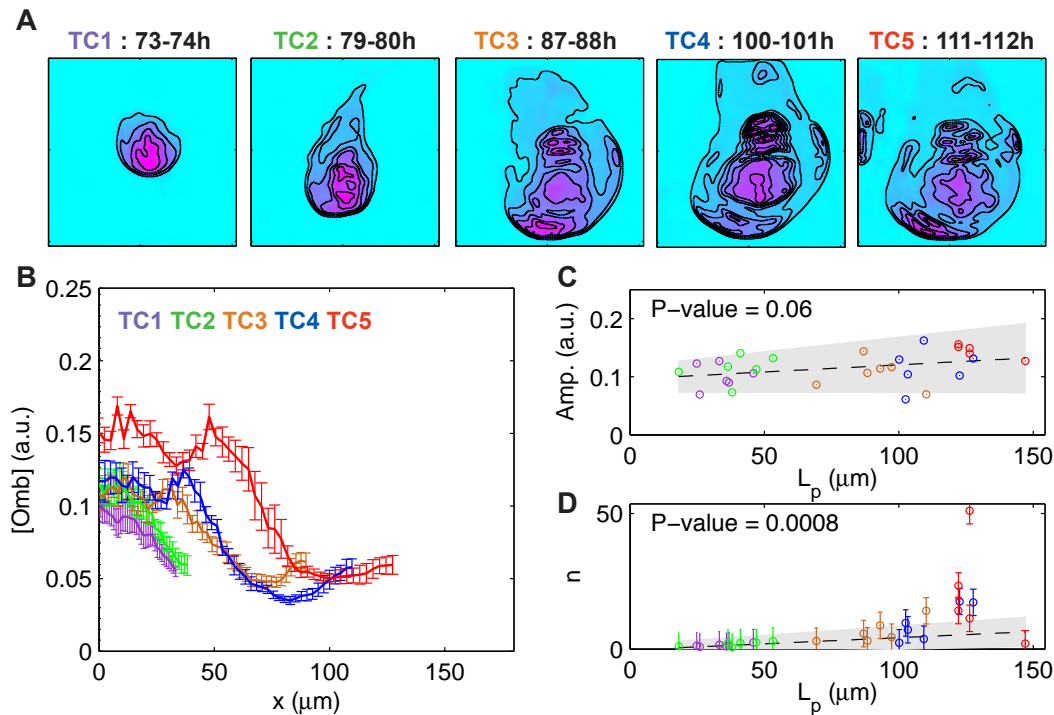
IV.9 *pentagone* function is essential for scaling properties of Dpp activity read-outs

Our collaborators recently identified a Pent-dependent feedback loop as a major modifier of the Dpp activity gradient (Vuilleumier et al., 2010). Here we repeated our analyses in a *pent* mutant background using a null allele, *pent 2-5*, in order to examine its potential involvement in scaling. While Pent function is absolutely essential for proper Dpp gradient formation and maintenance, in its absence flies are semi-viable and overall smaller with wings that lack L5, hence the name of the gene (Vuilleumier et al., 2010).

In the absence of *pent*, the defects in the Dpp activity gradient were visible very early on, with P-Mad and Brk forming very steep gradients that resembled sharp domains in all the TCs we examined (Figure IV.16A-D). P-Mad amplitudes were similar to wild-type levels in *pent* deficient discs and stayed rather constant across growth (Figure IV.17). We observed that the P-Mad domain expanded from TC1 to TC2, and after that did not expand any further and hence did not scale with tissue size (Figure IV.16C). Interestingly, none of the D/V related changes that took place in TC5 in the wild-type background were observed in the absence of *pent*. P-Mad levels were not suppressed significantly along the D/V and the profiles looked similar at different distances from the boundary, suggesting that Pent was a contributor to this effect (Figure IV.17). Indeed, Pent binds to Dally (Vuilleumier et al., 2010) which has a role in shaping gradients of both Wg and Dpp (Belenkaya et al., 2004; Han et al., 2005; Lin and Perrimon, 1999) and hence the D/V centered effects on the P-Mad profiles we described here were likely related to this connection.
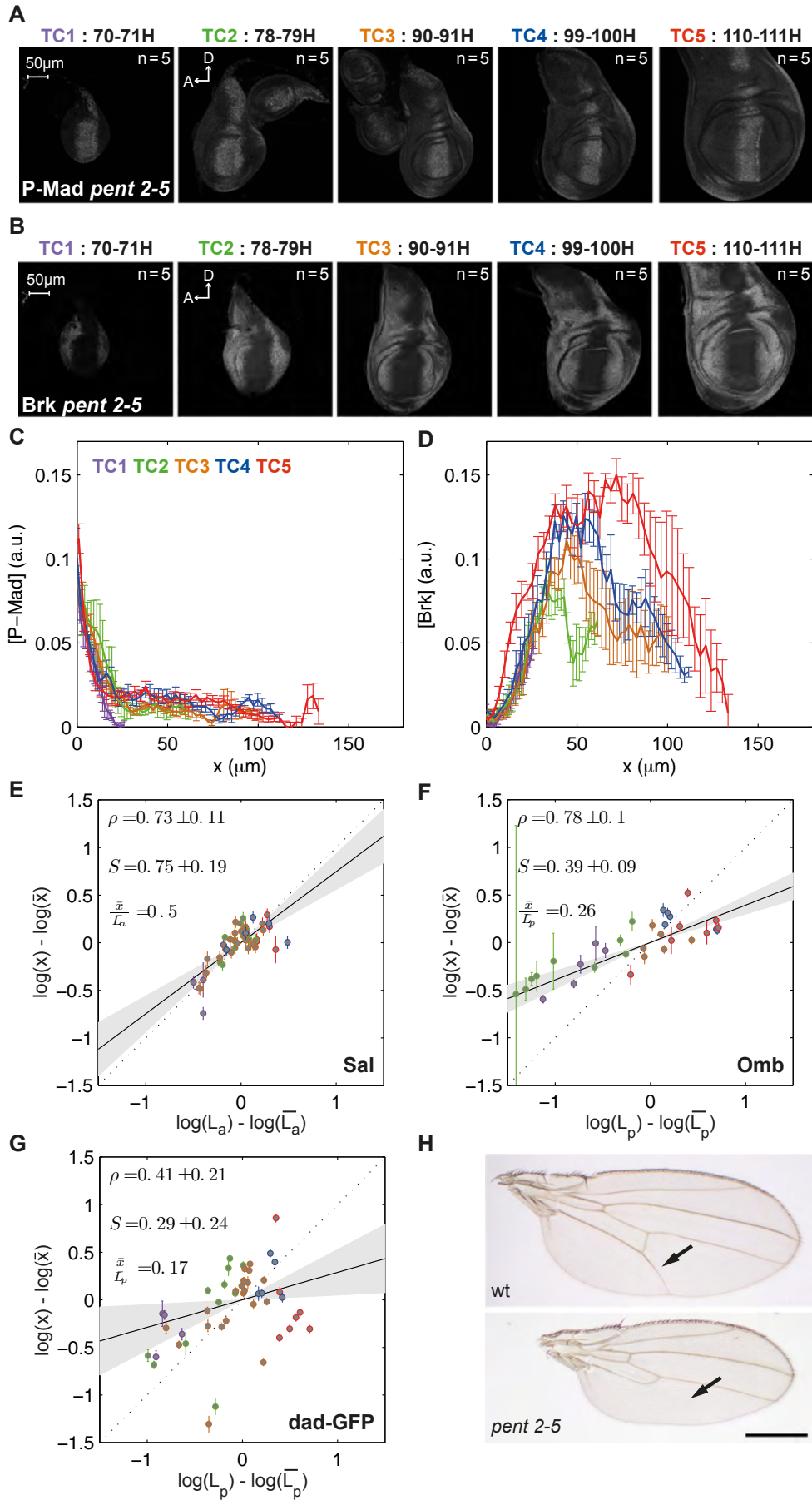
Figure IV.16: Scaling of the activity gradients requires Pentagone function. *(A-B) Representative images of P-Mad (A) and Brk (B) antibody stainings at each TC in pent 2-5 mutant background. (C-D) Average P-Mad (C) and Brk (D) profiles at 15% ventral offset in pent 2-5 display no adaptation to growing disc size. (E-G) Scaling and correlation of anterior Sal (E), posterior Omb (F) and posterior dad-GFP (G) domains in pent 2-5 mutant background. (H) Wild-type (wt) and pent 2-5 mutant adult fly wings. Arrows point to the position of L5, which is missing in pent 2-5 mutants. The scale bar is 500μm.*
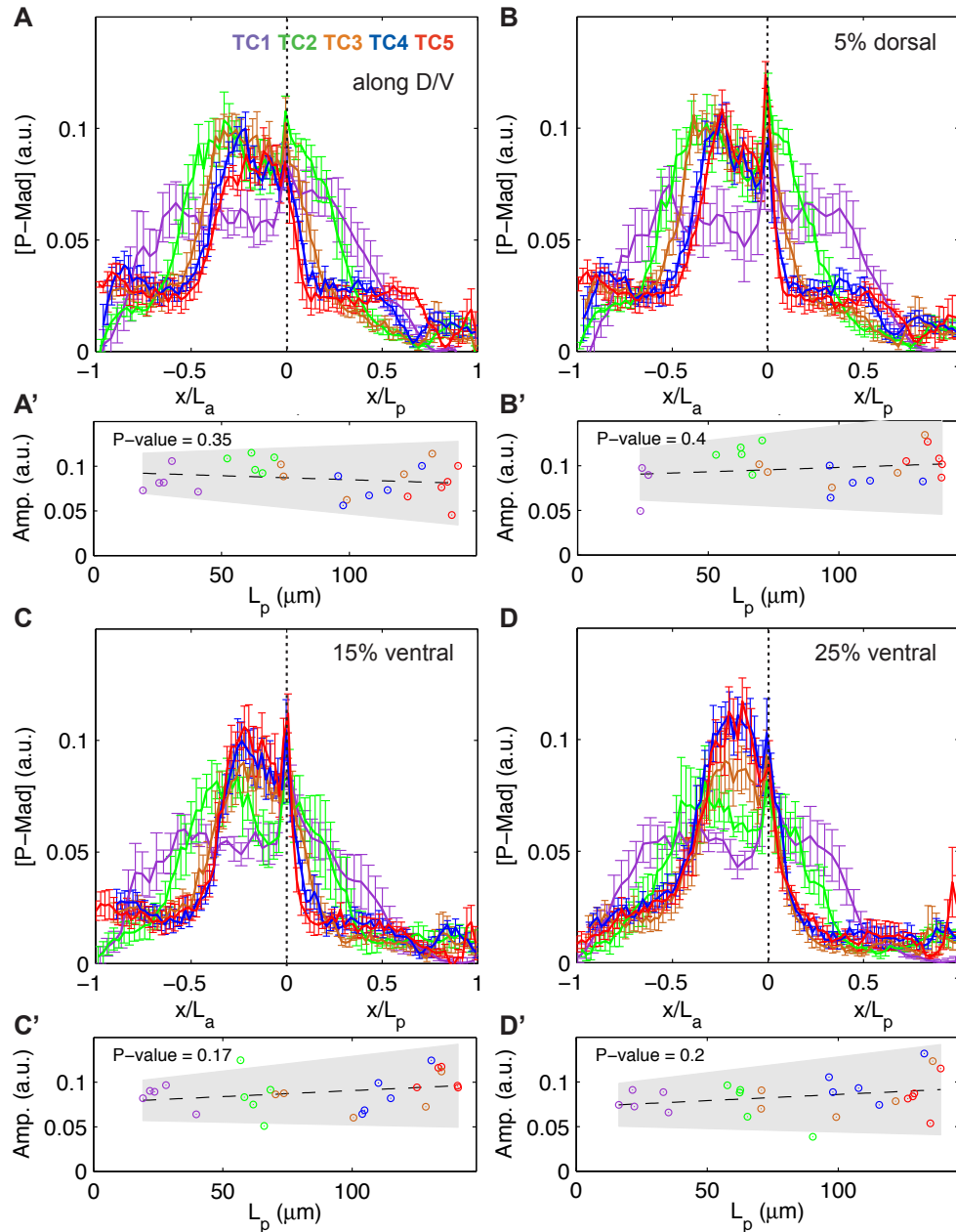
Figure IV.17: Lack of size adaptation of P-Mad profiles in *pent 2-5*. *(A-D) P-Mad profiles in pent 2-5 mutant background are shown in relative positions and averaged per TC along the D/V (A), and with 5% dorsal (B), 15% ventral (C), 25% ventral (D) offsets. Positions in the posterior compartment are normalized relative to the posterior compartment length L_p, while positions in the anterior compartment are normalized relative to the anterior compartment length L_a. (A'-D') The amplitude of the P-Mad profile in pent 2-5 mutant background (i.e. the concentration at A/P compartment boundary, $x = 0$) are plotted versus the posterior compartment length for each disc.*
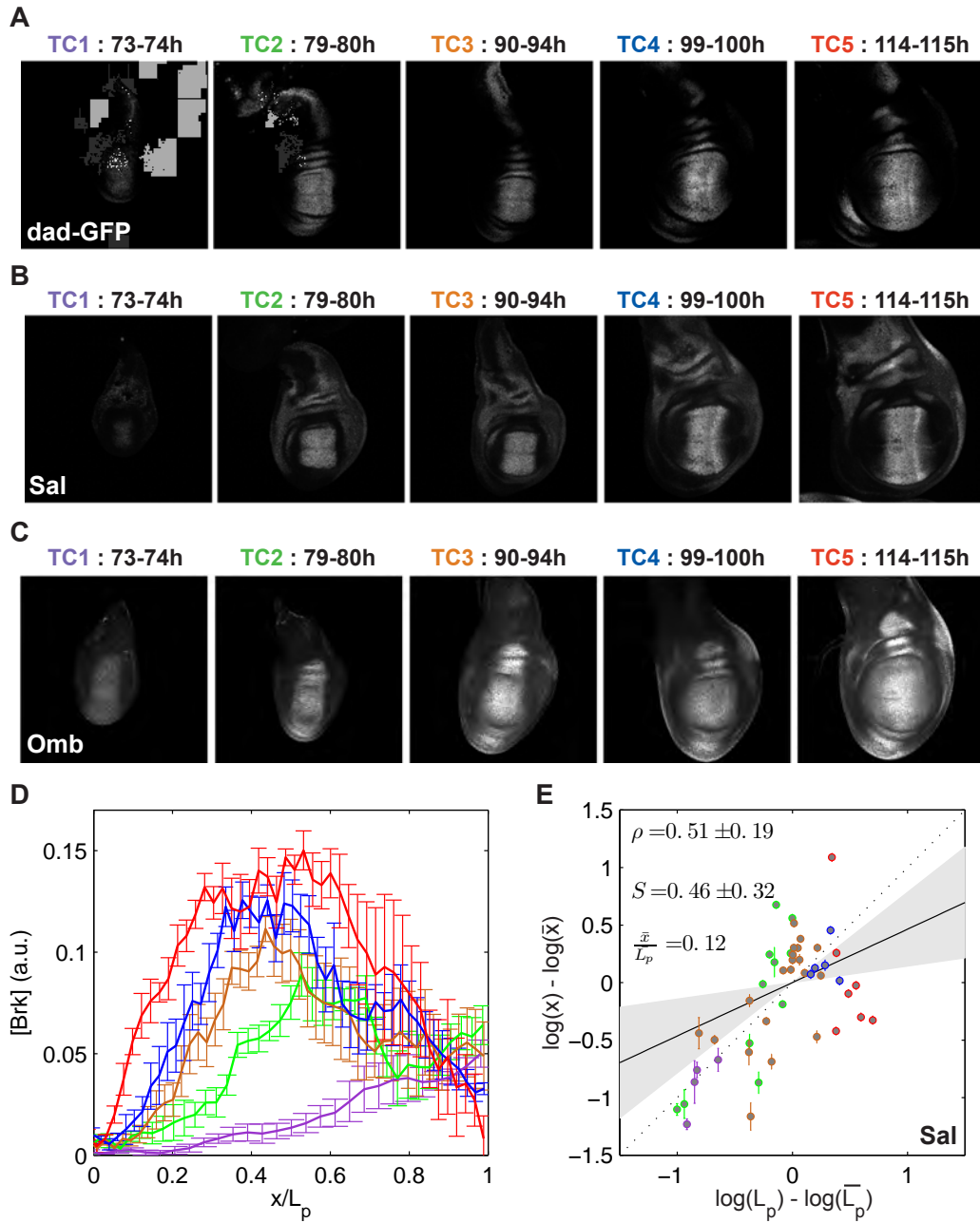
Figure IV.18: Expression patterns of target genes in *pent 2-5*. **(A-C)** Representative images of dad-GFP (A), Sal (B) and Omb (C) in pent 2-5 mutant background. **(D)** Brk profiles averaged per TC in relative positions at 15% ventral offset in pent 2-5 background. Brk levels still increase in pent mutants and the profiles move relatively inwards as the discs grow. **(E)** Posterior Sal domain boundary does not scale in pent 2-5 discs.

. . . .
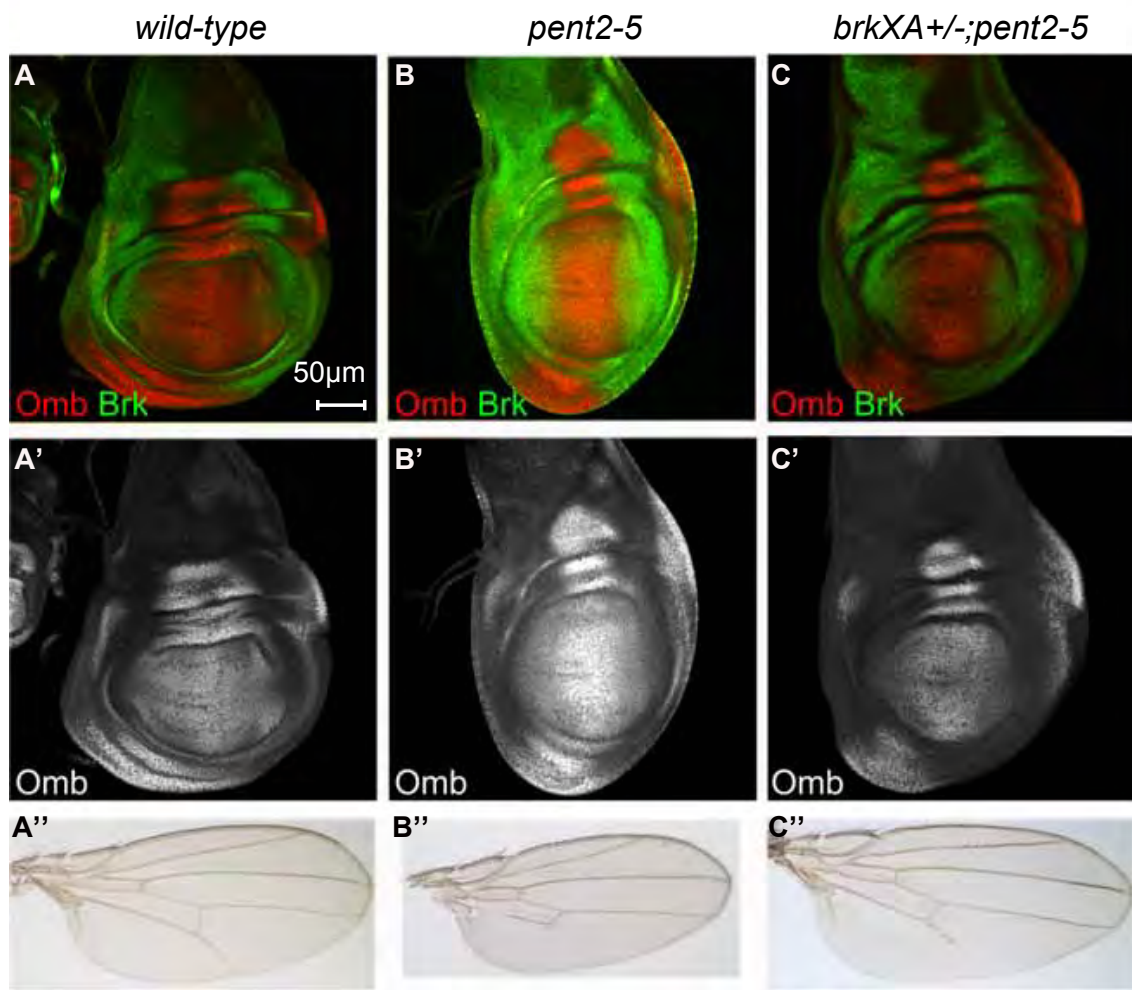
Figure IV.19: Omb and Brk domains overlap extensively in *pent 2-5*, and heterozygosity for *brk* rescues growth defects of *pent* mutants. *(A-A'')* wt *(B-B'')* pent2-5 *(C-C'')* brkXA +/-; pent2-5 (A, B, C) Third instar wing imaginal discs stained for Omb (red) and Brk (green). (A', B', C') Omb channel only (gray). (A'', B'', C'') Representative wings from female flies of the corresponding genotypes. Heterozygosity for brk rescues growth defects of pent mutants to a large extent. The slight rescue of L5 shown in C'' is highly variable and represents an average wing, some flies have an almost complete rescue of L5 while others display no rescue.

