THE UNIVERSITY OF CHICAGO

THE ETS FACTORS PNT AND YAN REGULATE DISTINCT CELL FATES IN *DROSOPHILA*

A DISSERTATION SUBMITTED TO

THE FACULTY OF THE DIVISION OF THE BIOLOGICAL SCIENCES

AND THE PRITZKER SCHOOL OF MEDICINE

IN CANDIDACY FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

COMMITTEE ON GENETICS, GENOMICS, AND SYSTEMS BIOLOGY

BY

CHUDONG WU

CHICAGO, ILLINOIS

MARCH 2021

Table of Contents

List of Figures

Acknowledgement

I am in deep appreciation for all the help and support from my mentors and colleagues at the University of Chicago, my friends and family. Completing my PhD work would not have been possible without you.

First of all, I would like to thank Ilaria Rebay, who is an exceptional scientist and a great mentor. As a scientist, Ilaria is incredibly intelligent and can always bring up insightful scientific ideas for the lab. More importantly, her excellent communication capabilities made the research ideas easily understandable and I think the whole lab benefited a lot from that. The clear communication style was also exploited in her mentorship. Ilaria's door was always open and frequently I got suggestions for key experiments and thoughts from her when I was struggled. Ilaria's dedication and patience steered me along this journey, and were absolutely essential to my success.

I thank committee for always helping me grow and intellectually challenging me. My committee members, Ed Munro, Mike Rust and Chip Ferguson, encouraged me to explore both the qualitative and quantitate aspects of the project, and illustrate the molecular mechanism in corresponding biological contexts. Their suggestions certainly increased the rigor and impact of my work.

Essential to my happiness, determination and success were my dear lab mates. I appreciate Jean-François Boisclair Lachance for building the basis of the Pnt work and providing me with indelible mentorship in my early days. I am grateful to Jemma Webber for giving me scientific

and life-wise suggestions, as she was a big sister for me. I admire Trevor Davis for his broad knowledge and capability. I could not emerge into the US culture without him. I thank Matt Hope for being considerate and always willing to share ideas openly in and outside the lab. I appreciate Xiao Sun for being sincere and authentic all the time. I also thank the past lab members, Charlene Hoi, Nicelio Sanches-Luege, and Juana Delao, and current lab members, Julio Miranda-Alban, Suzy Hur, Jacob Decker, Saman Tabatabaee, and Christine Cao for being with me along the journey.

Finally, I would love to thank my families. My parents and grandparents unconditionally love and support me from another country. They stand by my side in any situation. My dear husband and lifelong partner, Jiamin Yan, has always tried his best to help me. He helped me for all my talks and even learned fly work. I look forward to growing more mature and better together with him.

Abstract

Development of properly patterned tissues requires cells to transit from a multipotent state to diverse differentiated states in a precise spatiotemporal manner. These transitions are directed by a limited number of signaling pathways via their transcriptional effectors that regulate the necessary changes in gene expression. How repeated use of the same signaling machinery confers specificity to the spectrum of fate transitions it controls is still poorly understood. In my dissertation research, I focused on this broad question from the viewpoint of Receptor Tyrosine Kinase/Mitogen Activated Protein Kinase (RTK/MAPK) signaling pathway and its downstream transcriptional effectors, the repressor Yan and the activator Pointed (Pnt). My researched centered on two scales of this process: primarily how MAPK signaling regulate the activator Pnt for diverse cell fate transitions; secondly how Yan and Pnt control target gene expression. As RTK/MAPK signaling cascade is one of the most conserved pathways in multicellular animals and drives distinct cell fates in a broad range developmental contexts, the MAPK response becomes an ideal context in which to investigate this question.

I first studied the MAPK/Pnt response in the *Drosophila* eye, which consists of a highly ordered array of ommatidia. Each ommatidium contains eight distinct photoreceptors, R1-R8, specified sequentially in a precise spatial pattern. Recruitment of R1-R7 fates all require reiterative RTK/MAPK signaling-mediated activation of the transcriptional effector Pnt. However, R1-R7 can be divided into two rounds of specification defined by different MAPK activation levels, raising the question of how the Pnt response is tailored to the distinctive MAPK signaling contexts. I discovered a previously unstudied Pnt isoform, PntP3, and showed that its respective

inclusion or exclusion in the RTK transcriptional effector circuitry defines two distinct Pnt regulatory networks for the two rounds of specification. Using a combination of genetic loss-of-function experiments and analysis of isoform-specific expression patterns and mRNA levels, I mapped out the cross- and auto-regulatory interactions between the three Pnt isoforms that define the two networks. Furthermore, my results suggested the inclusion of PntP3 in the network conferred robustness to cell fate specification under conditions of compromised signaling. Considering the conservation of the RTK/MAPK/Pnt pathway, the interconnected transcriptional effector network that my work uncovered is likely to be exploited in diverse developmental scenarios.

As the recruitment of cell fates is driven directly by the activation of gene expression, I then studied Pnt/Yan regulation on gene transcription at the molecular level. Precise regulation of gene expression from transcription factors is organized by the DNA regulatory elements that they bind. Cis-regulatory modules contain clusters of transcription factor binding sites. The number, affinity and organization of the binding sites, termed the cis-regulatory syntax, coordinate transcription factors binding at the regulatory element and determine the transcription output. I used the Muscle Heart Enhancer (MHE), a well-defined enhancer that activates *even-skipped (eve)* expression for cardiac muscle precursor specification in the *Drosophila* embryo, as an example to study the cis-regulatory syntax organizing Yan and Pnt. There are eight ETS sites, which are recognized by Yan and Pnt, identified and characterized at MHE, with 3 strong and 4 weak affinity sites. Using individual and combinations of ETS site mutations, a previous study in the lab had suggested the respective Yan and Pnt's preference for strong and weak sites and

defined a pair of overlapping strong-affinity sites as critical to effective Yan repression. However the molecular mechanism behind these preferences was unknown.

I developed a thermodynamic model to study how this cis-regulatory syntax organizes the exact Yan and Pnt binding at each site in order to generate hypothesis for future experimental investigations. Using this model I demonstrated how three-dimensional long-range interactions among the three strong-affinity sites could facilitate Yan occupancy and transcription repression, and showed that such regulation would be dependent on the self-association ability of Yan. My results also predict that collaborative interactions between Pnt and other transcription activators bound at MHE is required to activate MHE expression. Thus my model proposes that the combination of long-range repression plus local collaborative activation contributes to precise gene activation. This work sets a theoretical basis for future experimental studies on a general principle for Yan and Pnt regulation on different enhancers.

Chapter 1 Introduction and Background

1.1 Signaling pathways regulate cell fate specification through transcription factions (TFs)

The proper development of multicellular organisms relies on the interpretation of many different signaling cues. The following introduction section focuses on RTK/MAPK signaling, a key signaling transduction pathway that has profound impact in many developmental processes. I will discuss its architecture, function in different developmental contexts, and the modes of regulation that allow it to achieve its specificity.

1.1.1 Transcription factors mediate signaling cues to direct cell fate specification

The development of multicellular organism at its core relies on the correct patterning of an array of organs and tissues, each of which are composed of cells adopting different identities in precise spatiotemporal sequences. Being the hotspot of developmental biology since 1900s, this cellular decision-making process has been revealed to be a sequence of cellular changes in response to extracellular inductive signals. Research efforts have thus focused on identifying the inductive signals and then investigating how they specify the hundreds of different cell fates needed to form a complex multicellular animal. Because these signaling mechanisms are evolutionarily conserved and work in essentially identical ways to specify fates in evolutionarily divergent metazoan species, studies in a genetically manipulable invertebrate like *Drosophila* have made huge contributions in uncovering developmental mechanisms at play in higher vertebrates.

A key realization that forms the core of the specificity conundrum, is that animals use a very limited number of signaling transduction pathways to generate the full diversity of cell fate adoptions; in other words, cellular diversity can not be simply explained by matching regulatory

diversity at the pathway level (Gerhart, 1999; Perrimon et al., 2012; Skov and Høegh, 2002). Rather, diversity must be achieved both through regulation of the cell-to-cell contacts that modulate the amplitude and duration of the transmitted signal and through adjustment of the downstream signaling effectors that receive, interpret and transduce the information into differential cellular responses. For example, a signaling pathways can act cooperatively or antagonistically depending on developmental context, and coordination with spatiotemporally restricted factors can further contribute to signaling specificity.

Signaling pathways act through their downstream transcriptional effectors to control the gene-expression programs which directly regulate cellular responses. Further increasing the diversity and specificity of gene regulation, transcription factors also act in combinatorial codes, just as we mentioned for the signaling pathways. More studies on the regulatory relationships between signaling pathways and on how the combinatorial action of their transcription effectors effectively integrates the upstream signaling cues is the key for better understanding the developmental mechanisms underlying context-dependent cellular responses.

1.1.2 The RAS-MAPK signal transduction pathway is essential to cell fate specification during diverse developmental processes in Drosophila

1.1.2.1 The architecture of the MAPK pathway in Drosophila

Mitogen-activated protein kinase (MAPK) pathway cascades have conserved roles in a wide range of cellular processes, including cell proliferation, differentiation and apoptosis (Cargnello and Roux, 2011; Chen and Thorner, 2007; Molina and Adjei, 2006). The MAPK protein was originally isolated from a mouse cell-line in the 1980s but then quickly found to be conserved

across organisms, all the way to yeast (Boulton et al., 1990; Ray and Sturgill, 1988; Rossomando et al., 1989; Widmann et al., 1999). Its broad functionality and evolutionary conservation, as well as the association of MAPK mutants with various diseases like cancer and neurological disorders (Bohush et al., 2018; Wagner and Nebreda, 2009), have stimulated extensive studies on the cellular mechanics underlying MAPK signaling cascades. Coordinated efforts in genetics and biochemistry across a range of *in vivo* and cultured cell systems have revealed deep understanding of precision and versatility with which MAPK signaling pathways convert extracellular stimuli into cellular responses.

MAPK signaling cascades transmit signals through sequential activation of layers of protein kinases. Defined by the type of MAPK, there are four conventional MAPK cascades comprising Extracellular Regulated Kinase (ERK1/2), c-Jun N-terminal kinase (JNK), p38 MAPK, and an additional newly evolved MAPK, ERK5, found only in vertebrates (Cargnello and Roux, 2011; Chen et al., 2001; Kyriakis and Avruch, 2001). For all MAPK cascades, sequential phosphorylation of upstream protein kinases introduces dual phosphorylation on the Thr-X-Tyr motif of MAPK. After that, MAPK phosphorylates specific serine and threonine residues of target substrates, which include other protein kinases and many transcription factors. The ERK cascade, being the first mammalian MAPK to be cloned and characterized, is the most thoroughly studied MAPK signaling pathway and plays pivotal roles in cell fate adoption (Burdon et al., 1999; Cooper et al., 1982). The core components of the ERK kinase cascade are Raf, a MAPKKK, MEK, a MAPKK, and ERK, the MAPK. The GTPase Ras serves as an upstream activating protein to initiate the sequential phosphorylation of this cascade (Molina and Adjei, 2006). For simplicity, throughout my thesis I will refer to this as either the Ras/MAPK pathway or the MAPK pathway.

The cytoplasmic Ras/MAPK pathway is activated in response to the association of extracellular activating ligands with integral membrane receptors (van der Geer P, Hunter T, 1994). In Drosophila, ligand binding to receptor tyrosine kinases (RTKs), including Torso, Epidermal Growth Factor receptor (EGFR, also known as DER, Drosophila EGF Receptor), Fibroblast Growth Factor Receptor (FGFR, known as Heartless (Htl) and Breathless (Btl)), Alk and Sevenless (Sev), triggers the activation of the Ras/ MAPK cascade (Basler and Hafen, 1988; Beiman et al., 1996; Shilo, 2005b; Xiangyi Lu, Lizabeth A. Perkins, 1993). As Fig 1.1 illustrates, association with ligand induces the dimerization of receptors, thereby enabling trans- and auto-phosphorylation of the conserved tyrosine residues. After phosphorylation, the RTKs are recognized by the adapter protein Grb2/ Drk. The protein complex then recruits the guanine nucleotide exchange factor Sos (Son of sevenless), which enables the activation of Ras. Once Sos is recruited by Grb2, all RTK/MAPK pathways utilize exactly the same downstream Ras/ Raf/(ERK)MAPK cascade (Sopko and Perrimon, 2013). MAPK then phosphorylates protein targets, including the ETS transcription effectors Yan and Pnt, to generate cellular responses (Brunner et al., 1994; Gabay et al., 1996; Rebay and Rubin, 1995).

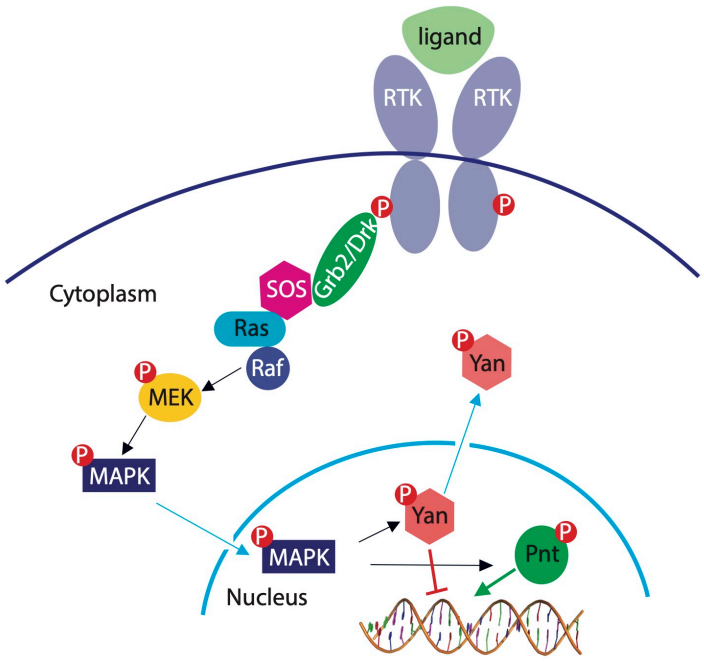


Fig 1.1 The Drosophila RTK signaling pathway

Linear flow from cell surface to the nucleus is depicted. On the cell surface, ligands bind RTK, which upon dimerization triggers the MAPK signaling cascade. Detailed flow of the pathway is described in the text. Yan and Pnt are ETS family transcription repressor and activator respectively. Upon phosphorylation, Yan is exported out of the nucleus and degraded while Pnt is activated.

1.1.2.2 The MAPK pathway establishes diverse cell fatse in a context dependent manner

RTK/Ras/MAPK signaling pathways establish cell fates in diverse developmental contexts in *Drosophila*. Although different RTK pathways all converge to the same Ras/Raf/MEK/MAPK cascade, the pathways direct distinct cell fates: the Torso pathway is the first RTK to be deployed in embryogenesis, defining the anterior and posterior embryonic termini (Xiangyi Lu, Lizabeth A. Perkins, 1993); the two FGF receptors are required for the patterning of mesoderm and tracheal system (Huang and Stern, 2005); the EGFR pathway controls muscle precursor specification during embryo development and wing and eye patterning in larval developmental stages (Perrimon and Perkins, 1997); graded activation of EGFR along the proximal–distal axis

of leg and wing discs specifies different fates (Campbell, 2002; Wang et al., 2000). The examples discussed below illustrate how the diverse outcomes depend on both the strength of the RTK/MAPK signaling and on cross-talk with other signaling cascades, and provide critical background information for interpretation of *pnt* mutant phenotypes described in Chapter 2.

1.1.2.3 Role of the MAPK pathway in specifying the wing field

The *Drosophila* wing originates from the embryonic ectoderm and develops into a disc shaped monolayer epithelium called the 'larval wing disc'. During early larval development, the distal and proximal regions of the wing disc are specified into two compartments that will give rise to the notum and the appendage (the wing field). By third instar, the wing field is subdivided into the proximal hinge and wing blade, and the wing blade is further divided into dorsal (D) and ventral (V) sides. The cells within each compartment are distinct precursors that will give rise to specific adult structures (Fig 1.2).

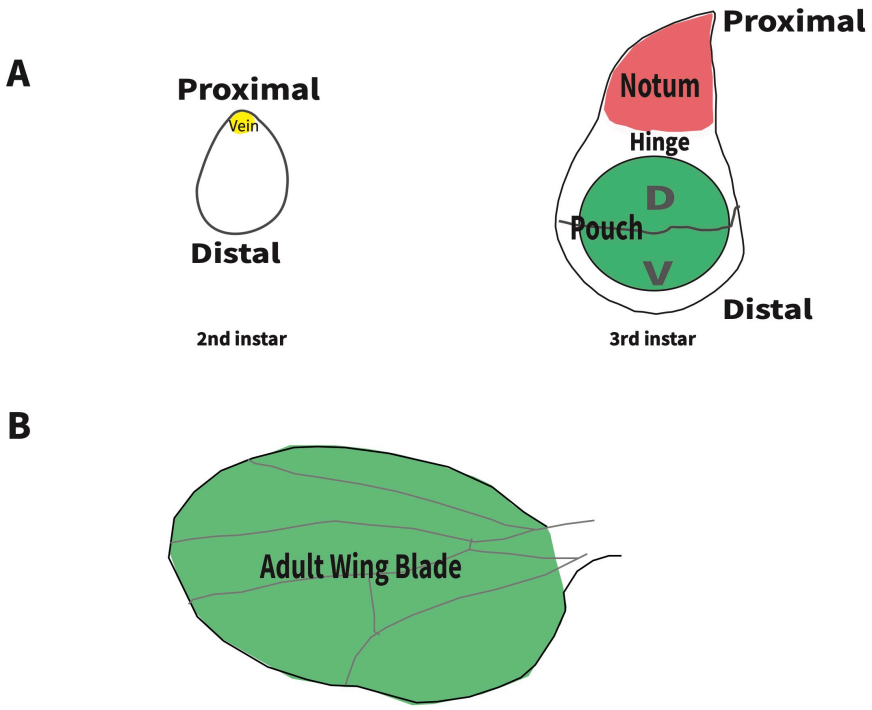


Fig 1.2 Wing development from 2nd instar larvae stage to adulthood

A) During the 2nd instar, the ligand Vein (yellow) is expressed at the proximal side of the disc and diffuses distally to generate an EGFR signaling gradient. During the 3rd instar, this gradient results in regional specification of the notum (red), hinge and wing pouch (green), which is further subdivided into dorsal (D) and ventral (V) side. C) The adult wing blade (green) is developed from the wing pouch.

EGFR signaling specifies different wing compartments in response to a gradient of its activating ligand, Vein (Vn). It is interesting to contrast Vein with another activating ligand, Spitz, that regulates EGFR signaling in most other embryonic and imaginal tissues. First, Vn possesses inherently weaker activation capacity comparing to Spitz. EGFR activation capacity was compared in cultured cells and in the embryo by measuring MAPK phosphorylation and induction of target gene expression (Golembo et al., 1999; Schnepp et al., 1996). Second, unlike Spitz which has to go through a tightly regulated cleavage process, Vn is freely secreted (Schnepp et al., 1996). In the early wing disc, Vn expressing cells in the proximal region secrete ligand which diffuses distally to generate a gradient of ligand (Shilo, 2005a; Wang et al., 2000).

EGFR is expressed in all cells, and so the ligand gradient results in a corresponding gradient of pathway activation. The gradient difference is increased by positive regulation on *vn* transcription from Vn/EGFR activation (Golembo et al., 1999; Wasserman and Freeman, 1998). At the proximal region, strong Vn/EGFR activation is required to specify the notum. Thus *vein* hypomorphs or compound heterozygosity for *egfr* and *vn* mutant both show loss of notum (Simcox et al., 1996; Wang et al., 2000). Interestingly, complete loss of Vn results in a tiny wing field, suggesting its role in the distal wing field where Vn levels are lowest (Clifford and Schüpbach, 1989; Simcox et al., 1996). Subsequent research showed that Vn/EGFR activation is required for the expression of Apterous, which is the dorsal selector gene of the wing field (Wang et al., 2000). Therefore, high levels of EGFR activation are required to specify the notum, while moderate Vn/EGFR activation is used to specify the dorsal wing field.

1.1.2.3 The MAPK pathway is required for the specification of cardiac cells during embryonic development

Heart development in *Drosophila melanogaster* exhibits remarkable similarities in terms of morphogenetic processes and regulatory mechanisms to heart development in vertebrates. For both *Drosophila* and vertebrates, proper heart development involves four major events: 1) the migration of a subset of mesodermal cells; 2) their specification as cardiac mesoderm, i.e. heart precursor cells; 3) the elaboration of different cardiac cell types; and 4) the final morphogenesis that sculpts the organ (Ahmad, 2017; Bodmer, 1993; Bodmer, 1995; Vogler and Bodmer, 2015). Multiple signaling pathways mediating these developmental processes are also conversed, with considerable attention focused on cardiac cell specification (Carmena et al., 1998; Liu et al., 1999; Liu et al., 2006; Lockwood and Bodmer, 2002; Park et al., 1996; Wu et al., 1995).

In *Drosophila*, the RTK/Ras/MAPK signaling pathway cooperates with the Wingless (Wg) and Decapentaplegic (Dpp) cascades for cardiac mesoderm specification in the embryo (Carmena et al., 1998; Halfon et al., 2000). Dpp is the primary factor in dorsal mesoderm specification. Loss-of-function mutations in any of the genes of the Dpp pathway, including *dpp* itself, its receptor and downstream transcription effectors, all lead to the failure to develop dorsal (including the cardiac) mesodermal derivatives (Frasch, 1995; Xu et al., 1998; Yin and Frasch, 1998). Wg signaling adds specificity to the more restricted dorsal cardiac mesoderm such that embryos lacking the two receptors of Wg or its downstream transcription effector are missing cardiac precursor cells (Bhanot et al., 1999; Chen and Struhl, 1999). Together, Wg and Dpp prepattern the dorsal cardiac mesoderm and establish the regional competency to respond to Ras/MAPK inductive signal (Lockwood and Bodmer, 2002).

Within the cardiac mesoderm, FGFR and EGFR signaling initiates cardiac cell fate specification through the downstream transcription effector Pnt. Overexpression of a dominant negative FGFR throughout the mesoderm resulted in a strong reduction of heart cells (Michelson et al., 1998). Similarly, loss-of-function mutations in *egfr* led to complete absence of cardiac cells (Grigorian et al., 2011). Furthermore, targeted overexpression of EGFR or FGFR ligands in the mesoderm increased the number of heart cells (Grigorian et al., 2011; Kadam et al., 2009; Klingseisen et al., 2009). In the following section I discuss the molecular mechanisms behind this signaling cooperativity.

1.1.3 MAPK coordinates with other signaling cascades through the regulation of transcription factors

1.1.3.1 Combinatorial action of transcriptional effectors of MAPK and Notch signaling

Cross-talk between MAPK and Notch signaling is a commonly used mechanism to control fate transitions in many different contexts during development, including the embryonic heart. Before returning to the discussion of heart development, I first summarize the interrelationships between MAPK and Notch in the *C. elegans* vulva and in the *Drosophila* eye. Development in both tissues requires a common pattern of sequential MAPK and Notch signals (Sundaram, 2005; Vivekanand and Rebay, 2006). For both systems as the illustration in Fig 1.3 shows, Notch signaling first contributes to the establishment of the "organizer cell" that produces the inductive signal that directs the specification of the subsequent cell fates, in *C. elegans* vulva this role is played by the Anchor Cell (AC) in the *Drosophila* eye by the R8 photoreceptor. In both systems the organizer cell expresses a short-range diffusible ligand, EGF in the vulva and Spitz in the eye, that activates the RTK/Ras/MAPK cascade in neighboring cells to initiate their specification. In turn, the activation of MAPK signaling stimulates the expression of the DSL (Delta/Serrate/LAG-2) protein, which acts as the ligand to activate Notch pathway in adjacent cells (Chen and Greenwald, 2004; Tsuda et al., 2002). The end result is that specific vulval or photoreceptor cell fates reflect the different combinations of high versus low MAPK and Notch signaling each cell experiences. Similar patterns of sequential signaling from MAPK to Notch, coupled with feedback between the pathways, have also been observed in the muscle and cardiac cell specification in *Drosophila* embryonic mesoderm and Drosophila tracheal tubulogenesis (Carmena et al., 2002; Ikeya and Hayashi, 1999).

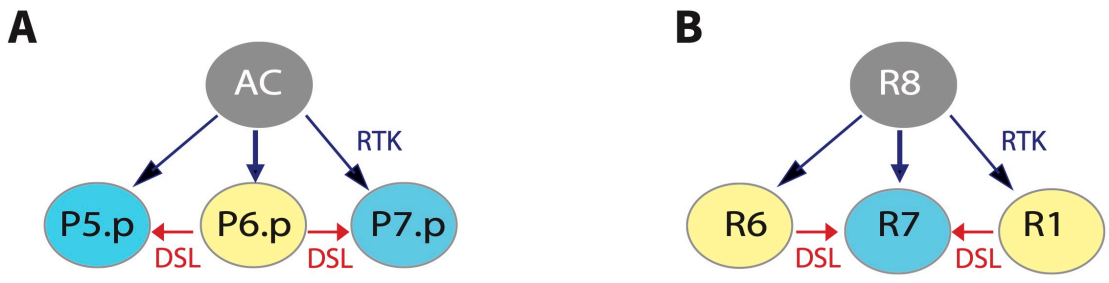


Fig 1.3 Combinatorial activation of RTK and Notch signaling specifies distinct cell fates in the C. elegans vulva and the Drosophila retina

The organizer cell is grey, and yellow cells are specified before the blue cells. Prior to their specification, the cells labled P5.p-P7.p in A and R6, R7, and R1 in B form an equivalence group; in both cases individual fates are specified by different combinations of RTK and Notch signaling. A) *C. elegans* vulval cell specification. The gonadal anchor cell (AC) expresses LIN-3, the activating ligand for the EGFR RTK LET-23. RTK signaling is activated most strongly in P6.p to specify it as the primary vulval fate and less strongly in P5.p and P7.p. RTK signaling in turn activates transcription of the DSL ligands for Notch, with highest DSL production occurring in P6.p. This inhibits Notch signaling in P6.p and activates it in P5.p and P7.p to induce the secondary vulval fates. Thus high RTK/ low Notch specifies the primary fate in P6.p while lower RTK / high Notch specifies the secondary fate in P5.p and p7.p. B) *Drosophila* photoreceptor specification. R8 expresses Spitz to activate EGFR RTK signaling in R1, R7 and R6. In contrast to the vulva where the cell with highest RTK activity turns on DSL ligand expression, in the eye the two cells with lower RTK activation, R1 and R6, express the DSL ligand to activate Notch signaling in the R7 cell where RTK activation is highest. The combinatorial action of high RTK and high Notch signaling specifies R7 while more modest levels of RTK and Notch pathway activity specifies R1 and R6.

Within single cells, when both MAPK and Notch signals are received, the two pathways can either antagonize or cooperate through the downstream transcription factors and other cofactors. In *Drosophila*, the transcription factor Suppressor of Hairless (Su(H)) acts as key mediator of Notch signaling (Schweisguth and Posakony, 1992). In some cases, Notch-activated Su(H) and MAPK-activated Pnt cooperate directly on gene enhancers to active expression. For example, the expression of D-Pax2, which is required for the specification of cone cells in the *Drosophila* eye, is regulated by the combined action of Su(H), Pnt and another transcription factor Lozenge (Lz) that all bind a *D-Pax2* regulatory enhancer (Flores et al., 2000). Loss of any of these three inputs,

either via loss-of-function mutations in the signaling factors or mutation of transcription factor binding sites in the enhancer, leads to loss of D-Pax2 expression. In other cases, the downstream transcriptional effectors of the Notch and Ras/MAPK pathways regulate important components of the other pathway to achieve antagonism. For example, in the undifferentiated precursor cells in *Drosophila* eye, Notch signaling, through its effector Su(H), promotes expression of the transcriptional repressor Yan (Rohrbaugh et al., 2002). Because Yan can bind to the same transcription factor binding site with Pnt, inappropriately high Yan levels allows it to outcompete Pnt and thereby reduce the response to RTK signaling. In addition to the pathway transcriptional effectors, other cofactors contribute to the RTK-Notch interrelationship. For example, Groucho functions as a corepressor for E(spl) proteins, which act downstream of Notch signaling. Groucho is also directly regulated by MAPK, with phosphorylation attenuating its function as a corepressor (Hasson et al., 2005). As Groucho is also a corepressor for the effectors of other signaling pathways, including Wg, Dpp and RTK, MAPK inactivating Groucho may modulate and coordinate the output of other cascades as well.

1.1.3.2 Cooperative binding with other factors on the regulatory elements of genes determines the spatiotemporal specificity of Ras/MAPK responses

Integration of signaling inputs often happens at the level of enhancers that contain multiple transcription factor binding sites to control the context-specific expression of the target gene. In the case of Ras/MAPK signaling, such combinatorial interactions between transcription factors are required for the regulation of spatiotemporal specificity. Homotypic clusters of transcription factor binding sites have been broadly studied and found to be essential for cooperative binding (Hertel et al., 1997; Lifanov et al., 2003; Wagner, 1997). Examples of cooperative binding of

transcription factors with DNA can be found from bacteriophage to mammalian cells (Johnson et al., 1979; Nie et al., 2020; Ravasi et al., 2010). Most regulatory regions also contain different types of binding motifs that promote recruitment of a different transcription factors. Although it is clear that heterotypic clustering of transcription factor binding sites contributes more than simply additive effects, how the various interactions produce the correct transcriptional output is not well understood. Some insight has come from synthetic transcriptional reporter assays that showed heterotypic clusters of binding sites conferred stronger activation than homotypic clusters (Fiore and Cohen, 2016; Reiter et al., 2017). These studies indicate that the different factors each contribute distinct regulatory cues and that their collaborative actions are required for robust transcription. In Drosophila, such collaboration enables the generic Ras/MAPK response to specify distinctive cell fates throughout development.

Tissue-restricted and pre-patterned factors are main contributors to the cell-type specificity of the Ras/MAPK response. A well-studied example is the *even-skipped* (*eve*) muscle and heart enhancer (MHE) that is necessary and sufficient to drive *eve* expression in the embryonic cardiac mesoderm. As introduced in section 1.1.2, the intersection of Dpp, Wg and MAPK signaling specifies Eve+ cardiac precursor cells (Carmena et al., 1998; Halfon et al., 2000),.Reflecting these signaling inputs, MHE activity is regulated by combinatorial binding of five transcription factors: the Wg-stimulated transcription factor dTCF, the Dpp-stimulated transcription factors Mad, Tin, and Twist and the MAPK-stimulated transcription factor Pnt (Halfon et al., 2000). For all five, mutation of individual transcription factor binding sites led to reduced MHE reporter expression whereas double mutants showed much stronger loss, indicating synergistic rather than additive contributions. Paralleling this result, ectopic expression of either Wg or activated Ras on their own had little or no effect on Eve expression, whereas coexpression led to strong expansion

of the Eve expression. Similar ectopic expression experiments uncovered analogous synergism between Twist and activated Ras.

Additional modulation comes from the transcriptional regulation among these five factors. For example, Wg signaling is required for the expression of Htl (FGF) and its downstream component Heartbroken (Hbr), and for Rhomboid, a protein involved in processing the EGFR ligand Spitz (Bang and Kintner, 2000; Halfon et al., 2000). In *wg* mutants, the expression of these proteins was missing. Dpp was also shown to be required for the expression of Htl, Hbr and Rhomboid. These layers of regulation ensure that MAPK activation downstream of FGFR and EGFR, and hence Eve expression, only occurs in cells at the intersection of the Wg and Dpp signaling domains. The two layers of regulation not only regulate the specificity of MAPK activation but also restrict the target gene expression through the collaborative actions between MAPK and the spatiotemporal specific factors. Subsequent studies described similar patterns of transcription factor binding on other mesodermal enhancers with additional cell-type specific factors (Carmena et al., 2002; Halfon et al., 2002), suggesting the mechanistic strategy is broadly applicable.

1.2 The ETS Transcription factors, Pointed (Pnt) and Yan

The following section focuses on Yan and Pnt, which are the key transcriptional effectors of MAPK signaling. The section first introduces the ETS factor superfamily, one of the most evolutionarily conserved families of transcription factors, that Yan and Pnt belong to, and then focuses on the function and regulation of Yan and Pnt.

1.2.1 ETS family transcription factors are essential for metazoan development

Transcription factors of the E-Twenty-Six specific (ETS) family, including the MAPK signaling effector Pnt, are evolutionarily conserved throughout metazoan lineages and work as critical components of signal transduction pathways to regulate diverse biological processes. Furthermore, de-regulation of the activities of several ETS factors has been implicated in human cancers. ETS factors thus serve as an important window to study how transcription factors integrate the inputs from signaling cascades to stimulate distinct cellular responses. Because of their importance, ETS factors are well characterized genetically, molecularly, and structurally.

1.2.1.1 The conservation of ETS factors in all metazoan species

Based on homology with the DNA-binding domain of ETS-1, which is the founding member of the ETS family originally discovered as part of the avian E26 retrovirus genome (Leprince et al., 1983), many ETS factors have been identified from various organisms. According to PCR and phylogenetic analysis, ETS members are found in early metazoans, such as sponges, cnidarians, flatworms and nematodes, suggesting the ETS family originated very early during metazoan evolution (Laudet et al., 1999). To date, identification of ETS genes from metazoan species has established the ETS family as one of the largest families of transcription factors, consisting of 28 members in humans, 27 in mice, 10 in *C. elegans* and 8 in *Drosophila* (Findlay et al., 2013). Based on sequence similarities in the ETS domain and the presence of further conserved domains, the ETS family members can be sub-classified into 10 groups (Sharrocks, 2001). One such conserved domain is the SAM domain, which was originally identified in the *Drosophila melanogaster* protein Pnt and is conserved among several subgroups (Klambt, 1993). Additional domains also include the OST domain of the GABPA subgroup gene and the B-box domain of

the TCF subgroup genes (Hollenhorst et al., 2011). The conservation of the protein domains is reflected in similar molecular functions of the ETS family and subfamily genes.

The ETS domain controls the DNA binding specificity. The ETS domain folds into a three-dimensional structure of winged helix-turn-helix (wHTH), which is composed of three alpha helices (H1, H2, and H3) and a four-stranded antiparallel beta sheet (Beta 1-4) (Liang et al., 1994; Mo et al., 2000). It recognizes a core GGAA/T DNA sequence (ETS binding site) (Graves and Petersen, 1998). Based on the structure of the Elk-1-DNA complex, the main DNA-protein contacts are from residues that are located in H3, which becomes the core recognition helix and forms sequence-specific contacts via two conserved arginine residues to recognize the two guanine nucleotides in the major groove (Mo et al., 2000). Interestingly, there is no direct DNA-protein contact outside the GGAA/T core even though individual ETS proteins can select a specific 11-base-pair sequence centered on the core motif. This suggests that DNA conformation dependent on the flanking sequences also contributes to binding specificity (Hollenhorst et al., 2011; Szymczyna and Arrowsmith, 2000).

Outside of the recognition helix, protein conformational changes triggered by either phosphorylation or interactions with co-regulatory factors modulate DNA-protein binding. A well characterized example is Ets1 protein, where helices N-terminal and C-terminal to the ETS domain serve as auto-inhibitory modules to block the DNA binding (Jonsen et al., 1996). Upon phosphorylation, the alpha helix switches from an unstructured state to folding against the core ETS domain and thus inhibits DNA binding. To relieve the autoinhibition, transcription factor AML-1/CBFα2 binding to Ets1 can disrupt the action of the inhibitory module (Kim et al., 1999). Similarly, interaction with another transcription factor Pax-5 also enhances Ets1 binding

to DNA (Fitzsimmons et al., 1996). Interactions with co-regulatory factors, the presence of which depends on cellular context, add additional layers of regulation on ETS factors and contribute to ETS factors achieving regulatory specificity.

1.2.1.2 ETS factors have major roles in hematopoietic and neuronal development in mammals

The requirements of ETS factors in mammalian development are diverse, but they have been primarily implicated in hematopoiesis and angiogenesis at the early stages of embryogenesis and later in nervous system development (Craig and Sumanas, 2016). In accordance with their involvement in diverse biological functions, the expression of several ETS factors, including Ets1, appears in a variety of tissues throughout development. Interestingly, most ETS factors are pervasively co-expressed, and despite minor differences in sequence-preference among the ETS factors, individual transcription factors exhibit distinctive functions depending on specific cellular context, such as the modulation from upstream signals and co-regulatory partner factors (Gille et al., 1992; Hollenhorst et al., 2011; Papoutsopoulou and Janknecht, 2000; Yang et al., 1998).