. . . .

Were these changes in P-Mad dynamics reflected in the expression pattern of its target genes? Repression of Brk via P-Mad did not seem to be affected by the removal of *pent* as Brk still closely followed P-Mad (Figure IV.16B/D). In the absence of *pent*, the Brk domain moved more interiorly, following the narrower P-Mad domain. Cells with highest levels of Brk were roughly in the middle or even more proximal, which represented a significant shift compared to being at $x/L_p \simeq 0.8$ in wild-type discs (Figures IV.16D and IV.18D). Movement of Brk more interiorly in the absence of *pent* was likely to be a major contributor to the growth defects of these mutants, as Brk is a well-established growth inhibitor (Martin et al., 2004; Schwank et al., 2008). Consistently, heterozygosity for *brk* was able to suppress these effects to a large extent (Figure IV.19). Like in the case of P-Mad, the graded expression pattern of Brk was lost in the absence of *pent*, and the Brk domain did not scale with tissue size (Figures IV.16D and IV.18D).

Next, we asked whether this failure of the Dpp activity gradients to adjust to tissue size led to narrower

expression domains of the downstream targets. We found that this was indeed the case, especially in the posterior compartment (Figures IV.16G and IV.18A). In wild-type discs, we observed very good scaling of Sal in the anterior compartment ($S = 0.88 \pm 0.11$ at 15% ventral offset), where it helps position L2, while it hyper-scaled in the posterior compartment. Interestingly, in *pent* mutant discs, scaling of Sal was less affected in the anterior compartment ($S = 0.75 \pm 0.19$ at 15% ventral offset) and the adult flies still had a properly positioned L2 (Figures IV.16E and IV.18E). However, scaling of Omb in the posterior compartment was reduced from almost perfect in wild-type ($S = 1.07 \pm 0.11$ at 15% ventral offset) to nearly lost ($S = 0.39 \pm 0.09$ at 15% ventral offset), while correlations were still very good (Figure IV.16F). In the absence of *pent*, the Omb domain boundary defined by the Hill fit shrinked to a quarter of the posterior compartment ($x/L_p = 0.26$). Despite this shrinkage, the Omb domain overlaped significantly with the Brk domain in this background, especially at the end of third instar stage, a phenomenon not observed to this extent in wild-type discs (Figure IV.19). A large stripe of cells expressed both Omb and Brk, raising the possibility that failure to define L5 might be due to this extensive overlap.

We conclude that the adaptation of the Dpp activity gradient to tissue size described in the first part requires Pent function.

IV.10 Scaling *during* growth

Size differences in the wing disc are due to growth, but also to natural fluctuations between discs. In the literature, scaling has been investigated at late 3rd instar by looking at various discs (Teleman and Cohen, 2000; Bollenbach et al., 2008), similar to the approach we undertook in the embryo (Part III). Using our scaling measure defined in Part II, we could also use our dataset to assess if the target gene domains (and hence the resulting veins) scale with the pouch at the end of the larval stages, rather than *during* growth. A major challenge is the difficulty to precisely determine the age of a disc, which was circumvented by our collaborators using CO_2 (cf. Subsection IV.4.1). Thus, our collaborators are now planning to investigate scaling of the veins in the adult wing, while we perform this "late stage" analysis.

In this work, we decided to first focus on scaling *during* growth, which might be even more important for morphogen gradient formation and growth coordination. Given the significant increase in size of the wing disc, scaling during growth is probably crucial to ensure robust patterning too. Ideally, we would have liked to perform live imaging on a single disc, in order to assess how a gene expression domain behaves as the disc grows. Rather, since the techniques for culturing wing discs *in vitro* or imaging them through the larval cuticle were (and are) not yet ready, we decided to measure scaling in a collection of discs of different sizes. One of the advantages of this technique is that we could use antibodies rather than UAS constructs, which can sometimes have unknown/unwanted effects (for example, we first used a UAS-GAL4 construct to visualize Wg and Ptc, which actually grew slower than wt discs). Also, live imaging requires dealing with the stability of the reporter, which can affect the real shape of the gene expression domain (see also Subsection IV.11.3).

The cell density and area of the posterior compartment of the wing imaginal disc were measured in a recent article over a period of 100h (Wartlick et al., 2011, Figure S5B and supplementary section 4 therein). This study showed that while the posterior compartment area increased by a factor 100 during this period, the

cell density only changed by a factor 2.5. In the 45h of interest in our study (corresponding to 45h-90h after hatching), the posterior compartment size increased by a factor ~12, while the cell density increased only by a factor ~1.5. Thus, the increase in amplitudes that we observed for some genes might in part be due to this effect. However, assuming that the increase in cell packing is uniform and since we measured positions in microns and not cell numbers, this should not affect our scaling results (except for P-Mad where we took into account the amplitudes). In our analyses, we generally assumed that the extracted profiles represented to a good approximation the nuclear protein levels and that the changes in cell density were negligible.

We measured pathway activity using an antibody specific to the phosphorylated form of Mad, and compared the P-Mad levels in space and time with the activity levels of direct target genes, such as *brk*, which plays key roles in both growth and patterning (Campbell and Tomlinson, 1999; Jazwinska et al., 1999; Martin et al., 2004; Schwank et al., 2008). We found that P-Mad levels scaled perfectly posterior to $0.4L_p$ with the exception of TC5 profiles near the D/V boundary. Previous studies that examined P-Mad scaling reached contradictory conclusions: the P-Mad gradients in late stage discs were reported to correlate with tissue size in a previous study (Teleman and Cohen, 2000) and to have no correlation in another (Bollenbach et al., 2008). Similarly, examination of P-Mad gradients across discs of different sizes led to the conclusion that the gradient did not expand (Hufnagel et al., 2007), and more recently that it actually scaled with the tissue size (Wartlick et al., 2011). It seems that most of the confusion can be attributed to different profile extraction protocols as well as various definitions of scaling, as discussed next.

Since P-Mad is an early signature of the activation of the Dpp signaling pathway, we wanted to find out how its scaling properties translated to its immediate key target, the *brk* gene. In addition to being directly repressed by P-Mad, the Brk protein itself is necessary for graded *brk* transcription (Moser and Campbell, 2005). We found that the range of Brk expression strictly followed P-Mad in both wild-type and *pent* mutant discs. Similar to what we observed for P-Mad, Brk also showed very good scaling for positions posterior to $0.4L_p$. By contrast, levels of Brk increased steadily as the discs grew and could not be explained by P-Mad dynamics alone. This increase in Brk could be due to the build-up of the unknown activator of *brk* transcription or, alternatively, the SEs in *brk* could become desensitized to the repressive input of P-Mad. Regardless of the cause of the increase, cells at a given relative position experienced increasing levels of the Brk repressor over time.

Different definitions of scaling

Traditionally, Dpp and P-Mad gradients have been described by a decaying exponential with characteristic decay length λ (Kicheva et al., 2007; Bollenbach et al., 2008; Wartlick et al., 2011). This decay length is different for each profile and corresponds to the position at which the protein levels have decreased by a factor $e \simeq 2.7$. The correlation between the decay length and the tissue size has been used as a proxy for scaling, e.g. in the work of Bollenbach et al., which found no significant scaling for the Dpp and P-Mad profiles at the end of 3rd instar (note that they used the length of the pouch along the A/P boundary as a measure of tissue size (Bollenbach et al., 2008)). In contrast, and consistent with our results, Wartlick et al. recently showed that the decay lengths of *dpp-GFP*, P-Mad, *brk-GFP* and *dad-RFP* do correlate with the length and the area of the posterior compartment during tissue growth (Wartlick et al., 2011). Importantly, they also assessed scaling *qualitatively* in the whole field by looking at the collapse of the profiles in relative
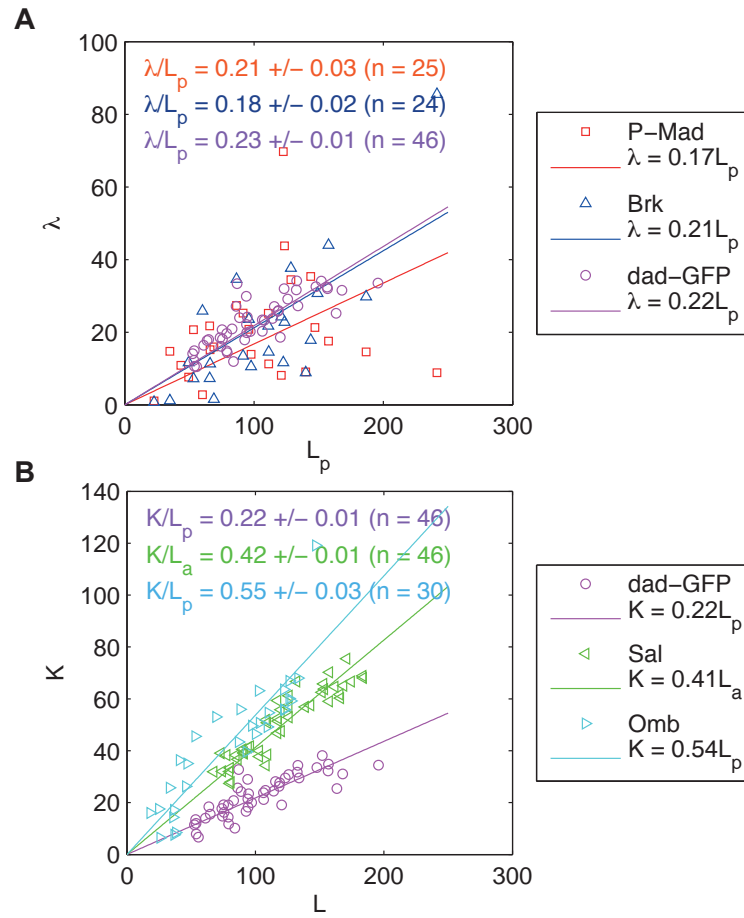
Figure IV.20: Decay length λ correlates with tissue size. *(A) P-Mad, Brk and dad-GFP profiles were fitted with an exponential and the resulting decay lengths λ are plotted as a function of the posterior compartment length L_p. The text in the plot shows the average λ/L_p ratio with its standard error. Relationships obtained from the linear regression are displayed in the boxes on the right. Note that L_p is measured along the D/V compartment boundary. (B) Same as A, but the dad-GFP, Sal and Omb profiles were fitted with a Hill function instead of an exponential.*

positions and normalized intensities (cf. Figure II.3). This method has two advantages: it does not require any fitting of the profiles, and it shows scaling at all positions and not just at the characteristic decay length position (i.e. $x = \lambda$).

In our work, we used the raw intensity measurements without fitting any function to the corresponding profiles, since the exponential was not the best fit at all time classes. Similarly, we wanted to assess scaling in the whole field and not just at one characteristic position. For this, in addition to looking at the collapse of the profiles, we used our measure of scaling (cf. Part II), which gave a *quantitative* measure for each position in the tissue. Lastly, we would like to highlight that we did not normalize the P-Mad intensities before measuring scaling, as the absolute protein levels are crucial for signal interpretation in the simplest scenario. In the case of Brk, however, whose amplitudes increased steadily, scaling only emerged after normalizing the intensities.

For comparison, we also showed in Figure IV.20 the ratios of the decay lengths over the pouch sizes for all the genes we investigated. Since the more downstream target genes were better fitted with a Hill function,

we also reported the ratio of the corresponding transition points as a function of pouch size for those genes. In agreement with (Wartlick et al., 2011), we found that P-Mad and Brk decay lengths correlated well with tissue size, with $\lambda_{P-Mad} = 0.21L_p$ and $\lambda_{Brk} = 0.18L_p$. We note that our estimate of λ_{P-Mad} was smaller than the previously reported value ($0.34L_p$), likely due to the fact that we measured tissue size along the D/V boundary, from the intersection of the A/P to the limits of the pouch, as opposed to the length of the posterior compartment at its widest position. Our measure of λ_{Brk} was very similar to that of (Wartlick et al., 2011). Hence, considering that we have a larger value for the tissue size, Brk protein must form a gradient with a larger decay length than the *brk-GFP* reporter that was used by (Wartlick et al., 2011).

IV.11 Our data in the light of three models

In this work, we carefully analyzed the dynamics and the scaling properties of the Dpp activity read-outs in the growing wing imaginal discs. We discuss here our findings with regard to three models that were put forward to explain scaling (the expansion-repression model, (Ben-Zvi and Barkai, 2010)), pattern formation (cf. Figure 1 in the *Prologue*) and uniform growth (Wartlick et al., 2011).

IV.11.1 Expander-repressor scaling model

A recent mathematical model termed expansion-repression integral feedback control suggested that scaling emerges as a natural consequence of a feedback topology (Ben-Zvi and Barkai, 2010). The hypothetical *expander* molecule facilitates the spread of the morphogen and in turn is repressed by it; scaling is achieved given that the expander is stable and diffusible. The known properties of Pent fit the requirements of this hypothetical agent: Pent is secreted, its transcription is directly inhibited by Dpp signaling and the protein is required for Dpp spreading. However, we do not know how stable Pent is, and *pent* transcription is never abolished in the entire field during the larval stages in which the gradient acts. To test whether Pent could be a key player involved in scaling of Dpp signaling during disc growth, we repeated our analyses at all time points in the absence of Pent. We found that the P-Mad and Brk gradients indeed failed to scale with the tissue size in this mutant background. Scaling of *dad*-GFP and Omb were also strongly affected, while Sal still exhibited some degree of Pent-independent scaling. Importantly, while the function of Pent was essential for proper scaling of the Dpp activity gradients, we noted that Pent alone could not account for the observed selective scaling of Omb and Sal domain boundaries. Scaling of these target genes specifically in those regions in which they have a patterning function pointed to the involvement of additional players, which will be the subject of future research. Still, our findings strongly suggested that Pent is a very good candidate as an expander of an "expansion-repression integral feedback control" model and therefore provided the first insights into the question of scaling in wing patterning. The exact biochemical functions of Pent have yet to be determined in order to get a more mechanical view of gradient scaling in the developing wing imaginal disc.

IV.11.2 Domain boundaries and the interpretation of the activity gradients

More than 40 years ago, Lewis Wolpert proposed the famous French flag model to explain pattern formation (Wolpert, 1969); here, we tested it with respect to the P-Mad and Brk gradients, the combined action of which defines the boundaries of the Dpp target genes. We measured average P-Mad and Brk concentrations at Dad, Omb and Sal boundaries across development. We found that the amount of P-Mad at these boundaries slightly increased (Dad), slightly decreased (Sal) or was constant throughout development (Omb). Among these three targets, the Omb domain was the widest and it corresponded to a region where the P-Mad gradients scaled perfectly; as a result P-Mad levels fluctuated very little at the Omb domain boundary. Interestingly, the domain boundary of Omb is thought to solely depend on Brk and hence constant P-Mad levels might be a mere coincidence. Remarkably, all the target genes we considered responded to significantly increasing levels of Brk suggesting that the target genes desensitized to Brk over time, so that more and more Brk could be tolerated at the domain boundary. Alternatively, if we consider that the domain boundaries of *dad-GFP* and Sal did not respond to constant P-Mad levels either (Figures IV.11D and IV.13G), another explanation would be that Brk and P-Mad signals are combined in a non-additive fashion in order to define the boundary position of the target genes. Following this assumption, we looked for a simple combination of these signals that was constant at the boundary of the target gene domains at all TCs. For *dad-GFP* in the posterior compartment, the ratio P-Mad^2/Brk was constant at the domain boundary (t-test P-value $= 0.13$ under the null hypothesis that the slope is equal to zero; in *pent 2-5*, t-test P-value $= 0.88$), while for Sal in the anterior compartment, the multiplicative combination P-Mad$^5 \cdot$ Brk^4 was constant (t-test P-value $= 0.42$; in *pent 2-5*, t-test P-value $= 0.91$). We propose that Brk and the unknown activator of Omb could be similarly combined in order to determine the Omb domain boundary.

IV.11.3 Uniform growth model

We used our data to further test a model that was recently proposed to explain the uniform growth in the wing imaginal disc (Wartlick et al., 2011). The model poses that the temporal changes in Dpp signaling levels drive tissue growth; cells divide when they experience a relative increase of 50% in the levels of Dpp signaling. Since it is the relative difference and not the absolute amount of Dpp signal that regulates cell division, the model can account for the uniform growth of the wing disc. As the relative increase in Dpp activity slows down, the cell cycles lengthen. Growth stops when the cell division time exceeds 30h. The model is based on the finding that Dpp activity scales with tissue size and that cells at a given relative position have increasing levels of Dpp signaling over time. In contrast, we did not observe a general temporal increase in the level of Dpp signaling at a given relative position in this study. P-Mad was the most upstream and the most dynamic read-out available for the activity of the Dpp pathway and we found that the relative increase in P-Mad levels throughout development was not significantly different from zero at most relative positions (in Figure IV.21A', almost all the error bars (95% confidence interval) cross the value $\Delta c/c = 0$). Why is the increase in Dpp-GFP levels not reflected in P-Mad levels? A potential explanation for this might be that the observed accumulation of Dpp-GFP was due to the stability and accumulation of Gal4 since Dpp-GFP was under UAS control (Wartlick et al., 2011). The authors showed that the half-life of the Dpp-GFP fusion protein is only 20min, but the Gal4 stability was not considered. Alternatively, the system could get desensitized over time and more and more Dpp would be required to lead to similar
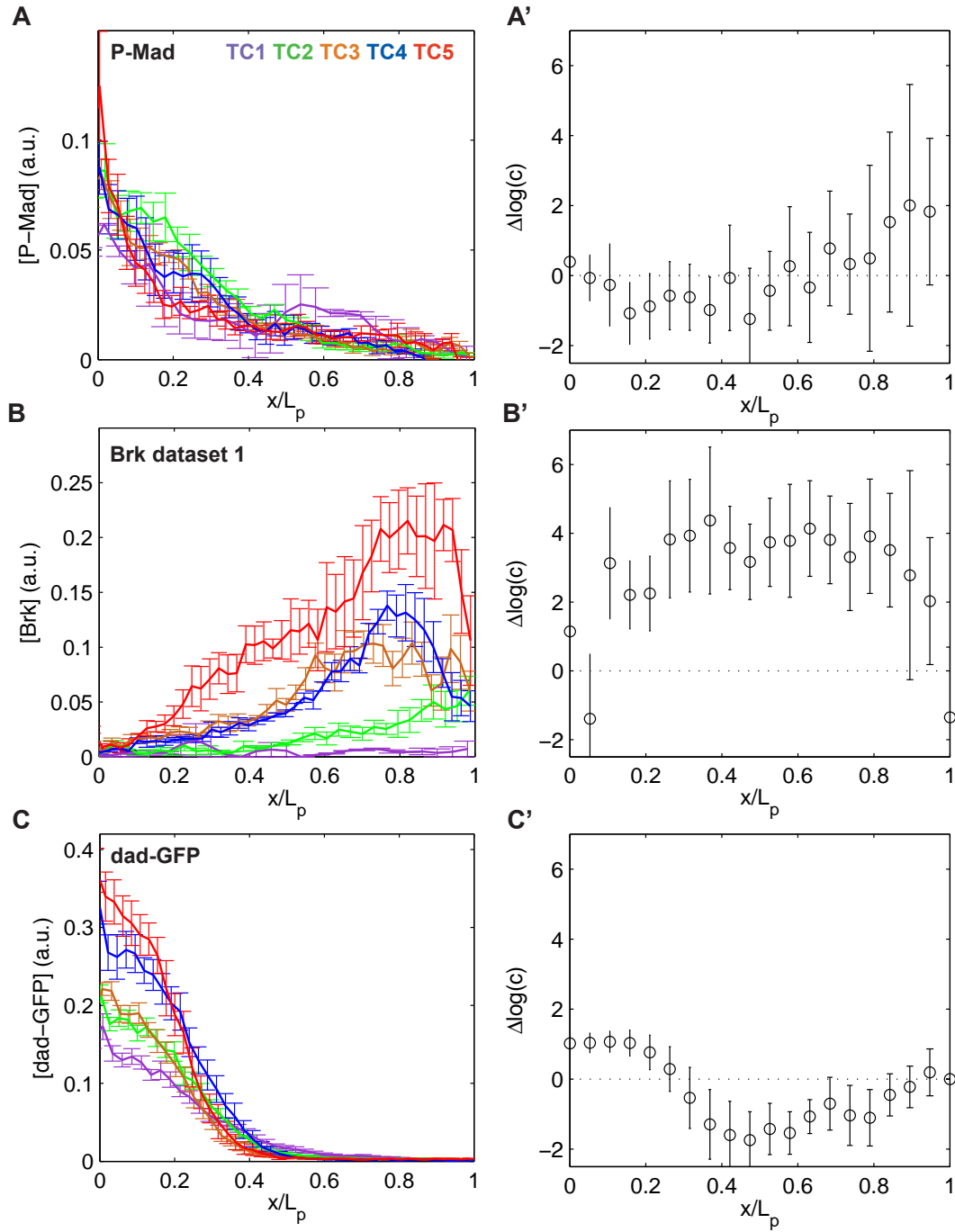
Figure IV.21: Changes in P-Mad, Brk and *dad-GFP* levels at relative positions. *(A-C) P-Mad (A), Brk (B) and dad-GFP (C) profiles averaged per TC in relative positions at 15% ventral offset. Error bars represent the standard error per TC at every relative position. (A'-C') Cells at a given relative position do not experience an increase in P-Mad levels over time, while Brk levels increase 4-5 folds in most of the field. A two-fold increase in dad-GFP levels is only seen in the medial 25% of the disc. Given a relative position x/L_p, we plot the log-concentration as a function of L_p for each disc (not shown). The linear regression yields an estimate of $dlog(c)/dL_p$, where c is the protein concentration. Here, we show the relative increase in protein concentration for each of these relative positions in the pouch ($\Delta log(c) = 1$ represents a 100% increase from TC1 to TC5): $\Delta log(c) \simeq dlog(c)/dL_p \cdot \Delta L_p \simeq \Delta c/c$, where $\Delta L_p = max(L_p) - min(L_p)$.*

P-Mad levels. Finally, increases in Dad levels could counteract the increase in Dpp levels, since Dad is an inhibitory Smad (Affolter and Basler, 2007; Tsuneizumi et al., 1997). In this sense, a recent publication shows that in *dad* mutants, P-Mad levels increase significantly when Tkv levels increase, while they are robust to changes in Tkv levels in wt (Ogiso et al., 2011).

Wartlick et al. monitored Dpp signaling levels using a *dad-RFP* reporter and found a 5-fold increase in the course of 36h (Wartlick, Mumcu et al. 2011). In our analysis, we used a similar tool, *dad-GFP*, and failed to fully reproduce their results. Though we also found that *dad-GFP* scaled with tissue size, its levels increased merely 2-fold over 40h, and this increase took place only in the medial 25% region of the disc, and cells in the lateral part experienced a decrease in *dad-GFP* levels (Figure IV.21C-C'). This disparity in the fold increases was likely due to the higher stability of RFP (Verkhusha et al., 2003) as the enhancer used, to our knowledge, was identical. Additionally, levels of Brk, another direct target of the pathway with a very well established role in suppressing growth in the lateral regions, increased in average 4-fold in most of the field in the interval studied, an observation not reported by (Wartlick et al., 2011) (Figure IV.21B-B'). In lateral areas, the increase in Dpp activity (if present) was below detection levels and would be opposed by increasing Brk levels. Importantly, increasing Brk levels, if they were to depend solely on Dpp, would suggest decreasing Dpp activity in lateral areas as Brk expression is directly suppressed by Dpp activity. Hence, our data raises serious questions about the validity of this uniform growth model, especially in the lateral regions of the pouch. We favor an alternative model that does not rely on Dpp activity alone to explain uniform growth in the wing disc (Schwank et al., 2011).

IV.12   Dpp gradient formation: another controversial topic

Dpp gradient formation has been very controversial lately, mostly opposing two views (not necessarily mutually exclusive). The first model assumes that Dpp diffuses extracellularly by *restricted* diffusion and binds to the receptors at the cells surface (Lander et al., 2002; Belenkaya et al., 2004). In the second model, Dpp travels intracellularly through cycles of endocytosis-exocytosis, named *transcytosis* (Entchev et al., 2000; Bollenbach et al., 2007). Dr Gerald Schwank (group of Prof. Konrad Basler, Uni. Zurich) has been seeking experimental evidence to discriminate between these two models. He also contacted me to build a model to help interpret the data, but at the time, I was too busy with the scaling project. Dr Sascha Dalessi from our group decided to take on this project and I assisted him on various questions, on the biological and modeling sides. Interestingly, these new data strongly favor the first model of extracellular Dpp diffusion (Schwank et al., 2011). In the following, I present yet another modeling project.

Dpp gradient formation by *restricted* diffusion

Experiments have shown that the heparan sulfate proteoglycan Dally acts as a co-receptor of Dpp (Fujise et al., 2003), while it also ensures that Dpp travels far enough from the source (Belenkaya et al., 2004). This dual role is however *a priori* paradoxal, since a co-receptor is believed to sequester the morphogen and thus usually prevents long-range gradient formation. The work of our collaborators on Pent also suggested that Pent is essential to yield Dpp long-range gradient formation and that it binds to Dally (Vuilleumier et al.,

2010). Before even building a model to explain the scaling of Dpp signaling, I wanted to make sense of the data on Dpp gradient formation. In this model, Dally is essential to keep Dpp close to the apical surface of the columnar cells, effectively enhancing Dpp diffusion rate by reducing the dimensionality of the diffusion space (from 3d to 2d). However, I also consider that Dpp binding and unbinding to Dally at the cell surface slows down its effective diffusion rate. Thus, some levels of Dally are required to build a proper long-range gradient, though overexpressing it yields to a short-range gradient due to Dpp sequestration by Dally.

The results are still preliminary but will be further investigated. In its simplest form, the equations of our model describe Dpp diffusion and binding/unbinding to Dally. Dpp diffusion constant increases as a function of Dally concentration up to some maximum (because the diffusion space reduces from 3d to 2d), and we assume that Dally is required for Dpp signaling (i.e. it is a co-receptor). Dpp is produced at the A/P compartment boundary, $[x_0; 0]$ with $x_0 = -0.1L$ (L is the pouch size). Dally is produced everywhere but repressed by Dpp signaling (Fujise et al., 2003), which we assume to be proportional to [Dpp-Dally], since Dally acts as a co-receptor. Also, experiments have shown that Dpp signaling is effectively reduced by Hh signaling at the Dpp production region (Funakoshi et al., 2001); since we do not model the Tkv receptor levels explicitly, we add an activation term $s_{DallyHh}$ that counter-effects Dpp repression at the A/P boundary. The equations read:

$$\frac{\partial Dpp}{\partial t} = \frac{\partial}{\partial x}\left(D_{Dpp}(Dally) \cdot \frac{\partial Dpp}{\partial x}\right) - \alpha_{Dpp} \cdot Dpp - k_+ Dpp \cdot Dally + k_- \cdot [Dpp - Dally] + s_{Dpp}\Theta(-x)\Theta(x - x_0)$$

$$\frac{\partial [Dpp - Dally]}{\partial t} = \frac{\partial}{\partial x}\left(D_{[Dpp-Dally]} \cdot \frac{\partial [Dpp - Dally]}{\partial x}\right) - \alpha_{Dpp-Dally} \cdot [Dpp - Dally] + k_+ Dpp \cdot Dally - k_- \cdot [Dpp - Dally]$$

$$\frac{\partial Dally}{\partial t} = \frac{\partial}{\partial x}\left(D_{Dally} \cdot \frac{\partial Dally}{\partial x}\right) - \alpha_{Dally} \cdot Dally - k_+ Dpp \cdot Dally + k_- \cdot [Dpp - Dally] + \frac{s_{Dally}}{1 + k_s[Dpp - Dally]} + s_{DallyHh}\Theta(-x)\Theta(x - x_0)$$

where $D_{Dpp}(Dally) = D_{Dpp}^{3d} + \frac{F \cdot Dally^n}{K_d^n + Dally^n}$, ensuring that the diffusion rate of Dpp increases with Dally levels (until saturation). Note however that because Dpp binds to Dally, its *effective* diffusion rate $\tilde{D}_{Dpp}$ is given by

$$\tilde{D}_{Dpp} = \frac{D_{Dpp}}{1 + \frac{Dally}{K}} = \frac{D_{Dpp}^{3d} + \frac{F \cdot Dally^n}{K_d^n + Dally^n}}{1 + \frac{Dally}{K}} \tag{IV.1}$$

where $K = \frac{k_-}{k_+}$.

In the simulations, we assume that Dally and [Dpp-Dally] are (almost) immobile, i.e. $D_{Dally}, D_{[Dpp-Dally]} \ll D_{Dpp}$.

We tested our model against five experimental results:

1. *dally* mutant disc (Fujise et al., 2003): Dpp signaling was no longer down-regulated at the A/P boundary and the Dpp gradient got steeper, showing lower concentrations than in wt away from the source.
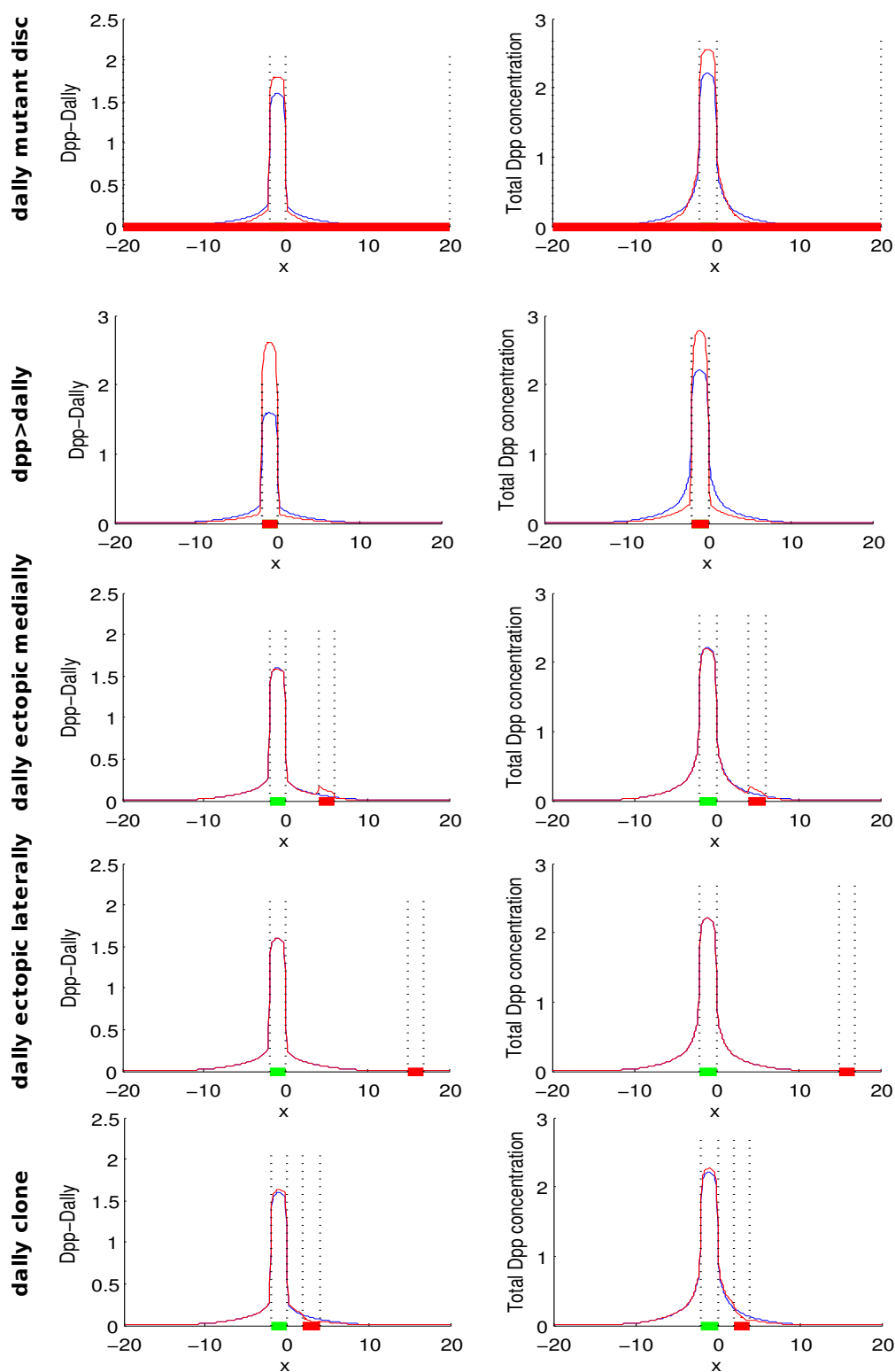
Figure IV.22: Total Dpp and [Dpp-Dally] ($\equiv$ P-Mad) levels in various mutant backgrounds (*wt* in blue, mutant in red). *The green area shows the dpp production region. The red area shows the region affected by the mutation.*

2. *dally* overexpression at the A/P compartment boundary (*dpp>dally*) (Fujise et al., 2003): phenotype similar to *dally* mutant disc.

3. *dally* ectopic overexpression in the medial regions (Fujise et al., 2003): the P-Mad levels increased in the dally$^+$ expressing clones[6].

4. *dally* ectopic overexpression in the lateral regions (Fujise et al., 2003): clones induced in the lateral regions did not show the elevated P-Mad levels of medial clones.

5. *dally* clone in the medial region (Belenkaya et al., 2004): extracellular Dpp levels were reduced within the mutant clone. Behind the clone, levels were also reduced compared to wt.

We reproduced these experiments *in silico* with our model. Results are shown in Figure IV.22 and agree well with all the experiments.

We also considered a more realistic model that takes into account the Dpp receptor Thickvein (Tkv). Tkv is down-regulated by Dpp signaling (proportional to [Dpp-Dally-Tkv]), Hh signaling and upregulated by Engrailed in the posterior compartment. The results presented in Figure IV.23 were obtained using the following model

$$\frac{\partial Dpp}{\partial t} = \frac{\partial}{\partial x}\left(D_{Dpp}(Dally) \cdot \frac{\partial Dpp}{\partial x}\right) - \alpha_{Dpp} \cdot Dpp - k_+ Dpp \cdot Dally + k_- \cdot [Dpp - Dally] + s_{Dpp}\Theta(-x)\Theta(x - x_0)$$

$$\frac{\partial [Dpp - Dally]}{\partial t} = -\alpha_{Dpp-Dally} \cdot [Dpp - Dally] + k_+ Dpp \cdot Dally - k_- \cdot [Dpp - Dally] - k^r_+ [Dpp - Dally] \cdot Tkv + k^r_- [Dpp - Dally - Tkv]$$

$$\frac{\partial Dally}{\partial t} = -\alpha_{Dally} \cdot Dally - k_+ Dpp \cdot Dally + k_- \cdot [Dpp - Dally] + \frac{s_{Dally}}{1 + k_{DallyDpp}[Dpp - Dally - Tkv]} + s_{DallyHh}\Theta(-x)\Theta(x - x_0)$$

$$\frac{\partial [Dpp - Dally - Tkv]}{\partial t} = -\alpha_{Dpp-Dally-Tkv} \cdot Dally + k^r_+ [Dpp - Dally] \cdot Tkv - k^r_- [Dpp - Dally - Tkv]$$

$$\frac{\partial Tkv}{\partial t} = -\alpha_{Tkv} \cdot Tkv - k^r_+ [Dpp - Dally] \cdot Tkv + k^r_- [Dpp - Dally - Tkv] + \frac{(s_{Tkv} + k_{TkvEn}\Theta(x))}{1 + k_{TkvDpp}[Dpp - Dally - Tkv] + \frac{k_{TkvHh}\Theta(-x)\Theta(x-x_0)}{1+|x|}}$$

and also agree well with the experiments. These preliminary results are a proof of principle that the same molecule, here Dally, can have *a priori* contradicting roles depending on its levels. However, since many parameters were unknown, we need to investigate better the robustness of our model and its validity to explain other types of experiments. Importantly, the effect on diffusion of the dimensionality reduction could be studied theoretically and numerically in 2D. Finally, since Pent has been shown to intract with Dally (Vuilleumier et al., 2010) and to be important for scaling, we could take Pent and Dally into account for a wing expander-repressor model of scaled Dpp gradient formation.