ETS factors, including Ets1, Ets2, Fli1 and Etv2, play critical roles in different stages of hemopoiesis, vasculogenesis and angiogenesis, and express in high levels in corresponding cells. At the first stage, Etv2 is expressed in the hemogenic endothelium and collaborates with the VEGF pathway member FLK1 to initiate the formation of hemangioblasts, the common precursor cells of blood and endothelial lineages (Lee et al., 2008; Wareing et al., 2012). Consequently, animals lacking Etv2 showed the most severe vascular defects as compared to other ETS transcription factors, and completely lack blood progenitors. Interestingly, the expression of Fli1 is also initiated in mesenchymal cells, most likely the hemangioblasts (Mélet

et al., 1996), but it mainly contributes to megakaryopoiesis and the formation of vasculature from hemogenic endothelial cells that are developed from the hemangioblasts (Hart et al., 2000; Spyropoulos et al., 2000).

Following another developmental trajectory of hemangioblasts, Ets1 and Ets2 mediate the differentiation of lymphoid lineages, such as B, T and nature killer cells, and myeloid lineages, such as macrophages (Maroulakou and Bowe, 2000). Functional redundancy between Ets1 and Ets2 has been implicated during embryonic angiogenesis. Animals losing either Ets1 or Ets2 appear normal for the formation of blood vessels, but losing both Ets1 and Ets2 results in significant apoptosis of angiogenic endothelial cells (Wei et al., 2009). More examples of redundant functionalities between other ETS factors have been revealed (Casie Chetty and Sumanas, 2020), indicating the importance of understanding the target specificity versus functional redundancy for ETS family factors.

The same ETS factors are also involved in the development of brain and central nervous system in mammals. High level expression of Ets1 occurs during embryonic vasculogenesis of the central nervous system and then is greatly reduced in fully differentiated tissues, suggesting its role for cell fate specification (Kola et al., 1993). Ets2 is expressed in the central nervous system during late embryogenesis. Specifically, it localizes in the cerebellum, hippocampus and cerebral cortex, indicating functions in multiple types of neuronal cells (Kola et al., 1993). There are additional ETS factors, such as Er81, PEA3, Erg and TEL1, that express and function in neurogenic processes (Lin et al., 1998; Vlaeminck-Guillem et al., 2000; Wang et al., 1997). Overall ETS factors are critical for embryogenesis and tissue formation in mammalian biology.

1.2.1.3 ETS factors in Drosophila

Eight ETS factors, all of which have corresponding human orthologs, have been identified in *Drosophila*. Specifically, the eight factors are Ecdysone-induced protein 74EF (Eip74EF), ETS21C, ETS65A, ETS96B, ETS97D, ETS98B, Pointed (Pnt), and Yan (also referred to as Anterior open, Aop), with Pnt and Yan being the best of characterized (Hsu and Schulz, 2000). Based on DNA binding specificities, *Drosophila* ETS proteins can be divided into four classes: Class I includes Pnt (human ortholog Ets1, Ets2), ETS21C (human ortholog Erg), ETS65A (human ortholog Fli1), ETS96B (human ortholog ETV5), and ETS97D (human ortholog GABP α); Class IIa includes Eip74EF (human ortholog ELF2); Class IIb includes Yan (human ortholog TEL); Class IV includes ETS98B (human ortholog PDEF) (Wei et al., 2010). To consider the conservation between the *Drosophila* protein and human ortholog, alignment of amino acids encompassing the ETS domain indicate that Pnt significantly resembles Ets1 with 96% identity and 98% similarity (Hsu and Schulz, 2000). Other ETS factors also show striking homology. Even the least similar pair, Yan and TEL still share 50% identity and 74% similarity in their ETS domains. Further comparisons on the domain structure, also taking into account the presence and position of the SAM domain, suggest that PntP2 (a Pnt protein isoform)/ETS1, Yan/TEL, D-Elg/GABPα and D-Ets-6/ERG pairs are most similar (Hsu and Schulz, 2000). The sequence and structure resemblance imply potential functional similarities.

Just as the ETS factors function throughout mammalian development, *Drosophila* ETS proteins have essential roles in diverse developmental processes. Pnt and Yan have roles in embryogenesis, as well as eye and wing patterning. The functions of Pnt and Yan will be discussed in more detail in the following sessions. The functions of other *Drosophila* ETS

protein are not as well characterized, but some contexts have been established. For example, ecdysone induces the expression of two protein isoforms of Eip74EF to carry out distinctive functions, despite the presence of an identical ETS DNA domain in the isoforms (Fletcher et al., 1997). Briefly, E74A is required for the metamorphosis of both larval and adult tissues and E74B contributes to puparium formation. Other studies have revealed the role of Ets21C in Drosophila innate immunity (Chambers et al., 2012) and Ets97D in oogenesis and egg chamber development (Gajewski and Schulz, 1995). To understand fully how the diverse functions of the ETS factors are achieved and regulated will require the focus of many *Drosophila* developmental biologists.

1.2.2 Pnt and Yan are critical downstream effectors of RAS-MAPK signaling cascade

As targets and effectors of Ras/MAPK signaling pathways, ETS factors function as critical nuclear integrators of the ubiquitous signaling cascade. For example, the ETS factor Pnt serves as an integration point for signals initiated from both EGFR and FGFR (Htl/ Hbr) stimuli for the specification of muscle heart cells (Dutta et al., 2005; Halfon et al., 2000). Similar signaling integration also occurs in mammalian cells via TCF subfamily factors responding to both ERK, JNK and p38 MAPK stimuli (Sharrocks et al., 1997). The integration relies on the fact that, despite different initiating stimuli, the signaling pathways regulate the ETS factors via MAPK stimulated phosphorylation of a single site. In many cases, MAPK initiated phosphorylation of the ETS factor greatly enhances DNA binding and transcriptional activities (Sharrocks, 2001). In some other cases, phosphorylation of the ETS protein, such as Yan in *Drosophila* and its mammalian homolog Tel, results in its degradation and thus causes a loss of its repressive activities (Arai et al., 2002; Rebay and Rubin, 1995).

The ETS activator Pnt and the repressor Yan are key downstream effectors of Ras/MAPK signaling in *Drosophila*. Both Pnt and Yan are direct targets of MAPK phosphorylation. Genetic interaction with components of the Ras/MAPK cascade, including EGFR and Sevenless, confirmed Yan being downstream of the Ras signaling and serving as a repressor for photoreceptor specification in the *Drosophila* retina (Lai and Rubin, 1992). Acting in opposition to Yan, Pnt promotes photoreceptor fate specification. Pnt has two well-characterized isoforms, PntP1 and PntP2, which share a common C-terminus containing the ETS domain (Klambt, 1993). PntP2 contains a MAPK-responsive SAM domain in the N-terminus, whereas PntP1 is constitutively active (O'Neill et al., 1994). MAPK mediated phosphorylation stimulates the transcriptional activity of PntP2, opposite to the effects of phosphorylation on Yan. Yan and Pnt, thus provide an opportunity to study how Ras/MAPK signaling modulates interactions between ETS factors to establish cell identities.

1.2.3 The two models of Pnt/Yan control of Drosophila development

1.2.3.1 The canonical bistable switch model

Pnt and Yan form a bistable switch, which is controlled by MAPK signaling, to regulate the transcription of downstream target genes associated with specific cell fate transitions. As the result of their homologous ETS DNA binding domain, Pnt and Yan compete for access to regulatory regions of common target genes, with Pnt activating and Yan repressing transcription (Tootle et al., 2003; Vivekanand and Rebay, 2006). The competition between Pnt and Yan is regulated by MAPK signaling. In the absence of the signaling, a high Yan state prevails in which Yan occupies the regulatory region and represses the gene expression program, maintaining the uncommitted state (O'Neill et al., 1994; Rebay and Rubin, 1995). Upon MAPK activation,

phosphorylated Yan is exported out of the nucleus and degraded whereas phosphorylated PntP2 converts into a potent activator and gains the access to the gene enhancers (Brunner et al., 1994; Gabay et al., 1996; O'Neill et al., 1994). This switch to a high Pnt state initiates the gene expression programs that drive the particular transition.

Further complexities of the bistable model have been revealed by the identification of Mae as a partner protein for both PntP2 and Yan. As mentioned above, some ETS factors, including Pnt and Yan, contain another conserved domain, the SAM domain. Mae, with a SAM domain, directly binds to the N-terminal SAM domains of both Yan and PntP2 (Baker et al., 2001; Qiao et al., 2006). Mae binding to Yan attenuates the repressive function of Yan by facilitating MAPK-mediated phosphorylation of Yan and promoting its depolymerization; together this increases the nuclear export and degradation of Yan (Baker et al., 2001). With respect to Pnt, the binding of Mae blocks the MAPK docking site on the SAM domain of PntP2, and thus negatively regulate the transcriptional activity of PntP2 (Qiao et al., 2006). Adding a further layer of control to the regulation, *mae* transcription is directly repressed by Yan and activated by Pnt (Vivekanand et al., 2004). Therefore, as Fig 1.4 illustrates, Mae forms feedback loops with Yan and Pnt to fine-tune the MAPK response. Specifically, the negative regulation from MAPK signal on Yan is further amplified by the upregulation of Mae when its repression from Yan is reduced. Thus even low levels of MAPK signal can relieve Yan-mediated repression on target genes. Meanwhile, MAPK-stimulated activation of Pnt transcriptional target genes, including *mae* itself, is limited by Mae. In this way the system achieves a Yan and Pnt mediated MAPK responsive switch which turns on as long as the MAPK signal is above the lowest threshold but which also sets the highest limit of the activity.

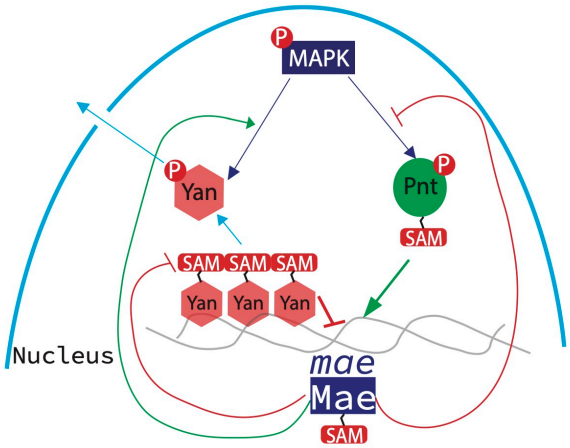


Fig 1.4 The feedback loop among Mae, Yan and Pnt

MAPK phosphorylates Yan to facilitate its export out of the nucleus, thereby relieving transcription repression of target genes, whereas MAPK phosphorylates and activates Pnt. The transcription of *mae* is repressed and activated by Yan and Pnt respectively. Once expressed, Mae interacts with Yan and Pnt through SAM-SAM interaction. Mae negatively regulates Yan mediated repression in two ways: facilitating the phosphorylation of Yan and disrupting Yan self-association. Mae-Pnt binding blocks phosphorylation of Pnt and thus prevents its activation. This feedback contributes to the bistability of the Yan-Pnt response to MAPK signaling.

1.2.3.2 Pnt recruits Yan and other corepressors to establish the initial repressed state

Recently Pnt has been implicated in facilitating collaborative actions of transcriptional repressors, including Yan, to confer the robustness of cell fate transitions. While the activating function of Pnt has been well-studied, Pnt was also shown to negatively regulate transcription of genes, such as *hid* in the embryo and *yan* in the eye discs (Kurada and White, 1998; Rohrbaugh et al., 2002). These studies raised the possibility of a role for Pnt as a transcriptional repressor, but did not investigate the regulatory mechanism. Only recently, a mechanism of the repressive regulation was revealed. Predicted from the canonical bistable model, Yan and Pnt binding at enhancers should be mutually exclusive, with high Yan occupancy when MAPK signaling is low and high Pnt binding when MAPK activation is high. However, motivated by the coexpression of Yan and Pnt in many tissues throughout *Drosophila* development (Boisclair Lachance et al.,

2014), the sequential ChIP-qPCR experiments showed co-occupancy of Yan and Pnt at a number of target gene loci (Webber et al., 2018). Furthermore, in *pnt* null mutant embryos, Yan DNA occupancy was reduced across the genome indicating Pnt is required for recruitment and/or stability of Yan DNA occupancy. In contrast, Pnt occupancy was broadly conserved in the absence of Yan. This unexpected co-occupancy, together with the extensive coexpression of Yan and Pnt in fly tissues (Boisclair Lachance et al., 2014), suggested the potential for collaborative functions between the two at the level of specific enhancers. Specifically, Pnt might be required for the assembly, and perhaps even the function, of Yan repressive complexes.

Regarding the molecular mechanism, the corepressor Groucho stood out as a likely player based on the result that in *pnt* null mutants Groucho occupancy was reduced in the exact same regions where Yan occupancy was reduced (Webber et al., 2018). Furthermore, of the genes associated with Yan and Groucho occupancy loss in the *pnt* mutant, more than 72 (including the validated Groucho target *E(spl)mbeta*-HLH) showed up-regulated mRNA expression in *pnt* mutants. Therefore, despite being well-characterized as an activator, Pnt can recruit and organize the binding of two repressors to negatively regulate gene expression. Motivated by the known relationship between Groucho and RNA polymerase II pausing (Kaul et al., 2014), through which transcription is paused until the appropriate signaling cues are received, further research revealed the genetic interaction between Pnt and the pausing machinery (Webber et al., 2018). These results together support the model that Pnt promotes the recruitment and stabilization of Yan and Groucho complexes, and coordinates interactions with the RNA polymerase II pausing machinery to maintain a transcriptionally repressed state in progenitor cells until the appropriate signaling cues stimulate cell fate transitions.

1.2.4 Pnt and Yan function in embryogenesis and eye development

Just like their mammalian counterparts, Pnt and Yan function to establish cell identities during embryogenesis and during the patterning of multiple imaginal tissues in *Drosophila*. As mentioned above, two Pnt isoforms have been studied. PntP1 is constitutively activate while PntP2 requires MAPK phosphorylation to be activated (O'Neill et al., 1994). Depending on the developmental context, either both PntP1 and PntP2 can be involved or only one isoform can take the lead. The following are two examples to demonstrate how the two isoforms work in distinct tissues to control cell fate decisions.

During embryogenesis, PntP1 and PntP2 are expressed in a distinct set of tissues and function differentially within these contexts. PntP1 is expressed and mediates EGFR/Ras/MAPK signaling in the embryonic ventral ectoderm whereas PntP2 is expressed in the mesoderm where it responds to both FGFR and EGFR signaling (Gabay et al., 1996; Klambt, 1993; Vivekanand, 2018). In the ventral ectoderm, the antagonistic functions of PntP1 and Yan establish an EGFR activation gradient to direct distinct cell fates (Hsu and Schulz, 2000). Specifically, high levels of EGFR activation initiated by the secreted ligand Spitz induces the expression of PntP1 in the ventral-most cells and phosphorylates and inactivates Yan (Gabay et al., 1996). PntP1 then activates the transcription of a set of secondary genes, including an antagonist of EGFR, Argos (Golembo et al., 1996). The expression and secretion of Argos lowers the level of EGFR activation in more lateral cells, allowing Yan to remain active and compete with PntP1to limit transcriptional activation of the target genes. In this way, a gradient of EGFR activation is formed from ventral to lateral. Of the two isoforms, only *pntP1* was the detected in this ventral ectoderm domain (Gabay et al., 1996; Klambt, 1993). Furthermore, ectopic expression of Spitz

induced a significant increase in *pntP1* transcripts, but did in *pntP2* transcripts, and the ectopic expression of PntP1 but not PntP2 was able to induce the expression of the ventral-most cell markers, indicating the sufficiency of PntP1 to mediate the ventral ectodermal gradient activation (Gabay et al., 1996). In contrast in the mesoderm, *pntP2* transcripts predominate and its overexpression induced extra cell fates that originated from the mesoderm (Halfon et al., 2000; Klambt, 1993). How different developmental contexts select which Pnt isoform will be expressed and how this choice informs the regulatory networks is still unclear, and is a topic worthy of future research.

In the retina, the sequential activation of the two Pnt isoforms initiates and sustains the MAPK response. As in other contexts, the antagonistic functions of Yan and Pnt control the retinal cell fate transitions. However, unlike the examples discussed in the embryo where only one Pnt isoform is involved, PntP1 and PntP2 work together in the retina. Mutants of either isoform including the *pntP1* specific null and *pntP2* specific hypomorph all show eye defects (Brunner et al., 1994; O'Neill et al., 1994; Scholz et al., 1993; Yang and Baker, 2001). More recently, *pntP1* transcription was found to be reduced in the pnt^{p2} mutant, indicating that PntP2 is required for *pntP1* transcription (Shwartz et al., 2013). This result, together with the findings that MAPK phosphorylation activates PntP2 and that *pntP1* transcription is dependent on MAPK activation (Brunner et al., 1994; Gabay et al., 1996; O'Neill et al., 1994; Shwartz et al., 2013; Tootle et al., 2003), suggested a sequential activation model in which PntP2 is activated by MAPK phosphorylation and then stimulates the transcription of *pntP1* (Shwartz et al., 2013). This sequential activation offers a key advantage over implementation of the individual isoform models, namely that sequential activation may be able to not only initiate the MAPK response fast but also sustain it (Shwartz et al., 2013). A simulation based on the transient kinetics of

MAPK activation and the stability of PntP1 protein showed that the phosphorylation of PntP2 initiates the response rapidly and then the induction of PntP1 expression will sustain the MAPK response for hours (Shwartz et al., 2013).

1.3 Drosophila eye development

Considering its highly ordered structure and the critical role of MAPK/Pnt activation during its development, the *Drosophila* retina is one of the best systems to explore MAPK/Pnt contributions to distinct cell fate specification. This section discusses the structure of the eye and how regulation from MAPK integrates with other signaling cues to specify distinct cell fates.

1.3.1 The structure and development of the compound eye

The *Drosophila* compound eye is highly organized and constitutes an excellent developmental system in which to study the molecular and cellular mechanisms that direct diverse cell fate transitions. The retina contains approximately 750 identical units, individually known as an ommatidium (Ready and Hanson, 1976; Tomlinson and Ready, 1987). Each ommatidium is composed of eight photoreceptor cells, from R1 to R8, that are surrounded by supporting cells, including four cone cells, and an array of pigment and bristle cells. Within an ommatidium, the cells have fixed organization with respect to each other due to the stereotyped order in which their fates are specified. The program of sequential cell fate specification that occurs in an individual ommatidial unit is repeated approximately 750 times in one retina tissue with almost perfect precision. This dissertation focuses on the specification of photoreceptor and cone cells in third instar larvae stage.

The highly ordered adult eye originates from an unpatterned epithelial monolayer referred to as the larval eye-antennal imaginal disc. The eye disc is specified in the embryonic and early larval stages by the actions of a network of transcription factors known as retinal determination genes. During embryogenesis, the expression of Eyeless and its paralog Twin of Eyeless determines the progenitor cells that will develop into the eye disc (Czerny et al., 1999; Davis and Rebay, 2017; Younossi-Hartenstein et al., 1993). Following the rounds of cell proliferation that expands the disc during the first two instar larval stages, the genes *eyes absent*, *sine oculis*, and *dachshund* induce the establishment of the presumptive eye field (Halder et al., 1998; Kenyon et al., 2003; Kumar and Moses, 2001). Then during the 3rd instar, a burst of Dpp and Hedgehog (Hg) signaling at the posterior edge of the eye field triggers cell cycle arrest and photoreceptor specification in a wave known as the morphogenetic furrow, to initiate the assembly of ommatidia (Wolff and Ready, 1991; Wolff and Ready, 1993).

1.3.2 The progression of the morphogenetic furrow (MF) initiates cell fate specification in the eye imaginal disc

The initiation and progression of the MF, which sweeps progressively across the eye disc from posterior to anterior, is driven by a network of signaling pathways. The initiator is the secreted Hedgehog (Hh) signal protein expressed in cells at the posterior margin of the eye disc prior to MF initiation (Heberlein and Mosest, 1995). Cells receiving the Hh signal anterior to the MF turn on the expression of a transcription factor Atonal (Ato), which leads to the establishment of the R8 photoreceptor as the founder cell of each ommatidium. R8 cells then secrete Spitz, which activates the EGFR/ Ras/ MAPK signaling cascade to turn on Pnt expression in the neighboring cells (Freeman, 1994; Tio et al., 1994). Pnt stimulates photoreceptor fate specification (O'Neill et

al., 1994) and simultaneously activates Hh expression (Rogers et al., 2005). Therefore, Hh and EGFR signaling via Pnt form a positive feedback loop to drive the progression of the MF.

The pace of MF progression is controlled by Decapentaplegic (Dpp) and Notch signaling pathways. Hh activates the expression of Dpp within the MF (Heberlein et al., 1993). Dpp limits the pace of MF progression through several repressors anterior to the MF. Firstly, Dpp acts at a long range signal to repress the expression of Homothorax, a transcription factor that prevents retinal cell specification in the anterior of the eye disc (Bessa et al., 2002). Such derepression allows cell to enter a preproneural state in which the cells are able to respond to Hh. Meanwhile, within a short range anterior to the MF, Dpp activates the expression of Hairy, which represses the expression of Ato and prevents premature differentiation of retinal cells (Greenwood and Struhl, 1999). To relieve the repression of Ato, Hh and Dpp together stimulate a ligand Delta to activate the Notch signaling pathway, which negatively regulates Hairy expression (Baonza and Freeman, 2001; Baonza and Freeman, 2005). Hh together with Dpp and Notch signaling therefore achieves precise spatial control of the anterior side of the MF and thus determines the onset of ommatidial cell fate specification.

1.3.3 The second mitotic wave (SMW) splits the two rounds of photoreceptor specification

Specification of the eight photoreceptors occurs in two rounds that are separated by a synchronized wave of cell proliferation referred to as the second mitotic wave (SMW) . Derived from a small number of embryonic retinal precursor cells, the *Drosophila* retina experiences two waves of cell proliferation, termed the first and second mitotic waves (Kumar, 2011). The first wave describes the continuous asynchronous cell proliferation starting from the first instar larval to the third instar larval stages that drives tissue growth. In early third instar eye disc, as the MF

initiates, cells arrest in G1, marking the cessation of the first mitotic wave (Ready and Hanson, 1976). As the MF passes, five-cell preclusters that will differentiate into R8, R2 and R5 pair, R3 and R4 pair emerge. This defines the first round of photoreceptor specification. After that, the undifferentiated precursor cells reenter the cell cycle and undergo a single synchronous round of division, the SMW. Upon resumption of G1 arrest the second round of specification initiates, starting with the R1, R6 and R7 photoreceptors and continuing with the cone, pigment and bristle cells. Despite their distinct biological functions in the retina, the precursors of the R1, R6, R7 photoreceptors and the four cone cells form a seven-cell equivalence group such that modulation of signals directing their cell fate specification can transform any one of them into any other (Tomlinson et al., 2011). Although disruption of the SMW does not cause photoreceptor specification defects (de Nooij and Hariharan, 1995), the transition may facilitate the resetting of signaling cues and profiles of gene expression, to confer spatiotemporal specific regulation.

1.3.4 MAPK signaling, via Pnt and Yan, directs both rounds of photoreceptor specification

The R8 photoreceptor serves as the organizer cell of each ommatidium and initiates the sequential recruitment of the other seven photoreceptors and the four cone cells. All the cell fates except R8 are induced by Spi-stimulated EGFR activation. Thus eye mosaic clones lacking Spitz or EGFR only contain the R8 cells (Freeman, 1994; Freeman, 1996; Tio et al., 1994), and the hyperactivation of EGFR signaling over-recruits all cell types except R8 (Freeman, 1996). Despite their dependence on a generic inductive signal, the precise temporal sequence of fate specification is achieved through the short-range action of Spitz, which is limited to 1-2 cell diameters. Although the inactive precursor of Spitz is expressed in all cells, its secretion requires processing by the chaperone protein Star and the protease Rhomboid (Freeman et al., 1992;

Wasserman et al., 2000), the expression of which is regulated by Ato in R8 cells. Then secreted and activated Spitz from R8 induces the specification of R2 and R5 in cells immediately neighboring to R8 cells. Specified R2 and R5 are able to express Star and Rhomboid and thus process Spitz to induce the specification of R3 and R4 photoreceptors, which will recruit the subsequent photoreceptors in the same way (Shilo, 2005b).

In additional to the EGFR RTK, Sevenless (Sev), another RTK that is activated by its ligand Bride of Sevenless (Boss) rather than Spitz, also contributes to photoreceptor recruitment. The expression of Sev is dynamic, such that it strongly expresses in R3, R4 and the precluster of R1, R6 and R7 and cone progenitor cells but then its expression is reduced to background level once R1 and R6 are specified and only remains high in R7 and cone cells (Tomlinson et al., 1987). R8 expresses the ligand Boss (Reinke and Zipursky, 1988). Restricted by the spatial arrangements of the photoreceptors, only R1, R6 and R7 progenitors are in contact with R8 and thus have the potential to activate Sev. However, Sev activation is required for R7 but not R1, R6 specification (Hafen et al., 1993; Reinke and Zipursky, 1988; Tomlinson and Ready, 1986). Mutants of Boss or Sev showed loss of R7 cells. Despite different ligands, Sev and EGFR share the downstream signal transduction machinery and are thus interchangeable. For example, constitutively active EGFR can substitute Sev to recruit R7 in the *sev* mutant (Freeman, 1996).

Yan, PntP1 and PntP2, in concert with other transcription factors, mediate MAPK activation to direct specific photoreceptor fates except R8. Yan has a predominant role in inhibiting over-recruitment of photoreceptor cells (Lai and Rubin, 1992). With respect to Pnt, the sequential activation of PntP1 and PntP2 stimulated by MAPK phosphorylation, as mentioned above, is

required for R1-R7 cell fate specification (Shwartz et al., 2013). Losing Spitz ligand, PntP1 or PntP2 all results in loss of R1-R7 photoreceptors.

How does this generic Pnt-mediated response to MAPK activation lead to different photoreceptor types? Great work has been done to demonstrate the cooperation between Pnt and other factors to regulate the expression of different target genes (Halfon et al., 2000; Hayashi et al., 2008; Tomlinson et al., 2011). The second-round photoreceptor specification serves a good example to illustrate the cooperative action. The precursors within the equivalence group of R1, R6, R7 and cone cells all express Lz and several eye selector geness, including So/Eya and Gl. However, on top of EGFR and these factors, only R7 cells contain both high levels of Notch and Sev (Basler and Hafen, 1988; Diaz-Benjumea and Hafen, 1994). Losing either Notch or Sev leads to fate switch transitions in which R7 precursor adopt ectopic R1/R6 or cone cell fates, respectively (Tomlinson et al., 2011). This collaborative action of EGFR, Sev, Notch plus the transcription effectors Lz, So/Eya and Gl takes effect at the enhancer of the *prospero* gene, whose expression marks the R7 and cone cell fate (Hayashi et al., 2008).

Adding to the complexity of such collaborative regulation, the organization of transcription factor binding sites within an enhancer also plays critical roles in transcriptional regulation. For example, reducing the spacing between two transcription factor binding sites may alter the transcriptional activity of the enhancer (Erceg et al., 2014). The pattern of the number, organization, spacing, affinity and orientation of transcription factor binding sites in an enhancer is termed the cis-regulatory logic. This will be discussed in more detail in Chapter 3.

While the importance of Pnt for photoreceptor specification has been clearly established, Pnt responses to regulate different transcriptional programs in accordance with distinctive

developmental contexts have not been fully investigated. In particular, the upstream signaling cues can be distinct, such as EGFR plus Sev regulating the second round photoreceptor fates while EGFR being the sole RTK source for the first round photoreceptor fates. Does RTK signaling regulate the Pnt response in an 'all-or-none' fashion, such that phosphorylation simply turns on the Pnt response? Or does Pnt respond to the distinct signaling environment in a cell type-specific manner. Interestingly, the hypomorphic allele of *pntP2* showed defects only for the second round photoreceptor fates R1, R6 and R7 despite the generic model of sequential activation from PntP2 to PntP1 that has been proposed for R1-R7 fates (Brunner et al., 1994; Shwartz et al., 2013). The sensitivity of the *pntP2* hypomorph may be an indication of differential Pnt response between the two rounds, an idea explored further in Chapter 2.

1.4 Scope of the thesis

This dissertation focuses on the regulatory interactions between Pnt isoforms that contribute to the specificity and robustness of RTK-directed cell fates, and on how interactions between Pnt and Yan at specific enhancers control gene transcription. The methods include both experiments on the bench and in silico calculations. Chapter 2 describes how PntP3, a Pnt isoform that had not been characterized before, enables the Pnt network to adopt distinct topological organization optimized for the unique signaling environments associated with each of the two rounds of photoreceptor specification. Chapter 3 describes the development of a thermodynamic model that I used to explore how the number, spacing and affinity of transcription factor binding sites within a gene enhancer organizes inputs from Yan and Pnt to achieve specific gene expression. Chapter 4 discusses the implications of my discoveries and offers suggestions for future investigation of diverse RTK/MAPK/Pnt/Yan-controlled cell fate transitions.

1.5 References

Ahmad, S. M. (2017). Conserved signaling mechanisms in Drosophila heart development. *Dev. Dyn.* **246**, 641–656.

Arai, H., Maki, K., Waga, K., Sasaki, K., Nakamura, Y., Imai, Y., Kurokawa, M., Hirai, H. and Mitani, K. (2002). Functional regulation of TEL by p38-induced phosphorylation. *Biochem. Biophys. Res. Commun.* **299**, 116–125.

Baker, D. A., Mille-Baker, B., Wainwright, S. M., Ish-Horowicz, D. and Dibb, N. J. (2001). Mae mediates MAP kinase phosphorylation of Ets transcription factors in Drosophila. *Nature* **411**, 330–334.

Bang, A. G. and Kintner, C. (2000). Rhomboid and Star facilitate presentation and processing of the Drosophila TGF-alpha homolog Spitz. *Genes Dev.* **14**, 177–186.

Baonza, A. and Freeman, M. (2001). Notch signalling and the initiation of neural development in the Drosophila eye. *Development* **128**, 3889 LP – 3898.

Baonza, A. and Freeman, M. (2005). Control of Cell Proliferation in the Drosophila Eye by Notch Signaling. *Dev. Cell* **8**, 529–539.

Basler, K. and Hafen, E. (1988). Sevenless and Drosophila eye development: a tyrosine kinase controls cell fate. *Trends Genet.* **4**,.

Beiman, M., Shilo, B. Z. and Volk, T. (1996). Heartless, a Drosophila FGF receptor homolog, is essential for cell migration and establishment of several mesodermal lineages. *Genes Dev.* **10**, 2993–3002.

Bessa, J., Gebelein, B., Pichaud, F., Casares, F. and Mann, R. S. (2002). Combinatorial control of Drosophila eye development by Eyeless, Homothorax, and Teashirt. *Genes Dev.*

2415–2427.

Bhanot, P., Fish, M., Jemison, J. A., Nusse, R., Nathans, J. and Cadigan, K. M. (1999). Frizzled and Dfrizzled-2 function as redundant receptors for Wingless during Drosophila embryonic development. *Development* **126**, 4175 LP – 4186.

Bodmer, R. (1993). The gene tinman is required for specification of the heart and visceral muscles in Drosophila. *Development* **118**, 719 LP – 729.

Bodmer, R. (1995). Heart development in Drosophila and its relationship to vertebrates. *Trends Cardiovasc. Med.* **5**, 21–28.

Bohush, A., Niewiadomska, G. and Filipek, A. (2018). Role of mitogen activated protein kinase signaling in parkinson's disease. *Int. J. Mol. Sci.* **19**,.

Boisclair Lachance, J.-F., Peláez, N., Cassidy, J. J., Webber, J. L., Rebay, I. and Carthew, R. W. (2014). A comparative study of Pointed and Yan expression reveals new complexity to the transcriptional networks downstream of receptor tyrosine kinase signaling. *Dev. Biol.* **385**, 263–278.

Boulton, T. G., Yancopoulos, G. D., Gregory, J. S., Slaughter, C., Moomaw, C., Hsu, J. and Cobb, M. H. (1990). An insulin-stimulated protein kinase similar to yeast kinases involved in cell cycle control. *Science (80-.).* **249**, 64 LP – 67.

Brunner, D., Dücker, K., Oellers, N., Hafen, E., Scholz, H. and Klämbt, C. (1994). The ETS domain protein pointed-P2 is a target of MAP kinase in the sevenless signal transduction pathway. *Nature* **370**, 386–389.

Burdon, T., Stracey, C., Chambers, I., Nichols, J. and Smith, A. (1999). Suppression of SHP-2 and ERK Signalling Promotes Self-Renewal of Mouse Embryonic Stem Cells. *Dev. Biol.*

210, 30–43.

Campbell, G. (2002). Distalization of the Drosophila leg by graded EGF-receptor activity. *Nature* **418**, 781–785.

Cargnello, M. and Roux, P. P. (2011). Activation and Function of the MAPKs and Their Substrates, the MAPK-Activated Protein Kinases. *Microbiol. Mol. Biol. Rev.* **75**, 50–83.

Carmena, A., Gisselbrecht, S., Harrison, J., Jiménez, F. and Michelson, A. M. (1998). Combinatorial signaling codes for the progressive determination of cell fates in the Drosophila embryonic mesoderm. *Genes Dev.* **12**, 3910–3922.

Carmena, A., Buff, E., Halfon, M. S., Gisselbrecht, S., Jiménez, F., Baylies, M. K. and Michelson, A. M. (2002). Reciprocal regulatory interactions between the Notch and Ras signaling pathways in the Drosophila embryonic mesoderm. *Dev. Biol.* **244**, 226–242.

Casie Chetty, S. and Sumanas, S. (2020). Ets1 functions partially redundantly with Etv2 to promote embryonic vasculogenesis and angiogenesis in zebrafish. *Dev. Biol.* **465**, 11–22.

Chambers, M. C., Song, K. H. and Schneider, D. S. (2012). Listeria monocytogenes Infection Causes Metabolic Shifts in Drosophila melanogaster. *PLoS One* **7**, e50679.

Chen, N. and Greenwald, I. (2004). The lateral signal for LIN-12/Notch in C. elegans vulval development comprises redundant secreted and transmembrane DSL proteins. *Dev. Cell* **6**, 183–192.

Chen, C. M. and Struhl, G. (1999). Wingless transduction by the Frizzled and Frizzled2 proteins of Drosophila. *Development* **126**, 5441 LP – 5452.

Chen, R. E. and Thorner, J. (2007). Function and regulation in MAPK signaling pathways: lessons learned from the yeast Saccharomyces cerevisiae. *Biochim. Biophys. Acta* **1773**,

1311–1340.

Chen, Z., Gibson, T. B., Robinson, F., Silvestro, L., Pearson, G., Xu, B., Wright, A., Vanderbilt, C. and Cobb, M. H. (2001). MAP Kinases. *Chem. Rev.* **101**, 2449–2476.

Clifford, R. J. and Schüpbach, T. (1989). Coordinately and differentially mutable activities of torpedo, the Drosophila melanogaster homolog of the vertebrate EGF receptor gene. *Genetics* **123**, 771–787.

Cooper, J. A., Bowen-Pope, D. F., Raines, E., Ross, R. and Hunter, T. (1982). Similar effects of platelet-derived growth factor and epidermal growth factor on the phosphorylation of tyrosine in cellular proteins. *Cell* **31**, 263–273.

Craig, M. P. and Sumanas, S. (2016). ETS transcription factors in embryonic vascular development. *Angiogenesis* **19**, 275–285.

Czerny, T., Halder, G., Kloter, U., Souabni, A., Gehring, W. J. and Busslinger, M. (1999). twin of eyeless, a Second Pax-6 Gene of Drosophila, Acts Upstream of eyeless in the Control of Eye Development. *Mol. Cell* **3**, 297–307.

Davis, T. L. and Rebay, I. (2017). Master regulators in development: Views from the Drosophila retinal determination and mammalian pluripotency gene networks. *Dev. Biol.* **421**, 93–107.

de Nooij, J. C. and Hariharan, I. K. (1995). Uncoupling cell fate determination from patterned cell division in the Drosophila eye. *Science* **270**, 983–985.

Diaz-Benjumea, F. J. and Hafen, E. (1994). The sevenless signalling cassette mediates Drosophila EGF receptor function during epidermal development. *Development* **120**, 569–

578.

Dutta, D., Shaw, S., Maqbool, T., Pandya, H. and VijayRaghavan, K. (2005). Drosophila Heartless Acts with Heartbroken/Dof in Muscle Founder Differentiation. *PLOS Biol.* **3**, e337.

Erceg, J., Saunders, T. E., Girardot, C., Devos, D. P., Hufnagel, L. and Furlong, E. E. M. (2014). Subtle Changes in Motif Positioning Cause Tissue-Specific Effects on Robustness of an Enhancer's Activity. *PLoS Genet.* **10**,.

Findlay, V. J., LaRue, A. C., Turner, D. P., Watson, P. M. and Watson, D. K. (2013). *Understanding the role of ETS-mediated gene regulation in complex biological processes*. 1st ed. Elsevier Inc.

Fiore, C. and Cohen, B. A. (2016). Interactions between pluripotency factors specify cis-regulation in embryonic stem cells. *Genome Res.* **26**, 778–786.