---

[6]It is experimentally possible to generate clones of mutated cells within a wild-type tissue. As a result, morphogen diffusion or degradation in the tissue can be perturbed if the local mutation involves a gene important for gradient formation (e.g. decreasing diffusion or participating in degradation). This can be modeled by considering that one or more parameters are no longer uniform in space (e.g. the diffusion or degradation rate).
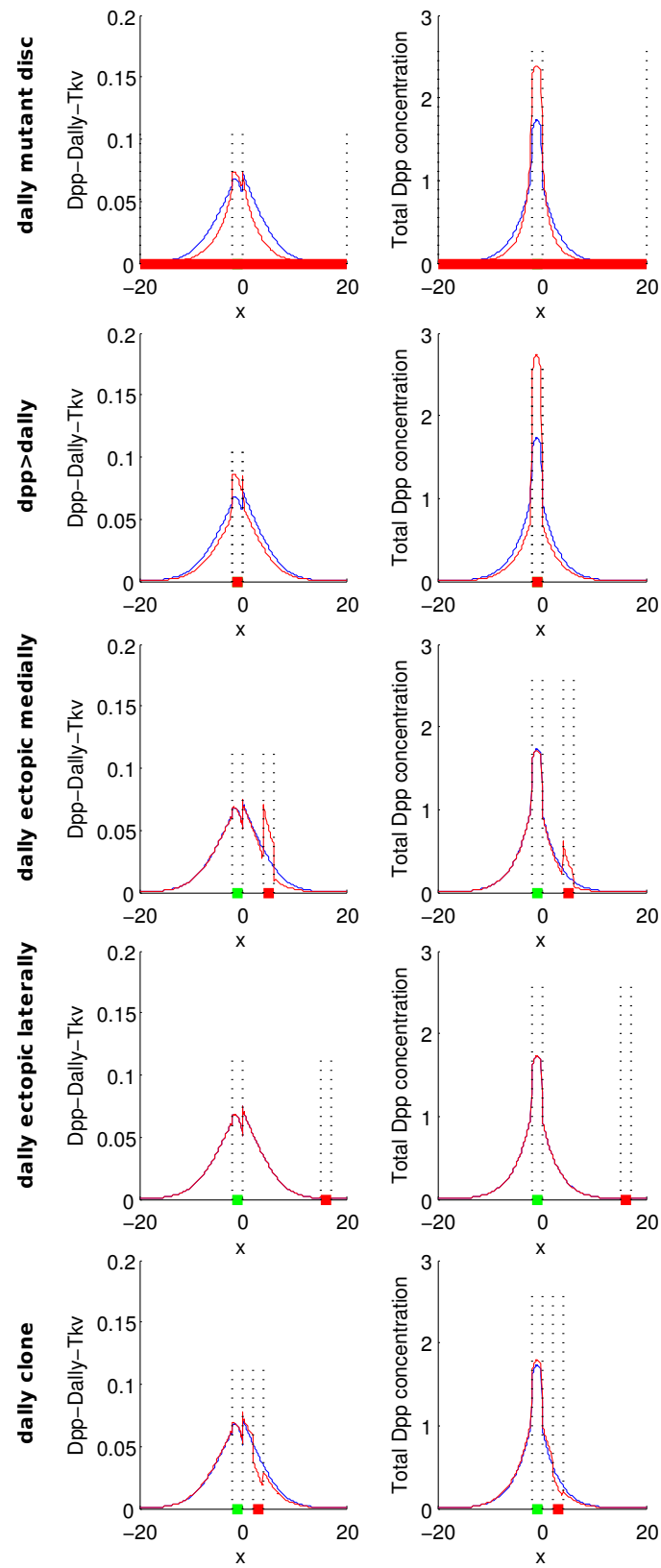
Figure IV.23: Total Dpp and [Dpp-Dally-Tkv] (≡ P-Mad) levels in various mutant backgrounds (*wt* in blue, mutant in red). *The green area shows the dpp production region. The red area shows the region affected by the mutation.*

IV.13   Perspectives

In the text, I have already mentioned several issues to be further investigated. While Pent seems to play an important role to ensure proper scaling of the Dpp targets, we still need to investigate further the potential mechanisms underlying scaling. These include the role of feedbacks from the signaling cascade for Dpp gradient formation (e.g. Pent that we studied here, but also Dally, Brk, Dad), the importance of cell division control (Hippo signaling pathway) and finally, cell size regulation (Insulin pathway). Also, live imaging in the future will certainly bring better insights into the questions of scaling and gene regulation.

In the near future, the plan is to continue analysing two other datasets that were produced by our collaborator Dr. Hamaratoglu using the same protocol. This time, the experiments were performed in *brk* mutants, *expanded* mutants and *pent 2-5/expanded* double mutants. The *expanded* gene[7] is part of the insulin signaling pathway and causes the discs to overgrow, while *pent 2-5* mutants have the opposite phenotype since the absence of Pent restricts the spread of Dpp (Vuilleumier et al., 2010). If Dpp scaling is solely controlled at the level of the Dpp morphogen gradient formation, acting on the insulin signaling pathway should not affect scaling. Otherwise, our data could also be used to better investigate growth defects. Indeed, we could follow the evolution of compartments area throughout development in the various mutant backgrounds to better understand how they affect growth.

Of importance too, future work could look more into the interpretation of the activity gradients by the target genes. This would require co-stainings of P-Mad and Brk with each target gene at various time classes, in order to properly determine what combination of P-Mad and Brk is read out by each target. These predictions should then be further tested by investigating the enhancer regions of these genes.

Lastly, the modeling that I presented in Section IV.12 requires more careful analysis of the parameters and robustness (results not published yet). Also, since our experiments raised many hypotheses on the regulation of the Dpp signaling pathway, we could build a model to test them. The model would be intended to show how increasing Dpp levels (Wartlick et al., 2011) can yield to nevertheless P-Mad levels of constant amplitude. P-Mad gradient should scale and translate into an inverse gradient of Brk whose spread scales but whose levels accumulate with time. P-Mad and Brk signals should be combined in order to yield the observed scaling of the target genes.

---

[7]Do not confuse with the *expander*, which is the hypothetical molecule in the expansion-repression integral feedback control mechanism by (Ben-Zvi and Barkai, 2010).

Part V

Other Contributions

. . . .

V.1   Iterative Signature Algorithm GUI

When I joined the computational biology group in February 2007, I started working on a graphical user interface (GUI) for biologists to explore transcription modules in gene expression data. Transcription modules can be obtained using different algorithms like hierarchical clustering or the Iterative Signature Algorithm (Bergmann et al., 2003; Ihmels et al., 2004) that can assign a gene to several groups of co-expressed genes. The general biological assumption is that genes that are co-expressed are potentially co-regulated and thus participate in a common biological pathway. Having a tool to look at transcription modules can therefore be very useful e.g. to speculate on the function of an unknown gene co-expressed with a group of genes with known function. In a case-control study, it can also highlight the pathways altered in the unhealthy tissue.

Although many algorithms existed for clustering, we felt the lack of a user-friendly and free interface for biologists (and not just programmers) to dig for biological information in those large datasets, in particular using the Iterative Signature Algorithm (Bergmann et al., 2003; Ihmels et al., 2004). Using the actionscript programming language, I therefore started developing a GUI for the ISA during 3-4 months. The display has two views: a global and a modular view.

V.1.1   Global view

The image in the global view represents all the modules found with the data for a given threshold (rows: samples or microarrays / columns: genes) (Figure V.1). Every module or overlapping module is assigned a number as well as a name, consistent with the gene ontology categories found for the module. Hovering over a module on the image displays the name of the module. Clicking on it switches to the modular view.

A basic 200% zoom is also available (upper left corner of the image). In order to use it, the user can click on the *on/off* text. The zoom can also be switched on by pressing the 'z' key. Releasing it will automatically turn off the zoom.

Figure V.1: Iterative Signature Algorithm Graphical User Interface. *Global View.*

. . . .

The panel on the right side allows the user to select a gene, a sample or a module. The search tool finds a gene or a sample in the list below. When a gene or sample is selected, extended information appears on the text area below. On the image, the gene/sample is located by a line and the modules it belongs to are highlighted (in yellow with an arrow in the upper left corner). When a module is selected in the combo box, the module is highlighted in the image. Since the modules in the image may be segmented (a gene can belong to several modules), a line on the left and lower side of the image indicates if other samples/genes belong to that module as well. To look at the module more precisely in the modular view, the user can click on it in the image or click on its link in the text area below the list of modules.

V.1.2   Modular view

The modular view is very similar to the global view, except that now only the genes and the samples that belong to the chosen module are visible (Figure V.2). Clicking somewhere in the image shows the name of the gene/sample that was hit (upper side of the image). In the combo box, either genes, samples or gene ontologies can be selected.

In the list below, the genes appear with their relative score in the module, their name and their long name (eg. (94%) EXO1 exonuclease 1). The genes are ordered according to their score, but they can also be sorted alphabetically by clicking on the "Sort items alphabetically" button on the top of the list. Samples in the list appear only with their relative score in the module and their name. Finally, gene onthologies appear with their $-log_{10}$(P-value) and name.
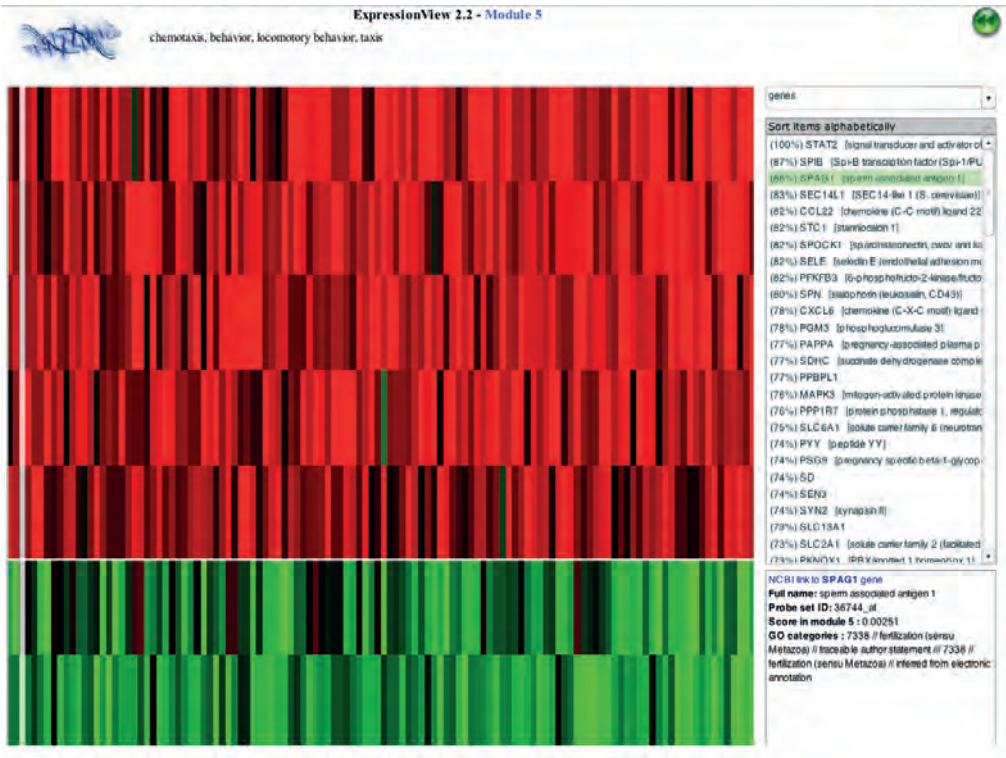
Figure V.2: Iterative Signature Algorithm Graphical User Interface. *Modular View.*
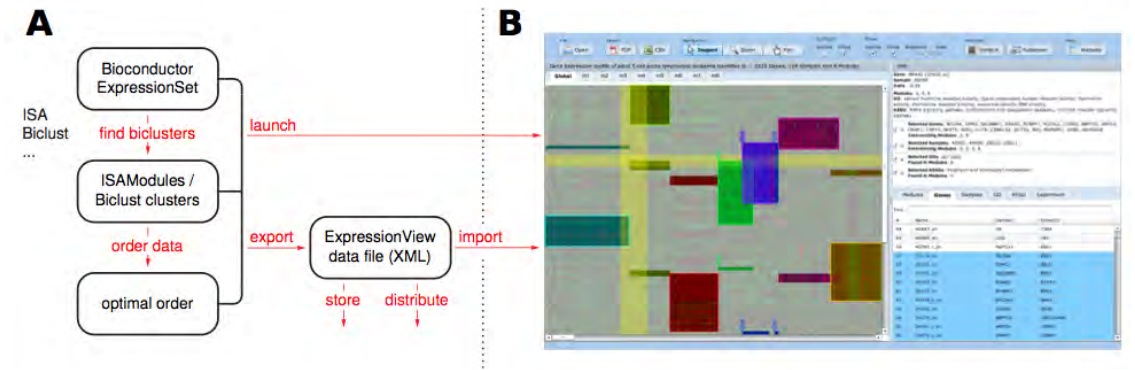
Figure V.3: Workflow of ExpressionView, showing the two parts of the analysis. *(A) These steps are performed by a bioinformatician, using GNU R. Starting from gene expression data in the form of a Bioconductor ExpressionSet, the first step is finding the modules. In a second step, the rows and columns of the gene expression matrix are rearranged to produce an easily readable overview of the results. The last step consists of combining the gene expression data and its associated metadata, possibly including results of enrichment analysis, with the results from the biclustering and produce an ExpressionView data file. (B) This file can be distributed and finally explored with the interactive Flash applet by the end-user. Please see the website for details on the data file format.* Reprinted from (Lüscher et al., 2010) by permission of Oxford University Press.

When clicking on a gene/sample/gene ontology, information appears in the text area below (the score is now absolute). On the image, a line locates the gene or the sample.

V.1.3 ExpressionView

This work was left aside when I started working on the *Drosophila* embryo project. In 2009, Andreas Lüscher joined the lab with the aim of retaking this project and making it into an R package. The ergonomy of the GUI remained, but it was reprogrammed in the Adobe Air package and greatly improved with additional features, becoming the *ExpressionView* (Lüscher et al., 2010).
URL: http://www.unil.ch/cbg/ExpressionView. Screenshots, tutorials and sample data sets can be found on the ExpressionView web site (Figure V.3).

V.2 Neuroenergetics

This work started in 2009 as a student project for the Bachelor course "Solving Biological Problems that Require Math". The biological problem was proposed by Prof. Luc Pellerin (Dpt Physiology, Uni. Lausanne), who works on neuroenergetics with a focus on lactate monocarboxylate transporters in the brain. I have been working on the mathematical aspects as a side project together with Dr Robert Costalat (Inserm, Uni. Paris) who is a long-time collaborator of Prof. Pellerin and who joined our effort soon after this project started. I will continue this project in Prof. Pellerin's group after my thesis, in order to finish wrapping up the results into a manuscript.

In the following, I present the main results of our current research. For this part however, I can hardly disentangle Dr Costalat contributions from mine, because we worked in a sequential manner: after a few months working on the project, Prof. Pellerin suggested that I update Dr Costalat on our results. So I sent him a progress report with my simulations and code, raising potential pitfalls and questions; he took this over, ran new simulations, suggested modifications and brought new questions; and this sequential exchange has continued until today, sharing countless e-mails, pieces of code and ideas! Prof. Pellerin followed all our exchanges, giving us significant biological input on the questions raised by our simulations and proposing new directions to investigate.

V.2.1 The astrocyte-neuron lactate shuttle model

Glucose has been considered the exclusive energy substrate of brain cells for a long time. Both neurons and astrocytes were thought to import glucose and oxidize it to sustain their energy needs. This view has however been challenged by the astrocyte-neuron lactate shuttle (ANLS) model (Pellerin and Magistretti, 1994; Pellerin and Magistretti, 2004). It proposes lactate as another important energy substrate for brain cells, together with glucose. In this model, astrocytes would mainly import glucose to convert it into lactate, producing enough ATP to sustain their activity. The produced lactate would be exported outside the cell and imported by the neurons to fuel oxidative phosphorylation, producing enough energy to sustain the demanding neuronal activity.

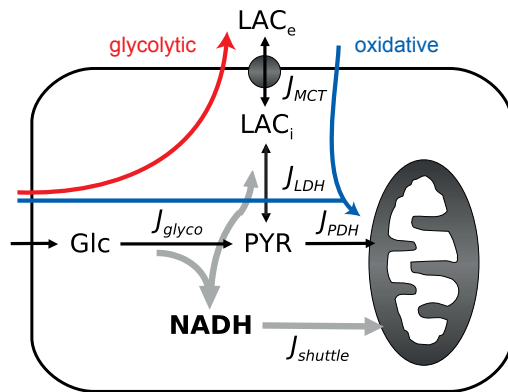### V.2.2   Defining brain cells metabolic profiles

Figure V.4: Model describing the simplified energetics of brain cells. *Glucose undergoes glycolysis and all the resulting pyruvate is either further metabolized by PDH or converted to lactate by LDH before being transported out of the cell. On the other hand, lactate can be transported into the cell and then metabolized into pyruvate by LDH. Because these processes require $NAD^+/NADH$, we also model the "recycling" shuttle of NADH to NAD^+ by mitochondria. Abbreviations: LAC_e, extracellular lactate concentration; LAC_i, intracellular lactate concentration; PYR, pyruvate concentration; NADH, NADH concentration; Glc, glucose concentration; J_{MCT}, transmembrane flux of lactate via MCTs; $J_{shuttle}$, flux of NADH to NAD^+ "recycling" by the mitochondria; J_{glyco}, glycolytic flux; J_{LDH}, metabolic flux via LDH; J_{PDH}, metabolic flux via PDH.*

Although all brain cells bear in principle a comparable capacity in terms of energetics, in reality they exhibit quite different metabolic profiles. The specific biochemical characteristics explaining such disparities and their relative importance are largely unknown. Figure V.4 shows a typical brain cell. Glucose (*Glc*) in the extracellular space can be taken up by the cell *via GLUT* transporters and metabolized into pyruvate (*PYR*) during glycolysis. This reaction requires NAD^+ which is converted into *NADH*. For glycolysis to proceed, the excess *NADH* should be "recycled" back into NAD^+. One way to do this is to convert pyruvate into lactate (LAC_i), a reversible reaction catalyzed by the lactate dehydrogenase (*LDH*). The produced lactate can then be transported outside the cell via the *MCT* transporters. The consumption of glucose for lactate production mainly is considered a *glycolytic* phenotype (red arrow in Figure V.4).

Another alternative for the cell is to use the mitochondrial shuttle to convert *NADH* into NAD^+ in order to proceed with glycolysis. This can happen provided that the cell has enough mitochondrial activity. In this scenario, the pyruvate produced by glycolysis can be further metabolized by the pyruvate dehydrogenase (*PDH*) into acetyl-coA and produce energy *via* oxidative phosphorylation in the mitochondria. And because *NADH* is recycled into NAD^+ by the mitochondrial shuttle, lactate can be imported into the cell and the *LDH* can perform the reverse reaction, converting it into pyruvate to feed oxidative phosphorylation as well. Such a phenotype is called *oxidative* (blue arrow in Figure V.4).