Fitzsimmons, D., Hodsdon, W., Wheat, W., Maira, S. M., Wasylyk, B. and Hagman, J. (1996). Pax-5 (BSAP) recruits ets proto-oncogene family proteins to form functional ternary complexes on a B-cell-specific promoter. *Genes Dev.* **10**, 2198–2211.

Fletcher, J. C., D'Avino, P. P. and Thummel, C. S. (1997). A steroid-triggered switch in E74 transcription factor isoforms regulates the timing of secondary-response gene expression. *Proc. Natl. Acad. Sci.* **94**, 4582 LP – 4586.

Flores, G. V., Duan, H., Yan, H., Nagaraj, R., Fu, W., Zou, Y., Noll, M. and Banerjee, U. (2000). Combinatorial signaling in the specification of unique cell fates. *Cell* **103**, 75–85.

Frasch, M. (1995). Induction of visceral and cardiac mesoderm by ectodermal Dpp in the early Drosophila embryo. *Nature* **374**, 464–467.

Freeman, M. (1994). The spitz gene is required for photoreceptor determination in the Drosophila eye where it interacts with the EGF receptor. *Mech. Dev.* **48**, 25–33.

Freeman, M. (1996). Reiterative Use of the EGF Receptor Trigger Differentiation of All Cell Types in the Drosophila Eye. *Cell* **87**, 651–660.

Freeman, M., Kimmel, B. E. and Rubin, G. M. (1992). Identifying targets of the rough homeobox gene of Drosophila: evidence that rhomboid functions in eye development. *Development* **116**, 335 LP – 346.

Gabay, L., Scholz, H., Golembo, M., Klaes, a, Shilo, B. Z. and Klämbt, C. (1996). EGF receptor signaling induces pointed P1 transcription and inactivates Yan protein in the Drosophila embryonic ventral ectoderm. *Development* **122**, 3355–3362.

Gajewski, K. M. and Schulz, R. A. (1995). Requirement of the ETS domain transcription factor D-ELG for egg chamber patterning and development during Drosophila oogenesis. *Oncogene* **11**, 1033—1040.

Gerhart, J. (1999). 1998 Warkany Lecture: Signaling pathways in development. *Teratology* **60**, 226–239.

Gille, H., Sharrocks, A. D. and Shaw, P. E. (1992). Phosphorylation of transcription factor p62TCF by MAP kinase stimulates ternary complex formation at c-fos promoter. *Nature* **358**, 414–417.

Golembo, M., Schweitzer, R., Freeman, M. and Shilo, B. Z. (1996). Argos transcription is induced by the Drosophila EGF receptor pathway to form an inhibitory feedback loop. *Development* **122**, 223 LP – 230.

Golembo, M., Yarnitzky, T., Volk, T. and Shilo, B. Z. (1999). Vein expression is induced by

the EGF receptor pathway to provide a positive feedback loop in patterning the Drosophila embryonic ventral ectoderm. *Genes Dev.* **13**, 158–162.

Graves, B. J. and Petersen, J. M. (1998). Specificity within the ets Family of Transcription Factors. In (ed. Vande Woude, G. F.) and Klein, G. B. T.-A. in C. R.), pp. 1–57. Academic Press.

Greenwood, S. and Struhl, G. (1999). Progression of the morphogenetic furrow in the Drosophila eye: the roles of Hedgehog, Decapentaplegic and the Raf pathway. *Development* **126**, 5795 LP – 5808.

Grigorian, M., Mandal, L., Hakimi, M., Ortiz, I. and Hartenstein, V. (2011). The convergence of Notch and MAPK signaling specifies the blood progenitor fate in the Drosophila mesoderm. *Dev. Biol.* **353**, 105–118.

Hafen, E., Dickson, B., Raabe, T., Brunner, D., Oellers, N. and van der Straten, A. (1993). Genetic analysis of the sevenless signal transduction pathway of Drosophila. *Development* **119**, 41 LP – 46.

Halder, G., Callaerts, P., Flister, S., Walldorf, U., Kloter, U. and Gehring, W. J. (1998). Eyeless initiates the expression of both sine oculis and eyes absent during Drosophila compound eye development. *Development* **125**, 2181 LP – 2191.

Halfon, M. S., Carmena, A., Gisselbrecht, S., Sackerson, C. M., Jiménez, F., Baylies, M. K. and Michelson, A. M. (2000). Ras pathway specificity is determined by the integration of multiple signal-activated and tissue-restricted transcription factors. *Cell* **103**, 63–74.

Halfon, M. S., Grad, Y., Church, G. M. and Michelson, A. M. (2002). Computation-based discovery of related transcriptional regulatory modules and motifs using a experimentally

validated combinatorial model. *Genome Res.* **12**, 1019–1028.

Hart, A., Melet, F., Grossfeld, P., Chien, K., Jones, C., Tunnacliffe, A., Favier, R. and Bernstein, A. (2000). Fli-1 is required for murine vascular and megakaryocytic development and is hemizygously deleted in patients with thrombocytopenia. *Immunity* **13**, 167–177.

Hasson, P., Egoz, N., Winkler, C., Volohonsky, G., Jia, S., Dinur, T., Volk, T., Courey, A. J. and Paroush, Z. (2005). EGFR signaling attenuates Groucho-dependent repression to antagonize Notch transcriptional output. *Nat. Genet.* **37**, 101–105.

Hayashi, T., Xu, C. and Carthew, R. W. (2008). Cell-type-specific transcription of prospero is controlled by combinatorial signaling in the Drosophila eye. *Development* **135**, 2787–2796.

Heberlein, U. and Mosest, K. (1995). Mechanisms of drosophila retinal morphogenesis: The virtues of being progressive. *Cell* **81**, 987–990.

Heberlein, U., Wolff, T. and Rubin, G. M. (1993). The TGFβ homolog dpp and the segment polarity gene hedgehog are required for propagation of a morphogenetic wave in the Drosophila retina. *Cell* **75**, 913–926.

Hertel, K. J., Lynch, K. W. and Maniatis, T. (1997). Common themes in the function of transcription and splicing enhancers. *Curr. Opin. Cell Biol.* **9**, 350–357.

Hollenhorst, P. C., McIntosh, L. P. and Graves, B. J. (2011). Genomic and Biochemical Insights into the Specificity of ETS Transcription Factors. *Annu. Rev. Biochem.* **80**, 437–471.

Hsu, T. and Schulz, R. A. (2000). Sequence and functional properties of Ets genes in the model organism Drosophila. *Oncogene* **19**, 6409–6416.

Huang, P. and Stern, M. J. (2005). FGF signaling in flies and worms: More and more relevant to vertebrate biology. *Cytokine Growth Factor Rev.* **16**, 151–158.

Ikeya, T. and Hayashi, S. (1999). Interplay of Notch and FGF signaling restricts cell fate and MAPK activation in the Drosophila trachea. *Development* **126**, 4455–4463.

Jean-François Boisclair Lachance, Nicolás Peláez, Justin J. Cassidy, J. L. and Webber, Ilaria Rebay, and R. W. C. (2014). A comparative study of Pointed and Yan expression reveals new complexity to the transcriptional networks downstream of receptor tyrosine kinase signaling. *Dev Biol.* **385**, 263–278.

Johnson, A. D., Meyer, B. J. and Ptashne, M. (1979). Interactions between DNA-bound repressors govern regulation by the λ phage repressor. *Proc. Natl. Acad. Sci.* **76**, 5061 LP – 5065.

Jonsen, M. D., Petersen, J. M., Xu, Q. P. and Graves, B. J. (1996). Characterization of the cooperative function of inhibitory sequences in Ets-1. *Mol. Cell. Biol.* **16**, 2065–2073.

Kadam, S., McMahon, A., Tzou, P. and Stathopoulos, A. (2009). FGF ligands in Drosophila have distinct activities required to support cell migration and differentiation. *Development* **136**, 739–747.

Kaul, A., Schuster, E. and Jennings, B. H. (2014). The Groucho Co-repressor Is Primarily Recruited to Local Target Sites in Active Chromatin to Attenuate Transcription. *PLOS Genet.* **10**, e1004595.

Kenyon, K. L., Ranade, S. S., Curtiss, J., Mlodzik, M. and Pignoni, F. (2003). Coordinating Proliferation and Tissue Specification to Promote Regional Identity in the Drosophila Head. *Dev. Cell* **5**, 403–414.

Kim, W. Y., Sieweke, M., Ogawa, E., Wee, H. J., Englmeier, U., Graf, T. and Yoshiaki, I. (1999). Mutual activation of Ets-1 and AML1 DNA binding by direct interaction of their autoinhibitory domains. *EMBO J.* **18**, 1609–1620.

Klambt, C. (1993). The Drosophila gene pointed encodes two ETS-like proteins which are involved in the development of the midline glial cells. *Development* **117**, 163–176.

Klingseisen, A., Clark, I. B. N., Gryzik, T. and Müller, H.-A. J. (2009). Differential and overlapping functions of two closely related Drosophila FGF8-like growth factors in mesoderm development. *Development* **136**, 2393–2402.

Kola, I., Brookes, S., Green, A. R., Garber, R., Tymms, M., Papas, T. S. and Seth, A. (1993). The Ets1 transcription factor is widely expressed during murine embryo development and is associated with mesodermal cells involved in morphogenetic processes such as organ formation. *Proc. Natl. Acad. Sci. U. S. A.* **90**, 7588–7592.

Kumar, J. P. (2011). My what big eyes you have: How the Drosophila retina grows. *Dev. Neurobiol.* **71**, 1133–1152.

Kumar, J. P. and Moses, K. (2001). EGF Receptor and Notch Signaling Act Upstream of Eyeless/Pax6 to Control Eye Specification. *Cell* **104**, 687–697.

Kurada, P. and White, K. (1998). Ras Promotes Cell Survival in Drosophila by Downregulating hid Expression. *Cell* **95**, 319–329.

Kyriakis, J. M. and Avruch, J. (2001). Mammalian Mitogen-Activated Protein Kinase Signal Transduction Pathways Activated by Stress and Inflammation. *Physiol. Rev.* **81**, 807–869.

Lai, Z. C. and Rubin, G. M. (1992). Negative control of photoreceptor development in Drosophila by the product of the yan gene, an ETS domain protein. *Cell* **70**, 609–620.

Laudet, V., Hänni, C., Stéhelin, D. and Duterque-Coquillaud, M. (1999). Molecular phylogeny of the ETS gene family. *Oncogene* **18**, 1351–1359.

Lee, D., Park, C., Lee, H., Lugus, J. J., Kim, S. H., Arentson, E., Chung, Y. S., Gomez, G., Kyba, M., Lin, S., et al. (2008). ER71 Acts Downstream of BMP, Notch, and Wnt Signaling in Blood and Vessel Progenitor Specification. *Cell Stem Cell* **2**, 497–507.

Leprince, D., Gegonne, A., Coll, J., De Taisne, C., Schneeberger, A., Lagrou, C. and Stehelin, D. (1983). A putative second cell-derived oncogene of the avian leukaemia retrovirus E26. *Nature* **306**, 395–397.

Liang, H., Mao, X., Olejniczak, E. T., Nettesheim, D. G., Yu, L., Meadows, R. P., Thompson, C. B. and Fesik, S. W. (1994). Solution structure of the ets domain of fli-1 when bound to DNA. *Nat. Struct. Biol.* **1**, 871–876.

Lifanov, A. P., Makeev, V. J., Nazina, A. G. and Papatsenko, D. A. (2003). Homotypic regulatory clusters in Drosophila. *Genome Res.* **13**, 579–588.

Lin, J. H., Saito, T., Anderson, D. J., Lance-Jones, C., Jessell, T. M. and Arber, S. (1998). Functionally Related Motor Neuron Pool and Muscle Sensory Afferent Subtypes Defined by Coordinate ETS Gene Expression. *Cell* **95**, 393–407.

Liu, P., Wakamiya, M., Shea, M. J., Albrecht, U., Behringer, R. R. and Bradley, A. (1999). Requirement for Wnt3 in vertebrate axis formation. *Nat. Genet.* **22**, 361–365.

Liu, J., Qian, L., Wessells, R. J., Bidet, Y., Jagla, K. and Bodmer, R. (2006). Hedgehog and RAS pathways cooperate in the anterior–posterior specification and positioning of cardiac progenitor cells. *Dev. Biol.* **290**, 373–385.

Lockwood, W. K. and Bodmer, R. (2002). The patterns of wingless, decapentaplegic, and

tinman position the Drosophila heart. *Mech. Dev.* **114**, 13–26.

Maroulakou, I. G. and Bowe, D. B. (2000). Expression and function of Ets transcription factors in mammalian development: A regulatory network. *Oncogene* **19**, 6432–6442.

Mélet, F., Motro, B., Rossi, D. J., Zhang, L. and Bernstein, A. (1996). Generation of a novel Fli-1 protein by gene targeting leads to a defect in thymus development and a delay in Friend virus-induced erythroleukemia. *Mol. Cell. Biol.* **16**, 2708–2718.

Michelson, A. M., Gisselbrecht, S., Zhou, Y., Baek, K.-H. and Buff, E. M. (1998). Dual functions of the heartless fibroblast growth factor receptor in development of the Drosophila embryonic mesoderm. *Dev. Genet.* **22**, 212–229.

Mo, Y., Vaessen, B., Johnston, K. and Marmorstein, R. (2000). Structure of the Elk-1-DNA complex reveals how DNA-distal residues affect ETS domain recognition of DNA. *Nat. Struct. Biol.* **7**, 292–297.

Molina, J. R. and Adjei, A. A. (2006). The Ras/Raf/MAPK Pathway. *J. Thorac. Oncol.* **1**, 7–9.

Nie, Y., Shu, C. and Sun, X. (2020). Cooperative binding of transcription factors in the human genome. *Genomics* **112**, 3427–3434.

O'Neill, E. M. D., Tjian, R. and Rubin, G. M. (1994). The Activities of Two Ets-Related Transcription Factors Required for Drosophila Eye Development Are Modulated by the Ras / MAPK Pathway. *Cell* **78**, 137–147.

Papoutsopoulou, S. and Janknecht, R. (2000). Phosphorylation of ETS transcription factor ER81 in a complex with its coactivators CREB-binding protein and p300. *Mol. Cell. Biol.* **20**, 7300–7310.

Park, M., Wu, X., Golden, K., Axelrod, J. D. and Bodmer, R. (1996). The Wingless Signaling

Pathway Is Directly Involved inDrosophilaHeart Development. *Dev. Biol.* **177**, 104–116.

Perrimon, N. and Perkins, L. A. (1997). There must be 50 ways to rule the signal: The case of the Drosophila EGF receptor. *Cell* **89**, 13–16.

Perrimon, N., Pitsouli, C. and Shilo, B.-Z. (2012). Signaling mechanisms controlling cell fate and embryonic patterning. *Cold Spring Harb. Perspect. Biol.* **4**, a005975–a005975.

Qiao, F., Harada, B., Song, H., Whitelegge, J., Courey, A. J. and Bowie, J. U. (2006). Mae inhibits Pointed-P2 transcriptional activity by blocking its MAPK docking site. *EMBO J.* **25**, 70–79.

Ravasi, T., Suzuki, H., Cannistraci, C. V., Katayama, S., Bajic, V. B., Tan, K., Akalin, A., Schmeier, S., Kanamori-Katayama, M., Bertin, N., et al. (2010). An atlas of combinatorial transcriptional regulation in mouse and man. *Cell* **140**, 744–752.

Ray, L. B. and Sturgill, T. W. (1988). Insulin-stimulated microtubule-associated protein kinase is phosphorylated on tyrosine and threonine in vivo. *Proc. Natl. Acad. Sci. U. S. A.* **85**, 3753–3757.

Ready, F. and Hanson, E. (1976). Developmen of the Drosophila Retina , a Neurocrystalline. *Dev. Biol.* **240**,.

Rebay, I. and Rubin, G. M. (1995). Yan functions as a general inhibitor of differentiation and is negatively regulated by activation of the Ras1/MAPK pathway. *Cell* **81**, 857–866.

Reinke, R. and Zipursky, S. L. (1988). Cell-cell interaction in the drosophila retina: The bride of sevenless gene is required in photoreceptor cell R8 for R7 cell development. *Cell* **55**, 321–330.

Reiter, F., Wienerroither, S. and Stark, A. (2017). Combinatorial function of transcription

factors and cofactors. *Curr. Opin. Genet. Dev.* **43**, 73–81.

Rogers, E. M., Brennan, C. A., Mortimer, N. T., Cook, S., Morris, A. R. and Moses, K. (2005). Pointed regulates an eye-specific transcriptional enhancer in the Drosophila hedgehog gene, which is required for the movement of the morphogenetic furrow. *Development* **132**, 4833–4843.

Rohrbaugh, M., Ramos, E., Nguyen, D., Price, M., Wen, Y. and Lai, Z. C. (2002). Notch activation of yan expression is antagonized by RTK/pointed signaling in the Drosophila eye. *Curr. Biol.* **12**, 576–581.

Rossomando, A. J., Payne, D. M., Weber, M. J. and Sturgill, T. W. (1989). Evidence that pp42, a major tyrosine kinase target protein, is a mitogen-activated serine/threonine protein kinase. *Proc. Natl. Acad. Sci. U. S. A.* **86**, 6940–6943.

Schnepp, B., Grumbling, G., Donaldson, T. and Simcox, A. (1996). Vein is a novel component in the Drosophila epidermal growth factor receptor pathway with similarity to the neuregulins. *Genes Dev.* **10**, 2302–2313.

Scholz, H., Deatrick, J., Klaes, A. and Klambt, C. (1993). Genetic dissection of pointed, a Drosophila gene encoding two ETS-related proteins. *Genetics* **135**, 455–468.

Schweisguth, F. and Posakony, J. W. (1992). Suppressor of Hairless, the Drosophila homolog of the mouse recombination signal-binding protein gene, controls sensory organ cell fates. *Cell* **69**, 1199–1212.

Sharrocks, A. D. (2001). The ETS-domain transcription factor family. *Nat. Rev. Mol. Cell Biol.* **2**, 827–837.

Sharrocks, A. D., Brown, A. L., Ling, Y. and Yates, P. R. (1997). The ETS-domain

transcription factor family. *Int. J. Biochem. Cell Biol.* **29**, 1371–1387.

Shilo, B. Z. (2005a). Regulating the dynamics of EGF receptor signaling in space and time. *Development* **132**, 4017–4027.

Shilo, B.-Z. (2005b). Regulating the dynamics of EGF receptor signaling in space and time. *Development* **132**, 4017 LP – 4027.

Shwartz, A., Yogev, S., Schejter, E. D. and Shilo, B. (2013). Sequential activation of ETS proteins provides a sustained transcriptional response to EGFR signaling. *Development* **140**, 2746–2754.

Simcox, A. A., Grumbling, G., Schnepp, B., Bennington-Mathias, C., Hersperger, E. and Shearn, A. (1996). Molecular, Phenotypic, and Expression Analysis ofvein,a Gene Required for Growth of theDrosophilaWing Disc. *Dev. Biol.* **177**, 475–489.

Skov, M. and Høegh, R. T. (2002). Three habits of highly effective signaling pathways: principles of transcriptional control by developmental cell signaling. *Genes Dev.* **16**, 1167–1181.

Sopko, R. and Perrimon, N. (2013). Receptor tyrosine kinases in Drosophila development. *Cold Spring Harb. Perspect. Biol.* **5**, 1–32.

Spyropoulos, D. D., Pharr, P. N., Lavenburg, K. R., Jackers, P., Papas, T. S., Ogawa, M. and Watson, D. K. (2000). Hemorrhage, Impaired Hematopoiesis, and Lethality in Mouse Embryos Carrying a Targeted Disruption of the Fli1Transcription Factor. *Mol. Cell. Biol.* **20**, 5643–5652.

Sundaram, M. V. (2005). The love-hate relationship between Ras and Notch. *Genes Dev.* **19**, 1825–1839.

Szymczyna, B. and Arrowsmith, C. (2000). DNA Binding Specificity Studies of Four ETS Proteins Support an Indirect Read-out Mechanism of Protein-DNA Recognition. *J. Biol. Chem.* **275**, 28363–28370.

Tio, M., Ma, C. and Moses, K. (1994). spitz, a Drosophila homolog of transforming growth factor-alpha, is required in the founding photoreceptor cells of the compound eye facets. *Mech. Dev.* **48**, 13–23.

Tomlinson, A. and Ready, D. F. (1986). Sevenless A Cell-Specific Homeotic Mutation of the Drosophila Eye. *Science (80-.).* **231**, 400 LP – 402.

Tomlinson, A. and Ready, D. F. (1987). Neuronal differentiation in the Drosophila ommatidium. *Dev. Biol.* **120**, 366–376.

Tomlinson, A., Bowtell, D. D. L., Hafen, E. and Rubin, G. M. (1987). Localization of the sevenless protein, a putative receptor for positional information, in the eye imaginal disc of Drosophila. *Cell* **51**, 143–150.

Tomlinson, A., Mavromatakis, Y. E. and Struhl, G. (2011). Three distinct roles for Notch in Drosophila R7 photoreceptor specification. *PLoS Biol.* **9**,.

Tootle, T. L., Lee, P. S. and Rebay, I. (2003). CRM1-mediated nuclear export and regulated activity of the receptor tyrosine kinase antagonist YAN require specific interactions with MAE. *Development* **130**, 845–857.

Tsuda, L., Nagaraj, R., Zipursky, S. L. and Banerjee, U. (2002). An EGFR/Ebi/Sno pathway promotes Delta expression by inactivating Su(H)/SMRTER repression during inductive Notch signaling. *Cell* **110**, 625–637.

van der Geer P, Hunter T, L. R. (1994). Receptor protein-tyrosine kinases and their signal

transduction pathways. *Annu. Rev. Cell Biol.* **10**, 251–337.

Vivekanand, P. (2018). Lessons from Drosophila Pointed, an ETS family transcription factor and key nuclear effector of the RTK signaling pathway. *Genesis* **56**, 1–13.

Vivekanand, P. and Rebay, I. (2006). Intersection of signal transduction pathways and development. *Annu. Rev. Genet.* **40**, 139–57.

Vivekanand, P., Tootle, T. L. and Rebay, I. (2004). MAE, a dual regulator of the EGFR signaling pathway, is a target of the Ets transcription factors PNT and YAN. *Mech. Dev.* **121**, 1469–1479.

Vlaeminck-Guillem, V., Carrere, S., Dewitte, F., Stehelin, D., Desbiens, X. and Duterque-Coquillaud, M. (2000). The Ets family member Erg gene is expressed in mesodermal tissues and neural crests at fundamental steps during mouse embryogenesis. *Mech. Dev.* **91**, 331–335.

Vogler, G. and Bodmer, R. (2015). Cellular Mechanisms of Drosophila Heart Morphogenesis. *J. Cardiovasc. Dev. Dis.* **2**, 2–16.

Wagner, A. (1997). A computational genomics approach to the identification of gene networks. *Nucleic Acids Res.* **25**, 3594–3604.

Wagner, E. F. and Nebreda, Á. R. (2009). Signal integration by JNK and p38 MAPK pathways in cancer development. *Nat. Rev. Cancer* **9**, 537–549.

Wang, L. C., Kuo, F., Fujiwara, Y., Gilliland, D. G., Golub, T. R. and Orkin, S. H. (1997). Yolk sac angiogenic defect and intra-embryonic apoptosis in mice lacking the Ets-related factor TEL. *EMBO J.* **16**, 4374–4383.

Wang, S. H., Simcox, A. and Campbell, G. (2000). Dual role for Drosophila epidermal growth

factor receptor signaling in early wing disc development. *Genes Dev.* **14**, 2271–2276.

Wareing, S., Eliades, A., Lacaud, G. and Kouskoff, V. (2012). ETV2 expression marks blood and endothelium precursors, including hemogenic endothelium, at the onset of blood development. *Dev. Dyn.* **241**, 1454–1464.

Wasserman, J. D. and Freeman, M. (1998). An Autoregulatory Cascade of EGF Receptor Signaling Patterns the Drosophila Egg. *Cell* **95**, 355–364.

Wasserman, J. D., Urban, S. and Freeman, M. (2000). A family of rhomboid-like genes: Drosophila rhomboid-1 and roughoid/rhomboid-3 cooperate to activate EGF receptor signaling. *Genes Dev.* **14**, 1651–1663.

Webber, J. L., Zhang, J., Massey, A., Sanchez-Luege, N. and Rebay, I. (2018). Collaborative repressive action of the antagonistic ETS transcription factors Pointed and Yan fine-tunes gene expression to confer robustness in Drosophila. *Dev.* **145**,.

Wei, G., Srinivasan, R., Cantemir-Stone, C. Z., Sharma, S. M., Santhanam, R., Weinstein, M., Muthusamy, N., Man, A. K., Oshima, R. G., Leone, G., et al. (2009). Ets1 and Ets2 are required for endothelial cell survival during embryonic angiogenesis. *Blood* **114**, 1123–1130.

Wei, G. H., Badis, G., Berger, M. F., Kivioja, T., Palin, K., Enge, M., Bonke, M., Jolma, A., Varjosalo, M., Gehrke, A. R., et al. (2010). Genome-wide analysis of ETS-family DNA-binding in vitro and in vivo. *EMBO J.* **29**, 2147–2160.

Widmann, C., Spencer, G., Jarpe, M. B. and Johnson, G. L. (1999). Mitogen-Activated Protein Kinase: Conservation of a Three-Kinase Module From Yeast to Human. *Physiol. Rev.* **79**, 143–180.

Wolff, T. and Ready, D. F. (1991). The beginning of pattern formation in the Drosophila compound eye: the morphogenetic furrow and the second mitotic wave. *Development* **113**, 841–50.

Wolff, T. and Ready, D. F. (1993). Pattern formation in the Drosophila retina. *Cold Spring Harb. Lab. Press* **2**, 1277–1325.

Wu, X., Golden, K. and Bodmer, R. (1995). Heart Development in Drosophila Requires the Segment Polarity Gene wingless. *Dev. Biol.* **169**, 619–628.

Xiangyi Lu, Lizabeth A. Perkins, N. P. (1993). The torso pathway in Drosophila: a model system to study receptor tyrosine kinase signal transduction. 47–56.

Xu, X., Yin, Z., Hudson, J. B., Ferguson, E. L. and Frasch, M. (1998). Smad proteins act in combination with synergistic and antagonistic regulators to target Dpp responses to the Drosophila mesoderm. *Genes Dev.* **12**, 2354–2370.

Yang, L. and Baker, N. E. (2001). Role of the EGFR/Ras/Raf pathway in specification of photoreceptor cells in the Drosophila retina. *Development* **128**, 1183–1191.

Yang, C., Shapiro, L. H., Rivera, M., Kumar, A. and Brindle, P. K. (1998). A role for CREB binding protein and p300 transcriptional coactivators in Ets-1 transactivation functions. *Mol. Cell. Biol.* **18**, 2218–2229.

Yin, Z. and Frasch, M. (1998). Regulation and function of tinman during dorsal mesoderm induction and heart specification in Drosophila. *Dev. Genet.* **22**, 187–200.

Younossi-Hartenstein, A., Tepass, U. and Hartenstein, V. (1993). Embryonic origin of the imaginal discs of the head of Drosophila melanogaster. *Roux's Arch. Dev. Biol.* **203**, 60–73.

Chapter 2 A context-dependent bifurcation in the Pointed transcriptional effector network contributes specificity and robustness to retinal cell fate acquisition

Chudong Wu, Jean-François Boisclair Lachance, Michael Z Ludwig, IlariaRebay

PLOS Genetics 16(11): e1009216

I.R., C.W., J-F.B.L. conceptualized the idea; J-F.B.L. carried out the preliminary analysis of Pnt isoform expression; M.Z.L. made the GFP-PntP3 construct; C.W. carried out all the experiments and made the figures; C.W. and I.R. wrote the draft.

2.1 Abstract

Spatiotemporally precise and robust cell fate transitions, which depend on specific signaling cues, are fundamental to the development of appropriately patterned tissues. The fidelity and precision with which photoreceptor fates are recruited in the Drosophila eye exemplifies these principles. The fly eye consists of a highly ordered array of ~750 ommatidia, each of which contains eight distinct photoreceptors, R1-R8, specified sequentially in a precise spatial pattern. Recruitment of R1-R7 fates requires reiterative receptor tyrosine kinase / mitogen activated protein kinase (MAPK) signaling mediated by the transcriptional effector Pointed (Pnt). However the overall signaling levels experienced by R2-R5 cells are distinct from those experienced by R1, R6 and R7. A relay mechanism between two Pnt isoforms initiated by MAPK activation directs the universal transcriptional response. Here we ask how the generic Pnt response is tailored to these two rounds of photoreceptor fate transitions. We find that during R2-

R5 specification PntP2 is coexpressed with a closely related but previously uncharacterized isoform, PntP3. Using CRISPR/Cas9-generated isoform specific null alleles we show that under otherwise wild type conditions, R2-R5 fate specification is robust to loss of either PntP2 or PntP3, and that the two activate *pntP1* redundantly; however under conditions of reduced MAPK activity, both are required. Mechanistically, our data suggest that intrinsic activity differences between PntP2 and PntP3, combined with positive and unexpected negative transcriptional auto- and cross-regulation, buffer first-round fates against conditions of compromised RTK signaling. In contrast, in a mechanism that may be adaptive to the stronger signaling environment used to specify R1, R6 and R7 fates, the Pnt network resets to a simpler topology in which PntP2 uniquely activates *pntP1* and auto-activates its own transcription. We propose that differences in expression patterns, transcriptional activities and regulatory interactions between Pnt isoforms together facilitate context-appropriate cell fate specification in different signaling environments.

2.2 Author Summary

Properly patterned tissues require cells transit from a multipotent state to diverse differentiated states in a precise spatiotemporal manner. Meanwhile the programs that direct cell fate adoption must be reliable despite genetic and nongenetic variation. In this study we use the Drosophila photoreceptors as a model system for understanding how specific and robust cell fate transitions are achieved. The specification of seven distinct photoreceptors R1-R7 requires repetitive inductive signaling from the receptor tyrosine kinase (RTK)/ mitogen-activated protein kinase (MAPK). The transcription factor Pointed (Pnt) operates downstream of MAPK to initiate the changes in gene expression appropriate to the particular transition. We asked how the generic MAPK/Pnt signal contributes to different photoreceptor fates. Previous work showed that R1-R7

photoreceptor specification can be subdivided into two rounds that experience different RTK signaling strengths. We find distinct Pnt regulatory networks operate in the two rounds, with the first round network incorporating a novel unstudied Pnt isoform, PntP3. Its inclusion stabilizes developmental transitions when signaling is reduced. We compare and contrast the expression patterns and transactivation potentials of the Pnt isoforms and uncover a web of transcriptional cross-regulation between them. Based on these explorations, we propose that the use of distinct Pnt network topologies provides an adaptive mechanism that permits reliable cell fate transitions under different MAPK signaling environments.

2.3 Introduction

Accurate and reliable transitions from a multipotent state to diverse differentiated states are critical to normal development. A small number of transcription factors acting downstream of an even smaller handful of signal transduction pathways coordinate the gene expression changes that drive cell fate acquisition (Flores et al., 2000; Halfon et al., 2000; Voas and Rebay, 2004). How these core transcriptional effectors confer both specificity, whereby cells adopt the correct fate in a precise spatiotemporal manner (Cagan, 2009; Guillemot, 2007; Wolpert, 1969), and robustness, whereby cells reliably execute the appropriate program despite genetic and nongenetic variations (Félix and Barkoulas, 2012; Liu et al., 2019), to the transitions they oversee remains poorly understood. In this paper we use photoreceptor specification in the developing retina of Drosophila as a model to explore these regulatory mechanisms.

The Drosophila retina is precisely patterned and highly organized. Each of the ~750 ommatidia that comprise the retina contains a core cluster of eight photoreceptors, R1-R8. These neurons

are specified in a stereotyped spatiotemporal sequence that is initiated repeatedly as the morphogenetic furrow (MF) travels anteriorly across the epithelial field (Wolff and Ready, 1991). Photoreceptor specification occurs in two distinct rounds that are spatially and temporally separated by a single synchronized cell division known as the second mitotic wave (SMW) (Ready and Hanson, 1976; Tomlinson and Ready, 1987). During the first round, R8 emerges from the morphogenetic furrow (MF)'s wake, followed by the R2/R5 and R3/R4 pairs. Ommatidial assembly then pauses for the SMW, after which the second round of specification recruits photoreceptors R1/R6 and finally R7 to the cluster. Recruitment of non-neuronal support cells to the ommatidia follows immediately, starting with the four lens-secreting cone cells.

Specification of all photoreceptors except R8 requires inductive signaling by the receptor tyrosine kinase (RTK) / Ras / mitogen-activated protein kinases (MAPK) pathway via the transcriptional effector Pointed (Pnt), the Drosophila homologue of the mammalian ETS family activators ETS1 and ETS2 (Freeman, 1996; Scholz et al., 1993). Multipotent retinal progenitors must therefore translate this generic RTK/Pnt signal into specific photoreceptor fates. Numerous studies have focused on combinatorial regulation to integrate the inputs from RTK/Pnt with specific inputs from regionally expressed transcription factors and other signaling pathway effectors. For example, RTK/Pnt, the Spalt transcription factors and Notch signaling collectively specify R4 fates in the first round (Domingos et al., 2004; Weber et al., 2008) whereas in the second round, RTK/Pnt and Notch signaling integrate with a different transcription factor, Lozenge, to regulate *prospero* transcription and R7 fates (Hayashi et al., 2008; Xu et al., 2000).

Increasing the complexity of these combinatorial codes, RTK signaling inputs are not identical during the two rounds of specification. Fate specification in the first round relies exclusively on signaling initiated by the epidermal growth factor receptor (EGFR), while specification of R1, R6 and R7 second round fates involves a second RTK, Sevenless (Sev) in addition to EGFR (Basler and Hafen, 1988; Reinke and Zipursky, 1988). Although only R7 fates are lost in a *sev* mutant, the R1 and R6 precursors express Sev, physically contact the Boss ligand-expressing R8 cell, and so are likely to have active Sev signaling (Tomlinson et al., 2011). Because both EGFR and Sev use the same Ras/MAPK/Pnt signaling cascade, it has been proposed that cells specified in the second round experience stronger MAPK activation than those in the first round (Freeman, 1996; Tomlinson et al., 2019). How the Pnt response is tailored to these two different signaling environments has not been explored.

Two Pnt isoforms, PntP1 and PntP2, were identified when the gene was first cloned and have been the focus of subsequent study. The two proteins share the C-terminal ETS DNA binding domain and so are thought to have identical target gene specificity (Klambt, 1993; Scholz et al., 1993). However distinct N-terminal transactivation domains confer both distinct activity and differential regulation by RTK/MAPK signaling (Brunner et al., 1994; O'Neill et al., 1994; Shwartz et al., 2013). Whereas MAPK activation promotes the transcription of *pntP1* (Shwartz et al., 2013) to produce the constitutively active PntP1 transcription factor, MAPK regulation of PntP2 occurs post-translationally by direct phosphorylation of a site within its unique N-terminal half (Brunner et al., 1994a; O'Neill et al., 1994). Unphosphorylated PntP2 binds target DNA but has very limited transactivation ability; thus phosphorylation by MAPK is required for its full activity (O'Neill et al., 1994; Qiao et al., 2006; Shwartz et al., 2013; Tootle et al., 2003).

Furthermore, expression of a PntP2 mutant in which the phosphoacceptor threonine within the MAPK consensus site was replaced with alanine produces dominant negative effects, consistent with unphosphorylated PntP2 binding and occluding target gene enhancers from appropriate activation (Brunner et al., 1994; O'Neill et al., 1994). Thus the final PntP2 transcriptional output within an individual cell reflects the sum of the low basal, or even the repressive, activity of unphosphorylated PntP2 plus the stronger activity of phosphorylated PntP2, with the availability of active MAPK determining the ratio between the two. The mammalian ETS1 and ETS2 proteins structurally and functionally resemble PntP2 (Wasylyk et al., 1997).

A previous sequential activation model posited that transient RTK/MAPK signaling activates PntP2, which in turn activates *pntP1* transcription, and that PntP1 then provides a stable, signaling-independent, transcriptional input to the combinatorial codes that initiate the specification of R1-R7 photoreceptor fates (Fig 1C; Shwartz et al. 2013). However the expression pattern of *PntP2* suggests further complexity, with lower levels in the region of R2-R5 specification and then higher levels in more posterior regions where R1, R6, R7 and cone fates are recruited (Shwartz et al., 2013). These differences parallel the differences in RTK signaling in the two rounds of photoreceptor specification and motivated us to explore how the Pnt response is tuned to these two distinct signaling environments.

In this study, we uncover distinct Pnt regulatory networks for the two rounds of specification that are distinguished by the inclusion/exclusion of another evolutionary conserved but previously unstudied Pnt isoform, PntP3. As predicted by its protein structure, PntP3 functions as a MAPK-responsive transcriptional activator, but with intrinsically higher activity than PntP2. In contrast

to R1, R6 and R7 specification where PntP2 is uniquely expressed and required, we find that PntP3 and PntP2 are coexpressed and under wild type conditions function redundantly during specification of first round R2-R5 photoreceptors. However under conditions of compromised signaling, the individual activities of PntP3 and PntP2 are essential to the robustness of these fate transitions. Mechanistically, we uncover distinct auto- and cross-regulatory transcriptional interactions between PntP2 and PntP3 during the two rounds of photoreceptor specification that likely optimize context-specific output, with the most striking a shift from PntP2 auto-repression during specification of first round fates to auto-activation during specification of second round fates. We conclude that a combination of functional redundancy, different transactivation potential and the reset of transcriptional regulatory interactions between Pnt isoforms adapts the transcriptional response to different RTK signaling environments.

2.4 Results

2.4.1 PntP3 is a MAPK-responsive transcriptional activator whose expression overlaps that of PntP2 in R2-R5 photoreceptors

Although the field has focused on the two isoforms identified when the *pnt* gene was first cloned (Brunner et al., 1994; Klambt, 1993; O'Neill et al., 1994; Scholz et al., 1993; Shwartz et al., 2013), the BDGP cDNA project together with subsequent high-throughput mRNA sequencing has revealed additional transcripts (S2.1A Fig; Celniker et al., 2009; Leader et al., 2018; Rubin et al., 2000). First, there is a second *pntP1* isoform (*pnt-E* in G-Browse) identical to "classic" *pntP1* (*pnt-C*) except for a longer 3'UTR and an encoded product with an extra two amino acids owing to the use of an alternate splice donor site at the 3' end of the first coding exon. Second, there is a

transcript identified as *pnt-D* that is closely related to but distinct from *pntP2* (*pnt-B*); we refer to this novel isoform as *pntP3*.

pntP3 is distinguished from *pntP2* by its unique transcription start site, 5'UTR and N-terminal coding exons, but then like *pntP2,* it splices into the exons encoding the sterile alpha motif (SAM) and adjacent MAPK consensus site, and the ETS domain (Fig 2.1A, B). To our knowledge there have not been any explicit studies of *pntP3*. However, just like *pntP1* and *pntP2, pntP3* was detected by RNA-Seq profiling in most developmental stages (Leader et al., 2018), suggesting it might contribute to the transcriptional response downstream of RTK signaling. Further, both PntP2 and PntP3 are conserved across *Drosophila* species from *D. melanogaster* to *D. virilis* (S2.1B Fig). Conservation across millions of years suggests strong evolutionary pressure for keeping both PntP2 and PntP3, implying essential functions.

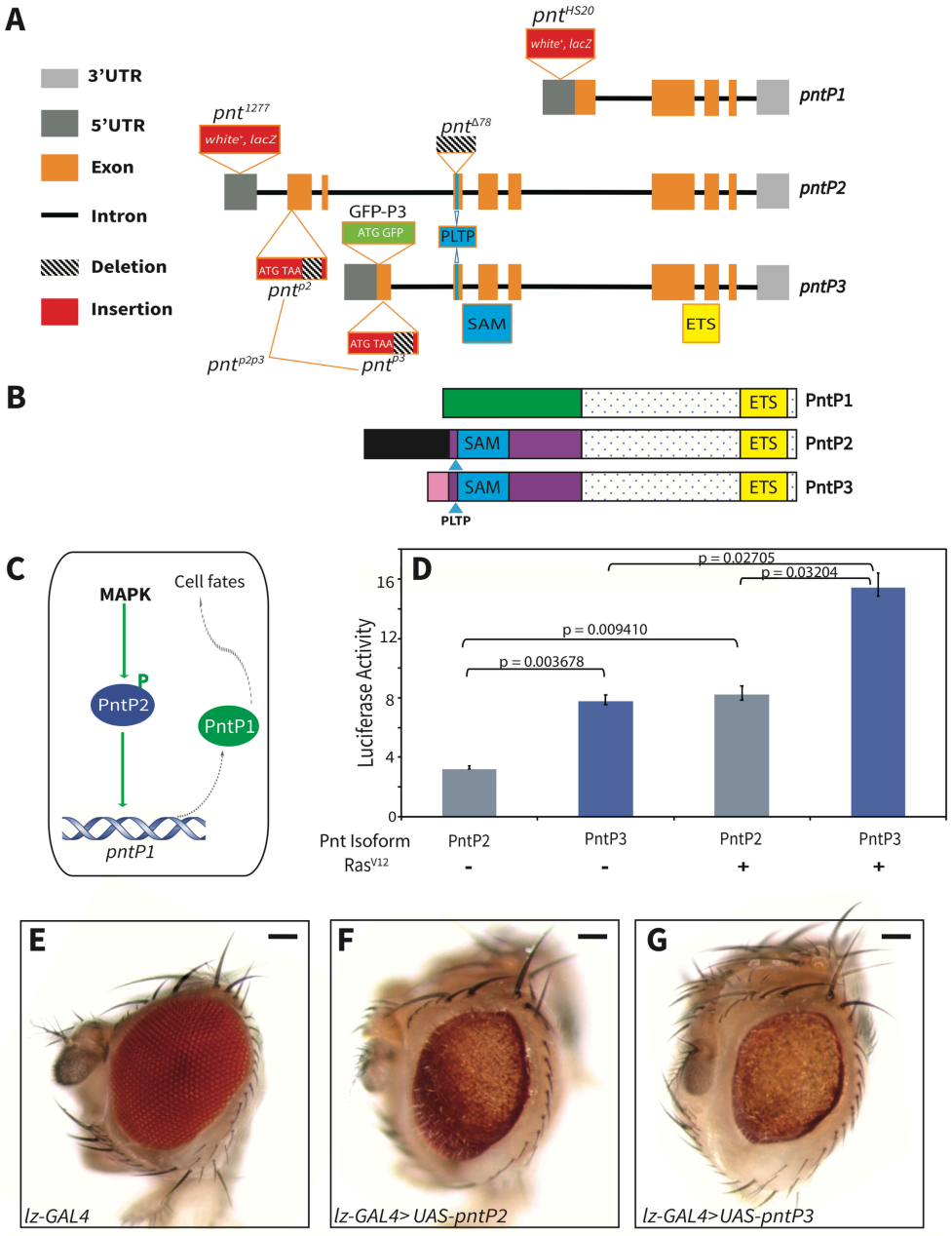


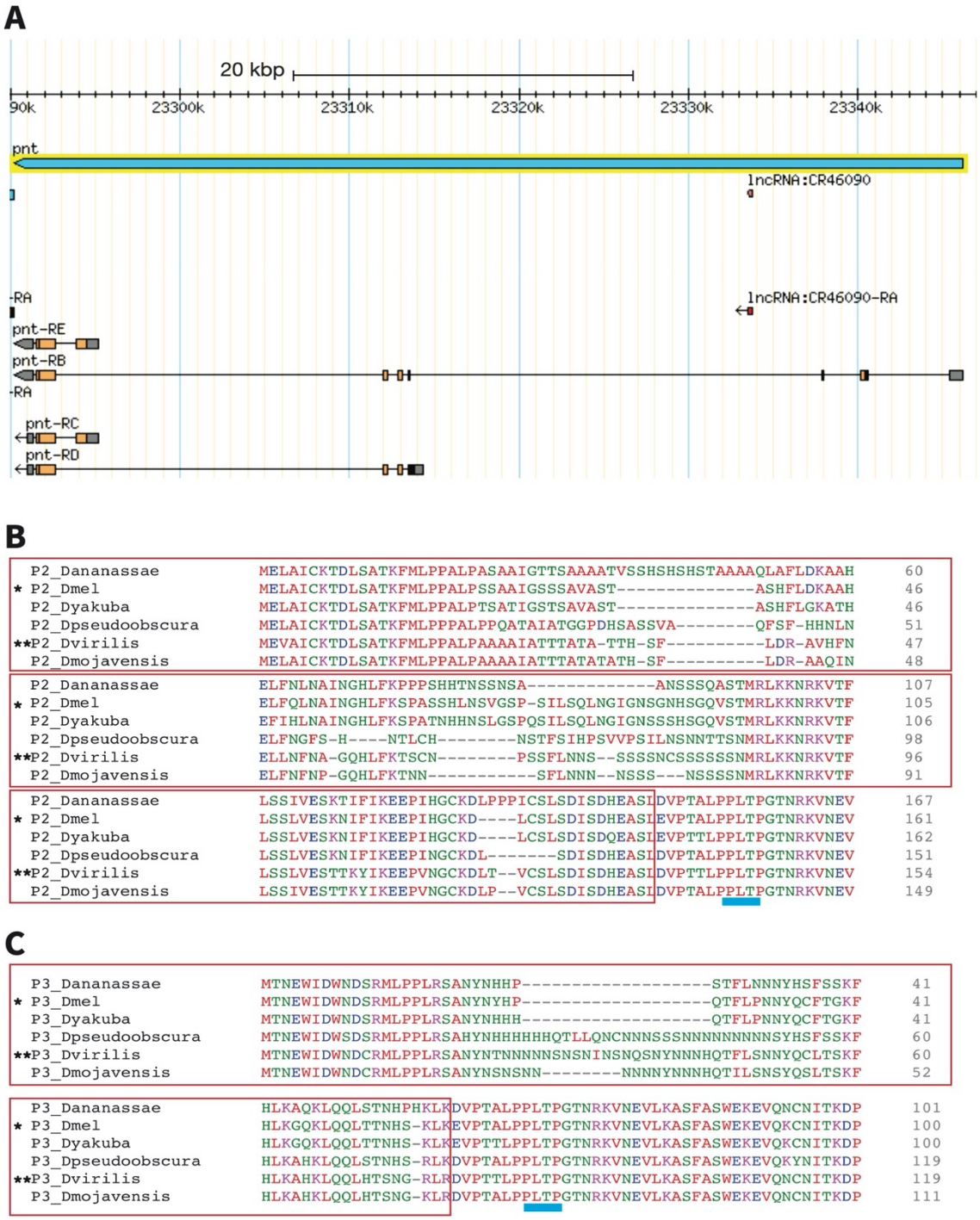
Fig 2.1 PntP3 is a stronger MAPK-responsive transcriptional activator than PntP2

A) A schematic, not to scale, of the ~55kb *pnt* locus. *pntP1*, *pntP2* and *pntP3* all splice into common 3' exons encoding the ETS DNA binding domain (yellow box). *pntP2* and *pntP3* share three internal exons encoding the SAM and PLTP MAPK phosphorylation site (blue boxes). Unique N-terminal exons encode isoform-specific sequences. Approximate insertion sites of key P-element-derived alleles are shown: the *white+, lacZ* enhancer trap insertions pnt^{1277} and pnt^{HS20} respectively report *pntP2* and *pntP1* expression (Scholz et al., 1993); the excision allele $pnt^{\Delta 78}$ disrupts the SAM-encoding exon common to *pntP2* and *pntP3* (O'Neill et al., 1994). Green box labeled ATG-GFP signifies the genomic BAC transgene in which PntP3 was N-terminally GFP tagged. Red boxes labeled ATG-TAA represent the CRISPR-generated null (Fig

2.1 continued) alleles of pnt^{p2} and pnt^{p3} that have stop codons immediately after the ATG and exonic deletions; pnt^{p2p3} carries identical stop codon insertions and deletions. B) A schematic of PntP1, PntP2 and PntP3 proteins highlights their distinct N-termini and common C-termini. The transactivation domains mapped to the distinct terminal regions of PntP1 and PntP2 (shown in green, purple, pink; Wasylyk et al., 1997; Watson et al., 1988). PntP2 and PntP3 differ only in sequence N-terminal to the MAPK site and SAM (141aa for PntP2; 59aa for PntP3). C) A schematic summarizing the sequential activation model: MAPK phosphorylation activates PntP2, phosphorylated PntP2 activates *pntP1* transcription and PntP1 protein drives cell fate specification (Shwartz et al., 2013). D) PntP3 has stronger activity but similar MAPK responsiveness than PntP2 in transcription assays using a reporter with 6 tandem high-affinity ETS sites (O'Neill et al., 1994). For each sample, activity was normalized to reporter alone control. Error bars are S.D. of three independent experiments. P-values were calculated using two tailed pair-wise Student T-tests. (E-G) *lz-GAL4*-driven overexpression of *UAS-pntP2* (F) and *UAS-pntP3* (G) disrupts external eye morphology, pigmentation and size relative to driver alone control (E). Scale bar: 50 μm.

Given the protein-level similarity, we began by asking whether PntP3, like PntP2, functions as a transcriptional activator positively regulated by MAPK phosphorylation. In transcriptional reporter assays in transiently transfected S2 cells, PntP3 was about two-fold more active than PntP2 in the absence of MAPK stimulation; MAPK stimulation induced a further ~two-fold activity increase for both isoforms (Fig 2.1D). When overexpressed in the developing eye, PntP3 also showed greater activity than PntP2, producing stronger disruptions of adult eye morphology with all Gal4 drivers tested (Fig 2.1E-G, S2.2 Fig). These phenotypes were associated with ectopic induction of photoreceptor fates in 3rd instar discs (S2.2 Fig), consistent with previous studies of Pointed overexpression (Shwartz et al., 2013). As predicted by their relative activities in S2 cells, ectopic expression of neuronal markers was more striking with *pntP3* overexpression than with *pntP2* (S2.2 A-F Fig). The expression level and subcellular localization to the nucleus were indistinguishable between the two isoforms (S2.2 G-I Fig), indicating differential transcriptional activity most likely underlies the phenotypic differences.

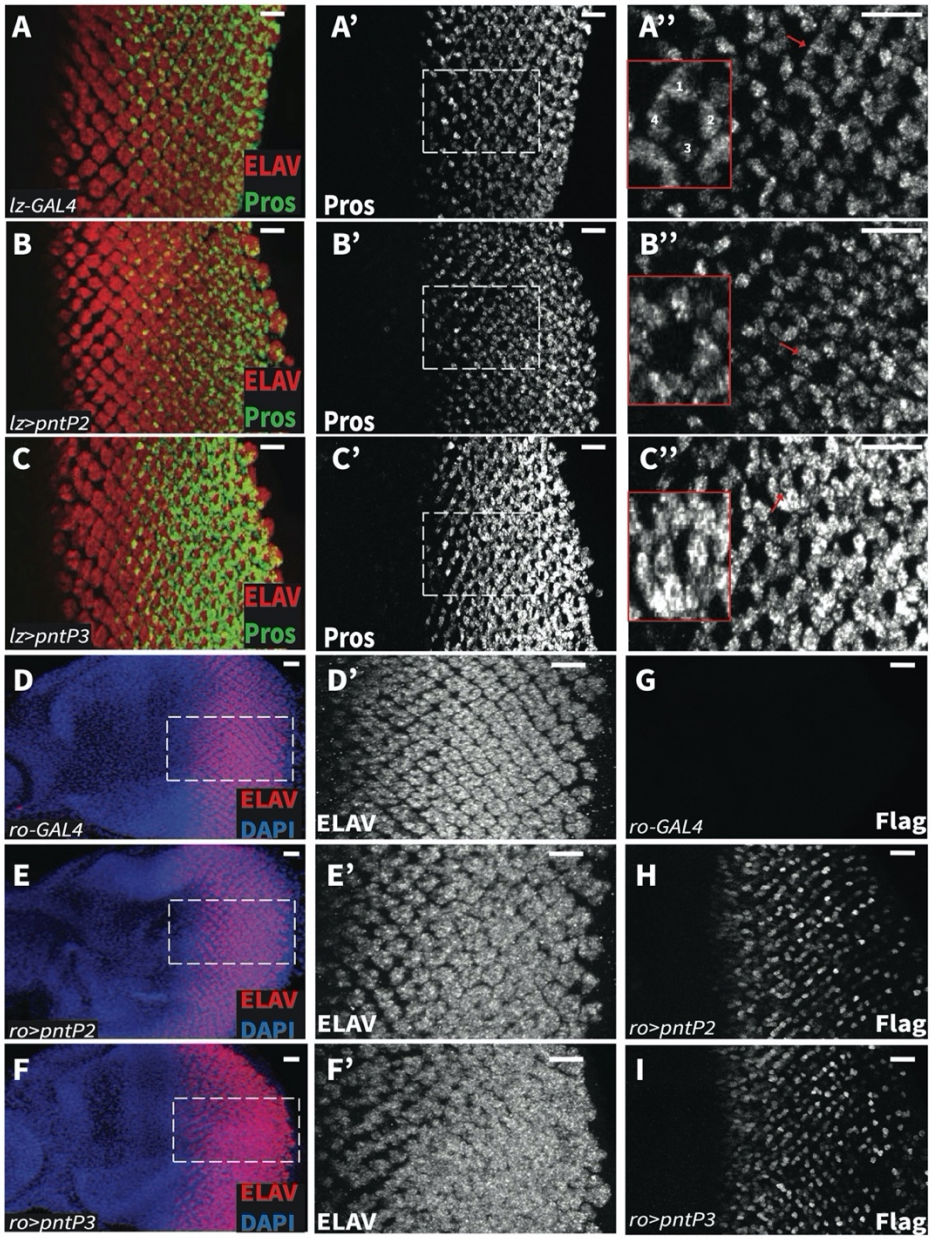
In addition to the activity differences between PntP2 versus PntP3 that we attribute to the unique sequences at the N-terminal ends of their transactivation domains (Figure 2.1B, S2.1B, C Fig), the use of separate 5' regulatory regions suggested that expression pattern differences might also distinguish their developmental roles. To explore this, we compared their endogenous expression in late 3rd instar eye discs where *pnt* function is essential for photoreceptor specification and has been well studied (Brunner et al., 1994a; Scholz et al., 1993; Yang et al., 2003). We relied on the *pntP2*-specific enhancer trap allele pnt^{1277} (Scholz et al., 1993; Shwartz et al., 2013) to report PntP2 expression. To visualize PntP3 expression, we inserted an N-terminal GFP tag in a genomic BAC that contains the entire *pnt* locus (GFP-P3, Fig 2.1A) and that we had previously shown to be fully functional (Lachance et al., 2014). The *GFP-PntP3* transgene fully complemented the lethality of $pnt^{\Delta 88}/Df(pnt)$ animals, a background null for all three isoforms, producing phenotypically wild type, fertile adults.



S 2.1 Fig Genomic organization and conservation of pnt

(A) A screen shot of the *pnt* locus from Flybase G-Browse, oriented 5' to 3' from right to left, summarizing the annotation of the B, C, D and E isoforms. (B, C) PntP2 (B) and PntP3 (C) are

(S 2.1 Fig continued) conserved from *D. melanogaster* (*) to *D. virilis* (**). Alignments show the unique N-terminal sequences (within the red rectangle) of PntP2 and PntP3. Conserved sequence begins immediately N-terminal to the PLTP (blue line)/SAM region (unboxed). The PntP2 N-terminus is 62% identical and 70.7% similar between *D. melanogaster* and *D. virilis*, while the PntP3 N-termini are 57% identical and 63% similar.



S 2.2 Fig Overexpressed PntP3 disrupts eye development more strongly than overexpressed PntP2

(S 2.2 Fig continued) All images show maximum projections of optical confocal sections of representative 3rd instar eye imaginal discs, oriented anterior left. Scale bars: 10 μm. (A-C) Overexpression of *UAS-pntP2* and *UAS-pntP3,* driven by *lz-Gal4*. Elav (red) marks all photoreceptors and Pros (green) marks R7 photoreceptors and cone cells. (A'-C') Overexpression of *pntP3* induced more ectopic Pros-positive cells than overexpression of *pntP2*. (A''-C'') Magnified views of boxed regions in A'-C' with further zoom in to a single ommatidium (red box, red arrow). The wild type pattern of four Pros positive cone cells is labeled in 2A". (D-I) Overexpression of *UAS-pntP2* and *UAS-pntP3*, driven by *ro-Gal4*. Elav (red) marks the photoreceptors and DAPI (blue) marks all nuclei. (D-F) Overexpression of *pntP3* induced more ectopic Elav expression than overexpression of *pntP2*. (D'-F') Magnified views of boxed regions in D-F. (G-I) Staining with anti-Flag to detect the epitope tag shows comparable levels and nuclear localization of PntP2 and PntP3.

Analysis of 3rd instar eye discs dissected from animals carrying both *GFP-PntP3* and the *pntP2*-specific enhancer trap allele revealed both distinct and overlapping patterns of expression (Fig 2.2). Lower magnification projections emphasized the complementary aspects of the two patterns, with GFP-PntP3 expression strongest in and immediately posterior to the MF and β-galactosidase (β-gal) reporting strongest *pntP2* expression in the posterior half of the eye field (Fig 2.2B). As the stereotyped differentiation sequence of ommatidial assembly means every cell can be unambiguously identified by its position and morphology (Ready et al., 1976; Tomlinson et al., 1987; Wolff et al., 1993; Pelaez et al., 2015), higher magnification views at different optical planes enabled cell type specific comparison of the two patterns (Fig 2.2C-F). Coexpression was detected in R2/R5 and R3/R4 photoreceptor pairs, in basal progenitors at the MF and in cone cells (Fig 2.2C, D and S3A). Complementary expression was detected posterior to the second mitotic wave (SMW) in photoreceptors R1, R6 and R7 where *pntP2* was high and GFP-PntP3 low, and in apically localized nuclei at the MF, including R8, where GFP-PntP3 was high and *pntP2* low (Fig 2.2E, F). The combined differences and similarities in cell type specific expression patterns raised the possibility of both distinct and overlapping functional requirements for PntP2 and PntP3 in the two rounds of photoreceptor specification (Fig 2.2A).

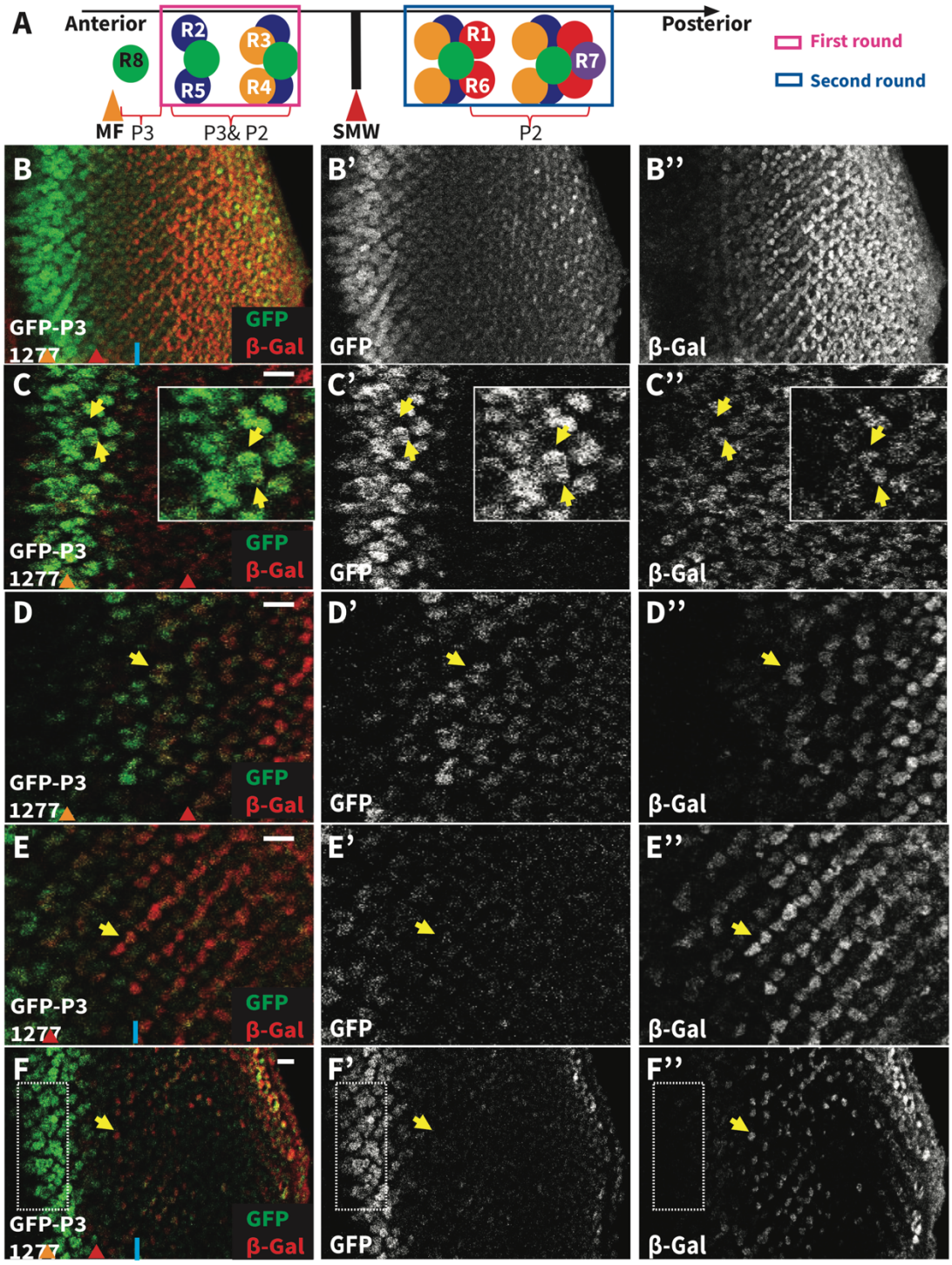
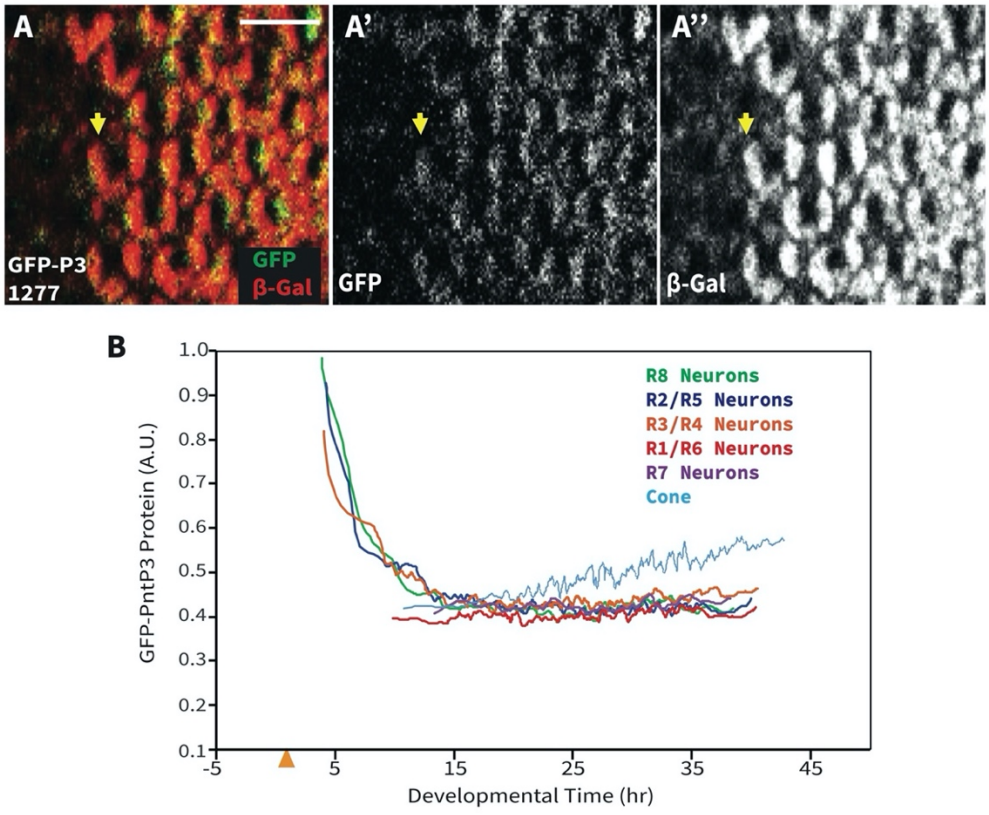


Fig 2.2 PntP3 and PntP2 show overlapping and complementary expression patterns

A) A schematic summarizing the sequential specification of photoreceptor fates (adapted from (Pelaez et al. 2015)) and the expression patterns of PntP3 and PntP2. R8 cells are specified first

(Fig 2.2 continued) near the MF (orange arrowhead) and express PntP3. R2/R5 and R3/R4 pairs are specified next (first round fates) and express both PntP3 and PntP2. After the SMW (red arrowhead), R1/R6 and R7 are specified (second round fates) and express PntP2. (B-F) Representative 3rd instar eye imaginal discs, oriented anterior left, comparing the pattern of *pntP2* transcription (red), as reported by pnt^{1277}, and PntP3 protein (green), as reported by a GFP-PntP3 genomic BAC transgene. (B) Maximum projection highlighting the complementary pattern of highest PntP3 in the MF region and highest *pntP2* posterior to the SMW (blue line). (B', B") Single channel images show that *pntP2* transcription starts anterior to the SMW in cells where GFP-PntP3 is expressed. (C-F) Single optical slices of the disc in (B) at different apical/basal planes. (C, D) Coexpression was detected in R2/R5 pairs (C, yellow arrows, insets show zoomed view) and in R3/R4 pairs (D, yellow arrows). (E-F) *pntP2* but not PntP3 was detected in R1/R6 pairs (E, yellow arrows) and in R7 (F, yellow arrows). PntP3 but not *pntP2* was detected in basal progenitors at the MF (F, boxed region). Scale bar: 10 μm.



References

1. Nicola´s Pela´ez, Arnau Gavalda-Miralles, Bao Wang, Heliodoro Tejedor Navarro, Herman Gudjonson, Ilaria Rebay, Aaron R Dinner, Aggelos K Katsaggelos, Lui´s AN Amaral RWC. Dynamics and heterogeneity of a fate determinant during transition towards cell differentiation. Elife. 2015;53: 160. doi:10.1017/CBO9781107415324.004

S 2.3 Fig PntP3 is expressed in multipotent cells, R2-R5 photoreceptors and cone cells

(S 2.3 Fig continued) (A) Higher magnification and single optical slice showed the overlapping expression patterns of pnt^{1277} (β-gal, red) and GFP-PntP3 (GFP, green) in cone cells (as indicated by the yellow arrows). Scale bar: 10 μm. (B) Moving averages of GFP-PntP3 levels highlight the peaks of expression at the MF (orange arrowhead) region in the first five photoreceptors recruited to each ommatidium, R8, R2, R5, R3 and R4 (green, blue, and orange lines). Expression in photoreceptors specified after the SMW, R1, R6, and R7, was not above baseline (red and purple lines. A slow increase in PntP3 levels was measured in the cone cells (blue line). Data plotted are from two discs from independent experiments and show the results from scoring 404 ommatidia with 404 R8 cells. Single cell measurements were made and plotted as described in *Pelaez et al*., 2015.

2.4.2 Redundant and non-redundant requirements for PntP2 and PntP3 during two distinct rounds of photoreceptor specification

Prior studies of *pnt* function during retinal development concluded that *pntP1* and *pntP2* are required non-redundantly to specify photoreceptors R1-R7 (O'Neill et al., 1994; Shwartz et al., 2013). However the *pntP2* allele used in the studies, $pnt^{\Delta 78}$ (O'Neill et al. 1994), was generated by imprecise excision of a P-element inserted into the first SAM-encoding exon, and so also disrupts *pntP3* (Fig 2.1A). This means that $pnt^{\Delta 78}$ phenotypes, in the eye loss of R1-R7, reflect the combined loss of *pntP2* and *pntP3*.

To reveal the individual requirements for the two isoforms we generated pnt^{p2} and pnt^{p3} specific mutants using CRISPR/Cas9 genome editing (Fig 2.1A). To confirm the effectiveness of the molecular strategy, we also engineered a pnt^{p2p3} double mutant allele. As reported for $pnt^{\Delta 78}$ (Morimoto et al., 1996; O'Neill et al., 1994), homozygous pnt^{p2p3} adults were never recovered, indicating that the combined function of the two isoforms is essential for viability. In contrast, homozygous pnt^{p3} animals were fully viable while homozygous pnt^{p2} animals occasionally survived (scoring 2058 3rd instar progeny from a cross between pnt^{p2}/*TM6B* parents found only

35 homozygous pnt^{p2} animals, a 1.7% survival rate). The differences in survival of the isoform specific mutants suggested both redundant and non-redundant requirements for PntP2 and PntP3 during development, with PntP2 playing the major role and PntP3 a more auxiliary one.

Focusing on photoreceptor specification, homozygous pnt^{p2p3} clones were missing all photoreceptors except R8 (Fig 2.3A-A''), consistent with published analysis of $pnt^{\Delta 78}$ (O'Neill et al., 1994; Shwartz et al., 2013). We reasoned that if the function of both PntP2 and PntP3 is required for photoreceptor specification, then neither single mutant should recapitulate the double mutant phenotype. If so, the requirement for PntP3 should manifest in the first round fates where it is strongly expressed but not in second round fates where its levels are low (Fig 2.2A, B and S2.3B Fig). Alternatively, if PntP3 does not contribute activity essential to photoreceptor specification, then the pnt^{p2} and pnt^{p2p3} mutants should show identical loss of R1-R7 phenotypes.

To test these predictions, we first assessed photoreceptor loss in adult eyes, using F-actin to highlight the number and spatial arrangement of the rhabdomeres. In a wild type ommatidium, the larger rhabdomeres of R1-R6 are arrayed in a trapezoidal-shaped ring around the smaller R7 rhabdomere (Fig 2.3B). Whereas ommatidia of homozygous pnt^{p3} adults had the full complement of photoreceptors (Fig 2.3C), those of the rare homozygous pnt^{p2} escapers did not (Fig 2.3D). Loss of R7 was fully penetrant (100%, n=204, white arrow), with most ommatidia missing two additional photoreceptors (73%, n=204, yellow circles), and some showing even greater loss (red arrow). The more modest photoreceptor loss seen in pnt^{p2} single mutants relative to pnt^{p2p3} double mutants indicates a functional requirement for PntP3.

Examination of 3rd instar discs confirmed that loss of *pnt*^{p2} resulted in loss of only the cell fates recruited during the second round of specification. First, only a few Pros positive cells remained in the posterior of discs from homozygous *pnt*^{p2} animals; a similar posterior scattering of Cut-positive cells suggested a complete failure to specify R7 photoreceptors and most cone cells (Fig 2.3I and S2.4A, B Fig). Second, normal expression of Sal and reduction of Svp expression to only two, rather than four cells per ommatidia, indicated correct specification of photoreceptors R3/R4 and a failure to specify R1/R6 (Fig 2.3J-L). Third, and confirming no other consistent photoreceptor specification defects, examination of Sens and Elav patterns in *pnt*^{p2} mosaic discs indicated normal recruitment of photoreceptors R8/R2/R5 (Fig 2.3M, L). Thus the complete loss of R1-R7 fates that occurs in *pnt*^{p2p3} double mutant ommatidia (Fig 2.3A) reflects the combined loss of redundant inputs to R2-R5 first found fates plus the PntP2-specific input to R1, R6 and R7 second round fates.