Using a modeling approach, we investigated under which conditions a cell is more oxidative (i.e. consumes glucose *and* lactate) or glycolytic (i.e. consumes glucose but exports lactate). We showed that modifying the kinetic parameters of the *PDH* and mitochondrial *NADH* shuttling within a realistic interval can yield to a striking switch in lactate flux direction. In this context, cells having a predominantly oxidative profile (exhibiting pronounced extracellular lactate uptake and consumption) could be turned into prominently glycolytic cells essentially releasing lactate by selectively reducing the aforementioned parameters. Importantly, our simulations show that these metabolic profiles are preserved upon stimulation of the cell. In the brain, cells only switch metabolic profile under extreme conditions, like hypoxia where an oxidative cell has no choice but become glycolytic. As we will see though, brain cells have a preferred metabolic profile under normal conditions, which seems to be oxidative for neurons and glycolytic for astrocytes (Pellerin, 2008). Indeed, astrocytes express the *LDH5* isoform while neurons express preferentially the *LDH1* iso-

form which facilitates the lactate conversion into pyruvate (Bittar et al., 1996). And in contrast to astrocytes, glucose uptake in neurons is reduced by glutamate (Loaiza et al., 2003; Porras et al., 2004). Astrocytes, which require much less energy than neurons, could therefore consume glucose and convert it into lactate, while neurons would import this lactate together with glucose to fuel oxidative phosphorylation (ANLS model (Pellerin and Magistretti, 2004)).

Thus, in the case of an oxidative cell, what is the importance of lactate vs glucose consumption in providing pyruvate to sustain oxidative phosphorylation? In order to test this, we fixed glycolysis to its basal level and showed that changes in lactate transport capacity and extracellular lactate concentration within the range described experimentally are sufficient to account for the described enhancement in oxidative metabolism upon activation.

V.2.3   Modeling the dynamics of glucose and lactate

We present a model that describes the dynamics of lactate and glucose in brain cells (cf. Figure V.4). As we will show, this model is very general and can be applied to both oxidative and glycolytic cells, by modifying some biochemical conditions.

In this model, extracellular lactate is taken up (or released) by brain cells via *MCT*s and further metabolized by *LDH* into pyruvate (or produced *via* this reaction). The intracellular lactate dynamics are therefore described by the following equation:

$$\frac{dLAC_i}{dt} = J_{MCT}(t) - J_{LAC_i,LDH}(t)$$

where J_{MCT} is the net transport of lactate inside the cell and $J_{LAC_i,LDH}$ the net consumption of intracellular lactate by the reaction catalyzed by *LDH*.

In parallel to the *LDH* reaction, glucose uptake and glycolysis contribute to the balance of pyruvate, which is further metabolized by *PDH* into acetyl-coA:

$$\frac{dPYR}{dt} = J_{PYR,LDH}(t) + J_{glyco}(t) - J_{PDH}(t)$$

where $J_{PYR,LDH}$ and J_{glyco} reflect the production of pyruvate by *LDH* and glycolysis, respectively, and J_{PDH} is the consumption of pyruvate by the reaction catalyzed by *PDH*. The latter reaction produces acetyl-coA which is used to fuel oxidative phosphorylation in the mitochondria.

Another important feature of our model is the role of *NADH*. Indeed, both the forward-*LDH* reaction ($LAC_i \rightarrow PYR$) and glycolysis produce *NADH* (from NAD^+), while it is "recycled" back into NAD^+ by the mitochondrial shuttle:

$$\frac{dNADH}{dt} = J_{NADH,LDH}(t) - J_{shuttle}(t) + J_{glyco}(t)$$

where $J_{NADH,LDH}$ and J_{glyco} are the net fluxes of *NADH* production by the forward-*LDH* reaction and the glycolysis, respectively, and $J_{shuttle} = k_{shuttle} \cdot NADH$ is the flux of mitochondrial shuttling. The mitochondrial shuttle therefore plays an essential role in determining the direction of the *LDH* reaction. (See

Subsection V.2.3.1 below for a detailed description of the model, since the equations above only describe a few components of the system, omitting among others all the bound states).

V.2.3.1 Detailed model

We consider a model where extracellular lactate LAC_e is taken up by the brain cells via the MCTs and further metabolized by the LDH and PDH into pyruvate PYR and acetyl-coA A, respectively (cf. Figure V.4). In parallel, glucose uptake and metabolism via the PDH also contribute to the production of acetyl-coA which fuels oxidative phosphorylation in the mitochondria. Note that LAC_i can only bind to LDH once the latter has been activated by NAD^+ (complex $[LDH - NAD^+]$); similarly, PYR can only bind to LDH once the latter has been activated by $NADH$ (complex $[LDH - NADH]$). We describe these processes with a system of coupled differential equations:

$$\frac{dLAC_e}{dt} = 0$$

$$\frac{dNAD^+}{dt} = k_{shuttle} \cdot NADH - l_1 \cdot NAD^+ \cdot LDH + l_2 \cdot [LDH - NAD^+] - J_{glyco}$$

$$\frac{d[LDH-NAD^+]}{dt} = l_1 \cdot LDH \cdot NAD^+ - l_2 \cdot [LDH - NAD^+] - l_3 \cdot LAC_i \cdot [LDH - NAD^+] + l_4 \cdot [LAC_i - LDH - NAD^+]$$

$$\frac{dLAC_i}{dt} = J_{MCT} - l_3 \cdot LAC_i \cdot [LDH - NAD^+] + l_4 \cdot [LAC_i - LDH - NAD^+]$$

$$\frac{d[LAC_i-LDH-NAD^+]}{dt} = l_3 \cdot LAC_i \cdot [LDH - NAD^+] - (l_4 + l_5) \cdot [LAC_i - LDH - NAD^+] + l_6 \cdot [PYR - LDH - NADH]$$

$$\frac{d[PYR-LDH-NADH]}{dt} = l_5 \cdot [LAC_i - LDH - NAD^+] - (l_6 + l_7) \cdot [PYR - LDH - NADH] + l_8 \cdot PYR \cdot [LDH - NADH] \cdot H_i^+$$

$$\frac{dPYR}{dt} = l_7 \cdot [PYR - LDH - NADH] - l_8 \cdot PYR \cdot [LDH - NADH] \cdot H_i^+ + J_{glyco} - p_1 \cdot PYR \cdot PDH + p_2 \cdot [PYR - PDH]$$

$$\frac{dNADH}{dt} = l_9 \cdot [LDH - NADH] - l_{10} \cdot LDH \cdot NADH - k_{shuttle} \cdot NADH + J_{glyco}$$

$$\frac{d[LDH-NADH]}{dt} = l_7 \cdot [PYR - LDH - NADH] - l_8 \cdot PYR \cdot [LDH - NADH] \cdot H_i^+ - l_9 \cdot [LDH - NADH] + l_{10} \cdot LDH \cdot NADH$$

$$\frac{d[PYR-PDH]}{dt} = p_1 \cdot PYR \cdot PDH - (p_2 + p_3) \cdot [PYR - PDH]$$

$$\frac{dA}{dt} = p_3 \cdot [PYR - PDH]$$

with $J_{MCT} = v_{max,MCT} \left(\frac{LAC_e H_e^+}{K_{Hie} + LAC_e H_e^+} - \frac{LAC_i H_i^+}{K_{Hie} + LAC_i H_i^+} \right)$ (Aubert et al., 2005),
$J_{glyco} = v_{max,glyco} \left(\frac{glc}{K_j + glc} \right) \cdot \left(\frac{NAD^+}{K_n + NAD^+} \right)$ and $J_{PDH} = \frac{dA}{dt}$. We are interested in stationary states.

V.2.3.2 Role of transport to fuel oxidative phosphorylation

We note that at steady state, the baseline values (subscript 0) obey to

$$J_{PDH,0} = J_{glyco,0} + J_{MCT,0}.$$

Let us define $r \doteq J_{MCT,0}/J_{glyco,0}$. If we assume glycolysis to be constant, we have upon activation:

$$(1 + \alpha) \cdot J_{PDH,0} = J_{glyco,0} + (1 + \gamma) \cdot J_{MCT,0}$$

where α is the increase in acetyl-coA production and γ the increase in transport flux, yielding:

$$\alpha = \frac{\gamma r}{1 + r}. \tag{V.1}$$

Thus, given the ratio r at basal state, we can determine what will be the increase in oxidative phosphorylation upon transport activation (in the case where glycolysis is fixed).

V.2.3.3 Choice of parameters

Some of the parameters in our equations, or a combination of them, have been measured or are known to lie within some range (Aubert et al., 2005; O'Brien et al., 2007; Seifert et al., 2006; Barros et al., 2005; Qiu et al., 2004; Itoh et al., 2003). The unknown parameters were set as to satisfy two physiological constraints: the ratio $r = J_{MCT}/J_{glyco} \in [0.75; 2.2]$ (Nehlig et al., 2004; Hyder et al., 2006; Bouzier-Sore et al., 2003; Bouzier-Sore et al., 2006) and the ratio $LAC_i/PYR \in [5; 40]$. In our simulations, we generally fixed $r = 0.75$ or $r = 2.2$ (see figure legends) and $LAC_{i,0}/PYR_0 \simeq 10$. We list below the parameters that were used to compute the basal state of a typical oxidative cell (in Subsection V.2.4, we discuss how these parameters can be modified to switch to a glycolytic-like cell). All the units are in mM, s or mM/s.

LDH metabolism					
l_1	10				
l_2	10				

MCT transport		l_3	0.58				
				PDH metabolism		**Glycolysis**	
$v_{max,MCT}$	0.45	l_4	1.1	p_1	10	$v_{max,glyco}$	0.15
LAC_e	1.2	l_5	0.08			**Mitoch. shuttle**	Glc	1
H_i^+	$10^{-4.1}$	l_6	0.7	p_2	0.02	$k_{shuttle}$	0.38
H_e^+	$10^{-4.3}$	l_7	1	p_3	0.05	K_j	5
K_{Hie}	0.7	l_8	34	PDH_{total}	0.9	K_n	0.001
		l_9	10				
		l_{10}	10				
		LDH_{total}	5				

V.2.4 Defining metabolic phenotypes (oxidative vs glycolytic)

In order to better understand the importance of distinct energy substrate sources for brain cell energetics, we considered a simple model taking into account glucose and lactate supply as well as their metabolism (cf. Figure V.4 and Section V.2.3). We used this model to determine the changes in lactate flux across the plasma membrane as well as in the glycolytic and pyruvate dehydrogenase-catalyzed fluxes, when switching conditions from a basal to a stimulated state. For the basal state, the contribution of lactate-derived and glycolysis-derived pyruvate was set to 43% and 57%, respectively (or $J_{MCT}/J_{glyco} = r = 0.75$), in accordance with the data of Nehlig and collaborators (Nehlig et al., 2004). For the stimulated state, we allowed for a 30% increase of the *PDH* maximal reaction rate $v_{max,PDH}$ and of the mitochondrial *NADH* shuttling rate $k_{shuttle}$, with a rather strong increase of 48.5% in the maximal glycolytic rate $v_{max,glyco}$ as well. In
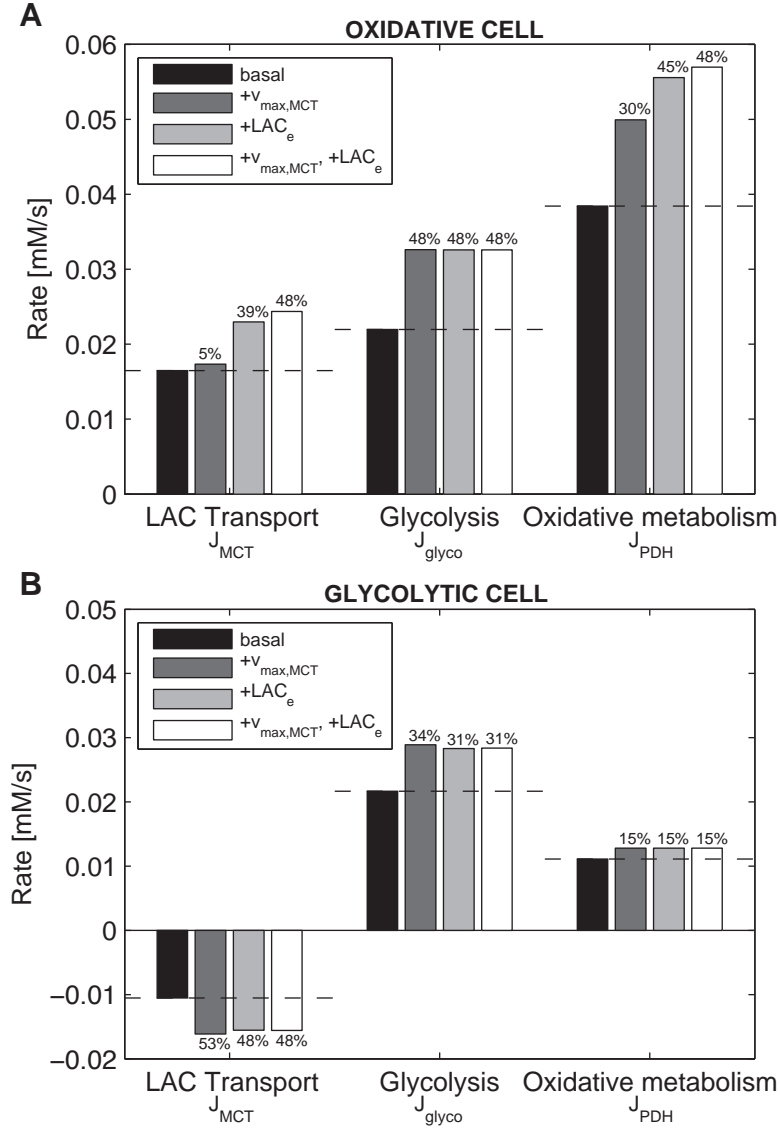
Figure V.5: $k_{shuttle}$ and $v_{max,PDH}$ determine the occurrence of oxidative vs. glycolytic phenotype. **(A) Oxidative phenotype:** In the basal state (black), both glycolysis and lactate transport contribute to J_{PDH} (57% and 43% respectively, (Nehlig et al., 2004)). We show the resulting transport, glycolytic and J_{PDH} fluxes upon stimulation (dark gray: +80% $v_{max,MCT}$, +30% $v_{max,PDH}$, +30% $k_{shuttle}$, +48.5% $v_{max,glyco}$; light gray: +80% LAC_e, +30% $v_{max,PDH}$, +30% $k_{shuttle}$, +48.5% $v_{max,glyco}$; white: +80% $v_{max,MCT}$, +80% LAC_e, +30% $v_{max,PDH}$, +30% $k_{shuttle}$, +48.5% $v_{max,glyco}$). **(B) Glycolytic phenotype:** In the basal state (black), parameters are chosen such that lactate is taken out of the cell ($v_{max,PDH}^{glycolytic} = 0.25 \cdot v_{max,PDH}^{oxidative}$, $k_{shuttle}^{glycolytic} = 0.07 \cdot k_{shuttle}^{oxidative}$). We show the resulting transport, glycolytic and J_{PDH} fluxes upon stimulation (dark gray: +80% $v_{max,MCT}$, +15% $v_{max,PDH}$, +0% $k_{shuttle}$, +48.5% $J_{max,glyco}$; light gray: +80% LAC_e, +15% $v_{max,PDH}$, +0% $k_{shuttle}$, +48.5% $v_{max,glyco}$; white: +80% LAC_e, +80% $v_{max,MCT}$, +15% $v_{max,PDH}$, +0% $k_{shuttle}$, +48.5% $v_{max,glyco}$). Cf. Section V.2.3 for model equations and parameters of a typical oxidative cell.

addition, three scenarios were considered for lactate transport stimulation : 1) An increase in lactate transport rate of 80% based on the data of Pierre et al. (Pierre et al., 2009), due to *MCT*s translocation to the cell membrane 2) An increase in the extracellular lactate concentration LAC_e of 80% as reported by Hu and Wilson (Hu and Wilson, 1997) 3) A concomitant increase in $v_{max,MCT}$ and LAC_e of the same magnitude. Simulations allowed to determine the resulting fluxes for lactate transport J_{MCT}, glycolysis J_{glyco}, and oxidative metabolism which we assume proportional to J_{PDH}. As can be observed in Figure V.5A, a significant enhancement of 30% in oxidative metabolism is obtained when increasing lactate transport capacity, $v_{max,MCT}$, in a stimulated cell. Note however that the increase in lactate flux remains modest (+5%) in this condition. If we instead increase the extracellular lactate concentration LAC_e, oxidative metabolism is further enhanced (+45%), with a significant increase in lactate flux (+39%). A combination of enhanced lactate transport capacity and extracellular lactate concentration leads to a 48% increase in oxidative metabolism with a concomitant 48% increase in lactate flux. It is noteworthy that in this last case, the contribution of each metabolic pathway has remained the same as in the basal state ($r = 0.75$). Based on these metabolic responses, such a behavior appears characteristic of oxidative cells.

It is also known that a certain fraction of brain cells do not respond to a stimulation by importing lactate but rather by producing and releasing lactate. In order to reproduce this specific behavior, some of the parameters had to be modified. We found that reducing both the initial *PDH* reaction rate and the activity of the mitochondrial *NADH* shuttle leads to a striking different type of response. Figure V.5B presents the result of simulations performed when the basal values for these parameters were set to $0.25 \cdot v_{max,PDH}$ and $0.07 \cdot k_{shuttle}$ as compared to the simulations in Figure V.5A. Moreover, in the stimulated cell, the increase in *PDH* activity was limited to 15% and no increase in mitochondrial *NADH* shuttle was allowed (instead of +30% for both parameters in Figure V.5A). In such case, a stimulation causing an activation of glycolysis and lactate transport produced a more limited increase in oxidative metabolism (+15%), but importantly yielded a prominent release of lactate (-53%), as indicated by the negative sign of the flux. Increasing the extracellular lactate concentration had no further impact on oxidative metabolism and only a marginal effect on lactate efflux and on the glycolytic flux. In this case, the observed metabolic behavior rather corresponds to a typically predominantly glycolytic cell.