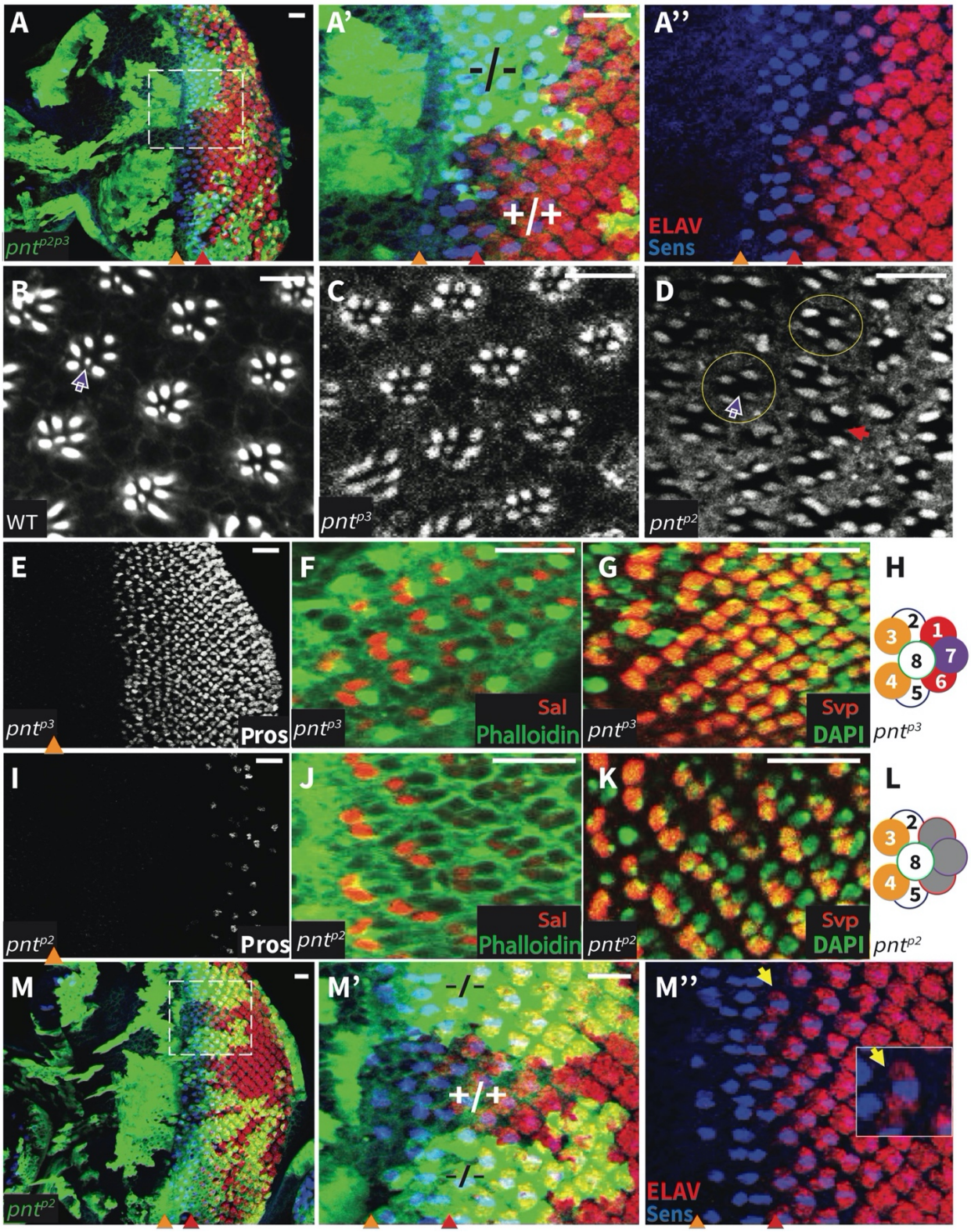
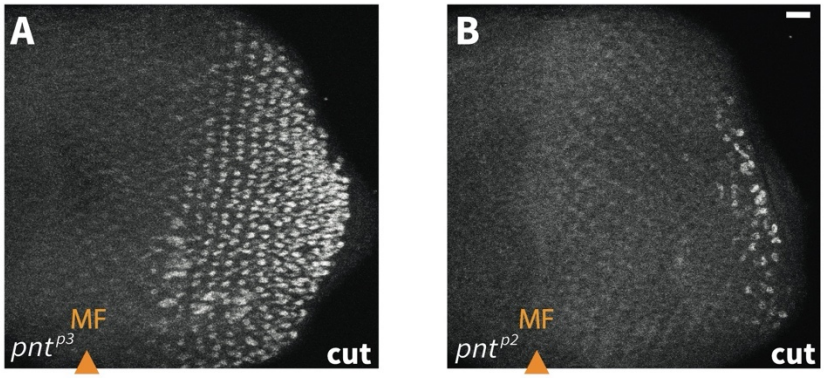


Fig 2.3 Redundant and unique requirements for PntP2 and PntP3 in photoreceptor specification

(Fig 2.3 continued) A) Representative 3rd instar eye disc of genotype *eyFLP/+;act-Gal4,UAS-GFP;FRT82B,pnt^{p2p3}/tub-Gal80,FRT82B,* oriented anterior left. Elav (red) marks all photoreceptors and anti-Sens (blue) marks R8. Homozygous *pnt^{p2p3}* mutant clones, positively marked by GFP (green), lack all photoreceptors except R8. A' and A'' show zoomed views of boxed region in A. MF, orange arrowhead. SMW, red arrowhead. Scale bar: 10 µm. (B-D) Adult retinas stained with phalloidin to mark the rhabdomeres. (B, C) Both wild type and *pnt^{p3}* mutants have regularly arrayed rows of ommatidia; white arrow points to the small R7 rhabdomere at the center of the outer trapezoid formed by the larger R1-R6 rhabdomeres. (D) All ommatidia of homozygous *pnt^{p2}* mutants lack R7, most also lack two outer rhabdomeres (yellow circles), and some show even greater loss (red arrow). Scale bars: 5 µm. (E-L) Representative 3rd instar eye discs, oriented anterior left. R7 and cone cells are marked by Pros (white), R3/R4 pairs are marked by Sal (red) or Svp (red), and R1/R6 pairs are marked by Svp (red). (E-G) *pnt^{p3}* mutants appear wild type. (I-K) *pnt^{p2}* mutants lack R7 and most cone cells (I), have normal R3/R4 specification (J, K), and lack R1/R6 (K). (H, L) Schematic summaries of photoreceptor specification patterns in *pnt^{p3}* and *pnt^{p2}* mutants. Scale bars: 10 µm. (M) Representative 3rd instar eye disc of same genotype as in (A). Homozygous *pnt^{p2}* clones, positively marked by GFP (green), show normal R8/R2/R5 specification as marked by Elav and Sens. K' and K'' show zoomed views of boxed region in K; yellow arrow points to a newly specified R2/R5 pair. Scale bar: 10 µm.



S 2.4 Fig Cone cell loss in pntp2 mutants

(A-B) Representative third instar eye discs oriented anterior left, stained with anti-Cut to mark the cone cells. Homozygous *pnt^{p3}* mutants appear fully wild type (A) while only a few scattered Cut-positive cells remain in *pnt^{p2}* mutants (B). Scale bar: 10 µm.

2.4.3 PntP2 and PntP3 provide robustness through redundant activation of pntP1 transcription

A central tenet of the current model of *pnt* function during photoreceptor specification is that PntP2 activates *pntP1* transcription (Shwartz et al. 2013; Fig 2.1C) . Given the partial genetic redundancy between PntP2 and PntP3, we asked whether PntP3 also contributes to this activation.

To start, we used reverse transcription quantitative polymerase chain reaction (RT-qPCR) to measure *pntP1* transcript levels in pnt^{p2} and pnt^{p3} mutant tissues. In both mutants, decreases in *pntP1* transcripts were measured in eye discs across three independent biological replicates, although the changes were not statistically significant ($p = 0.1$; Fig 2.4A; S 2.1 Table). This suggests either redundancy between PntP2 and PntP3 with respect to activating *pntP1* transcription, inadequate sensitivity in the RT-qPCR assay, or that PntP2 and PntP3 are not the primary activators of *pntP1*.

We were concerned that by grinding up whole tissue we were destroying spatial information and therefore missing locally significant changes in *pntP1* levels. Also, the animals lacking both PntP2 and PntP3 do not survive to 3rd instar, precluding RT-qPCR analysis of the double mutant. Thus we turned to fluorescence in situ hybridization (FISH) to ask whether the two isoforms work redundantly to active *pntP1* expression. FISH probes targeting the *pntP1* isoform-specific exons revealed an expression pattern consistent with that of the *pntP1* enhancer trap allele (Scholz et al., 1993; Shwartz et al., 2013). Specifically, we detected peak *pntP1* transcription in a periodic pattern at the MF, lower levels of expression in the zone between the MF and SMW region and then lowest levels posterior to the SMW (Fig 2.4B, D, S2.5B Fig). In pair-wise comparisons of wild type vs. pnt^{p2} and wild type vs. pnt^{p3}, no changes in *pntP1* transcription

were noted (Fig 2.4B-E), consistent with the RT-qPCR results. However in *pnt*^{p2p3} mutant clones, *pntP1* expression at the MF was strongly reduced (Fig 2.4F-F'). We conclude that PntP2 and PntP3 redundantly activate *pntP1*.

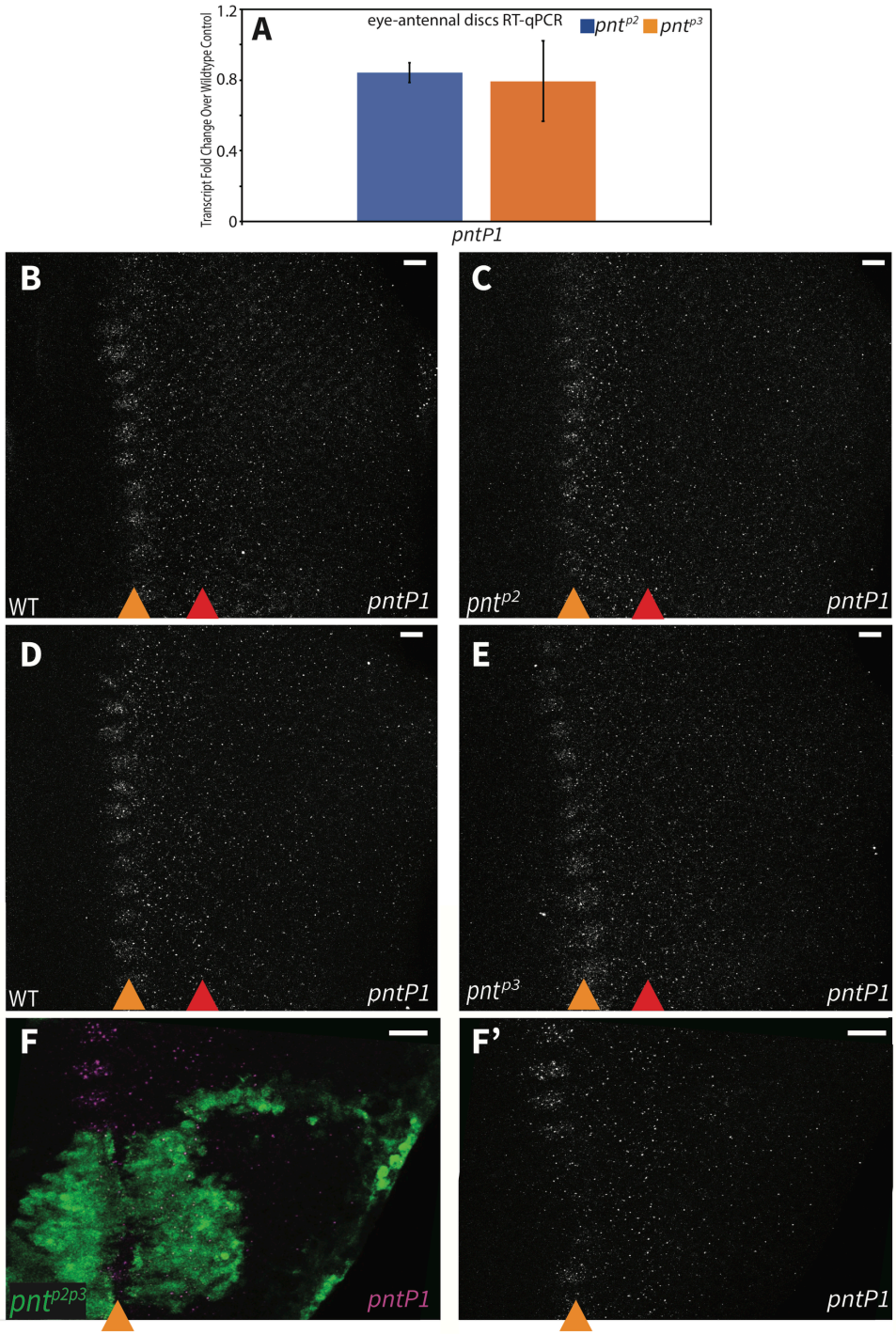
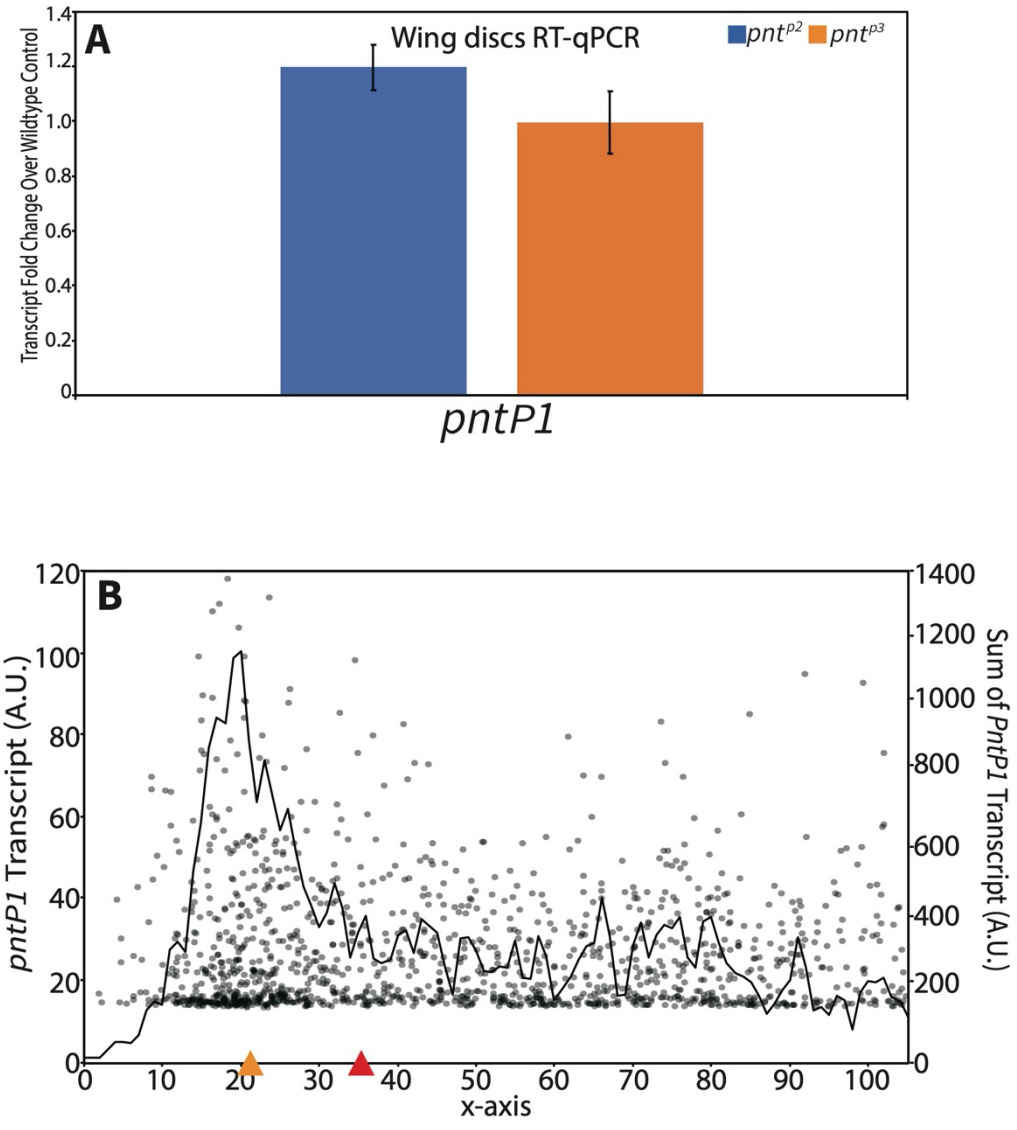


Fig 2.4 PntP2 and PntP3 redundantly activate pntP1 transcription

(Fig 2.4 continued) (A) RT-qPCR comparison of *pntP1* transcript levels in wild type versus pnt^{p2} (blue bars) and pnt^{p3} (orange bars) 3rd instar eye-antennal discs. No significant changes were detected. Error bars represent S.D. of three independent experiments. Significance was calculated via pair-wise Student T-tests between the mutant sample and the control gene. (B-F) *pntP1* FISH in 3rd instar eye imaginal discs, oriented anterior left, with MF marked by orange arrowhead and SMW (Fig 2.4 continued) marked by red arrowhead. (B-E) Maximum projections, (F) partial projections. Scale bars: 5 µm.
(B, C) *pntP1* transcripts patterns were comparable between wild type (B) and pnt^{p2} (C). (D, E) *pntP1* transcript patterns were comparable between wild type (D) and pnt^{p3} (E). (F) Homozygous pnt^{p2p3} mutant clones, positively marked with GFP (green), show reduced *pntP1* at the MF. Consistent results were obtained from analyzing 12 clones from 9 discs across 3 independent experiments.



S 2.5 Fig pntP1 transcription did not change in pntp2 and pntp3 mutant wing imaginal discs

(S 2.5 Fig continued) (A) RT-qPCR comparison of *pntP1* transcript levels in wild type versus pnt^{p2} (blue bars) and pnt^{p3} (orange bars) null mutant 3rd instar wing imaginal discs. No significant change was detected. Error bars represent standard deviations of three independent experiments. (B) Quantification of *pntP1* FISH in the wildtype disc of Fig 2.5B from maximum projections. *pntP2* levels begin to rise to peak at the MF (yellow arrow), quickly decrease between MF and SMW (S 2.5 Fig continued) (red arrow) and slowly decrease to a steady state posterior to the SMW. Each dot plots the product of the fluorescent intensity and the size of an individual *pntP1* FISH focus, representing the relative amount of *pntP1* transcript (y-axis on the left) The line connects the moving average of the sum of all foci within one-pixel windows along the x-axis (y-axis on the right).

2.4.4 Context specific auto- and cross-regulation of pntP2 transcription

Having established the functional redundancy of PntP2 and PntP3 with respect to induction of *pntP1*, we next investigated how the system tunes these two parallel inputs. In particular we wondered whether cross-regulatory feedback might coordinate and optimize PntP2/PntP3 expression levels, and ultimately their activity. To test this, we used RT-qPCR to measure changes in *pntP2* and *pntP3* transcript levels in eye imaginal discs dissected from pnt^{p2} and pnt^{p3} homozygous mutant 3rd instar larvae.

Two findings emerged. Most striking, and unexpectedly, the experiment uncovered negative auto-regulation for both isoforms (Fig 2.5A). Thus, *pntP2* transcript levels were significantly increased in pnt^{p2} mutant tissue ($p < 0.01$) and *pntP3* transcripts were significantly increased in pnt^{p3} mutant tissue ($p < 0.05$). Given the surprising nature of this result, we repeated the experiment using wing imaginal discs, and again found significant increases in transcript levels in the respective mutant (Fig 2.5B). Thus both isoforms negatively regulate their own transcription, either directly or indirectly.

Second, evidence of cross-regulation emerged from the eye disc experiments (Fig 2.5A), with a modest, but reproducible and significant 12% average increase in *pntP3* transcript levels measured in pnt^{p2} mutant tissue ($p < 0.05$; S2.1 Table). In the converse experiment, the increases in *pntP2* transcript levels measured in pnt^{p3} mutant tissue were more variable across the three biological replicates (S1 Table), resulting in a statistically insignificant 47% average increase ($p = 0.18$). Thus the possibility of bidirectional inhibitory cross-regulation remains an open question. Cross-regulatory interactions were not detected in the wing disc (Fig 2.5B).

The coexpression of PntP2 and PntP3 in the anterior half of the disc where first round photoreceptor fates are specified predicted that the regulatory interactions uncovered by RT-qPCR were occurring in this context. We therefore turned to FISH to corroborate the negative auto-regulation and to assess further the possibility of cross-regulatory interactions. We found that *pntP2* transcription initiated at the MF, peaked in the region of the second mitotic wave (SMW), and then continued at a more moderate level across the posterior half of the disc (Fig 2.5C, E). This pattern was consistent with that reported by the enhancer trap pnt^{1277} although the prolonged perdurance of beta-galactosidase likely over-reports *pntP2* levels in the posterior (Fig 2.2). Unfortunately our FISH protocol was not able to detect *pntP3*, presumably because its specific exon is too short for adequate numbers of probes (see Methods).

We next compared *pntP2* transcript levels in wildtype versus pnt^{p2} null mutant retinal tissue. In both whole mutant eye discs (Fig 2.5C-F) and in null mutant clones (Fig 2.5G), increased *pntP2* transcription was evident in the MF and in the adjacent region where *pntP2* levels normally peak (Fig 2.5C and S2.6). We also noticed a change not predicted by the RT-qPCR analysis, namely a

decrease in *pntP2* transcripts in the posterior of the disc (Fig 2.5C-F and S2.6). This suggests *pntP2* transcription is regulated differently in anterior versus posterior regions of the developing eye field.

Because PntP3 expression is strongest anteriorly (Fig 2.2), we wondered whether the increase in *pntP2* transcripts detected in pnt^{p2} mutant tissue reflected cross-regulatory activation by PntP3. To test this we examined *pntP2* transcript levels in pnt^{p2p3} double mutant clones (Fig 2.5H). No increase was detected at the MF or in the adjacent region of peak expression. In more posterior pnt^{p2p3} mutant clones, *pntP2* transcript levels were lower than in adjacent wild type clones, exactly as seen in pnt^{p2} single mutant clones (Fig 2.5G). Thus in anterior regions where PntP3 expression is strong, loss of PntP2 results in a PntP3-dependent increase in *pntP2* transcription whereas in posterior regions where PntP3 expression is normally low, loss of PntP2 results in a PntP3-independent reduction in *pntP2* transcription. Thus the regulatory relationships between Pnt isoforms established at the MF are reset after the SMW.

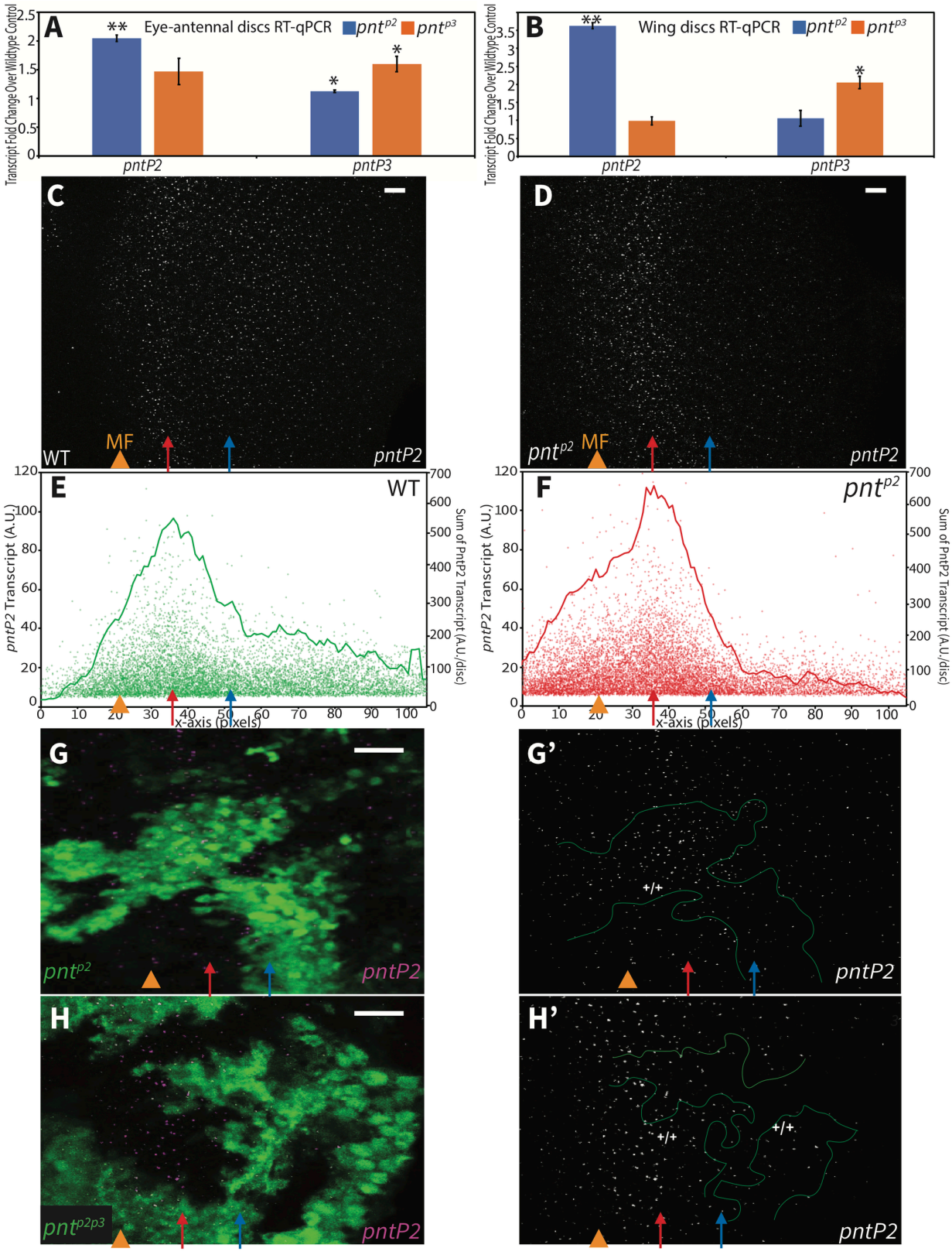
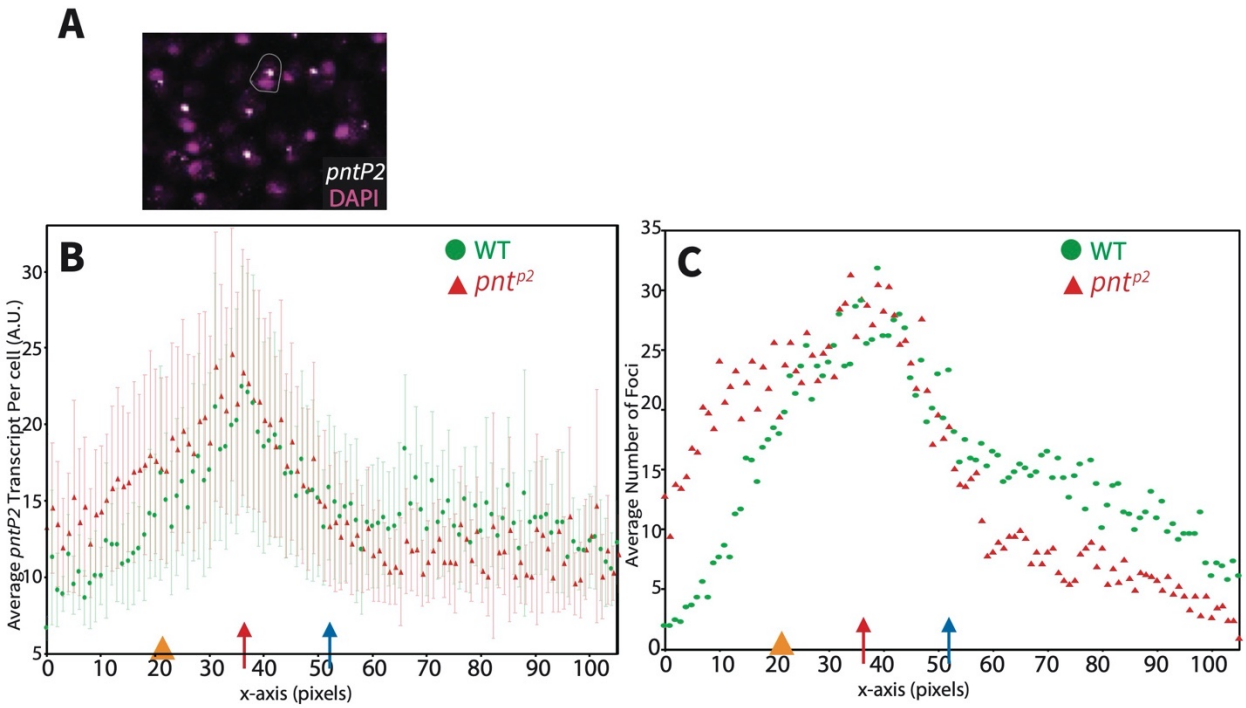


Fig 2.5 Distinct context-specific interactions regulate pntP2 transcription across the eye field

(Fig 2.5 continued) (A, B) RT-qPCR comparison of *pntP2* and *pntP3* transcript levels in wild type versus pnt^{p2} (blue bars) and pnt^{p3} (orange bars) 3rd instar eye-antennal (A) and wing (B) discs. Significant increases were detected. Error bars represent S.D. of three independent experiments. Significance was calculated via pair-wise Student T-tests between the mutant sample and the control. **, $p < 0.01$; *, $p< 0.05$. (C, D) Maximum projection images of *pntP2* FISH in representative wild type and pnt^{p2} 3rd instar eye imaginal discs, oriented anterior left. Orange arrowheads mark the MF, red arrows mark the peak of *pntP2* expression and blue arrows mark the start of lower expression in the posterior half of the disc; the three can be mapped to correspondingly colored marks in Fig. 2.2A based on the pixel distances. In pnt^{p2} discs (D) relative to wild type (C), an increased and broader peak of *pntP2* transcripts was detected in and immediately posterior to the MF while a decrease was seen in the posterior half of the disc. Scale bar: 5 μm. (E, F) Quantification of *pntP2* FISH in 6 wildtype (E) and 6 pnt^{p2} mutant (F) discs from maximum projections. In wild type, *pntP2* levels begin to rise anterior to the MF, peak and decrease to a steady state. In pnt^{p2} discs, *pntP2* levels were higher than normal in the anterior half (left of blue arrow) but lower in the posterior (right of blue arrow). Each dot plots the product of the fluorescent intensity and the size of an individual *pntP2* FISH focus, representing the relative amount of *pntP2* transcript (y-axis on the left) The line connects the moving average of the sum of all foci within one-pixel windows along the x-axis (y-axis on the right). (G) Homozygous pnt^{p2} clones in a 3rd instar eye disc, positively marked with GFP (green). Clone boundary is circled with green line (G'). *pntP2* levels in the mutant clones appeared higher in the anterior region but decreased in the posterior relative to adjacent wild type tissue. Examination of 8 clones in 7 discs from 3 independent experiments showed consistent changes. Images are partial projections. Scale bar: 5 μm. (H) *pntP2* FISH in homozygous pnt^{p2p3} clones, positively marked with GFP (green). *pntP2* levels in anterior mutant clones were indistinguishable from wild type but appeared decreased in more posterior clones. Examination of 9 clones in 6 discs from 2 independent experiments showed consistent changes. Images are partial projections. Scale bar: 5 μm.



S 2.6 Fig pntP2 transcriptional autoregulation

(A) Single optical slice from a 3rd instar eye disc shows individual *pntP2* FISH foci (white dots) in each cell. DAPI (magenta) marks the nuclei.(B-C) Additional quantitative analysis of pntP2 FISH in wild type (green dots) versus pnt^{p2} mutants (red triangles) using maximal projections of the same set of eye imaginal discs used for the analysis in Fig. 2.7C, D. (B) Each dot/triangle plots the average *pntP2* transcript per cell (quantified as the product of focus intensity and area) for each pixel window along the x-axis. Error bars depict standard deviation. Consistent with the data and analysis in Fig. 2.7A-D, the average *pntP2* transcript level per cell is higher in the peak region and lower in the posterior in pnt^{p2} mutant discs than in corresponding regions in wild type discs. (C) Quantification of the number of *pntP2* FISH foci counted. The dots/triangles represent the total number of foci counted for each pixel window along the x-axis. In pnt^{p2} mutant discs, more foci were counted in cells leading up to the peak region and fewer were counted in the posterior (blue arrow and to the right). Together with Fig. S2.6B, this analysis suggests that the increase in *pntP2* transcript in pnt^{p2} mutant is a compound effect of increased transcription in cells normally transcribing *pntP2* at a detectable level, and increased transcription in cells that normally transcribe *pntP2* below a detectable level. Conversely, decreased transcription in cells in the posterior results in both lower average transcript per cell and fewer cells with detectable levels.

2.4.5 PntP2 and PntP3 buffer developmental transitions against compromised RTK Signaling

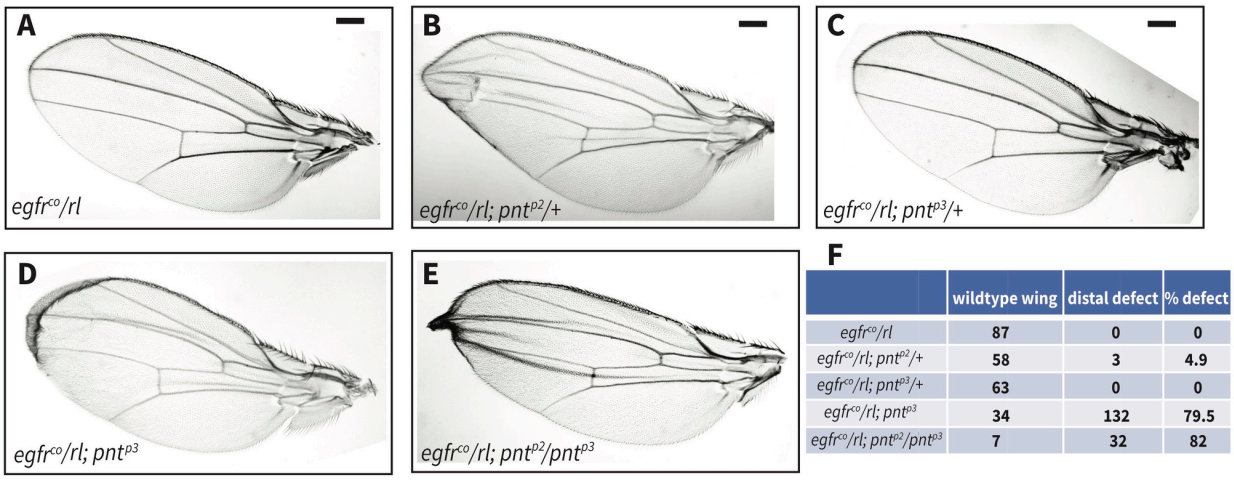
Functional redundancy can provide robustness not only toward loss-of-function mutations in each gene, but also toward variation in the signaling environment in which the gene products function; such variation can result from either genetic or environmental stress. Given that Pnt mediates RTK/MAPK signaling, we asked whether the inclusion of PntP3 in the network protects development from being perturbed when RTK signaling is compromised. We first reduced pathway activity by removing one copy each of *egfr* and of the MAPK encoding gene *rolled (rl)*. We grew the animals either at constant optimal temperature (25ºC) or subjected them to repeated 18ºC to 31ºC temperature shifts. Indicating the remaining activity was sufficient to support normal development, adult *egfr/rl* heterozygotes appeared fully wild type (Fig 2.6A and S2.7 Fig). In the absence of temperature stress, reducing the dose of either *pntp2* or *pntp3* did not produce retinal defects, but in the wing, another context in which PntP2 has been implicated in EGFR-mediated regulation of patterning (Paul et al., 2013), defects were noted (Fig 2.6B, C). Most striking, removal of either both copies of pnt^{p3} or one copy each of pnt^{p2} and pnt^{p3} in the *egfr/rl* background resulted in an 80% penetrant disruption in pattern (Fig 2.6D-F). This suggests that PntP3 confers robustness in situations when RTK/Pnt signaling levels fall below a certain threshold. When temperature stress was added, *egfr/rl;* pnt^{p3} animals died as pharate adults, but again without significant photoreceptor loss (S2.7 Fig). However in retinas dissected from pharate *egfr/rl;* pnt^{p2}/pnt^{p3} quadruple heterozygotes, 12% of ommatidia were missing photoreceptors (S2.7 Fig). Thus the strong genetic synergy noted between *pntP2* and *pntP3* during wing patterning (Figure 2.6F) is also important for retinal development.

Encouraged by these findings, we turned to a more extreme sensitized background to test further the robustness hypothesis in the retina. Specifically we crossed the pnt^{p2} and pnt^{p3} alleles to flies carrying a *Sev-*Yan^{ACT} transgene, a genetic background in which constitutive activity of the RTK

antagonist Yan blocks specification of the photoreceptor fates in which it is expressed (Rebay and Rubin, 1995; Rebay et al., 2000). The *sev* regulatory sequences drive expression mainly in the R3/ R4 pair, in R7 photoreceptors and in cone cells (Basler et al., 1989; Tomlinson et al., 1987). We took advantage of the dose-sensitivity of *Sev-Yan^{ACT}* transgenes (Rebay and Rubin, 1995; Rebay et al., 2000) and selected a line that causes fully penetrant loss of R7 but only partial loss of R3/R4 photoreceptors; as a result, overall disruption to the adult eye pattern is modest (Fig 2.7A, D, G; Rebay and Rubin, 1995). Zooming in and quantifying the number of rhabdomeres in over a thousand ommatidia showed that about ~50% of *Sev-Yan^{ACT}* ommatidia were missing either an R3 or R4 rhabdomere (Fig 2.7D, yellow star and Fig 2.7G, orange bar); and so in contrast to those only missing R7 but retaining the normal complement of 6 large rhabdomeres representing the R1-R6 photoreceptors (Fig 2.7D, red arrow and Fig 2.7G, grey bar), these only had 5 rhabdomeres. Thus, in this genetic background, RTK signaling is just barely sufficient to support R3/R4 specification. Because PntP2 and PntP3 are co-expressed and functionally redundant in normal R3/R4 fate specification (Fig 2.2 and 2.3), this provided an ideal context to assess whether this redundancy provides robustness to compromised signaling

Removal of one copy of either *pntP2* or *pntP3* dominantly enhanced the *Sev-Yan^{ACT}* rough eye phenotype, producing visibly stronger disruptions in the adult eye pattern (Fig 2.7A-C). Quantification of rhabdomere numbers across thousands of individual ommatidia supported this qualitative impression and revealed a shift toward more penetrant photoreceptor loss (Fig 2.7E-G). Thus introducing heterozygosity for either *pnt^{p2}* or *pnt^{p3}* reduced the frequency of ommatidia with the full complement of R1-R6 rhabdomeres to only ~10-20% (Fig 2.7G, grey bars) and increased the frequency of R3/R4 loss (Fig 2.7G. orange and blue bars). Loss of either both copies of *pnt^{p3}* or one copy each of *pnt^{p2}* and *pnt^{p3}* enhanced even further, with quantification

showing similar patterns of increased photoreceptor loss (Fig 2.7G; S2.8A, B Fig). In these enhanced backgrounds, ommatidia with fewer than four rhabdomeres were occasionally found (Fig 2.7F, red star); this could reflect the additional loss of R1 or R6 cells, where Sev drives expression at much lower levels (Tomlinson et al., 1987; Tomlinson et al., 2011) or a later consequence of cone cell loss (Rebay and Rubin, 1995) on overall photoreceptor survival. Overall these results reveal non-redundant contributions of both PntP2 and PntP3 to R3/R4 photoreceptor specification under conditions of reduced RTK signaling.