V.2.5   Importance of lactate metabolism in sustaining oxidative metabolism

We have seen in the previous simulations that lactate metabolism can significantly contribute to an enhancement in oxidative metabolism in cells with a predominantly oxidative phenotype. In these simulations, we also considered the presence of a strong glycolytic component in parallel with lactate metabolism. However, evidence from the literature suggests that this is not the case in certain brain cell types (Herrero-Mendez et al., 2009). Thus, we investigated the capacity of lactate metabolism to sustain the enhanced oxidative metabolism typically observed in oxidative brain cells under fixed glycolytic conditions. We studied four scenarios combining two lactate/glycolysis-supplied pyruvate ratios, $J_{MCT}/J_{glyco} = r$, and two distinct levels of oxidative rate, $v_{max,PDH}$: $r = 0.75$, +30% $v_{max,PDH}$ (Figure V.6A); $r = 0.75$, +70% $v_{max,PDH}$ (Figure V.6B); $r = 2.2$, +30% $v_{max,PDH}$ (Figure V.6C); and $r = 2.2$, +70% $v_{max,PDH}$ (Figure V.6D). In each case, a series of incremental oxidative fluxes were delineated, allowing to determine the combination of increased lactate transport and extracellular lactate concentration that can sustain each oxidative flux. Non-shaded
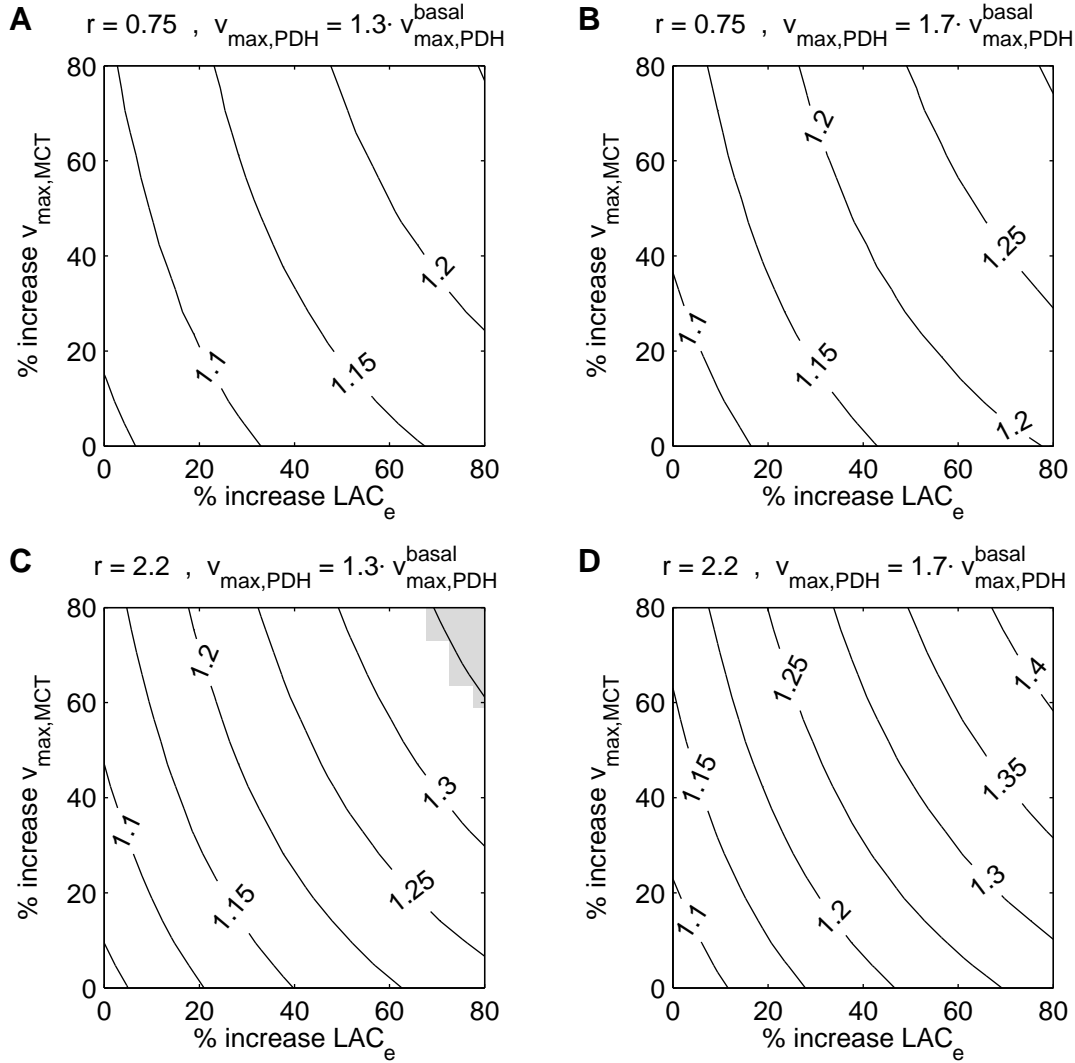
Figure V.6: Importance of lactate transport for oxidative phosphorylation in oxidative cells. *(A) As a basal point, it is assumed that lactate transport and glycolysis contribute 43% and 57%, respectively, to oxidative phosphorylation ($r = 0.75$; (Nehlig et al., 2004)). Lactate transport is then stimulated by increasing $v_{max,MCT}$ and LAC_e, so that oxidative phosphorylation (that we assume proportional to J_{PDH}) is multiplied by a factor α (iso-curves). In parallel, $v_{max,PDH}$ is increased by 30%. Note that the glycolytic flux remains fixed to its basal level and that the NADH "recycling" rate, $k_{shuttle}$, is also multiplied by α. Non-shaded areas represent physiologically plausible parameter regions where the ratio $r \in [5;40]$. (B) As in (A), but $v_{max,PDH}$ is increased by 70%. (C) As in (A), but $r = 2.2$ (Hyder et al., 2006). (D) As in (C), but $v_{max,PDH}$ is increased by 70%. Cf. Section V.2.3 for model equations and parameters numerical values.*

. . . .

$r \setminus \gamma$	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8
0.1	0.0091	0.018	0.027	0.036	0.046	0.055	0.064	0.073
0.4	0.029	0.057	0.086	0.11	0.14	0.17	0.20	0.23
0.7	0.041	0.082	0.12	0.16	0.21	0.25	0.29	0.33
1.0	0.050	0.10	0.15	0.20	0.25	0.30	0.35	0.40
1.3	0.057	0.11	0.17	0.23	0.28	0.34	0.40	0.45
1.6	0.062	0.12	0.18	0.25	0.31	0.37	0.43	0.49
1.9	0.066	0.13	0.20	0.26	0.33	0.39	0.46	0.52
2.2	0.069	0.14	0.21	0.28	0.34	0.41	0.48	0.55
2.5	0.071	0.14	0.21	0.29	0.36	0.43	0.50	0.57
2.8	0.074	0.15	0.22	0.29	0.37	0.44	0.52	0.59
3.1	0.076	0.15	0.23	0.30	0.38	0.45	0.53	0.60
3.4	0.077	0.15	0.23	0.31	0.39	0.46	0.54	0.62

Table 1: Increase of oxidative metabolism J_{PDH} obtained with distinct ratios r of glycolysis/lactate-derived pyruvate and different increases in lactate transport flux γ. *White: below 10%; light gray: between 10% and 29%; dark gray: above 29%.*

. . . .

areas represent plausible parameter regions for which the intracellular lactate/pyruvate ratio remains within physiological limits ([5;40]). As can be seen, a rather wide range of increased oxidative rates can be accounted for by only altering lactate transport and extracellular lactate concentration.

In order to better delineate under which range of basal conditions a change in lactate metabolism alone can satisfy physiological increases in oxidative rate, we varied the lactate/glycolysis-supplied pyruvate ratio r between extreme cases (0.1 to 3.4 with increments of 0.3) together with increases γ in lactate flux of up to 80% (see Equation V.1 and Table 1). As can be observed, increases in oxidative rate J_{PDH} of up to 10% (white), between 10% and 29% (light gray), or even above 29% (dark gray) can be easily accounted by physiological increases in lactate flux for a wide range of initial conditions which encompasses most published values (Nehlig et al. 2004; Hyder et al. 2006; Bouzier-Sore et al. 2006; see also below).

V.2.6 Assessing the validity of the model with experimental data

It is important to assess the validity of the model by testing whether it can account for published data. First, a J_{MCT}/J_{glyco} ratio $r = 0.69/0.31 = 2.2$ as determined in vivo by (Hyder et al., 2006) was modeled with plausible flux rates for lactate transport J_{MCT}, glycolysis J_{glyco} and oxidative metabolism J_{PDH} (Figure V.7A in black). Of interest, it gave rise to physiological intracellular lactate and pyruvate concentrations of 0.39mM and 0.024mM respectively, resulting in a LAC_i/PYR ratio of 16.6. When considering *in vitro* data of (Bouzier-Sore et al., 2006) in cultured neurons, we used a ratio $J_{MCT}/J_{glyco} = r = 0.76/0.24 = 3.2$ with an extracellular lactate concentration of 1.1mM as in the experimental condition (Figure V.7A, dark gray). Note that compared to *in vivo* data, a reduction in glycolysis flux of 29% and a decrease of oxidative flux of 9% was necessary to account for the data. Interestingly, increasing the extracellular lactate concentration to 5.5mM, as used in (Bouzier-Sore et al., 2003), led to an improbable intracellular pyruvate concentration of 88mM (Figure 4A, light gray). However, it was possible to account for the data if the glycolytic flux was reduced by 67% (Figure V.7A, white). In such case, we obtained a ratio $r = 0.92/0.08 = 11.5$, which
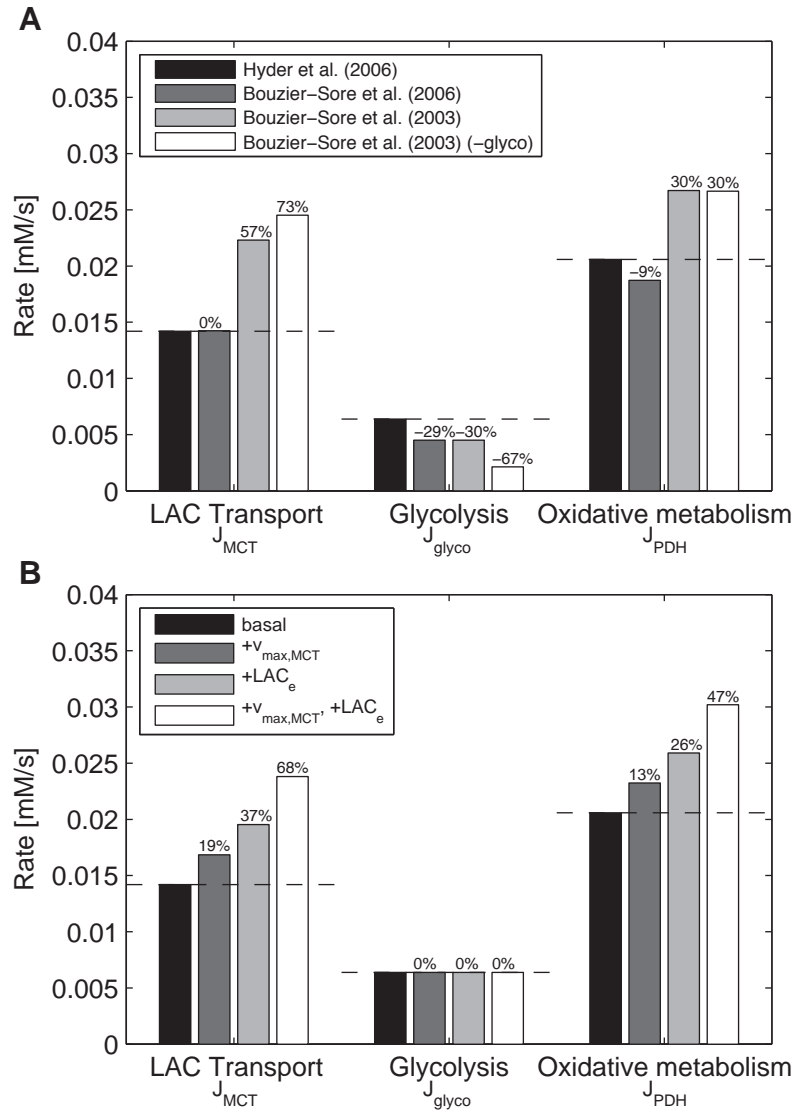
Figure V.7: Model validity tested for different experimental conditions. *(A) Black: $r = J_{MCT}/J_{glyco} = 0.69/0.31 = 2.2$, as evaluated in vivo by (Hyder et al., 2006), with physiological values for intracellular lactate (LAC_i) and pyruvate (PYR) concentrations (0.39 and 0.024 mM, respectively). Dark gray: $r = J_{MCT}/J_{glyco} = 0.76/0.24 = 3.2$, as observed in vitro by (Bouzier-Sore et al., 2006), with the experimentally used $LAC_e = 1.1mM$. Light gray: LAC_e is increased to 5.5mM, as described in the experiment by (Bouzier-Sore et al., 2003), resulting in a non-plausible value of 88mM for intracellular pyruvate. White: Same conditions as in light gray, but the glycolytic rate was lowered by 67%, resulting in a more likely intracellular pyruvate concentration of 9.0mM and a J_{MCT}/J_{glyco} ratio equal to $0.92/0.08 = 11.5$, which matches experimental results by (Bouzier-Sore et al., 2003). (B) Effect of lactate transport enhancement in the case of a $r = J_{MCT}/J_{glyco} = 0.69/0.31 = 2.2$, as evaluated in vivo by (Hyder et al., 2006). Basal state: $LAC_e = 1.1mM$ (black). Stimulations: +0% LAC_e, +80% $v_{max,MCT}$, +70% $v_{max,PDH}$ (dark gray); +100% LAC_e, +0% $v_{max,MCT}$, +70% $v_{max,PDH}$ (light gray); +100% LAC_e, +80% $v_{max,MCT}$, +70% $v_{max,PDH}$ (white).*
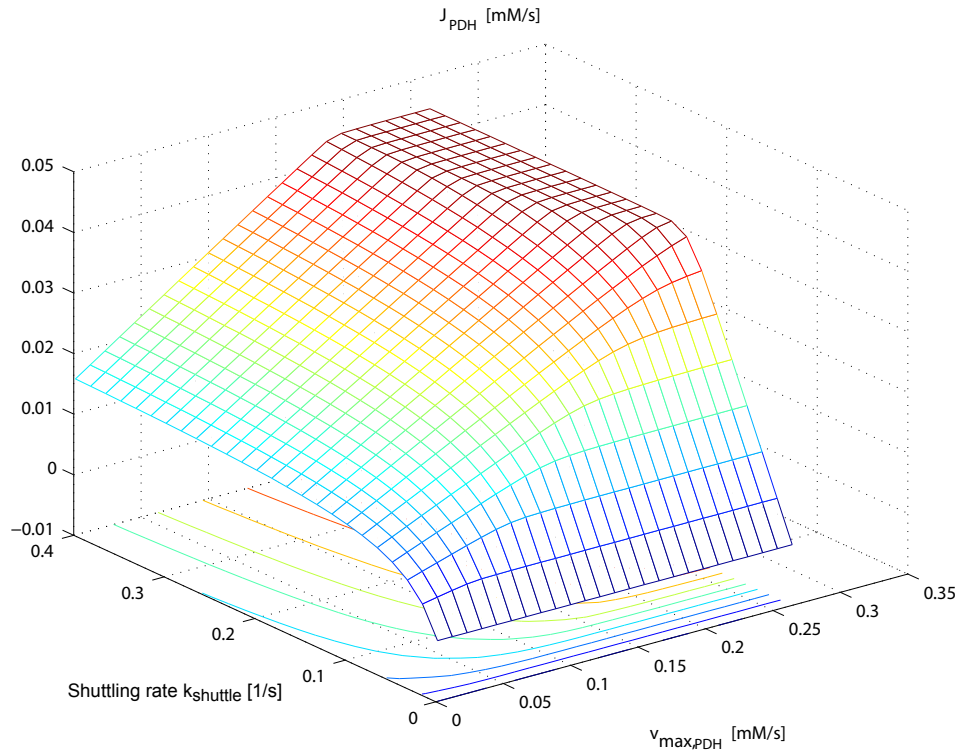
Figure V.8: PDH-metabolic flux as a function of the mitochondrial NADH shuttling rate $k_{shuttle}$ and the maximal speed of the PDH reaction $v_{max,PDH}$, which is proportional to the total amount of active PDH.

matches the experimental results. In the last two conditions, the oxidative flux was increased by 30% as compared to the *in vivo* situation.

Finally, we decided to assess under conditions close to those observed *in vivo* whether lactate metabolism can entirely account for the enhancement in oxidative rates, known to occur following brain activation. Changes in lactate, glycolytic and oxidative fluxes following a stimulation were determined with the model using the experimental *in vivo* data of (Hyder et al., 2006), i.e. a J_{MCT}/J_{glyco} ratio $r = 0.69/0.31 = 2.2$ and a resting extracellular lactate concentration of 1.1mM (Figure V.7B). Three stimulated conditions were considered: increased lactate transport rate of 80% (Figure V.7B, dark gray), increased extracellular lactate concentration of 100% (Figure V.7B, light gray), and a combination of both (Figure V.7B, white). An increase in oxidative capacity of 70% was applied in parallel for all three stimulated conditions, as proposed by (Hyder et al., 2006). As can be seen, the increase in lactate transport rate or the extracellular lactate concentration led to an increase in oxidative flux of 13% and 26% respectively. But even more strikingly, a combination of the two parameters gave rise to a significant 47% increase in oxidative rate.

V.2.7   Robustness of the model

Since various parameters in our model were arbitrarily set (provided some constraints were met, cf. Section V.2.3), we investigate here the robustness of our results when these parameters are varied over a wide range, ensuring they still satisfy those physiological constraints.

LDH_{total} [mM]	PDH_{total} [mM]	$v_{max,glyco}$ [mM/s]
0.005	5.1	0.75
0.05	4.9	0.74
0.1	4.8	0.72
0.5	3.7	0.58
1	2.8	0.47
5	0.9	0.20
10	0.5	0.13

Table 2: Parameters used when varying the total amount of *LDH. The total amount of PDH and maximum speed of glycolysis* $v_{max,glyco}$ *were also tuned in order to preserve the constraints* $r = 0.75$ *and* $LAC_i/PYR \simeq 10$.

....

V.2.7.1   Effect of mitochondrial *NADH* shuttle vs. *PDH* parameters

The mitochondrial shuttle has been shown to be mainly active in neurons and a recent study (Halim et al., 2010) shows that total amount of active *PDH* is also higher in these cells as compared to astrocytes. We thus want to investigate how these two processes contribute to oxidative phosphorylation and modulate it.

As shown in Figure V.8, an increase in the mitochondrial shuttling rate yields an increase in oxidative phosphorylation (which we assume proportional to J_{PDH}). Interestingly, small changes in the mitochondrial shuttling rate induce a greater oxidative response than small changes in the total amount of active *PDH* (proportional to v_{maxPDH}). However, the effect due to an increase in mitochondrial shuttle also saturates faster. It therefore appears that modulating the mitochondrial shuttling rate is an efficient means to increase oxidative phosphorylation upon high demand, while modulating the total amount of active *PDH* determines the maximal oxidative capacity of the cell (i.e. the saturation level of J_{PDH} when changing the mitochondrial shuttling rate). This suggests that oxidative cells should have higher levels of active *PDH* than their glycolytic counterparts, as was indeed recently observed (Halim et al., 2010).

V.2.7.2   Effect of LDH parameters

We want to investigate the robustness of the model to changes in *LDH*-related parameters, since these were kept constant for our simulations in the main text. There, we considered an *LDH* concentration of the order of the mM, corresponding to about 35g/L (*LDH* molecular weight was taken as ~35000Da). This high concentration of *LDH* is most probably not uniform, but found locally in the cell, in regions where the protein might be confined. This is in agreement with the idea of "enzyme organization": cells are not homogeneous entities containing metabolites and enzymes but rather, enzymes are (i) confined to organelles, (ii) bound to membranes, (iii) bound to the cytoskeleton or (iv) associated to other proteins for metabolic channeling (Mendes et al., 1995). Therefore, though measured *in vitro* enzyme concentrations are often of the order of the μM for metabolite concentrations of the order of the mM, *in vivo* enzyme concentrations might actually be of the order of the mM locally, just like metabolite concentrations.

However, it is not yet clear if "enzyme organization" indeed applies to the *LDH* or if it is distributed (more or less) uniformly into the cell. In the muscle, some data suggest that the *LDH* might be localized in
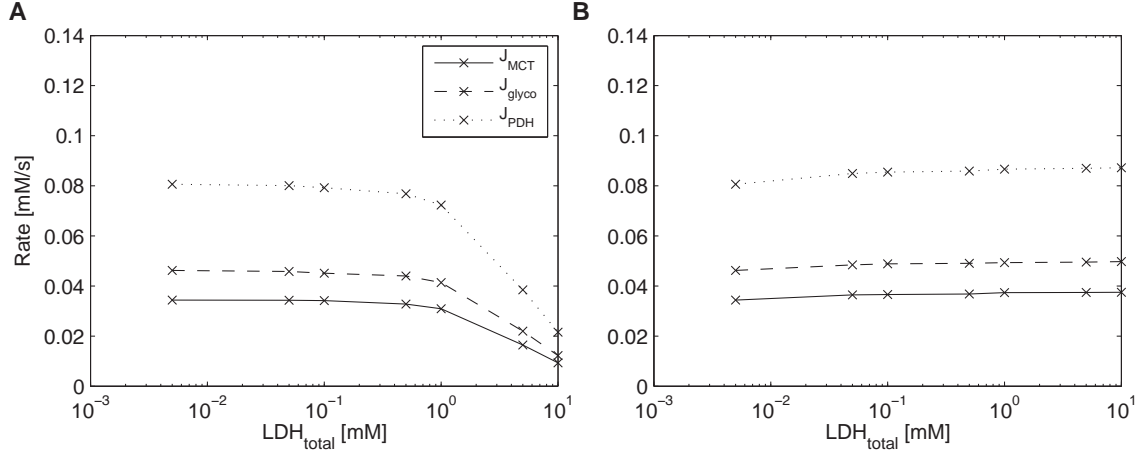
Figure V.9: Effect of the total *LDH* pool on transport, glycolysis and *PDH*-metabolic flux. *(A) Parameters are tuned to ensure that $r = 0.75$ and $LAC_i/PYR \simeq 10$, cf. Table 2. (B) Same as (A), but the mitochondrial shuttling rate $k_{shuttle}$ and the total pool of $NAD^+ + NADH$ are also tuned to get a monotonic increase of the fluxes as a function of LDH_{total}.*

"subcellular particulate fractions" (Nitisewojo and Hultin, 1976), but no clear evidence has yet been found in the brain to our knowledge. Since in vitro studies report the maximal rate normalized per mg protein and that, to our knowledge, in vivo measurements of the LDH pool have not been reported, we consider in our simulations several pools of total *LDH*, ranging from 0.5μM to 10mM. When the total *LDH* pool is divided by a factor *x*, all the *LDH* kinetic parameters are multiplied by *x*, to preserve $v_{max,LDH}$ and the Michaelis-Menten constant. Also, in order to preserve physiological conditions, we impose that the ratio of transport to glycolysis $r = 0.75$ and that the ratio of intracellular lactate to pyruvate $LAC_i/PYR \simeq 10$, like in our previous simulations. Thus, when we modify the total amount of *LDH*, we also need to tune the total amount of *PDH* and the glycolytic flux J_{glyco} to satisfy these constraints (cf. Table 2).

In Figure V.9A, we see that despite changing the total amount of *LDH* by a factor 2000, the mitochondrial flux J_{PDH} only varies by a factor ~4, showing great robustness. Interestingly, it appears that increasing the total amount of *LDH* disfavors lactate transport into the cell, glycolysis and *PDH*-metabolic flux, due to the limited capacity of the cell to "recycle" *NADH* into NAD^+ and the fact that a significant pool of *LDH* does not get activated by NAD^+ to participate in the forward *LDH* reaction, since the available pool of NAD^+ is too small (~80% of free (i.e. not-activated) *LDH* in the case $LDH_{total} = 10$mM). As a consequence, the *LDH* reaction slows down. It follows that the net transport inside the cell is reduced and thus, to preserve the transport to glycolysis ratio as well as the intracellular lactate to pyruvate ratio, both glycolysis and the *PDH*-metabolic flux decrease. As we show in Figure V.9B, increasing both the mitochondrial shuttling rate (up to a factor 100) and the total pool of $NAD^+ + NADH$ (up to a factor 5) yields a monotonic increase in all the fluxes as a function of total *LDH*.

Importantly too, as we see in Figure V.10, the model responds very similarly upon activation at low and high concentrations of *LDH* ($LDH_{total} = 0.5$μM, compare with Figure V.5A where $LDH_{total} = 5$mM). Thus, although the value of the *in vivo LDH* concentration remains debated, our study shows that if we preserve some physiological constraints (ratio LAC_i/PYR and transport over glycolysis ratio *r*), the cell ends up responding in a very similar fashion upon activation, almost independently of the total *LDH* pool.
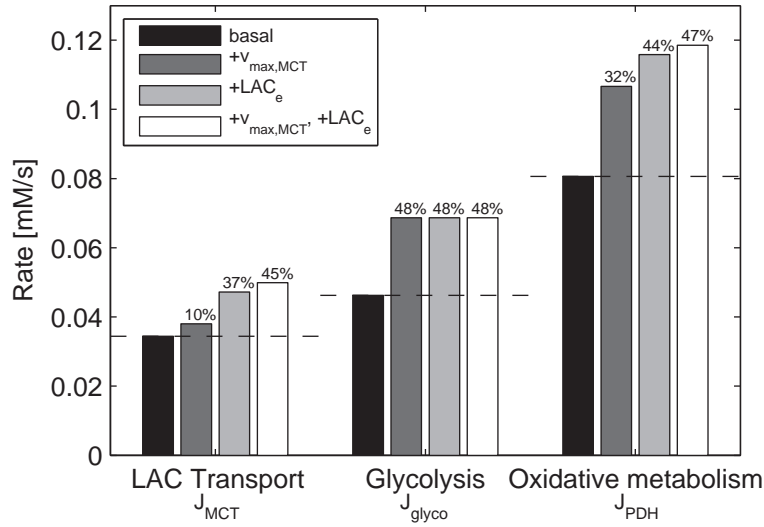
Figure V.10: Effect of LDH parameters upon stimulation. *In the basal state (black), both glycolysis and lactate transport contribute to J_{PDH} (57% and 43% respectively). We show the resulting transport, glycolytic and J_{PDH} fluxes upon stimulation (dark gray: +80% $v_{max,MCT}$, +30% $v_{max,PDH}$, +30% $k_{shuttle}$, +48.5% $v_{max,glyco}$; light gray: +80% LAC_e, +30% $v_{max,PDH}$, +30% $k_{shuttle}$, +48.5% $v_{max,glyco}$; white: +80% $v_{max,MCT}$, +80% LAC_e, +30% $v_{max,PDH}$, +30% $k_{shuttle}$, +48.5% $v_{max,glyco}$). $LDH_{total} = 0.5 \mu M$ (compare with Figure V.5A where $LDH_{total} = 5mM$).*

. . . .

V.2.8   Perspectives

This work provides key elements to understand why certain brain cell types constitutively adopt a particular metabolic profile and how specific features can be altered under different physiological and pathological conditions in order to face evolving energy demands. However, before we can submit it for publication, we need to further investigate two important issues. First, pH values can be very important for lactate transport since protons are co-transported by the *MCT*s. Thus, we want to assess the effect of intracellular and extracellular pH values on the transport, glycolytic and oxidative phosphorylation rates. Second, the equation for lactate transport that we used (Aubert et al., 2005) has been debated. We are planning to investigate the effect of *MCT*-related parameters as well as alternative equations for *MCT* transport to test the robustness of our findings.

Conclusion

....

A PhD thesis is with no doubt a great experience. Apart from the obvious technical and biological knowledge that I have gained, it has taught me a lot about myself and how to deal with ideas, people, frustration and criticism. I think I have become much more open to discussing my projects with people and accepting criticism as a constructive way to improve the results. I have also learned to better communicate my results and share knowledge. And most importantly, I realized how amazing Nature can be.

Retrospectively, I really enjoyed working on various different projects. Starting with the embryo was very helpful, because the system is not so complicated and the players are pretty well known. Moreover, since I did not participate in the design of the experiments, I could focus more on the modeling and image analysis side, which were quite new to me (during my Master, I had been working on structural biology, which involves modeling at the atomistic/electronic level using quantum mechanics). However, I realised with time, and in particular during the reviewing process of our manuscript (de Lachapelle and Bergmann, 2010a), that I should have understood better the biology and experimental setup of this study. Although the issues raised by the reviewers and (Jaeger, 2010) did not invalidate the results that we obtained (cf. Section III.8), it became clear to me that being part of the experimental design is essential when investigating a biological question, even for someone working on the modeling/data analysis level. Indeed, biological systems are noisy enough without adding any supplementary sources of imprecision.

When my supervisor suggested that I join the SystemsX.ch wingX consortium, I therefore spent several months reading applied biology literature in order to understand the system. I realised that the wing imaginal disc is by far less well understood than the early embryo development, and often came across contradicting publications, as discussed in Part IV. Since scaling issues were also being debated in the field, it felt natural to investigate this question in this new system. After spending a few months trying to model the system in order to reproduce results from the literature (Subsection IV.12), we set up a collaboration with Fisun Hamaratoglu from Prof. Affolter group. Being able to give feedback on the experimental design was a rewarding experience, though it took some time until Fisun and I agreed on the protocol for data collection, imaging, image analysis and quantification. Today, I feel confident about the results we obtained, because even though I did not perform the experiments myself, I could discuss them extensively with Fisun. And it was also very interesting to get feedback from a biologist on my image analysis, quantifications and modeling.

In the neuroenergetics project, I was also in close contact with a biologist, Prof. Pellerin, though we based our modeling on existing experiments. This experience also showed me that more than working on the experimental design, it is important to collaborate with people that have complementary skills and to share knowledge (both ways). After I started working with Prof. Pellerin, I therefore decided to invite 5-6 students/post-docs from the wingX consortium to a "crash course on modeling in biology (for biologists)", organised over 2.5 days in 2009 and repeated in 2010 upon demand. I really enjoyed discussing modeling issues with experimentalists and sharing points of view. I believe this was extrememly helpful when collaborating with Fisun on the wing project. This course was also the starting point to our collaboration with Gerald Schwank (cf. Section IV.12).

During my thesis, I tried to get more involved in the experiments to get a real "wet lab" experience. I had the opportunity to participate in a "hands on" summer school (within the FACEiT consortium) where I learned many techniques used in molecular biology. I also participated in a one week pratical course at Uni. Basel on experimental techniques in developmental biology. However, I never got to perform any experiment for my research projects, which I somehow miss. In the future, I will try to get even closer to the bench to understand better biology. I believe this will help me to better formulate research questions and envision experimental possibilities... *PhD was just the appetizer!*

References

Aegerter-Wilmsen, T., Aegerter, C., Hafen, E. and Basler, K. (2007). Model for the regulation of size in the wing imaginal disc of Drosophila. Mech Dev *124*, 318–326.

Aegerter-Wilmsen, T., Aegerter, C. M. and Bisseling, T. (2005). Model for the robust establishment of precise proportions in the early Drosophila embryo. J Theor Biol *234*, 13–19.

Affolter, M. and Basler, K. (2007). The Decapentaplegic morphogen gradient: from pattern formation to growth regulation. Nat Rev Genet *8*, 663–74.

Affolter, M. and Basler, K. (2011). Cell biology. Cytonemes show their colors. Science *332*, 312–3.

Aubert, A., Costalat, R., Magistretti, P. and Pellerin, L. (2005). Brain lactate kinetics: Modeling evidence for neuronal lactate uptake upon activation. Proc Natl Acad Sci U S A *102*, 16448–16453.

Aza-Blanc, P. and Kornberg, T. (1999). Ci: a complex transducer of the hedgehog signal. Trends Genet *15*, 458–462.

Barrio, R. and de Celis, J. F. (2004). Regulation of spalt expression in the Drosophila wing blade in response to the Decapentaplegic signaling pathway. Proceedings of the National Academy of Sciences of the United States of America *101*, 6021–6026.

Barros, L., Porras, O. and Bittner, C. (2005). Why glucose transport in the brain matters for PET. Trends Neurosci *28*, 117–119.

Belenkaya, T., Han, C., Yan, D., Opoka, R., Khodoun, M., Liu, H. and Lin, X. (2004). Drosophila Dpp morphogen movement is independent of dynamin-mediated endocytosis but regulated by the glypican members of heparan sulfate proteoglycans. Cell *119*, 231–244.

Ben-Zvi, D. and Barkai, N. (2010). Scaling of morphogen gradients by an expansion-repression integral feedback control. Proc Natl Acad Sci U S A *107*, 6924–9.

Bergmann, S., Ihmels, J. and Barkai, N. (2003). Iterative signature algorithm for the analysis of large-scale gene expression data. Phys Rev E Stat Nonlin Soft Matter Phys *67*, 031902.

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.

Bergmann, S., Tamari, Z., Schejter, E., Shilo, B.-Z. and Barkai, N. (2008). Re-examining the Stability of the Bicoid Morphogen Gradient. Cell *132*, 15–17.

Berleth, T., Burri, M., Thoma, G., Bopp, D., Richstein, S., Frigerio, G., Noll, M. and Nüsslein-Volhard, C. (1988). The role of localization of bicoid RNA in organizing the anterior pattern of the Drosophila embryo. Embo J *7*, 1749–56.

Biehs, B., Sturtevant, M. A. and Bier, E. (1998). Boundaries in the Drosophila wing imaginal disc organize vein-specific genetic programs. Development (Cambridge, England) *125*, 4245–4257.

Birdsall, K., Zimmerman, E., Teeter, K. and Gibson, G. (2000). Genetic variation for the positioning of wing veins in Drosophila melanogaster. Evol Dev *2*, 16–24.

Bittar, P. G., Charnay, Y., Pellerin, L., Bouras, C. and Magistretti, P. J. (1996). Selective distribution of lactate dehydrogenase isoenzymes in neurons and astrocytes of human brain. J Cereb Blood Flow Metab *16*, 1079–89.

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

Bollenbach, T., Kruse, K., Pantazis, P., González-Gaitán, M. and Jülicher, F. (2007). Morphogen transport in epithelia. Phys Rev E Stat Nonlin Soft Matter Phys *75*, 011901.

Bollenbach, T., Pantazis, P., Kicheva, A., Bökel, C., González-Gaitán, M. and Jülicher, F. (2008). Precision of the Dpp gradient. Development *135*, 1137–1146.

Bouzier-Sore, A.-K., Voisin, P., Bouchaud, V., Bezancon, E., Franconi, J.-M. and Pellerin, L. (2006). Competition between glucose and lactate as oxidative energy substrates in both neurons and astrocytes: a comparative NMR study. Eur J Neurosci *24*, 1687–1694.

Bouzier-Sore, A.-K., Voisin, P., Canioni, P., Magistretti, P. J. and Pellerin, L. (2003). Lactate is a preferential oxidative energy substrate over glucose for neurons in culture. J Cereb Blood Flow Metab *23*, 1298–306.

Burke, R. and Basler, K. (1996). Dpp receptors are autonomously required for cell proliferation in the entire developing Drosophila wing. Development *122*, 2261–9.

Burz, D. S., Rivera-Pomar, R., Jäckle, H. and Hanes, S. D. (1998). Cooperative DNA-binding by Bicoid provides a mechanism for threshold-dependent gene activation in the Drosophila embryo. EMBO J *17*, 5998–6009.

Campbell, G. and Tomlinson, A. (1999). Transducing the Dpp morphogen gradient in the wing of Drosophila: regulation of Dpp targets by brinker. Cell *96*, 553–62.

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.

Coppey, M., Berezhkovskii, A. M., Kim, Y., Boettiger, A. N. and Shvartsman, S. Y. (2007). Modeling the bicoid gradient: diffusion and reversible nuclear trapping of a stable protein. Dev Biol *312*, 623–630.

Crauk, O. and Dostatni, N. (2005). Bicoid determines sharp and precise target gene expression in the Drosophila embryo. Curr Biol *15*, 1888–1898.

Crick, F. (1970). Diffusion in embryogenesis. Nature *225*, 420–422.

de Celis, J. and Bray, S. (2000). The Abruptex domain of Notch regulates negative interactions between Notch, its ligands and Fringe. Development *127*, 1291–302.

de Lachapelle, A. M. and Bergmann, S. (2010a). Precision and scaling in morphogen gradient read-out. Mol Syst Biol *6*, 351.

de Lachapelle, A. M. and Bergmann, S. (2010b). Pre-steady and stable morphogen gradients: can they coexist? Molecular Systems Biology *6*, 428.

Dequéant, M.-L. and Pourquié, O. (2008). Segmental patterning of the vertebrate embryonic axis. Nat Rev Genet *9*, 370–382.

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.

Driever, W. and Nüsslein-Volhard, C. (1988a). The bicoid protein determines position in the Drosophila embryo in a concentration-dependent manner. Cell *54*, 95–104.

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

Driever, W. and Nüsslein-Volhard, C. (1989). The bicoid protein is a positive regulator of hunchback transcription in the early Drosophila embryo. Nature *337*, 138–143.

Driever, W., Thoma, G. and Nüsslein-Volhard, C. (1989). Determination of spatial domains of zygotic gene expression in the Drosophila embryo by the affinity of binding sites for the bicoid morphogen. Nature *340*, 363–367.

Dubrulle, J. and Pourquié, O. (2004). fgf8 mRNA decay establishes a gradient that couples axial elongation to patterning in the vertebrate embryo. Nature *427*, 419–22.

Edgar, B. A. (2006). How flies get their size: genetics meets physiology. Nat Rev Genet *7*, 907–16.

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.

Entchev, E., Schwabedissen, A. and Gonzalez-Gaitan, M. (2000). Gradient formation of the TGF-beta homolog Dpp. Cell *103*, 981–91.

Esser, A. T., Smith, K. C., Weaver, J. C. and Levin, M. (2006). Mathematical model of morphogen electrophoresis through gap junctions. Dev Dyn *235*, 2144–59.

Flenner, E., Marga, F., Neagu, A., Kosztin, I. and Forgacs, G. (2008). Relating biophysical properties across scales. Curr Top Dev Biol *81*, 461–83.

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

Foronda, D., Pérez-Garijo, A. and Martín, F. A. (2009). Dpp of posterior origin patterns the proximal region of the wing. Mech Dev *126*, 99–106.

Fowlkes, C. C., Hendriks, C. L. L., Keränen, S., Weber, G. H., RÂžbel, O., Huang, M.-Y., Chatoor, S., DePace, A. H., Simirenko, L., Henriquez, C., Beaton, A., Weiszmann, R., Celniker, S., Hamann, B., Knowles, D. W., Biggin, M. D., Eisen, M. B. and Malik, J. (2008). A quantitative spatiotemporal atlas of gene expression in the Drosophila blastoderm. Cell *133*, 364–374.