F

	wildtype wing	distal defect	% defect
$egfr^{co}/rl$	87	0	0
$egfr^{co}/rl; pnt^{p2}/+$	58	3	4.9
$egfr^{co}/rl; pnt^{p3}/+$	63	0	0
$egfr^{co}/rl; pnt^{p3}$	34	132	79.5
$egfr^{co}/rl; pnt^{p2}/pnt^{p3}$	7	32	82

Fig 2.6 PntP2 and PntP3 buffer against the reduction of MAPK activation during wing patterning

Adult wings are oriented anterior up and distal left. **(A-F)** Representative adult wings showing the effects of reduced pnt^{p2} and pnt^{p3} dose in a sensitized $egfr^{co}$ */rl* background. (A) $egfr^{co}/rl$ trans-heterozygotes appeared wild type. (B) Loss of one copy of pnt^{p2} produced occasional distal margin defects. (C) Loss of one copy of pnt^{p3} did not disrupt patterning. (D) Loss of both copies of pnt^{p3} resulted in penetrant distal margin defects. (E) Simultaneous reduction in dose of pnt^{p2} and pnt^{p3} synergistically increased wing margin defects. (F) Quantification of wings as either wild type or with margin defects for each genotype in (A-E). Scale bar: 0.1 mm.

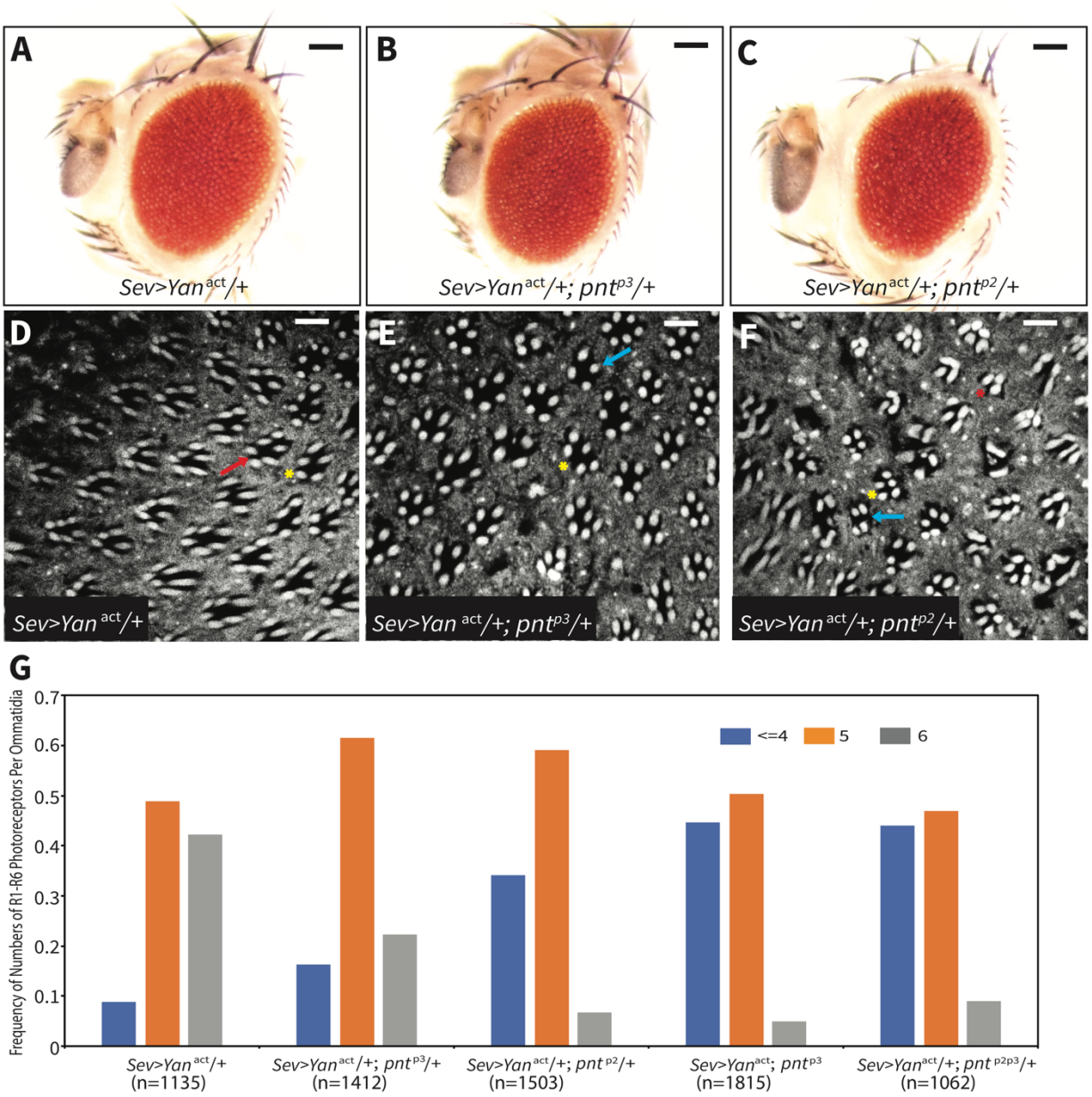
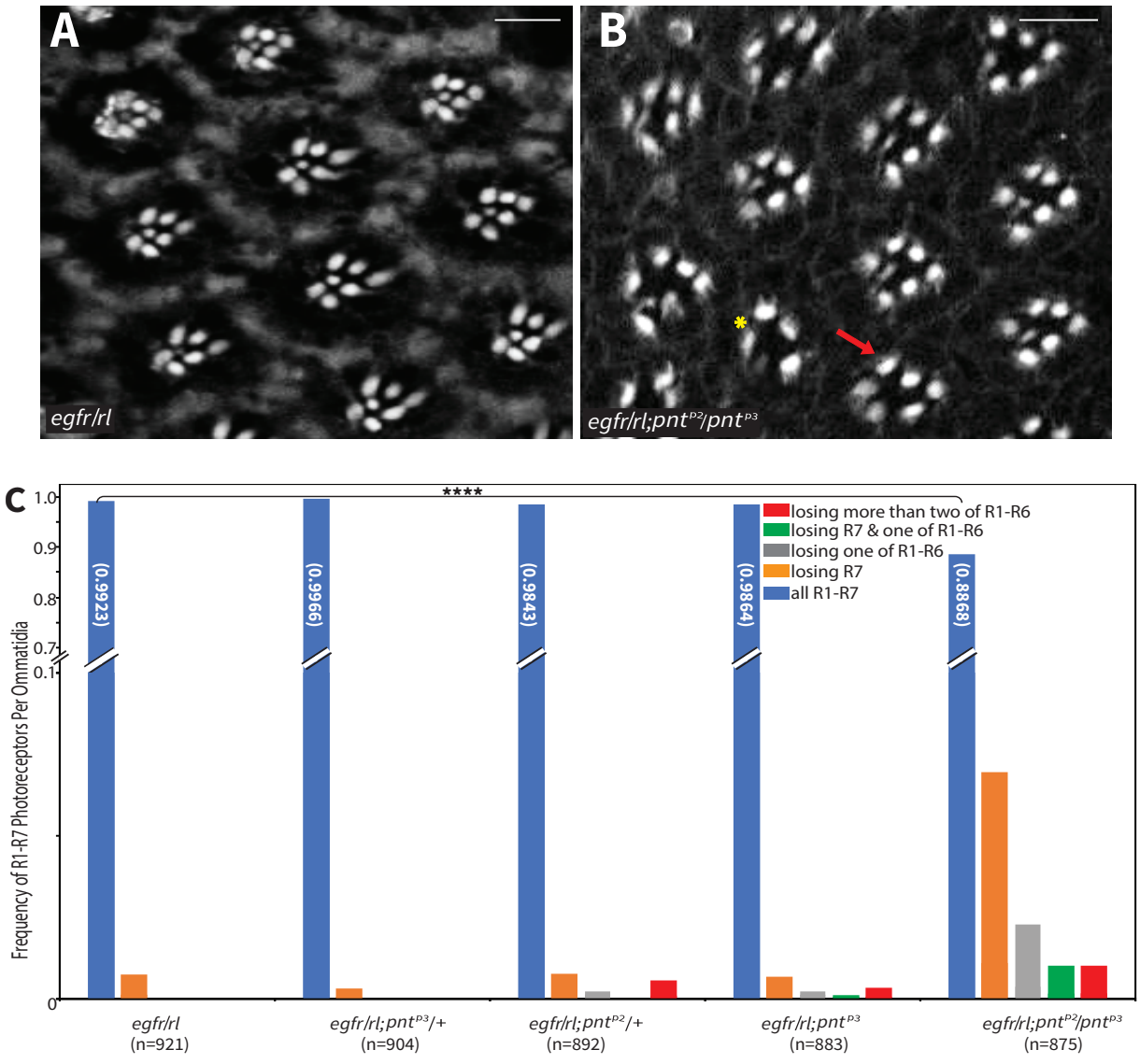


Fig 2.7 PntP2 and PntP3 stabilize R3/R4 fate transitions against compromised RTK signaling

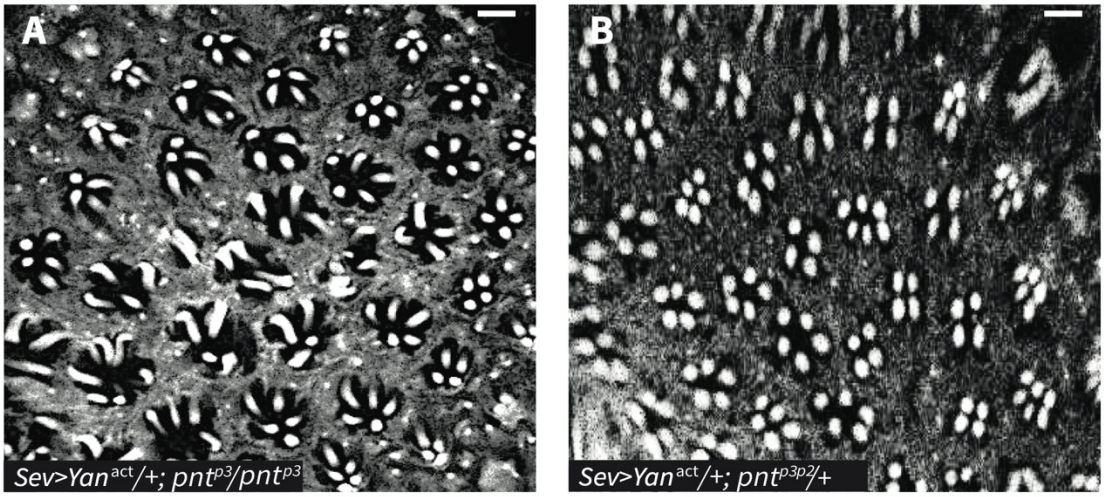
(A-F) Heterozygosity for either *pnt*^{p3} or *pnt*^{p2} dominantly enhances the *Sev-Yan*^{act} induced disruption in external eye morphology and photoreceptor loss. (A-C) Representative adult eyes of indicated genotype. Scale bar: 50 µm. (D-F) Phalloidin staining of representative adult eyes of indicated genotype. Red arrow points to the ommatidium lacking the R7 rhabdomere, yellow stars indicate ommatidia with five outer rhabdomeres, blue arrows point to ommatidia with four outer rhabdomeres and red star indicates the ommatidium with three outer rhabdomeres. Scale bar: 5 µm. (G) Quantification of photoreceptor loss in phalloidin-stained adult eyes expressed as the frequency of ommatidia with 6, 5, or <=4 remaining outer rhabdomeres. n represents the number of ommatidia scored for each genotype. In the *Sev-Yan*^{act} controls, ommatidia with 5 or 6

(Fig 2.7 continued) outer rhabdomeres were present at a roughly 1:1 ratio and less than 10% of ommatidia had fewer than 4. Heterozygosity for either *pnt*^{p3} or *pnt*^{p2} shifted this distribution: for both, the majority of ommatidia (~60%) had 5 outer rhabdomeres. However reduction in *pnt*^{p2} caused a larger increase in fraction of ommatidia with 4 or fewer outer rhabdomeres than did reduction in *pnt*^{p3}. When both copies of *pnt*^{p3} were removed, ommatidia with <=4 or 5 outer rhabdomeres were present at a roughly 1:1 ratio and less than 10% of ommatidia had 6 rhabdomeres. A similar enhancement pattern was scored in animals heterozygous for the double mutant *pnt*^{p2p3}.



S 2.7 Fig Combining genetic with temperature stress reveals contributions of pntP2 and pntP3 to developmental robustness

(S 2.7 Fig continued) (A-B) Single optical slice of phalloidin staining of representative adult eyes of indicated genotype. The animals had been subjected to temperature stress during larval development. *egfr/rl* (A) , *egfr/rl; pnt^{p3}/pnt^{p2}* (B). Red arrow points to an ommatidium lacking the central R7 rhabdomere and yellow star points to an ommatidium that also lost outer R1-R6 rhabdomeres. Scale bars: 5 μm. (C) Quantification of photoreceptor loss in phalloidin-stained adult eyes expressed as the frequency of ommatidia with all seven rhabdomeres (R1-R7), losing R7 only, losing one of the outer R1-R6, or losing both R7 and more than one outer rhabdomeres. n represents the number of ommatidia scored for each genotype. In the *egfr/rl* controls, more than 99% ommatidia had all R1-R7 rhabdomeres. Heterozygosity for either pnt^{p3} or pnt^{p2} or homozygous loss of pnt^{p3} were essentially indistinguishable from control. In contrast, in *egfr/rl; pnt^{p3}/pnt^{p2}* retinas, only 88% of ommatidia had the full complement of R1-R7 rhabdomeres. Of the 12% that lost photoreceptors, 7% lost only R7, 2% lost only one outer, 1% lost both R7 and one outer and 1% lost more than one outer. Significance was calculated via two tailed Student T-tests between each mutant and the *egfr/rl* control. Only *egfr/rl; pnt^{p3}/pnt^{p2}* showed significant photoreceptor loss, ****, $p < 0.0001$.



S 2.8 Fig Reducing the dose of pntp3 and pntp2 enhances Sev-Yanact induced photoreceptor loss

(A-B) Phalloidin staining of representative adult eyes of indicated genotype showing that homozygous pnt^{p3} (A) or heterozygous pnt^{p2p3} (B) enhances the *Sev-Yan^{act}* induced photoreceptor loss. Scale bar: 5 μm.

2.5 Discussion

In this study we explore the contributions of a previously uncharacterized Pointed isoform, PntP3, to the transcriptional effector network that directs developmental transitions downstream of receptor tyrosine kinase signaling. We show that PntP3, like PntP2, functions as a MAPK responsive transcription factor, but that despite their molecular and functional similarities, PntP3 and PntP2 have distinct expression patterns, transcriptional activities and mutant phenotypes. Together our results suggest that essential regulatory responsibilities previously attributed solely to PntP2, are actually distributed between PntP2 and PntP3, and that depending on context, the two work redundantly, uniquely or synergistically. We speculate that a network of auto- and cross-regulatory interactions between the isoforms fine-tunes Pnt transcriptional output to confer specificity and robustness to the developmental transitions it directs.

Our investigation of the PntP3 isoform has uncovered a context-dependent bifurcation in the transcriptional effector network that transduces RTK/MAPK signaling. In doing so, it has also corrected an erroneous assumption regarding the role of the closely related PntP2 isoform. Prior to our study, the accepted model was that MAPK phosphorylation of PntP2, followed by PntP2p-mediated induction of *pntP1* transcription, provided the essential activating input for RTK-dependent transitions (Fig 2.1C; Shwartz et al., 2013). As exemplified by studies in the eye, the genetic cornerstone of this model was that null alleles of either *pntP2* or *pntP1* produce identical phenotypes of failing to specify R1-R7 fates (O'Neill et al., 1994; Shwartz et al., 2013; Yang and Baker, 2003). However the allele $pnt^{\Delta 78}$ (O'Neill et al., 1994), previously misinterpreted as a *pntP2*-specific null, actually disrupts the exon common to *pntP2* and *pntP3*. Thus the failure to specify R1-R7 fates reflects the compound loss of PntP2 and PntP3. We note that an earlier study using hypomorphic truly *pntP2*-specific alleles concluded correctly that there is an “absolute

requirement for *pntP2* function in R1, R6 and R7" but did not detect the requirement in R2-R5 (Brunner et al., 1994a).

A schematic summarizing the combined contributions of PntP3 and PntP2 to photoreceptor fate specification is presented in Figure 2.8 as a framework for considering some of the mechanistic implications of our work. To recap briefly the key phenotypes and regulatory interactions on which the model is based, our study revealed redundant functional requirements for PntP2 and PntP3 in specifying the first round fates R2/R5/R3/R4; thus either single mutant recruits wild type 5-cell ommatidial clusters, while only in the double mutant are R2-R5 fates lost. Molecularly, PntP2 and PntP3 redundantly activate *pntP1* transcription (Fig 2.8A) with significant reduction in *pntP1* levels detected only in the double mutant. In contrast, only PntP2 is required during the second round of photoreceptor specification and so eyes from isoform-specific pnt^{p2} null mutants lack R1, R6, R7 fates whereas pnt^{p3} mutant ommatidia are wild type. Because *pntP1* transcript levels posterior to the SMW are already quite low in wild type discs, our FISH experiments were unable to detect the presumed reduction in *pntP1* in pnt^{p2} mutant discs.

As a general developmental strategy, the redundant use of PntP2 and PntP3 may provide an effective buffer against genetic perturbations that reduce RTK signaling. Using R2-R5 photoreceptor specification as a specific example, the presence of redundant MAPK effectors in the early stages of ommatidial assembly may maximize overall robustness by minimizing early "mistakes" that would derail the entire process. Supporting this idea, we found that in a genetically sensitized background with reduced MAPK signaling output in R3, R4 precursors

(Rebay and Rubin, 1995; Rebay et al., 2000), loss or reduction in dose of either *pntP2* or *pntP3,* which in otherwise wild type discs did not compromise patterning, now resulted in loss of these cell fates. Analogous results were obtained in the wing, and when temperature stress was added on top of genetic stress, animals lacking PntP3 failed to eclose. Thus redundant use of PntP2 and PntP3 can confer developmental robustness.

Just as inadequate signaling compromises developmental transitions, so will excessive, oncogenic-levels of pathway activation. For example, genetic perturbations that enhance RTK pathway output, such as increased Pnt expression or activity, severely disrupt ommatidial assembly and wing patterning (Brunner et al., 1994b; Karim et al., 1996; Prober and Edgar, 2000). Therefore to prevent redundant use of PntP2 and PntP3 from overactivating transcriptional programs, the Pnt output needs to be fine-tuned.

The negative auto-regulation of both *pntP2* and *pntP3* transcript levels uncovered in our study may serve this purpose (Fig 2.8A). Although Pnt is well-established as a transcriptional activator, a handful of studies have implicated Pnt in negative regulation of gene expression (Kurada and White, 1998; Rohrbaugh et al., 2002; Webber et al., 2018; Zhu et al., 2011).The underlying molecular mechanisms are still under investigation, but based on our prior work showing extensive Pnt chromatin occupancy across the *pnt* locus (Webber et al., 2018), we favor a mechanism in which direct auto-repression keeps PntP2 and PntP3 levels in check. However an indirect mechanism involving Pnt-mediated transcriptional activation of a repressive factor is equally plausible. If direct auto-regulation is used, the ability of Pnt to recruit and co-occupy

enhancers with the ETS family repressor Yan and the corepressor Groucho uncovered in a recent study (Webber et al., 2018) could provide the repressive mechanism.

Counteracting the negative auto-regulation at *pntP2* and *pntP3*, we also uncovered positive transcriptional cross-regulation whereby PntP3 can activate *pntP2*. Thus in pnt^{p2p3} double mutant clones, the increase in *pntP2* transcript levels that occurs in pnt^{p2} single mutants was no longer observed. Again, we favor the simplest model of direct activation of *pntP2* by PntP3 (Fig 2.8A), but cannot rule out more complicated indirect regulatory relays. Whether the converse cross-regulation of *pntP3* transcription by PntP2 occurs, and whether PntP2 and/or PntP3 positively auto-regulate their transcription anterior to the SMW remains to be assessed. The decrease in *pntP2* transcript levels measured posterior to the SMW in pnt^{p2} mutant tissue argues that positive auto-regulation is possible, making it plausible that such regulation could also fine-tune PntP2/PntP3 levels and output during specification of first round fates.

How specific PntP2:PntP3 ratios influence the acquisition of different photoreceptor cell fates will be in an interesting focus for future work. Numerous studies have shown that regulatory networks can either amplify or suppress both the intrinsic noise (i.e. the randomness associated with transcription and translation) and extrinsic noise (i.e. the fluctuations in cellular processes or environment) of protein levels to influence cell fate decisions (Chang et al., 2008; Singh and Hespanha, 2009; Voliotis and Bowsher, 2012). Very speculatively, perhaps the network of auto-repressive and cross-activating interactions between PntP2 and PntP3 also tunes the cell-to-cell variation in Pnt isoform or Pnt target gene expression, thereby influencing the response to inductive signaling.

Another intriguing feature of the network of transcriptional interactions uncovered in our study is that corresponding to the switch from redundancy between PntP2 and PntP3 to the uniqueness of PntP2, the balance of PntP2 autoregulation shifts from repression during first round fate specification to activation during the second round (Fig 2.8A). Fig 2.8B offers speculation on how the distinct RTK signaling environments anterior vs. posterior to the SMW, combined with intrinsic differences in PntP2 vs. PntP3 activity, could produce this shift. Briefly, we propose that the level of MAPK activity determines the ratios between the unphosphorylated forms of PntP2 and PntP3 and the phosphorylated forms, PntP2p and PntP3p, and that these ratios in turn dictate specific transcriptional output.

For example, during photoreceptor specification, R2-R5 first round fates rely exclusively on EGFR signaling while the R1, R6, R7 photoreceptors specified during the second-round experience additional RTK signaling through Sevenless (Sev); studies focused on R7 specification have highlighted the requirement for both EGFR and Sev (Basler and Hafen, 1989; Stark et al., 1976; Tomlinson et al., 2019). Using the same Ras/MAPK/Pnt pathway, EGFR and Sev-initiated signals can be considered interchangeable (Fortini et al., 1992; Freeman, 1996), with lower pathway activity required in the first round and higher activity needed in the second (Tomlinson et al., 2011; Tomlinson et al., 2019). Because both PntP2 and PntP3 are direct MAPK substrates whose transactivation potential is increased by phosphorylation, their combined transcriptional output will be sensitive to the abundance of activated MAPK. Under conditions of lower signaling and when both isoforms are co-expressed, as occurs anterior to the SMW, competition for the limited pool of activated MAPK will lead to domination by the

unphosphorylated, less active forms. The presence of PntP3, whose unphosphorylated form has equivalent activity to PntP2p, and whose phosphorylated form has twice the activity of PntP2p (Fig 2.1D), may be important to make sure pathway output remains above a certain threshold in situations with lower levels of signaling. Although at first glance this might predict that the system would not tolerate loss of PntP3, because loss of PntP3 also reduces MAPK substrate competition, this would shift the distribution of PntP2 protein toward the phosphorylated more active form, thereby ensuring a robust transcriptional response.

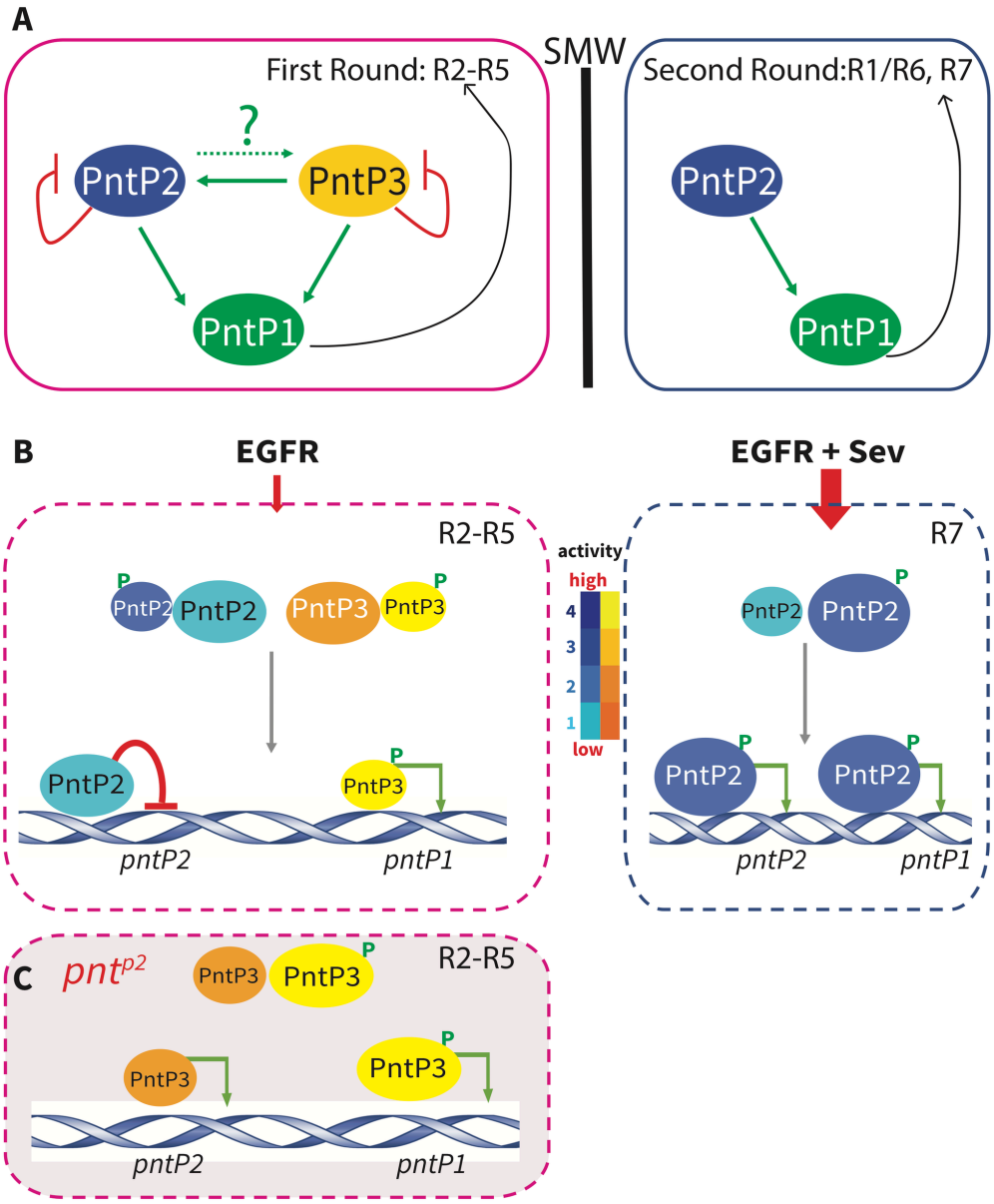


Fig 2.8 Model: Context-specific topology and function of the Pnt network

(A) A schematic summary of the Pnt network. During first-round specification, both PntP2 and PntP3 auto-repress their transcription, PntP3 activates *pntP2*, and PntP2 and PntP3 redundantly activate *pntP1*. During the second-round, PntP2 auto-activates its own transcription and activates *pntP1*; PntP3 does not contribute. (B) Proposed Pnt network functions in the different signaling environments of first and second round specifications. The color scheme illustrates the range of transactivation activity: for PntP3, yellow indicates high activity and orange low; for PntP2 dark blue indicates high activity and light blue low. As depicted for PntP2, the low transactivation potential of the unphosphorylated forms may allow the proteins to operate as transcriptional (Fig

2.8 continued) repressors. Different sized ovals depict relative abundance of the phosphorylated versus unphosphorylated forms. See Discussion for details.

How might these relationships manifest at the level of target gene enhancers? Given that PntP2 and PntP3 have the same ETS DNA binding domain and are identical except for the sequences N-terminal to the SAM, we expect they recognize the same DNA binding sites. Thus in the simplest scenario in which the phosphorylated and unphosphorylated forms of both PntP2 and PntP3 compete equally for enhancer occupancy, situations in which the unphosphorylated forms predominate would prevent excess activation of target genes. Much greater regulatory complexity is possible if modest enhancer-specific preferences between PntP2 and PntP3 and between the phosphorylated and unphosphorylated forms bias the competition. We suggest such biased competition will be essential to achieving limited activation, or even repression, of target genes such as *pntP2*, while allowing strong induction of others, such as *pntP1,* in the same cell. Based on a large-scale interactome study that reported closely related isoform pairs often have distinct protein-protein interaction patterns (Yang et al., 2016), it is possible that association with distinct cofactors also contributes to Pnt isoform enhancer occupancy bias.

The substrate competition-based model also readily explains the transcriptional shifts that may occur in the individual pnt^{p2} and pnt^{p3} mutants (Fig 2.8C). If one removes either PntP2 or PntP3, then overall competition for activated MAPK is eased, resulting in domination by the phosphorylated form of the remaining protein to boost transcriptional output. This would derepress targets like *pntP2*, as detected in our experiments, while activation of targets like *pntP1* would continue at physiologically functional levels. This same scenario plays out in an even stronger form in the wild type disc during specification of second round photoreceptor

fates, where the combination of only PntP2 plus twice the RTK pathway input would result in phosphorylation of an even greater proportion of total PntP2 protein (Fig 2.8B). Because PntP2 appears to have intrinsically weaker transactivation potential than PntP3, ensuring full phosphorylation in situations where it is the sole MAPK effector may be critical to activating the transcriptional program.

Besides its role in the specification of first round photoreceptor fates, the expression of PntP3 suggests it might also contribute to the patterning of R8 and the cone cells. For R8, although its specification does not require EGFR/Pnt activation, EGFR signaling is essential for the proper spacing between R8 cells (Domínguez et al., 1998; Frankfort and Mardon, 2002; Kumar et al., 1998; Yang and Baker, 2001). As GFP-PntP3 and *pntP1* both express at high levels at the MF, they may have redundant functions in R8 spacing. For cone cells, previous work has suggested that as with first round photoreceptor fates, EGFR is the sole RTK involved in their specification (Freeman, 1996). Our finding that GFP-PntP3 is coexpressed with *pntP2* in cone cells, combined with the partial loss of Cut-positive cells that we noted in pnt^{p2} mutant discs, suggests PntP3 and PntP2 may together ensure robust specification of cone cell fates. Further investigation of the function and regulation of Pnt isoforms in a broad range of developmental contexts will be an interesting direction for further studies.

Our study adds to the growing appreciation of the enormous regulatory potential available to developing tissues through the combinatorial expression and use of different protein isoforms, and also offers insights beyond the *Drosophila* arena. The human homologs ETS1 and ETS2, although encoded by separate genes, show intriguing structural and functional parallels to the *Drosophila* PntP2 and PntP3 isoforms (Wasylyk et al., 1997; Watson et al., 1988). For example,

ETS1 and ETS2 have distinct sequences at the N-terminal end of the conserved transactivation domain, similar MAPK responsiveness, and overlapping but not identical functions and expression patterns (Albagli et al., 1994; Wasylyk et al., 1997). Given these striking parallels, continued exploration of the molecular mechanisms underlying Pnt-mediated transcriptional responses may provide new insight into signaling robustness and specificity in mammalian systems.

2.6 Materials and Methods

2.6.1 Drosophila strains

From the Bloomington *Drosophila* Stock Center: *Lz-Gal4*, $pnt^{\Delta 88}$, pnt^{1277}, *FRT82b,* $egfr^{co}$. Additional strains: *ro-GAL4* (Mavromatakis and Tomlinson, 2013), $Sev\text{-}yan^{ACT}$ (Rebay et al., 2000), rl^{S135} (Karim et al., 1996), *ey-FLP; act-Gal4, UAS-GFP/CyO; FRT82b, tub-GAL80/TM6B* (a gift of Wei Du, University of Chicago, IL, USA), *UAS-flag-pntP2*, *UAS-flag-pntP3*, GFP-PntP3, pnt^{p2}, pnt^{p3}, pnt^{p2p3} (this work). Flies were cultured at 25°C unless otherwise indicated. For further details of *Drosophila* strains and genetics, see the supplementary Materials and Methods.

2.6.2 Temperature stress

Following the experiment published by Li et al 2009 (Li et al. , 2009), flies of the relevant genotypes were crossed at 25°C and transferred to fresh bottles daily. Cultures were maintained at 25°C until larvae reached early third instar, shifted to 31°C for 16–24 hr, and then subjected to

seven to ten rounds of temperature cycles. Each round consisted of a shift to 18°C for 2 hours followed by 31°C for 1.5-2 hours. At the completion of the final round, the late third-instar larvae were returned to 25°C and then retinas were dissected from eclosed or pharate adults. Bottles were incubated in air-circulating incubators for each temperature step.

2.6.3 Transcription assays

2.25×10^6 *Drosophila* S2 cells were transfected with 100 ng of 6X-ETS luciferase reporter construct, 100 ng of PntP2/pMTHA or 100 ng of PntP3/pMTHA, 20 ng of actin >Renilla luciferase, and if applicable, 5 ng of Ras^{V12}/pMT. For further details of transcription assay experiments, see the supplementary Materials and Methods.

2.6.4 Immunohistochemistry and microscopy

Eye-antennal and wing imaginal discs were fixed in 4% paraformaldehyde (PFA) for 10 min. For adult retina staining, decapitated adult heads were fixed in 4% PFA for 20 min, and then the dissected retinas were post-fixed in 4% PFA for 10 minutes. Primary and secondary antibodies were incubated overnight at 4°C. For antibody details, see supplementary Materials and Methods. Imaging was performed with a Zeiss LSM 880 confocal microscope, using 0.8 to 1.0 µm steps and projecting maximally through the desired tissue unless otherwise noted. To image adult eyes and wings, decapitated heads and dissected wings were imaged with a Canon EOS Rebel camera fitted to a Leica stereo microscope. Individual slices were merged using iSolution-Lite software (IMT-Digital).

2.6.7 Quantitative reverse transcription PCR (RT-qPCR)

Total RNA was extracted from 60 pairs of late 3rd instar eye-antennal or wing discs for cDNA synthesis. qPCR was performed using the iTaq Universal SYBR Green Supermix (Bio-Rad Laboratories) on a 7300 Real-time PCR Machine (Applied Biosystems). Subsequent disassociation analysis was performed with 7300 system software to confirm the sequence specificity of the reaction. For further details of RT-qPCR and primers used, see the supplementary Materials and Methods.

2.6.8 CRISPR/Cas9-mediated generation of pnt mutants

Genomic DNA containing the *pntP2* or *pntP3* specific exon regions was amplified by PCR. After insertion of stop codons and deletion of coding sequence, the fragments were inserted into pHD-Scarless (generated by O'Connor-Giles laboratory, *Drosophila* Genomics Resource Center, 1364; Gratz et al., 2013). The resulting templates were confirmed by sequencing. Guide RNAs were subcloned into the pU6-Bbs1 chiRNA plasmid (Addgene, 45946; Gratz et al., 2013). Each template (300 ng/µL) and the two guide RNAs (75 ng/µL), were injected into a GFP/ RFP-negative *vasa-Cas9* strain (a gift from Rick Fehon). To generate pnt^{p2p3}, the template and guide RNAs used to generate pnt^{p2} were injected with *nanos-Cas9* plasmid (a gift from Rick Fehon, (Ren et al., 2013)) into the pnt^{p3} strain. G_0 adults were crossed individually to w^{1118}, and transformants were identified by 3X-Pax-RFP expression in the eyes of the F1 progeny. The 3X-Pax-RFP piggyBac cassette was excised, and RFP-negative progeny were crossed to *TM6B* to establish stocks. The alleles were confirmed by restriction digest and sequencing. For further details and sequences of the CRISPR templates and guide RNAs, see the supplementary Materials and Methods.

2.6.9 Fluorescence In situ hybridization (FISH)

DNA probes conjugated with NHS ester-ATT0 633 fluorophore were used to target specific exons of *pnt* isoforms. Dissected eye discs from white pre-pupae were fixed in 1% PFA, dehydrated with methanol, and incubated with probes at 62°C. Imaging was performed with a Zeiss LSM 880 confocal microscope, using 0.8 μm steps and projecting maximally through the tissue. Maximum projected images were analyzed with Fiji and Microsoft Excel. In all pair-wise comparison of wild type vs. *pnt* mutant, discs from the two strains were dissected, processed and imaged in parallel. For details of probes used, FISH protocol and image analysis, see the supplementary Materials and Methods.

2.7 Acknowledgements

We thank Jemma Webber and Matt Hope for initial analysis of PntP3 transcriptional activities; Matt Hope for making expression constructs; Wei Du for fly strains; Jiajie Xu, Will Yee and Hitoshi Matakatsu in Rick Fehon's lab for advice on RT-qPCR and CRISPR; Rachael Bakker in Rich Carthew's lab and Jiacheng Zhang in Jingyi Fei's lab for advice on FISH; Nicolas Pelaez in Rich Carthew's lab for advice on GFP-PntP3 quantification; Claude Desplan for Salm antibody; Hugo Bellen for Sens antibody; Rick Fehon for the equipment to photograph adult eyes; Ed Munro, Chip Ferguson, Mike Rust, Rich Carthew, and Rebay lab members Trevor Davis, Nicelio Sanchez-Luege, Xiao Sun, Julio Miranda-Alban, and Suzy Hur for helpful discussions. Chip Ferguson, Matt Hope and Jemma Webber provided helpful comments on the manuscript. The work was supported by R01GM080372 to IR. Further support from NIH R01 EY025957 to IR and from the Genomics Core Facility through a UChicago Cancer Center Support Grant P30 CA014599 is acknowledged.

2.8 References

Albagli, O., Soudant, N., Ferreira, E., Dhordain, P., Dewitte, F., Bégue, A., Flourens, A., Stéhelin, D. and Leprince, D. (1994). A model for gene evolution of the ets-1/ets-2 transcription factors based on structural and functional homologies. *Oncogene* **9 11**, 3259–3271.

Basler, K. and Hafen, E. (1988). Sevenless and Drosophila eye development: a tyrosine kinase controls cell fate. *Trends Genet.* **4**,.

Basler, K. and Hafen, E. (1989). Dynamics of Drosophila eye development and temporal requirements of sevenless expression. *Development* **107**, 723–731.

Basler, K., Siegrist, P. and Hafen, E. (1989). The spatial and temporal expression pattern of sevenless is exclusively controlled by gene-internal elements. *EMBO J.* **8**, 2381–2386.

Brunner, D., Dücker, K., Oellers, N., Hafen, E., Scholz, H. and Klämbt, C. (1994a). The ETS domain protein pointed-P2 is a target of MAP kinase in the sevenless signal transduction pathway. *Nature* **370**, 386–389.

Brunner, D., Oellers, N., Szabad, J., Biggs, W. H., Zipursky, S. L. and Hafen, E. (1994b). A gain-of-function mutation in Drosophila MAP kinase activates multiple receptor tyrosine kinase signaling pathways. *Cell* **76**, 875–888.

Cagan, R. (2009). *Principles of Drosophila Eye Differentiation*. 1st ed. Elsevier Inc.

Celniker, S. E., Dillon, L. A. L., Gerstein, M. B., Gunsalus, K. C., Henikoff, S., Karpen, G. H., Kellis, M., Lai, E. C., Lieb, J. D., MacAlpine, D. M., et al. (2009). Unlocking the secrets of the genome. *Nature* **459**, 927–930.

Chang, H. H., Hemberg, M., Barahona, M., Ingber, D. E. and Huang, S. (2008). Transcriptome-wide noise controls lineage choice in mammalian progenitor cells. *Nature*

453, 544–547.

Domingos, P. M., Mlodzik, M., Mendes, C. S., Brown, S., Steller, H. and Mollereau, B. (2004). Spalt transcription factors are required for R3/R4 specification and establishment of planar cell polarity in the Drosophila eye. *Development* **131**, 5695–5702.

Domínguez, M., Wasserman, J. D. and Freeman, M. (1998). Multiple functions of the EGF receptor in Drosophila eye development. *Curr. Biol.* **8**, 1039–1048.

Félix, M. A. and Barkoulas, M. (2012). Robustness and flexibility in nematode vulva development. *Trends Genet.* **28**, 185–195.

Flores, G. V., Duan, H., Yan, H., Nagaraj, R., Fu, W., Zou, Y., Noll, M. and Banerjee, U. (2000). Combinatorial signaling in the specification of unique cell fates. *Cell* **103**, 75–85.

Fortini, M. E., Simon, M. A. and Rubin, G. M. (1992). Signalling by the sevenless protein tyrosine kinase is mimicked by Has1 activation. **355**, 559–561.

Frankfort, B. J. and Mardon, G. (2002). R8 development in the Drosophila eye: A paradigm for neural selection and differentiation. *Development* **129**, 1295–1306.

Freeman, M. (1996). Reiterative Use of the EGF Receptor Trigger Differentiation of All Cell Types in the Drosophila Eye. *Cell* **87**, 651–660.

Gratz, S. J., Cummings, A. M., Nguyen, J. N., Hamm, D. C., Donohue, L. K., Harrison, M. M., Wildonger, J. and O'connor-Giles, K. M. (2013). Genome engineering of Drosophila with the CRISPR RNA-guided Cas9 nuclease. *Genetics* **194**, 1029–1035.

Guillemot, F. (2007). Spatial and temporal specification of neural fates by transcription factor codes. *Development* **134**, 3771–3780.

Halfon, M. S., Carmena, A., Gisselbrecht, S., Sackerson, C. M., Jiménez, F., Baylies, M. K. and Michelson, A. M. (2000). Ras pathway specificity is determined by the integration of

multiple signal-activated and tissue-restricted transcription factors. *Cell* **103**, 63–74.

Hayashi, T., Xu, C. and Carthew, R. W. (2008). Cell-type-specific transcription of prospero is controlled by combinatorial signaling in the Drosophila eye. *Development* **135**, 2787–2796.

Jean-François Boisclair Lachance, Nicolás Peláez, Justin J. Cassidy, J. L. and Webber, Ilaria Rebay, and R. W. C. (2014). A comparative study of Pointed and Yan expression reveals new complexity to the transcriptional networks downstream of receptor tyrosine kinase signaling. *Dev Biol.* **385**, 263–278.

Karim, F. D., Chang, H. C., Therrien, M., Wassarman, D. A., Laverty, T. and Rubin, G. M. (1996). A screen for genes that function downstream of Ras1 during Drosophila eye development. *Genetics* **143**, 315–329.

Klambt, C. (1993). The Drosophila gene pointed encodes two ETS-like proteins which are involved in the development of the midline glial cells. *Development* **117**, 163–176.

Kumar, J. P., Tio, M., Hsiung, F., Akopyan, S., Gabay, L., Seger, R., Shilo, B. Z. and Moses, K. (1998). Dissecting the roles of the Drosophila EGF receptor in eye development and MAP kinase activation. *Development* **125**, 3875–3885.

Kurada, P. and White, K. (1998). Ras Promotes Cell Survival in Drosophila by Downregulating hid Expression. *Cell* **95**, 319–329.

Leader, D. P., Krause, S. A., Pandit, A., Davies, S. A. and Dow, J. A. T. (2018). FlyAtlas 2: A new version of the Drosophila melanogaster expression atlas with RNA-Seq, miRNA-Seq and sex-specific data. *Nucleic Acids Res.* **46**, D809–D815.

Li, X., Cassidy, J.J., Reinke, C.A., Fischboeck, C. and Carthew, R.W. (2009). A microRNA imparts robustness against environmental fluctuation during development. *Cell* **137**, 273–83.

Liu, K., Xu, K. and Song, Y. (2019). Faster, higher, stronger: Timely and robust cell fate/identity commitment in stem cell lineages. *Open Biol.* **9**,.

Mavromatakis, Y. E. and Tomlinson, A. (2013). Switching cell fates in the developing Drosophila eye. *Development* **140**, 4353–61.

Morimoto, A. M., Jordan, K. C., Tietze, K., Britton, J. S., O'Neill, E. M. and Ruohola-Baker, H. (1996). Pointed, an ETS domain transcription factor, negatively regulates the EGF receptor pathway in Drosophila oogenesis. *Development* **122**, 3745–3754.

Nicola´s Pela´ez, Arnau Gavalda-Miralles, Bao Wang, Heliodoro Tejedor Navarro, Herman Gudjonson, Ilaria Rebay, Aaron R Dinner, Aggelos K Katsaggelos, Luı´s AN Amaral, R. W. C. (2015). Dynamics and heterogeneity of a fate determinant during transition towards cell differentiation. *Elife* **53**, 160.

O'Neill, E. M. D., Tjian, R. and Rubin, G. M. (1994). The Activities of Two Ets-Related Transcription Factors Required for Drosophila Eye Development Are Modulated by the Ras / MAPK Pathway. *Cell* **78**, 137–147.

Paul, L., Wang, S., Manivannan, S. N., Bonanno, L., Lewis, S., Austin, C. L. and Simcox, A. (2013). Dpp-induced Egfr signaling triggers postembryonic wing development in Drosophila. *Proc. Natl. Acad. Sci.* **110**, 5058–63.

Prober, D. A. and Edgar, B. A. (2000). Ras1 promotes cellular growth in the Drosophila wing. *Cell* **100**, 435–446.

Qiao, F., Harada, B., Song, H., Whitelegge, J., Courey, A. J. and Bowie, J. U. (2006). Mae inhibits Pointed-P2 transcriptional activity by blocking its MAPK docking site. *EMBO J.* **25**, 70–79.

Ready, F. and Hanson, E. (1976). Developmen of the Drosophila Retina , a Neurocrystalline.

Dev. Biol. **240**,.

Rebay, I. and Rubin, G. M. (1995). Yan functions as a general inhibitor of differentiation and is negatively regulated by activation of the Ras1/MAPK pathway. *Cell* **81**, 857–866.

Rebay, I., Chen, F., Hsiao, F., Kolodziej, P. a, Kuang, B. H., Laverty, T., Suh, C., Voas, M., Williams, A. and Rubin, G. M. (2000). A Genetic Screen for Novel Components of the Ras/Mitogen-Activated Protein Kinase Signaling Pathway That Interact With the. *Genetics* **154**, 695–712.

Reinke, R. and Zipursky, S. L. (1988). Cell-cell interaction in the drosophila retina: The bride of sevenless gene is required in photoreceptor cell R8 for R7 cell development. *Cell* **55**, 321–330.

Ren, X., Sun, J., Housden, B. E., Hu, Y., Roesel, C., Lin, S., Liu, L. P., Yang, Z., Mao, D., Sun, L., et al. (2013). Optimized gene editing technology for Drosophila melanogaster using germ line-specific Cas9. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 19012–19017.

Rohrbaugh, M., Ramos, E., Nguyen, D., Price, M., Wen, Y. and Lai, Z. C. (2002). Notch activation of yan expression is antagonized by RTK/pointed signaling in the Drosophila eye. *Curr. Biol.* **12**, 576–581.

Rubin, G. M., Hong, L., Brokstein, P., Evans-Holm, M., Frise, E., Stapleton, M. and Harvey, D. A. (2000). A Drosophila Complementary DNA Resource. *Science (80-.).* **287**, 2222 LP – 2224.

Scholz, H., Deatrick, J., Klaes, A. and Klambt, C. (1993). Genetic dissection of pointed, a Drosophila gene encoding two ETS-related proteins. *Genetics* **135**, 455–468.

Shwartz, A., Yogev, S., Schejter, E. D. and Shilo, B. (2013). Sequential activation of ETS proteins provides a sustained transcriptional response to EGFR signaling. *Development* **140**,

2746–2754.

Singh, A. and Hespanha, J. P. (2009). Optimal feedback strength for noise suppression in autoregulatory gene networks. *Biophys. J.* **96**, 4013–4023.

Stark, W. S., Walker, J. A. and Harris, W. A. (1976). Genetic dissection of the photoreceptor system in the compound eye of Drosophila melanogaster. *J. Physiol.* **256**, 415–439.

Tomlinson, A. and Ready, D. F. (1987). Neuronal differentiation in the Drosophila ommatidium. *Dev. Biol.* **120**, 366–376.

Tomlinson, A., Bowtell, D. D. L., Hafen, E. and Rubin, G. M. (1987). Localization of the sevenless protein, a putative receptor for positional information, in the eye imaginal disc of Drosophila. *Cell* **51**, 143–150.

Tomlinson, A., Mavromatakis, Y. E. and Struhl, G. (2011). Three distinct roles for Notch in Drosophila R7 photoreceptor specification. *PLoS Biol.* **9**,.

Tomlinson, A., Mavromatakis, Y. E. and Arias, R. (2019). The role of Sevenless in Drosophila R7 photoreceptor specification. *Dev. Biol.* **454**, 181–189.

Tootle, T. L., Lee, P. S. and Rebay, I. (2003). CRM1-mediated nuclear export and regulated activity of the receptor tyrosine kinase antagonist YAN require specific interactions with MAE. *Development* **130**, 845–857.

Voas, M. G. and Rebay, I. (2004). Signal Integration during Development: Insights from the Drosophila Eye. *Dev. Dyn.* **229**, 162–175.

Voliotis, M. and Bowsher, C. G. (2012). The magnitude and colour of noise in genetic negative feedback systems. *Nucleic Acids Res.* **40**, 7084–7095.

Wasylyk, C., Bradford, a P., Gutierrez-Hartmann, a and Wasylyk, B. (1997). Conserved mechanisms of Ras regulation of evolutionary related transcription factors, Ets1 and Pointed

P2. *Oncogene* **14**, 899–913.

Watson, D. K., Mcwilliams, M. J., Lapis, P., Lautenberger, J. A., Schweinfest, C. W. and Papas, T. S. (1988). Mammalian ets-1 and ets-2 genes encode highly conserved proteins. **85**, 7862–7866.

Webber, J. L., Zhang, J., Massey, A., Sanchez-Luege, N. and Rebay, I. (2018). Collaborative repressive action of the antagonistic ETS transcription factors Pointed and Yan fine-tunes gene expression to confer robustness in Drosophila. *Dev.* **145**,.

Weber, U., Pataki, C., Mihaly, J. and Mlodzik, M. (2008). Combinatorial signaling by the Frizzled/PCP and Egfr pathways during planar cell polarity establishment in the Drosophila eye. *Dev. Biol.* **316**, 110–123.

Wolff, T. and Ready, D. F. (1991). The beginning of pattern formation in the Drosophila compound eye: the morphogenetic furrow and the second mitotic wave. *Development* **113**, 841–50.

Wolff, T. and Ready, D. F. (1993). Pattern formation in the Drosophila retina. *Cold Spring Harb. Lab. Press* **2**, 1277–1325.

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

Xu, C., Kauffmann, R. C., Zhang, J., Kladny, S. and Carthew, R. W. (2000). Overlapping Activators and Repressors Delimit Transcriptional Response to Receptor Tyrosine Kinase Signals in the Drosophila Eye. *Cell* **103**, 87–97.

Yang, L. and Baker, N. E. (2001). Role of the EGFR/Ras/Raf pathway in specification of photoreceptor cells in the Drosophila retina. *Development* **128**, 1183–1191.

Yang, L. and Baker, N. E. (2003). Cell cycle withdrawal, progression, and cell survival

regulation by EGFR and its effectors in the differentiating Drosophila eye. *Dev. Cell* **4**, 359–369.

Yang, X., Coulombe-Huntington, J., Kang, S., Sheynkman, G. M., Hao, T., Richardson, A., Sun, S., Yang, F., Shen, Y. A., Murray, R. R., et al. (2016). Widespread Expansion of Protein Interaction Capabilities by Alternative Splicing. *Cell* **164**, 805–817.

Zhu, S., Barshow, S., Wildonger, J., Jan, L. Y. and Jan, Y.-N. (2011). Ets transcription factor Pointed promotes the generation of intermediate neural progenitors in Drosophila larval brains. *Proc. Natl. Acad. Sci.* **108**, 20615 LP – 20620.

Chapter 3 Thermodynamic Modeling and Quantitative Measurements Support 3 Layers of Cooperativity in Cis-regulatory Logic

Chudong Wu, Zhiyue Lu, Aaron Dinner and Ilaria Rebay

CW, ZL, AD, and IR conceptualized the idea. ZL and CW worked out the equations and code together. CW carried out all the tests and made the figures. CW and IR wrote the draft.

3.1 Introduction

3.1.1 Cis-regulatory modules (CRMs) organize activating and repressing transcriptional inputs

For reliable recruitment of diverse cell fates during development, gene expression must be regulated in a spatiotemporally precise manner that is robust to environmental or genetic variation. Robust and precise gene expression is achieved through the combined activating and repressing inputs from TFs organized by gene regulatory elements; I will refer to these as cis-regulatory modules (CRMs). Each CRM is defined by a unique cluster of TF binding sites, the sequence of which determines the specific cis-regulatory syntax (namely, the number, affinity, position, spacing and orientation of the binding sites). Despite the ease with which CRMs can be identified, the intrinsic logic of the cis-regulatory syntax and how it organizes the interactions between TFs and DNA, and between TFs, is poorly understood.

The direct interaction between TFs and DNA is through the DNA binding domain of the TFs, and thus TFs are grouped into different families based on type of DNA binding domain (Lambert et al., 2018). The DNA binding domains in TFs from the same family all share a conserved

structural fold and can recognize similar, if not identical, DNA sequence *in vitro*. Despite this shared DNA binding capability, protein paralogs from the same family nevertheless execute distinct functions *in vivo* (Weirauch et al., 2014). For example, swapping the DNA binding domains between two closely related homeobox TFs led to severe homeotic transformation (Mann et al., 2009). Thus, studies have focused on resolving the paradox of how TFs with similar DNA binding capability exert different biological functions.

Clusters of TF binding sites with differential binding affinities combined with homo- and heterotypic protein interactions contribute to resolving the specificity paradox. In eukaryotes, the length of individual TF binding sites is relatively short with an average of 10 bp (Stewart et al., 2012). Thus a cis-regulatory module ranging from 100-1000 bp always contains clusters of TF binding sites (Gupta and Liu, 2005). Besides the optimal strong affinity binding sites, high throughput experiments and computational modeling have identified weak affinity sites (Le et al., 2018; Ruan et al., 2017; Stormo et al., 2015), which are suggested to be essential to modulate TF specificity (Kribelbauer et al., 2019). For example, an inverse correlation between binding site affinity and specificity was revealed for Hox protein binding sequences such that clusters of weak affinity sites were best discriminated, while strong-affinity sites could be bound by different Hox factors (Crocker et al., 2015). Using experimental manipulations, the authors found that loss of individual weak-affinity sites resulted in reduced and more variable expression under conditions of genetic and environmental variations, while mutations converting the weak-affinity sites to strong-affinity sites led to ectopic expression (Crocker et al., 2015). Thus clustered multiple weak-affinity sites permits the cooperative and additive interactions needed for robust expression without compromising specificity.

As another example, multiple weak-affinity Cubitus interruptus (Ci) sites enabled gene activation in response to low Hedgehog (Hh) signaling whereas engineered strong-affinity sites caused Ci to behave as a repressor, such that gene activation became restricted to regions of maximal Hh signaling (Ramos and Barolo, 2013). Hh signaling forms a gradient in the *Drosophila* wing disc; in the presence of high Hh signaling, its transcriptional effector Ci functions as an activator, while in regions of low signaling, Ci is cleaved to be a transcriptional repressor (Méthot and Basler, 1999). A repressor cooperativity model in which Ci protein cooperatively interacts when it acts as a repressor but not when it is an activator, was developed to explain the transition from repression in the region of low signaling to activation at the region with high signaling (Parker et al., 2011; Ramos and Barolo, 2013). With cooperativity, Ci repressor may still gain binding preference over Ci activator even when it is at the same or lower concentration than the activator. Activation can only happen when the concentration of Ci activator reaches the threshold. At the boundary of threshold, although there may be substantial amounts of Ci repressor, exploiting weak affinity sites instead of high affinity sites can attenuate the effects of cooperative repression and permit gene activation (Parker et al., 2011).

Besides homotypic protein cooperation, heterotypic protein cooperation or competition also contributes to expression specificity. The enhancers, and the TFs that bind them to control anterior-posterior, dorsal-ventral patterning during early embryonic development are well characterized and have provided good systems to study how CRM syntax influences expression regulation (Papatsenko et al., 2002; Papatsenko et al., 2009). The work has defined broad general requirements for the cis-regulatory syntax to organize a) the interactions between strong/weak binding sites and TFs, b) the cooperativity between TFs, and c) the competition between TFs. How these principles are applied and interpreted at enhancers that regulate specific signaling-

induced cell fate transitions later in development needs further exploration. The example my thesis work has focused on is an enhancer within the *eve* locus that relies on cooperative activating inputs from Pnt, Mad, Tin, Twist and dTCF to induce a specific Eve+ cardiac cell type (Halfon et al., 2000).

3.1.2 The muscle heart enhancer (MHE) organizes the interactions of Yan, Pnt and other TFs to regulate even-skipped (eve) expression

The Drosophila ETS repressor Yan and activator Pnt, which recognize and bind competitively to similar DNA sequences but exhibit different binding site preferences, are great example to investigate the syntax rules. The core motif recognized by ETS factors is 5′-GGAA/T-3′. According to the different flanking sequences around the core motif, the ETS binding sites can be separated into strong- and weak-affinity sites in accordance with the binding preferences identified for fly, human, and mouse ETS family TFs (Boisclair Lachance et al., 2018; Charlton-Perkins et al., 2011; Nitta et al., 2015; Webber et al., 2013a; Wei et al., 2010). As ETS factors, Yan and Pnt recognize the same core motif and thus directly compete for the binding sites. However, their preferences for the flanking sequences around the core motif are different (Fig 3.1) based on high throughput SELEX (Nitta et al., 2015). These distinct preferences explain in vitro gel shift assay results that showed that Yan and Pnt could bind to the same DNA probes but with different binding strengths (Halfon et al., 2000; Xu et al., 2000). This suggests that the respective Yan and Pnt preference for strong vs. weak sites may be different, offering a mechanism for cis-regulatory syntax to bias Yan-Pnt competition to achieve correct transcriptional output.

In vivo, RTK signaling modulates Yan and Pnt competition by directly regulating Yan and Pnt activity. When pathway activity is high, Pnt activity is high and Yan activity is low, whereas when RTK activity is low, Yan activity is low and Pnt activity is high (O'Neill et al., 1994; Rebay and Rubin, 1995). Thus, in addition to the different intrinsic DNA binding preferences influencing their activity at specific CRMs, the competition between Yan and Pnt is also sensitive to the level of RTK activation.

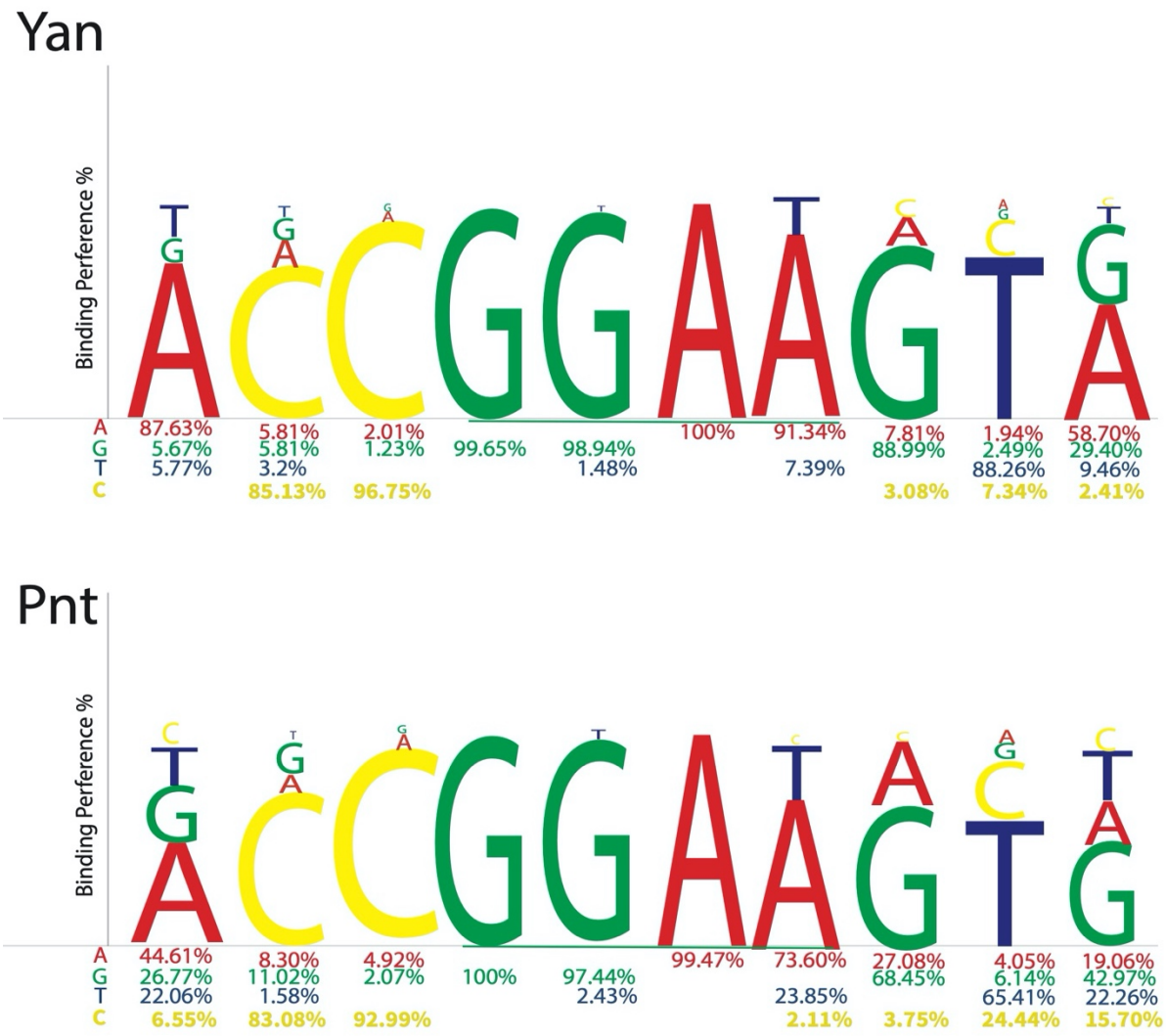


Fig 3.1 Different sequence preferences for Yan and Pnt

(Fig 3.1 continued) Based on high throughput SELEX (Nitta et al., 2015), the image indicates the core motif of the ETS sites (green line underlined) and the flanking sequences around the core motif. The sizes of AGCT represent their relative preferences at the position. The percentage preferences of AGCT at each position are labeled. Only the ones > 1% are labeled. The core sequence is the same for Yan and Pnt but their flanking sequence preferences are different.

In addition to different sequences biasing Yan-Pnt competition at each individual ETS site, the cis-regulatory syntax must also accommodate Yan self-association, which will further influence Yan and Pnt occupancy patterns. Unlike Pnt, Yan and its human counterpart, TEL1, self-associate via their sterile α motifs (SAMs) and this SAM-SAM interaction is critical for transcriptional repression (Qiao et al., 2004; Zhang et al., 2010). Although the exact mechanism of how this interaction contributes to repressive function is uncertain, it has been suggested that self-association could mediate cooperative binding to the ETS sites in CRMs (Green et al., 2010; Hope et al., 2017). Using gel shift assays, SAM-SAM interactions were shown to mediate cooperative binding of TEL1 at paired ETS sites in vitro (Green et al., 2010). Theoretical modeling demonstrated a similar mechanism whereby the self-association ability promotes the preferential cooperative recruitment of Yan to tandem ETS sites (Hope et al., 2017). Therefore strong affinity ETS sites in close proximity may bias Yan-Pnt competition toward Yan, leading to repression of target gene transcription, while multiple weaker affinity sites spaced at a distance may bias toward Pnt and transcriptional activation.

The *eve* muscle heart enhancer (MHE), which contains a cluster of strong and weak ETS sites bound by Yan and Pnt, as well as binding sites for other transcriptional activators, is well characterized. The MHE is both necessary and sufficient to induce *even-skipped (eve)* expression during cardiac muscle precursor specification at stage 11 of embryogenesis (Halfon et al., 2000; Knirr and Frasch, 2001). Two other enhancers, D1 and D2, contribute to robustness Webber et al., 2013b), but their cis-regulatory syntax has not yet been characterized.

Focusing on the MHE, besides the eight ETS sites, there are 13 Mad, Twist, Tin and dTCF sites, which organize the interactions of the TFs (Carmena et al., 1998; Carmena et al., 2002; Halfon et al., 2000). Among the eight ETS sites, a pair of strong affinity sites strongly bias toward Yan, such that mutating them significantly elevated the MHE reporter expression, as a result of compromised repression (Boisclair Lachance et al., 2018). Mutation of this strong affinity pair in the endogenous *eve* MHE reduced the recruitment of Yan not only at MHE but also at D1 and D2, suggesting loss of long-range cooperative Yan interactions across the *eve* locus (Boisclair Lachance et al., 2018).

In additional to the cooperative Yan-Yan and antagonistic Yan-Pnt interactions, there is also evidence for heterotypic cooperativity among other activating TFs in MHE regulation. For example, mutating an ETS site along with a dTCF site led to a synergistic decrease in MHE reporter expression compared to individual mutation of the two sites (Halfon et al., 2000), suggesting cooperativity between the two. While experimental manipulation of a CRM coupled with quantitative measurement of expression changes is well-suited to defining the critical TF inputs, uncovering the "rules" that integrate the different combinations of TF-DNA and TF-TF interactions to produce the final expression output is a much more challenging endeavor, and better suited to a mathematical modeling approach.

Therefore to investigate how the CRM syntax in the *eve* MHE organizes the repressive and activating inputs from Yan and Pnt, and their interactions with other transcription factors, we developed a thermodynamic model that relates transcription factor occupancy to gene expression output. The modeling effort was inspired by predictions of multiple scales of Yan-Yan cooperativity that emerged from our earlier *in vivo* quantification of the changes in MHE driven

expression that result from mutation of individual and combinations of ETS sites (Boisclair Lachance et al., 2018). By comparing the simulation outputs with the experimental measurements, we tested multiple hypotheses about how the MHE cis-regulatory syntax organizes local and long-range transcription factor interactions to achieve the final transcriptional output. We uncovered that the respective contributions of strong- and weak-affinity sites to repression and activation. Both are mediated by long-range interactions directly or indirectly. On top of these ETS site based interactions, our model also suggested that the heterotypic interactions between Pnt and the other transcription activators permits more precise gene activation, such that activation only occurs within a narrow range of Yan and Pnt concentrations. We propose that the cis-regulatory syntax of the MHE encodes two layers of regulation. The first layer encompasses the short, medium and long-range interactions by which the ETS sites organize the activating and repressive inputs from Pnt and Yan, and then the second layer consists of heterotypic interactions with the other transcription effectors that modulates output to confer precise gene activation.

3.2 Results

3.2.1 Formulation of the model

We set out to develop a thermodynamic model of transcriptional output dictated by the recruitment of Yan and Pnt to MHE. There were two major requirements of the model: firstly, it needed to consider all possible occupancy patterns of Yan and Pnt at chemical equilibrium on the MHE; and secondly, it needed to calculate transcriptional output based on occupancy state. The first step was developed from a previous model of Yan binding to an artificial enhancer composed of a lattice of ETS sites (Hope et al., 2017). We revised the model to represent the ETS syntax of the 312bp *eve* MHE. Thus the eight identified ETS binding sites were defined as

specific sites and all other sequence was defined as non-specific binding sites for an ETS factor. Based on the footprint of the ETS DNA binding domain of TEL, we defined each binding site, specific or non-specific, as ~10bp (De et al., 2014; Fig 3.2A, B). As the spacing between ETS8 and ETS7 is 153bp, or ~15 non-specific binding sites, to simplify the calculation, we considered the MHE as two separate elements: the sequences from ETS1 to ETS7 defined a consecutive 14 site element with 7 specific and 7 non-specific sites while ETS8 defined a second simpler element with one specific site ETS8 isolated from the first element.

To calculate the fractional occupancies of Yan and Pnt, the model considered the specific and non-specific protein-DNA interactions for both Yan and Pnt, and SAM-mediated protein-protein interactions for Yan (Fig 3.2C; 3.4 Materials and methods). α_{yan} and α_{pnt} represented the respective free energy value of Yan and Pnt specific binding to ETS sites, β_{yan} and β_{pnt} represented the respective free energy value of Yan and Pnt binding to non-specific sites, and γ_{yan} represented the free energy of a Yan-Yan interaction via the SAM domains when two Yan molecules were bound on adjacent sites. As Yan self-associated, each adjacent Yan molecule contributed an additional increment of γ_{yan}. As a function of the free energy change, and Yan and Pnt concentrations, the fractional occupancy of a microstate, which is a pattern of Yan and Pnt occupancy across MHE, could be calculated.

The second step was to translate fractional occupancy to gene expression. To do this we incorporated the concept of a basal transcription complex (BTC) (Parker et al., 2011), in which BTC represents a factor whose occupancy at the gene is necessary and rate-limiting for transcription but is not related to any specific molecule or complex (Fig 3.2C; 3.4 Materials and methods). In this way gene expression output was considered proportional to the probability of

BTC binding, which in turn was determined by the Yan-Pointed occupancy of each microstate. As Yan is a transcriptional repressor, the interaction between Yan and BTC reduced the probability of BTC recruitment. Therefore, each Yan molecule bound at MHE contributed a positive free energy $\omega_{\mathbf{yan}}$ to BTC binding. In contrast, each bound molecule of the activator Pnt contributed a negative free energy $\omega_{\mathbf{pnt}}$ to enhance BTC binding. The occupancy of BTC as a function of the free energy represented proportional gene expression. Using this framework, we compared gene expression driven by wildtype MHE versus MHE carrying specific ETS site mutations (further details in Materials and methods).