Fujise, M., Takeo, S., Kamimura, K., Matsuo, T., Aigaki, T., Izumi, S. and Nakato, H. (2003). Dally regulates Dpp morphogen gradient formation in the Drosophila wing. Development *130*, 1515–1522.

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, Q. and Finkelstein, R. (1998). Targeting gene expression to the head: the Drosophila orthodenticle gene is a direct target of the Bicoid morphogen. Development *125*, 4185–4193.

Gao, S., Steffen, J. and Laughon, A. (2005). Dpp-responsive silencers are bound by a trimeric Mad-Medea complex. J Biol Chem *280*, 36158–64.

Gilbert, S. (2006). Developmental biology.

González-Gaitán, M. (2003). Signal dispersal and transduction through the endocytic pathway. Nat Rev Mol Cell Biol *4*, 213–224.

Greco, V., Hannus, M. and Eaton, S. (2001). Argosomes: a potential vehicle for the spread of morphogens through epithelia. Cell *106*, 633–45.

Gregor, T., Bialek, W., de Ruyter van Steveninck, R. R., Tank, D. W. and Wieschaus, E. F. (2005). Diffusion and scaling during early embryonic pattern formation. Proc Natl Acad Sci U S A *102*, 18403–18407.

Gregor, T., McGregor, A. P. and Wieschaus, E. F. (2008). Shape and function of the Bicoid morphogen gradient in dipteran species with different sized embryos. Dev Biol *316*, 350–358.

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

Gregor, T., Wieschaus, E. F., McGregor, A. P., Bialek, W. and Tank, D. W. (2007b). Stability and nuclear dynamics of the bicoid morphogen gradient. Cell *130*, 141–152.

Halim, N. D., Mcfate, T., Mohyeldin, A., Okagaki, P., Korotchkina, L. G., Patel, M. S., Jeoung, N. H., Harris, R. A., Schell, M. J. and Verma, A. (2010). Phosphorylation status of pyruvate dehydrogenase distinguishes metabolic phenotypes of cultured rat brain astrocytes and neurons. Glia *58*, 1168–76.

Han, C., Yan, D., Belenkaya, T. and Lin, X. (2005). Drosophila glypicans Dally and Dally-like shape the extracellular Wingless morphogen gradient in the wing disc. Development *132*, 667–679.

He, F., Ren, J., Wang, W. and Ma, J. (2011). A multiscale investigation of bicoid-dependent transcriptional events in Drosophila embryos. PLoS ONE *6*, e19122.

He, F., Saunders, T. E., Wen, Y., Cheung, D., Jiao, R., ten Wolde, P. R., Howard, M. and Ma, J. (2010). Shaping a morphogen gradient for positional precision. Biophys J *99*, 697–707.

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.

Herrero-Mendez, A., Almeida, A., Fernández, E., Maestre, C., Moncada, S. and Bolaños, J. P. (2009). The bioenergetic and antioxidant status of neurons is controlled by continuous degradation of a key glycolytic enzyme by APC/C-Cdh1. Nat Cell Biol *11*, 747–52.

Hoch, M., Gerwin, N., Taubert, H. and Jäckle, H. (1992). Competition for overlapping sites in the regulatory region of the Drosophila gene Krüppel. Science *256*, 94–97.

Hoch, M., Seifert, E. and Jäckle, H. (1991). Gene expression mediated by cis-acting sequences of the KrÂžppel gene in response to the Drosophila morphogens bicoid and hunchback. EMBO J *10*, 2267–2278.

Holloway, D. M., Harrison, L. G., Kosman, D., Vanario-Alonso, C. E. and Spirov, A. V. (2006). Analysis of pattern precision shows that Drosophila segmentation develops substantial independence from gradients of maternal gene products. Dev Dyn *235*, 2949–2960.

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

Howard, M. and ten Wolde, P. R. (2005). Finding the center reliably: robust patterns of developmental gene expression. Phys Rev Lett *95*, 208103.

Hu, Y. and Wilson, G. S. (1997). A temporary local energy pool coupled to neuronal activity: fluctuations of extracellular lactate levels in rat brain monitored with rapid-response enzyme-based sensor. J Neurochem *69*, 1484–90.

Hufnagel, L., Teleman, A., Rouault, H., Cohen, S. and Shraiman, B. (2007). On the mechanism of wing size determination in fly development. Proc Natl Acad Sci U S A *104*, 3835–40.

Huxley, J. S. (1972). Problems of Relative Growth (2nd ed.). *New York: Dover.*

Hyder, F., Patel, A. B., Gjedde, A., Rothman, D. L., Behar, K. L. and Shulman, R. G. (2006). Neuronal-glial glucose oxidation and glutamatergic-GABAergic function. J Cereb Blood Flow Metab *26*, 865–77.

Ibañes, M. and Belmonte, J. (2008). Theoretical and experimental approaches to understand morphogen gradients. Mol Syst Biol *4 SP* -, 176.

Ibañes, M., Kawakami, Y., Rasskin-Gutman, D. and Belmonte, J. (2006). Cell lineage transport: a mechanism for molecular gradient formation. Mol Syst Biol *2 SP* -, 57.

Ihmels, J., Bergmann, S. and Barkai, N. (2004). Defining transcription modules using large-scale gene expression data. Bioinformatics *20*, 1993–2003.

Itoh, Y., Esaki, T., Shimoji, K., Cook, M., Law, M. J., Kaufman, E. and Sokoloff, L. (2003). Dichloroacetate effects on glucose and lactate oxidation by neurons and astroglia in vitro and on glucose utilization by brain in vivo. Proc Natl Acad Sci USA *100*, 4879–84.

Jaeger, J. (2010). A matter of timing and precision. Mol Syst Biol *6*, 427; author reply 428.

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

Jaeger, J. and Reinitz, J. (2006). On the dynamic nature of positional information. Bioessays *28*, 1102–1111.

Jaeger, J., Sharp, D. H. and Reinitz, J. (2007). Known maternal gradients are not sufficient for the establishment of gap domains in Drosophila melanogaster. Mech Dev *124*, 108–128.

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

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*, 3323–34.

Kicheva, A., Pantazis, P., Bollenbach, T., Kalaidzidis, Y., Bittig, T., JÂžlicher, F. and González-Gaitán, M. (2007). Kinetics of morphogen gradient formation. Science *315*, 521–525.

Kondo, S. and Miura, T. (2010). Reaction-diffusion model as a framework for understanding biological pattern formation. Science *329*, 1616–20.

Kutejova, E., Briscoe, J. and Kicheva, A. (2009). Temporal dynamics of patterning by morphogen gradients. Curr Opin Genet Dev *19*, 315–22.

Lander, A., Nie, Q. and Wan, F. (2002). Do morphogen gradients arise by diffusion? Dev Cell *2*, 785–796.

Lawrence, P. and Struhl, G. (1996). Morphogens, compartments, and pattern: Lessons from Drosophila? Cell *85*, 951—961.

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

Levine, M. and Davidson, E. H. (2005). Gene regulatory networks for development. Proc Natl Acad Sci U S A *102*, 4936–4942.

Lin, X. and Perrimon, N. (1999). Dally cooperates with Drosophila Frizzled 2 to transduce Wingless signalling. Nature *400*, 281–4.

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 biology *9*, e1000596.

Liu, J. and Ma, J. (2011). Fates-shifted is an F-box protein that targets Bicoid for degradation and regulates developmental fate determination in Drosophila embryos. Nat Cell Biol *13*, 22–9.

Loaiza, A., Porras, O. H. and Barros, L. F. (2003). Glutamate triggers rapid glucose transport stimulation in astrocytes as evidenced by real-time confocal microscopy. J Neurosci *23*, 7337–42.

Lopes, F. J. P., Vieira, F. M. C., Holloway, D. M., Bisch, P. M. and Spirov, A. V. (2008). Spatial bistability generates hunchback expression sharpness in the Drosophila embryo. PLoS Comput Biol *4*, e1000184.

Lott, S. E., Kreitman, M., Palsson, A., Alekseeva, E. and Ludwig, M. Z. (2007). Canalization of segmentation and its evolution in Drosophila. Proc Natl Acad Sci U S A *104*, 10926–10931.

Lucchetta, E. M., Lee, J. H., Fu, L. A., Patel, N. H. and Ismagilov, R. F. (2005). Dynamics of Drosophila embryonic patterning network perturbed in space and time using microfluidics. Nature *434*, 1134–1138.

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.

Lüscher, A., Csárdi, G., de Lachapelle, A. M., Kutalik, Z., Peter, B. and Bergmann, S. (2010). ExpressionView–an interactive viewer for modules identified in gene expression data. Bioinformatics *26*, 2062–3.

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

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. and Reinitz, J. (2009a). Canalization of gene expression and domain shifts in the Drosophila blastoderm by dynamical attractors. PLoS Comput Biol *5*, e1000303.

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. and Reinitz, J. (2009b). Canalization of gene expression in the Drosophila blastoderm by gap gene cross regulation. Plos Biol *7*, e1000049.

Martin, F., Perez-Garijo, A., Moreno, E. and Morata, G. (2004). The brinker gradient controls wing growth in Drosophila. Development *131*, 4921–30.

Martin-Castellanos, C. and Edgar, B. (2002). A characterization of the effects of Dpp signaling on cell growth and proliferation in the Drosophila wing. Development *129*, 1003–13.

Meinhardt, H. and Gierer, A. (2000). Pattern formation by local self-activation and lateral inhibition. Bioessays *22*, 753–60.

Mendes, P., Kell, D. B. and Welch, G. R. (1995). Metabolic Channeling in Organized Enzyme Systems: Experiments and Models. vol. 11, pp. 1 – 19.

Milan, M., Campuzano, S. and Garcia-Bellido, A. (1996). Cell cycling and patterned cell proliferation in the wing primordium of Drosophila. Proc Natl Acad Sci U S A *93*, 640–5.

Minami, M., Kinoshita, N., Kamoshida, Y., Tanimoto, H. and Tabata, T. (1999). brinker is a target of Dpp in Drosophila that negatively regulates Dpp-dependent genes. Nature *398*, 242–6.

Morata, G. (2001). How Drosophila appendages develop. Nature reviews Molecular cell biology *2*, 89–97.

Morgan, T. H. (1901). REGENERATION AND LIABILITY TO INJURY. Science *14*, 235–48.

Moser, M. and Campbell, G. (2005). Generating and interpreting the Brinker gradient in the Drosophila wing. Dev Biol *286*, 647–658.

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*, 221–33.

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

Nahmad, M. and Stathopoulos, A. (2010). Establishing positional information through gradient dynamics: a lesson from the Hedgehog signaling pathway. Fly (Austin) *4*, 273–7.

Nehlig, A., Wittendorp-Rechenmann, E. and Lam, C. D. (2004). Selective uptake of [14C]2-deoxyglucose by neurons and astrocytes: high-resolution microautoradiographic imaging by cellular 14C-trajectography combined with immunohistochemistry. J Cereb Blood Flow Metab *24*, 1004–14.

Niessing, D., Dostatni, N., Jäckle, H. and Rivera-Pomar, R. (1999). Sequence interval within the PEST motif of Bicoid is important for translational repression of caudal mRNA in the anterior region of the Drosophila embryo. EMBO J *18*, 1966–1973.

Nijhout, H. F. and Grunert, L. W. (2010). The cellular and physiological mechanism of wing-body scaling in Manduca sexta. Science *330*, 1693–5.

Nitisewojo, P. and Hultin, H. O. (1976). A comparison of some kinetic properties of soluble and bound lactate dehydrogenase isoenzymes at different temperatures. Eur J Biochem *67*, 87–94.

O'Brien, J., Kla, K., Hopkins, I., Malecki, E. and McKenna, M. (2007). Kinetic parameters and lactate dehydrogenase isozyme activities support possible lactate utilization by neurons. Neurochem Res *32*, 597–607.

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.

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. Development, growth & differentiation *53*, 668–678.

Panáková, D., Sprong, H., Marois, E., Thiele, C. and Eaton, S. (2005). Lipoprotein particles are required for Hedgehog and Wingless signalling. Nature *435*, 58–65.

Pellerin, L. (2008). Brain energetics (thought needs food). Curr Opin Clin Nutr Metab Care *11*, 701–5.

Pellerin, L. and Magistretti, P. J. (1994). Glutamate uptake into astrocytes stimulates aerobic glycolysis: a mechanism coupling neuronal activity to glucose utilization. Proc Natl Acad Sci USA *91*, 10625–9.

Pellerin, L. and Magistretti, P. J. (2004). Neuroenergetics: calling upon astrocytes to satisfy hungry neurons. Neuroscientist *10*, 53–62.

Pierre, K., Chatton, J.-Y., Parent, A., Repond, C., Gardoni, F., Luca, M. D. and Pellerin, L. (2009). Linking supply to demand: the neuronal monocarboxylate transporter MCT2 and the alpha-amino-3-hydroxyl-5-methyl-4-isoxazole-propionic acid receptor GluR2/3 subunit are associated in a common trafficking process. Eur J Neurosci *29*, 1951–63.

Porras, O. H., Loaiza, A. and Barros, L. F. (2004). Glutamate mediates acute glucose transport inhibition in hippocampal neurons. J Neurosci *24*, 9669–73.

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*, 229–40.

Qiu, L., Zhao, W. and Sick, T. (2004). Spectroscopic studies of mitochondrial NADH fluorescence signals in brain slices. Conf Proc IEEE Eng Med Biol Soc *7*, 4986–4989.

Ramírez-Weber, F. A. and Kornberg, T. B. (1999). Cytonemes: cellular processes that project to the principal signaling center in Drosophila imaginal discs. Cell *97*, 599–607.

Reeves, G. T., Muratov, C. B., Schupbach, T. and Shvartsman, S. Y. (2006). Quantitative models of developmental pattern formation. Developmental cell *11*, 289–300.

Salazar-Ciudad, I., Garcia-Fernández, J. and Solé, R. V. (2000). Gene networks capable of pattern formation: from induction to reaction-diffusion. J Theor Biol *205*, 587–603.

Saunders, T. and Howard, M. (2009). When it pays to rush: interpreting morphogen gradients prior to steady-state. Phys Biol *6*, 046020.

Schwank, G., Dalessi, S., Yang, S.-F., Yagi, R., Lachapelle, A. M. d., Affolter, M., Bergmann, S. and Basler, K. (2011). Formation of the Long Range Dpp Morphogen Gradient. PLoS biology *9*, e1001111.

Schwank, G., Restrepo, S. and Basler, K. (2008). Growth regulation by Dpp: an essential role for Brinker and a non-essential role for graded signaling levels. Development *135*, 4003–13.

Schwank, G., Tauriello, G., Yagi, R., Kranz, E., Koumoutsakos, P. and Basler, K. (2011). Antagonistic growth regulation by Dpp and Fat drives uniform cell proliferation. Dev Cell *20*, 123–30.

Seifert, F., Golbik, R., Brauer, J., Lilie, H., Schröder-Tittmann, K., Hinze, E., Korotchkina, L., Patel, M. and Tittmann, K. (2006). Direct kinetic evidence for half-of-the-sites reactivity in the E1 component of the human pyruvate dehydrogenase multienzyme complex through alternating sites cofactor activation. Biochemistry *45*, 12775–12785.

Shingleton, A. W., Frankino, W. A., Flatt, T., Nijhout, H. F. and Emlen, D. J. (2007). Size and shape: the developmental regulation of static allometry in insects. Bioessays *29*, 536–48.

Sivasankaran, R., Vigano, M., 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*, 6162–72.

Slack, J. M. W. (2006). Essential Developmental biology.

Small, S., Arnosti, D. N. and Levine, M. (1993). Spacing ensures autonomous expression of different stripe enhancers in the even-skipped promoter. Development *119*, 762–772.

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

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

Smith, J., Theodoris, C. and Davidson, E. H. (2007). A gene regulatory network subcircuit drives a dynamic pattern of gene expression. Science *318*, 794–7.

Spirov, A., Fahmy, K., Schneider, M., Frei, E., Noll, M. and Baumgartner, S. (2009). Formation of the bicoid morphogen gradient: an mRNA gradient dictates the protein gradient. Development *136*, 605–614.

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

Stern, D. L. and Emlen, D. J. (1999). The developmental basis for allometry in insects. Development *126*, 1091–101.

Strigini, M. (2005). Mechanisms of morphogen movement. J Neurobiol *64*, 324–33.

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

Sturtevant, M., Biehs, B., Marin, E. and Bier, E. (1997). The spalt gene links the A/P compartment boundary to a linear adult structure in the Drosophila wing. Development *124*, 21–32.

Surkova, S., Kosman, D., Kozlov, K., Manu, Myasnikova, E., Samsonova, A. A., Spirov, A., Vanario-Alonso, C. E., Samsonova, M. and Reinitz, J. (2008). Characterization of the Drosophila segment determination morphome. Dev Biol *313*, 844–862.

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*, 703–12.

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. and Cohen, S. (2000). Dpp gradient formation in the Drosophila wing imaginal disc. Cell *103*, 971–980.

Teleman, A., Strigini, M. and Cohen, S. (2001). Shaping morphogen gradients. Cell *105*, 559–562.

Thomson, D. (1917). On Growth and Form. *Cambridge: Cambridge University Press*.

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. (1952). The Chemical Basis of Morphogenesis. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences *237*, 37–72.

Verkhusha, V. V., Kuznetsova, I. M., Stepanenko, O. V., Zaraisky, A. G., Shavlovsky, M. M., Turoverov, K. K. and Uversky, V. N. (2003). High stability of Discosoma DsRed as compared to Aequorea EGFP. Biochemistry *42*, 7879–84.

Vuilleumier, R., Affolter, M. and Pyrowolakis, G. (2011). Pentagone: patrolling BMP morphogen signaling. Fly (Austin) *5*.

Vuilleumier, R., Springhorn, A., Patterson, L., Koidl, S., Hammerschmidt, M., Affolter, M. and Pyrowolakis, G. (2010). Control of Dpp morphogen signalling by a secreted feedback regulator. Nat Cell Biol *12*, 611–7.

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 *331*, 1154–9.

Weiss, A., Charbonnier, E., Ellertsdottir, E., Tsirigos, A., Wolf, C., Schuh, R., Pyrowolakis, G. and Affolter, M. (2010). A conserved activation element in BMP signaling during Drosophila development. Nat Struct Mol Biol *17*, 69–76.

Wolpert, L. (1969). Positional information and the spatial pattern of cellular differentiation. J Theor Biol *25*, 1–47.

Yamaguchi, M., Yoshimoto, E. and Kondo, S. (2007). Pattern regulation in the stripe of zebrafish suggests an underlying dynamic and autonomous mechanism. Proc Natl Acad Sci USA *104*, 4790–3.

Yao, L.-C., Blitz, I. L., Peiffer, D. A., Phin, S., Wang, Y., Ogata, S., Cho, K. W. Y., Arora, K. and Warrior, R. (2006). Schnurri transcription factors from Drosophila and vertebrates can mediate Bmp signaling through a phylogenetically conserved mechanism. Development *133*, 4025–34.

Zecca, M., Basler, K. and Struhl, G. (1995). Sequential organizing activities of engrailed, hedgehog and decapentaplegic in the Drosophila wing. Development *121*, 2265–78.