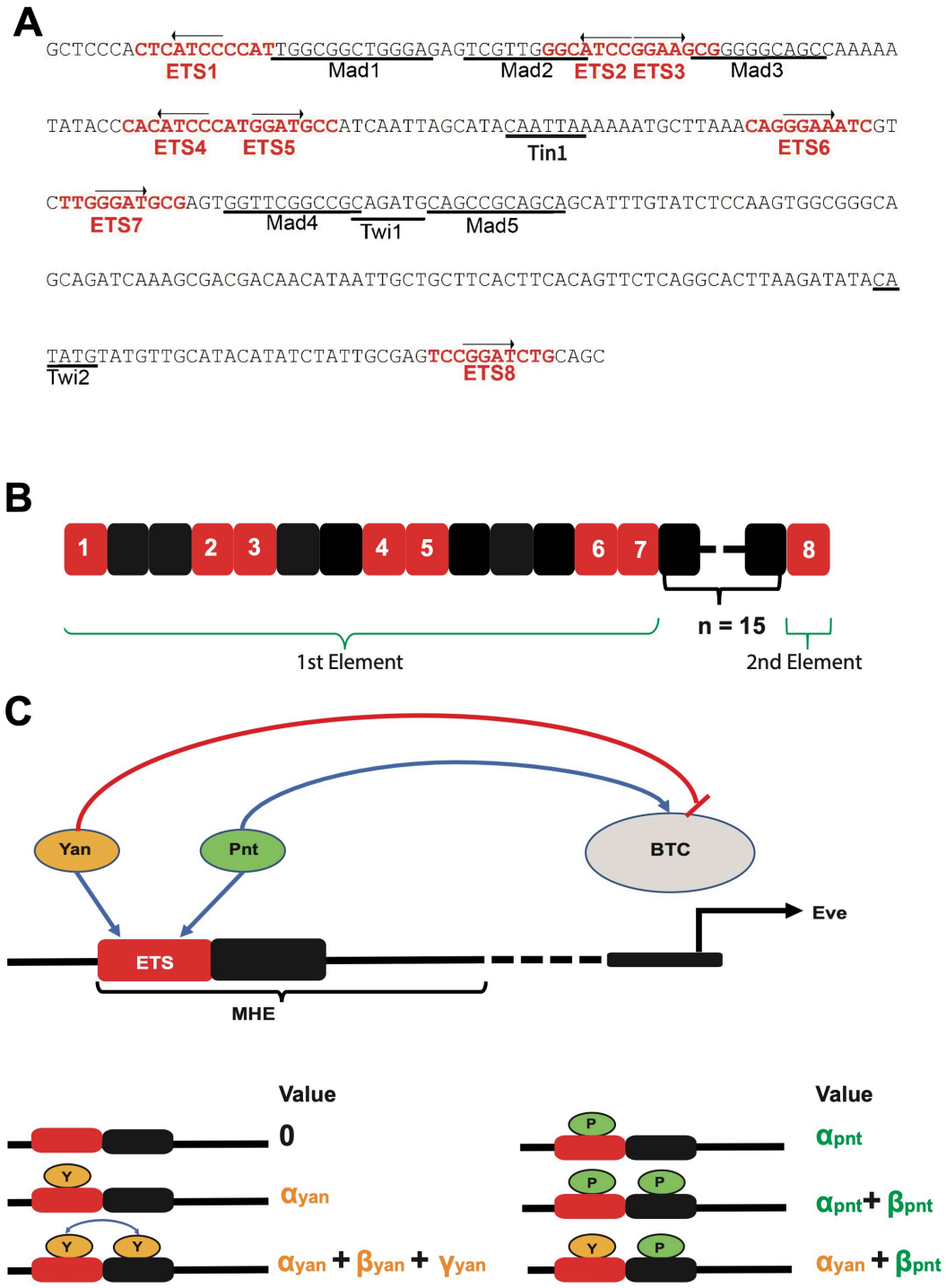
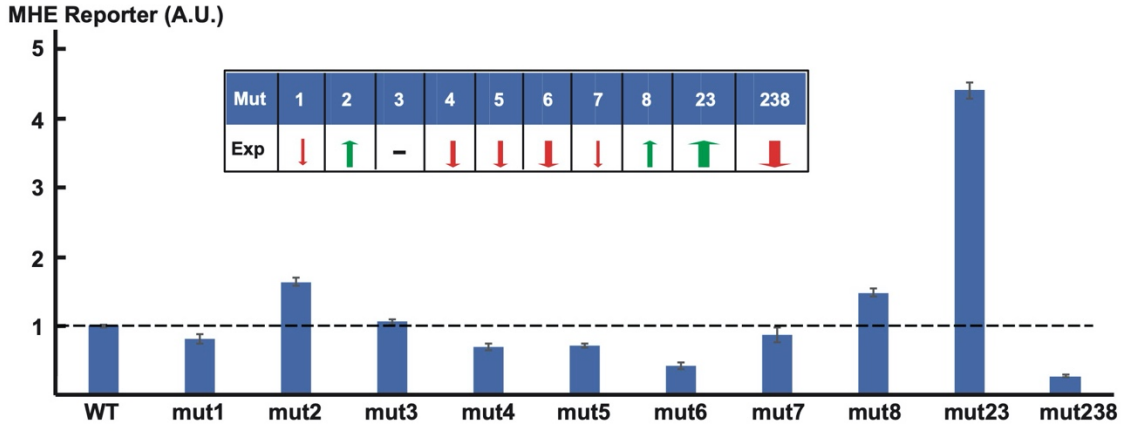


Fig 3.2 Lattice thermodynamic model for Yan and Pnt binding at MHE at equilibrium

A) MHE sequence with putative ETS sites in red. Black arrows indicate site orientation. Previously characterized binding sites for the activators Mad, Twi, Tin and dTCF sites (Halfon et al., 2000) are underlined. B) Conceptualized lattice representing MHE. Each square represents an

(Fig 3.2 continued) ETS specific/nonspecific binding site with a footprint of ~10bp. Red ones represent the specific sites and the numbers correspond to the ones in A. Black ones represent the ETS non-specific DNA. The two elements in the linear model were indicated. C) Yan and Pnt molecules compete for individual binding sites, and BTC binds to a separate position to initiate *eve* transcription. The binding of BTC is stabilized by Pnt and destabilized by Yan, and the fractional occupancy of BTC represents relative transcription. The bottom panel demonstrates microstates of Yan/Pnt binding at the ETS specific/non-specific sites and the corresponding free energy for the microstate to the right.



S 3.1 Fig A summary of the changes of MHE reporter expression driven by ETS site mutations

The expression of MHE wildtype and ETS mutant reporters from previous published experiments are summarized (Boisclair Lachance et al., 2018). The table on the top indicates either increased (green arrows) or decreased (red arrows) transcription of the corresponding mutants comparing to the MHE wildtype reporter expression. The width of the arrow indicates the amount of decrease/increase of transcription. MHE^{mut23} showed the largest increase of transcription, while MHE^{mut2} and MHE^{mu8} showed smaller increase; MHE^{mut238} and MHE^{mut6} showed relatively large decrease of transcription while MHE^{mut1}, MHE^{mut4}, MHE^{mut5}, and MHE^{mut7} showed smaller decrease.

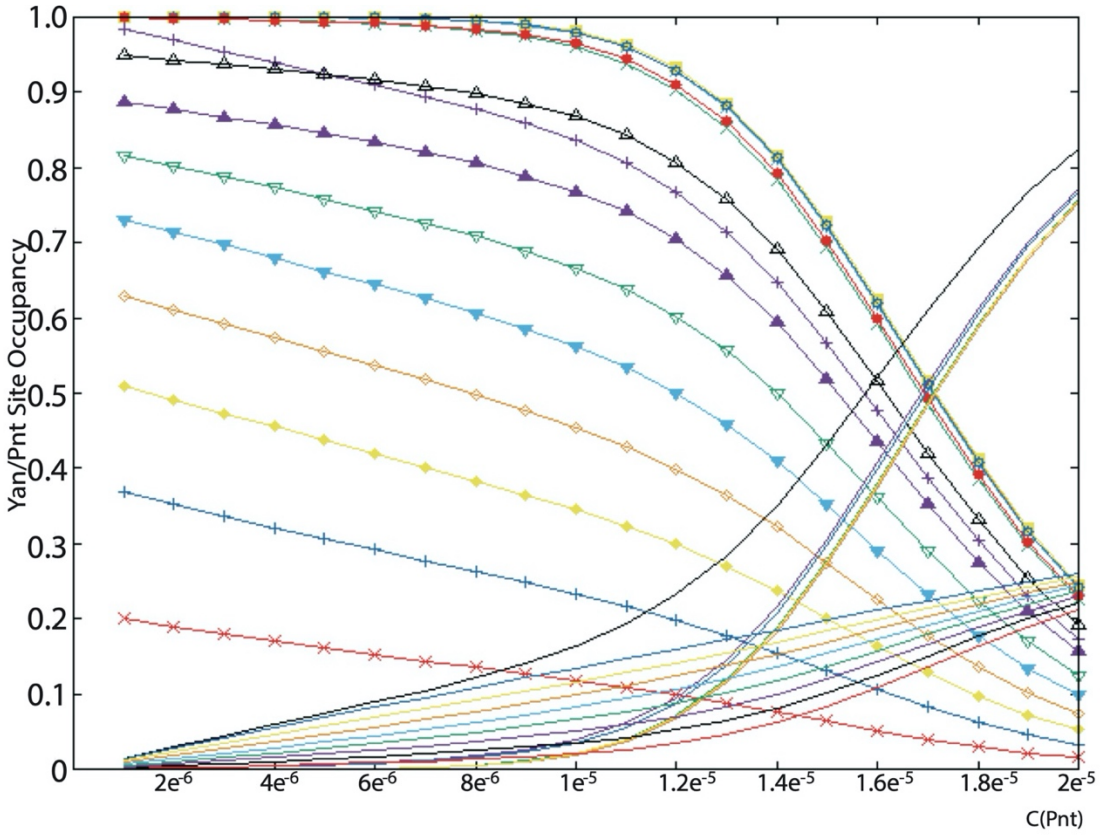
3.2.2 Parameterization and implementation of the model

To calculate the fractional occupancy of both Yan and Pnt, we upgraded the binary notation scheme for ordering microstates from our earlier Yan occupancy study (Hope et al., 2017). Originally when only Yan was considered, each microstate was assigned an integer index k, where k is a value from 0 to 2^n-1. n represents the total number of binding sites including the

ETS sites and the non-specific ones. This binary representation reflected the molecular configuration of the microstates and also increased the computation efficiency. Thus we continued the utilization of the binary system, but to include Pnt we doubled the number of microstates to 2^{2n}, from 0-2^{2n-1}. The first half were microstates with only Yan and the other half added Pnt. To consider the molecular configuration, the $n+1^{th}$ to $2n^{th}$ binary digits also represented the n binding sites, just as the 1^{th} to n^{th} binary digits. The $n+1^{th}$ to $2n^{th}$ binary digits indicate the binding of Pnt whereas as the 1^{th} to n^{th} binary digits indicate the binding of Yan. Because Yan and Pnt cannot occupy the same site at the same time, the microstates with the binary digits showing co-occupation of Yan and Pnt were assigned an infinite positive energy, such that the fraction of those microstates were zero.

To begin our calculation, we used the literature to parameterize the model with reasonable values. Although there are no biochemical measurements of Yan's and Pnt's specific or non-specific binding to DNA, considering the conservation of the ETS domain, we firstly assigned the α_{yan}, β_{yan}, α_{pnt}, and β_{pnt} values according to the affinities measured for the human homolog of Yan, TEL1 (De et al., 2014). Thus α_{yan} and β_{yan} were set to -9.955 and -5.797 kcal/mol, respectively. α_{pnt} and β_{pnt} were adjusted to be specific folds of α_{yan} and β_{yan} respectively. γ_{yan} was set to -7.000 kcal/mol according to previous direct measurement of Yan's SAM-SAM affinity (Qiao et al., 2004). Regarding the nuclear concentration of Yan and Pnt, as there were no published measurements, we set the concentration range to be 0.1 to 1000nM based on measurements of other TFs (Biggin, 2011). A table to summarize the parameter values is in 3.4 Materials and methods.

Interestingly, our initial calculations using the simplest model in which α_{yan} equals to α_{pnt}, and β_{yan} equals to β_{pnt}, showed that when the concentration of Pnt was at the physiological range (0.1-10^4nM) and the concentration of Yan was low (0.5nM), the fractional occupancy of Yan at all sites on MHE was always much higher than that of Pnt. This meant that expression would always be repressed (S3.2 Fig), an output that was obviously physiologically incorrect. This suggested that our calculation for γ_{yan} was over-emphasizing the contributions of Yan-Yan interactions. Because the ETS non-specific DNA contains specific binding sites for other transcription factors that are highly expressed in Eve-positive cells (Boisclair Lachance et al., 2018; Halfon et al., 2000), we reasoned that these DNA sites were likely inaccessible to Yan and Pnt. Therefore, we assigned β_{yan} and β_{pnt} to be equally so small that the fraction of microstates with Yan or Pnt bound at the nonspecific DNA would be minimal. This permitted gene activation at the high concentrations of Pnt and low concentrations of Yan within the physiological range (S3.3 Fig B).



S 3.2 Fig. Yan and Pnt fractional occupancy map for the simplest linear model

Yan and Pnt's fractional occupancy at ETS specific/nonspecific sites at a broad range of Pnt concentrations from 5e-10 to 2e-5 M and Yan concentration of 5e-10. All plain lines represent the fractional occupancy of Pnt and lines with indicators represent the occupancy of Yan. Different colors represent ETS specific/nonspecific binding sites in Fig 3.2B. As Pnt concentration increased, the fractional occupancy of Pnt at each site also increased accordingly whereas that of Yan decreased. Only when Pnt concentration reached ~8e-6 and above, there were sites at which the occupancy of Pnt outcompetes that of Yan.

3.2.3 Long-range cooperativity of strong-affinity ETS sites confers robust MHE transcriptional repression

3.2.3.1 Sufficiency of the long-range cooperation for MHE repression

We then used the model to explore how the cis-regulatory syntax of MHE organizes the repressive and activating transcriptional inputs from Yan and Pnt. We modeled different

combinations of ETS site mutations and compared the transcription output predictions from our simulations to the experimental measurements (Boisclair Lachance et al., 2018). To start we tested the importance of binding affinity differences between the ETS sites by considering an MHE with eight identical ETS sites as a linear array. Suggesting the importance of affinity differences, the homogeneous model failed to predict the experimental results for most ETS site mutants (Fig 3.3A, S3.3A Fig). For example, zooming in to the concentration range of low Yan and high Pnt that permits *eve* gene activation (S3.3B Fig), the model predicted that individual mutation of MHE ETS sites 2-7, referred to as MHE^{mut2} – MHE^{mut7}, would increase transcriptional output relative to MHE^{WT} while individual mutation of ETS1 or ETS8 would reduce output. Comparison to prior *in vivo* reporter measurements showed agreement only with MHE^{mut2}. Further discordancy between model prediction and experimental measurement was noted for multi-site mutations. Most striking was with $MHE^{mut2,3}$, where the model predicted minimal change but experimental measurements had found a strong increase in expression. In considering why the model failed, we noted than any single ETS site mutation that disrupted one of the three ETS pairs (ETS2-ETS3, ETS4-ETS5, and ETS6-ETS7) equally increased transcriptional output. However *in* vivo, only disruption of the ETS2-ETS3 pair derepresses. This suggested that by considering all ETS sites identical in a simple linear array, the model was overestimating the contribution of cooperative repression at ETS4-ETS5 and ETS6-ETS7 and underestimating the contribution at ETS2-ETS3. In addition, mutation of the non-paired sites, MHE^{mut1} and MHE^{mut8}, produced similar reductions in transcriptional output *in silico*, but opposing reduction or increase in outputs *in vivo*.

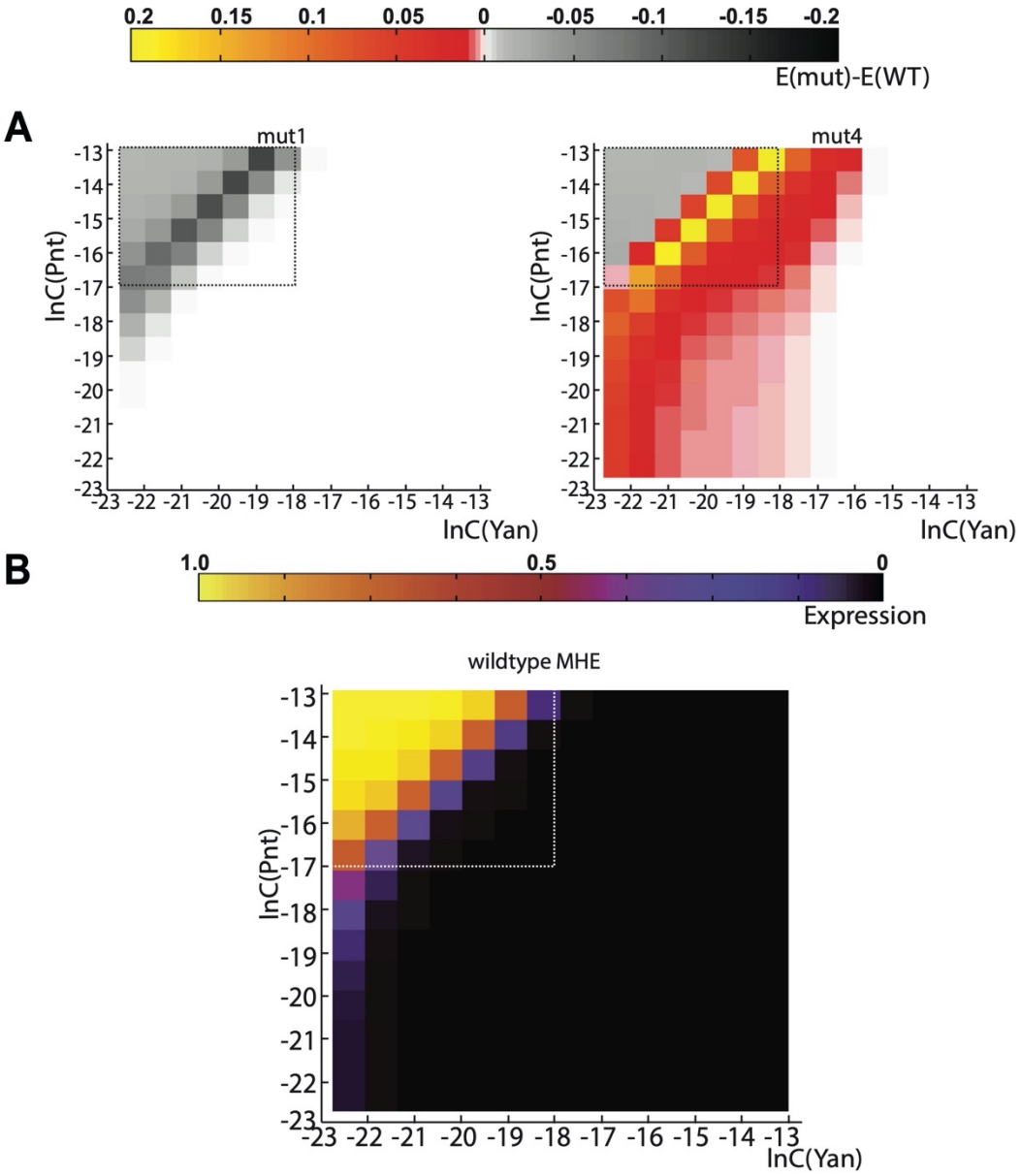
We attempted to correct these problems by categorizing sites as strong-affinity (ETS2, 3 and 8) and weak-affinity (ETS1, 4, 5, 6, 7) in the model. To start we assigned the DNA binding affinity

for the weak sites (α_w) to be 1/10 that of the strong sites (α_S) and then repeated the calculations with the same set of ETS mutations. The results showed that only the pair of strong sites (ETS2-ETS3) now contributed to repression, bringing simulation and experiment in closer agreement. Thus either MHE^{mut2} or MHE^{mut3} increased transcription while mutation of any individual weak site participating in the ETS4-ETS5 or ETS6-ETS7 pairs had no effect on transcriptional output (Fig 3.3B). However ETS8, now a strong affinity site, still contributed to activation, not to repression as predicted by in vivo measurement. Making α_w to be 1/5 and 1/100 folds of the strong sites resulted in the same trend of transcriptional changes (data not shown).

Up until this point our simulations assumed identical DNA binding affinities between Yan and Pointed with respect to all sites. We therefore asked whether biasing Yan and Pnt binding towards sites with differential affinities might improve the model's ability to discriminate between strong and weak site contributions to gene transcription. Based on prior work suggesting that strong sites contribute more to repression and weak sites more to activation, we assigned Yan stronger affinity toward the strong sites and Pnt stronger affinity toward the weak sites ($\alpha_{\mathbf{yan\text{-}S}} > \alpha_{\mathbf{pnt\text{-}S}} > \alpha_{\mathbf{pnt\text{-}W}} > \alpha_{\mathbf{yan\text{-}W}}$), with the strong-affinity of Yan 5 fold that of Pnt at strong sites, and the weak-affinity of Pnt 5 fold that of Yan at weak sites. Here strong-affinity of Yan ($\alpha_{\mathbf{yan\text{-}S}}$) and weak-affinity of Pnt ($\alpha_{\mathbf{pnt\text{-}W}}$) equaled to the respective strong and weak affinities assigned to binding sites in the Yan and Pnt non-biased model described in the last paragraph. We introduced the two new parameters $\alpha_{\mathbf{pnt\text{-}S}}$ and $\alpha_{\mathbf{pnt\text{-}W}}$ to fit the 10 MHE mutants. The results of our calculations showed that with the introduction of Yan-Pnt bias toward strong versus weak sites, mutation of any individual weak site now produced the experimentally predicted reduction in transcriptional activation, although the strength of deactivations produced by the mutation of

weak sites in simulation were weak comparing to the strong derepression from the strong site mutations (Fig 3.3C).

This final model (Fig 3.3C) left us with three remaining points of disagreement between simulation and experiment (Boisclair Lachance et al., 2018), all centered around mutation of the strong sites ETS2, 3 and 8. First, the model predicted that mutation of ETS3 should increase output, whereas *in vivo* no significant change was measured. Second, the model predicted that mutation of ETS8 should decrease output, whereas *in vivo* this causes a significant increase. Third, although the model correctly predicted the increased expression resulting from the double mutant $MHE^{mut2,3}$, it failed to predict the reduced expression measured for the triple mutant $MHE^{mut2,3,8}$. Because two of these failures involved ETS8, we wondered whether by conceptualizing MHE into two separate linear elements, with ETS8 completely isolated in the second element, we were missing cooperative regulatory interactions critical to repression.



S 3.3 Fig Expression change and wildtype MHE expression across the whole range of Yan and Pnt concentrations

A) The color indicates the expression of mutant reporter (E(mut)) subtracting the expression of wildtype reporter (E(WT)), and thus the color from grey to black indicates reduced transcription whereas the color from red to yellow indicates increased transcription. The expression change across the whole tested physiological concentration of Yan and Pnt for mut1 and mut4 as representatives. The white squares are the zoomed-in crop of high concentrations of Pnt and low concentrations of Yan in Fig 3.3 and the following main figures. B) The expression levels of MHE wildtype reporter across the whole tested concentrations. The color from black to yellow represents the gene transiting from fully repressed state to fully activated state. The white square

(S 3.3 Fig continued) indicates the range of Pnt and Yan concentrations where the gene is likely to be activated, in the zoomed-in pictures.

The discrepancy for MHE^{mut8} between the model predictions and the experimental results motivated us to consider the spatial arrangement of MHE ETS sites from a 3D perspective. The two-element models above took MHE as a linear DNA sequence in which cooperative interactions can only occur between proteins bound to adjacent sites. However in reality DNA/chromatin is not a simple linear structure. The persistence length of DNA, within which the molecule is rigid and unlikely to bend, is 150bp (Manning, 2006). The MHE sequence spanning the first seven ETS sites, ETS1-7, is less than 150 bp and so can be considered as a linear element. However the spacing between ETS7 and ETS8 is 153 bp, meaning that with DNA bending, interactions between one Yan molecule bound to ETS8 and another bound to a site within ETS1-7 element are theoretically possible.

We considered two possibilities for such non-linear interactions. In the first scenario we assumed DNA bending would put ETS8 in close proximity to the strong affinity ETS2,3 pair. Because Yan occupancy dominates at strong affinity sites, this physical configuration will introduce additional γ energy from Yan self-association. For example, in a microstate in which ETS2 and ETS8 are both bound by Yan, in a purely linear model they would be too far apart for SAM-SAM interactions and so γ would be set to zero. However with DNA bending, SAM-SAM interactions become possible, and so γ now contributes to the free energy calculation of this microstate. Under these assumptions, individual, double or triple mutations involving ETS sites 2, 3 and 8 all increased transcription output (Fig 3.4A). For the cooperativity among 2, 3 and 8, considering that from the experimental results ETS2 contributed strongly to repression while individual ETS3 had nonsignificant impact, we expected that ETS8 was biased towards

interacting with ETS2. With the bias, the simulation showed that MHE^{mut2} had more increased expression than that of MHE^{mut3}. To further distinguish the impact of the characterization of strong and weak sites, and the long-range cooperativity among ETS2, 3 and 8, we only kept the cooperativity but considered all ETS sites containing the same binding affinities. The cooperativity was sufficient for ETS2, 3 and 8 contributing to repression and ETS1, 4-5, 6-7 contributing to activation, even without the distinct between strong and weak sites (Fig 3.4B). To consider the strength of activation and repression, the introduction of the nonlinear cooperativity weakened the repression but changed the strength of activation minimally, adjusting the balance between repression and activation (Fig 3.3BC, 3.4AB).

In the second scenario we considered the possibility that ETS8 interacts with weak-affinity sites. We found that if ETS8 was in close proximity to ETS4 and ETS5, the two weak-affinity sites now contributed to transcriptional repression (Fig 3.4C) in contradiction to the experimental results. The same results were obtained when ETS8 interacts with the pair of ETS6 and ETS7. Summarizing from our tests of different spatial interactions between the ETS sites, we conclude that the physical proximity of ETS8 to ETS2 and ETS3 allows these three strong-affinity sites to form a strong repressive complex across the enhancer.

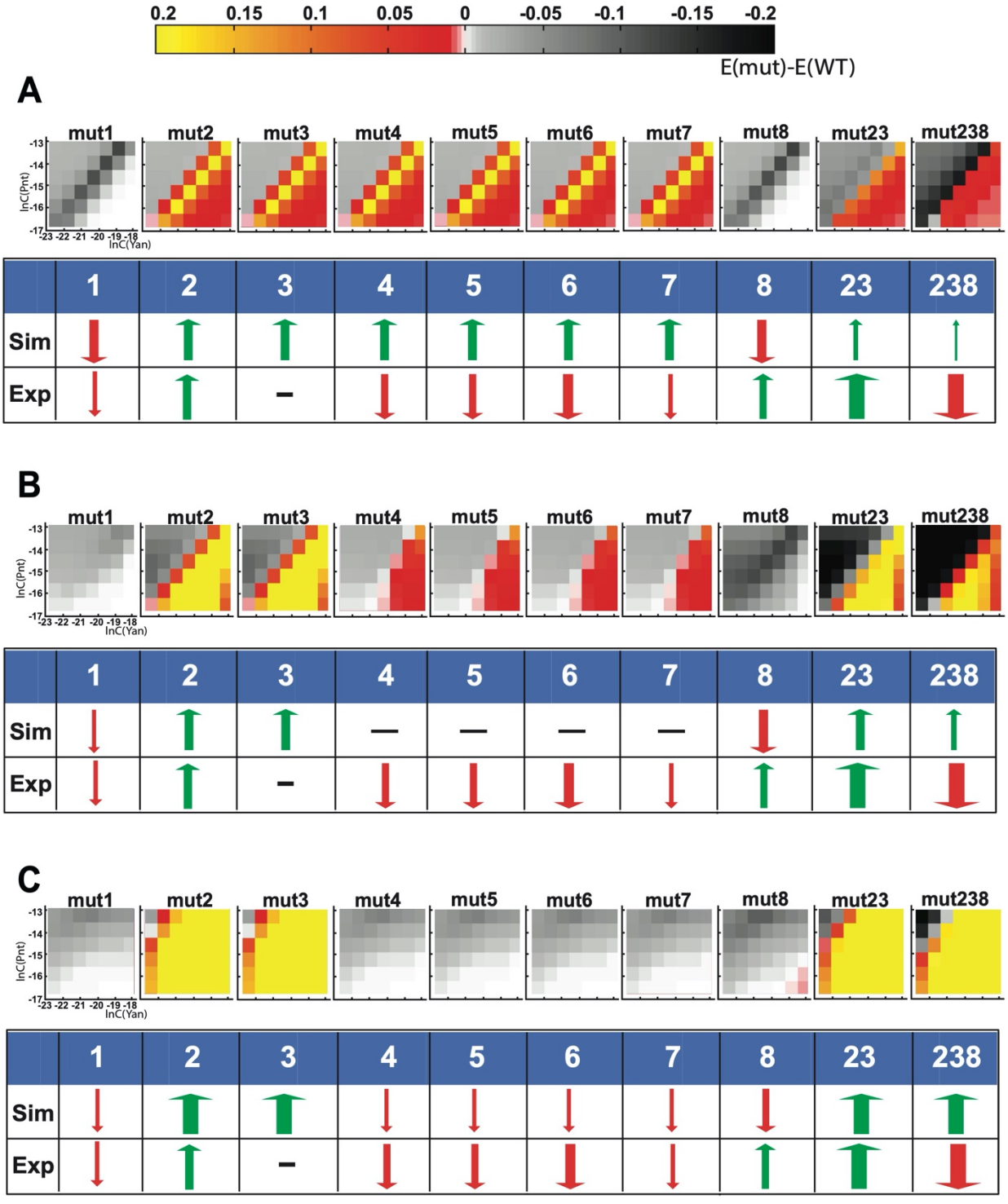


Fig 3.3 Expression for mutant in the background of ETS8 is distant from ETS1-ETS7

(Fig 3.3 continued) All images represent the relative expression change between wildtype and mutant MHE reporters. The color indicates the expression of mutant reporter (E(mut)) subtracting the expression of wildtype reporter (E(WT)), and thus the color from grey to black indicates reduced transcription whereas the color from red to yellow indicates increased transcription. All images are zoomed-in to the high Pnt concentration and low Yan concentration state of the full range of tested concentrations (S3A Fig). All the tables below the images indicate either increased or reduced transcription of the corresponding mutants, and compare the simulation results to experimental ones from previous publication (S1 Fig;(Boisclair Lachance et al., 2018)). All the images are for the linear configuration model in which ETS8 was isolated from the element containing ETS1 to ETS7. A) When all ETS sites were assigned identical Yan/Pnt DNA binding affinities, all paired ETS sites (2 and 3, 4 and 5, 6 and 7) contributed to repression while the isolated siteds (1 and 8) contributed to activation. B) When there was a distinction between strong-affinity sites (2, 3 and 8) and weak-affinity sites (1, 4, 5, 6 and 7), the paired strong sites ETS2, 3 contributed strongly to repression, the isolated strong site ETS8 and weak site ETS1 contributed to activation at different strengths, and the paired weak sites ETS4, 5, 6 and 7 contributed neither activation nor repression. C) When Pnt and Yan preferred weak and strong sites respectively, the paired strong sites ETS2, 3 contributed strongly to repression, the weak sites ETS1, 4, 5, 6 and 7 contributed to activation, and the isolated ETS8 contributed to activation weakly.

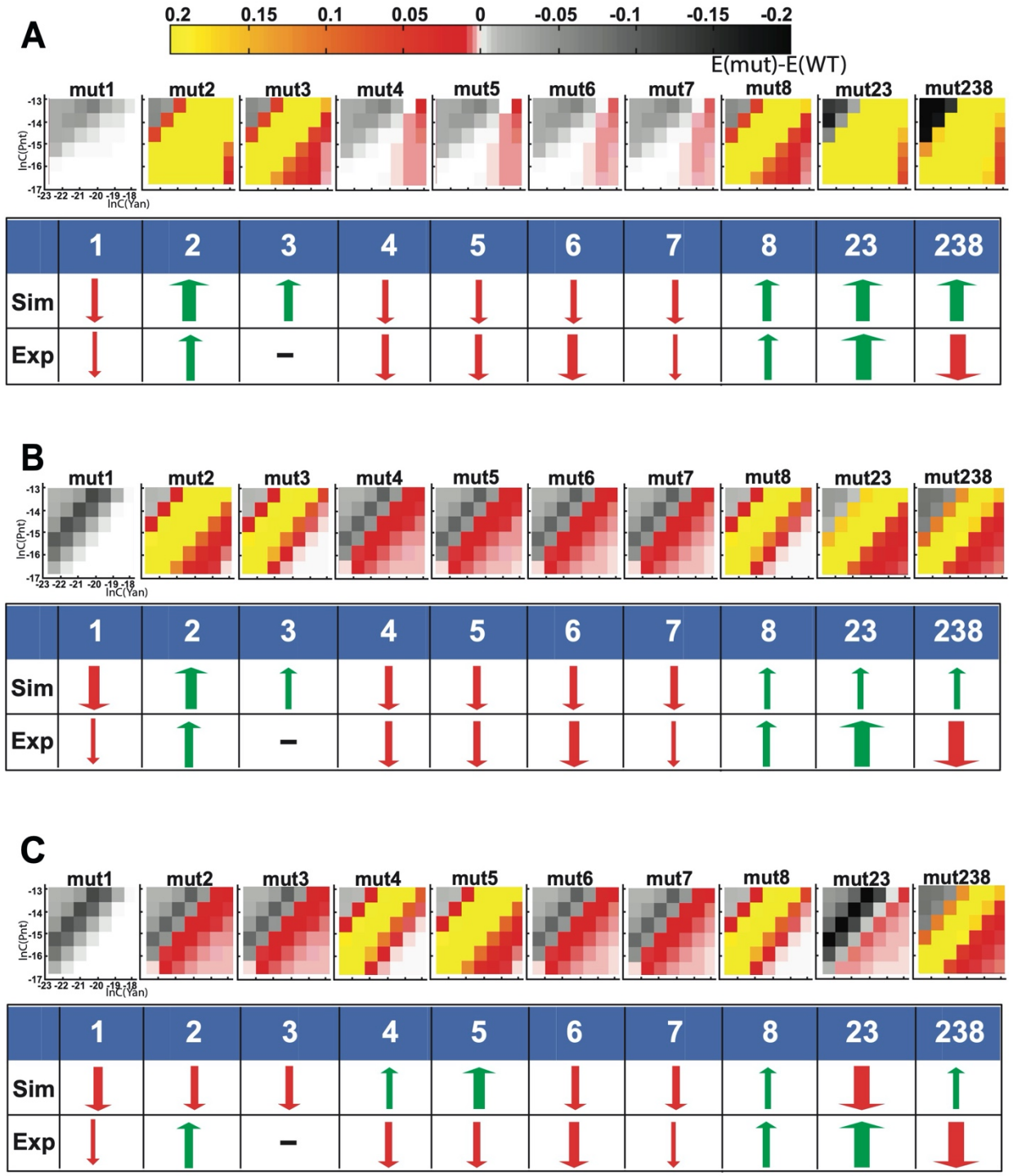


Fig 3.4 Cooperativity between high-affinity ETS sites 2, 3, 8 may produce strong MHE transcriptional repression

All images represent the relative expression change between MHE wildtype and mutant reporters. The representations of the color and the tables are as in Fig 3.3. A and B are images for

(Fig 3.4 continued) the spatial interaction model with ETS8 in close proximity with the ETS23 pair, with bias towards ETS2. C are images for the model with ETS8 in close proximity with the ETS 45. A) When all the three strong sites ETS2, 3 and 8 were in close proximity, all strong sites contributed strongly to repression and all weak sites contributed to activation. B) When there was no distinction between strong and weak sites, with ETS2, 3 and 8 in close proximity, ETS2, 3 and 8 contributed to repression and ETS1, 4, 5, 6 and 7 contributed to activation. C) When there was no distinction between strong and weak sites, and ETS8 was in proximity with ETS4, 5, we found ETS4, 5, 8 contributed to repression while the rest of the sites contributed to activation.

3.2.3.2 The pair of overlapping MHE ETS2, 3 sites interact with ETS8 through the self-association of Yan

Our model emphasized the importance of collaborative interactions among the strong-affinity sites ETS2, 3 and 8, prompting us to investigate further the molecular mechanism of the interactions. Unlike the other ETS site pairs in the MHE, the ETS2 and ETS3 pair is composed of two antiparallel GGAA/T cores immediately juxtaposed (Fig 3.1A). Both in vitro gel shift assays and structural modeling showed that ETS2 and 3 were unlikely to be co-occupied (Boisclair Lachance et al., 2018). If only one Yan molecule at a time can occupy the ETS2, 3 pair, how are the synergistic repressive functions of ETS2 and 3 achieved?

We hypothesized that physical proximity between Yan molecules occupying ETS8 and the ETS23 pair facilitates synergistic repression. To test this, we set a limit in our simulation that forbade simultaneous occupancy of ETS2 and ETS3, i.e. only a single molecule of Yan or Pnt was allowed to bind the ETS23 pair in any given microstate. If the hypothesis was true, we expected that the one molecule limit would not change the repressive function of the strong sites. Indeed, simulations showed that ETS2, 8 and 23 as a pair contributed to the transcriptional repression of MHE expression, even with only one molecule at the ETS2, 3 pair (Fig 3.5A). The

pairing of strong-affinity sites ETS23 may increase the probability of Yan occupation at the pair so that there is always a Yan molecule available to interact with the Yan bound on ETS 8.

However, compared to the model when two Yan molecules were allowed to co-occupy ETS2, 3 (Fig 3.4A), the limit of one Yan molecule attenuated both the repression from ETS23 pair and activation from ETS4, 5, 6, and 7 (Fig 3.5A). We speculated that the exact number of Yan molecules forming a repressive complex at ETS2, 3 and 8 in vivo depends on the spatial position of ETS8, such that although there is only one Yan molecule bound at ETS23 pair and another at ETS8, there could be additional unbound Yan molecules associated with the two DNA-bound Yan molecules. Further speculation about how this might contribute to formation of effective repressive complexes is presented in the Discussion.

We hypothesized that SAM-mediated self-association of Yan mediates the non-linear collaborative interactions among the three strong-affinity ETS sites. Self-association is critical for the repressive function of Yan (Webber et al., 2013a), and that motivated us to test if Yan self-association contributed to the repressive functions of ETS2, 3 and 8. To study this we set $\gamma = 0$, effectively allowing only Yan monomers to occupy ETS site across the MHE. As a result, the respective transcriptional repression and activation from the strong-affinity sites and weak affinity sites both were lost (Fig 3.5B), indicating the importance of Yan self-association to the transcriptional regulation mediated by strong- and weak-affinity sites. Furthermore, MHE^{mut8} and MHE^{mut23} showed reduced transcription in the Yan monomer background, indicating that losing Yan self-association disrupted the repressive complex at ETS2, 3 and 8. This simulated result was in contradiction to the increased expression of MHE^{mut8} and $MHEmut^{23}$ in experiments undertaken in a Yan wildtype background (Boisclair Lachance et al., 2018).

We also compared the transcription driven by the wildtype MHE when Yan self-association was allowed or prevented (Fig 3.5C). When Yan self-association was not allowed in the model, transcription was activated at a broader range of Yan and Pnt concentrations, consistent with a reduction in Yan-mediated repression. To validate these *in silico* results experimentally, I used CRISPR/Cas9 engineering to introduce specific missense mutations in endogenous *yan* that prevent SAM-SAM association. I generated two different Yan monomer alleles, yan^{V105R} and yan^{A86D}. Then the intensity of GFP expression driven by the MHE^{WT} and ETS site mutant reporters in stage 11 embryos can be measured in these monomer mutants. Further discussion of these planned experiments and their expected results can be found in Chapter 4.

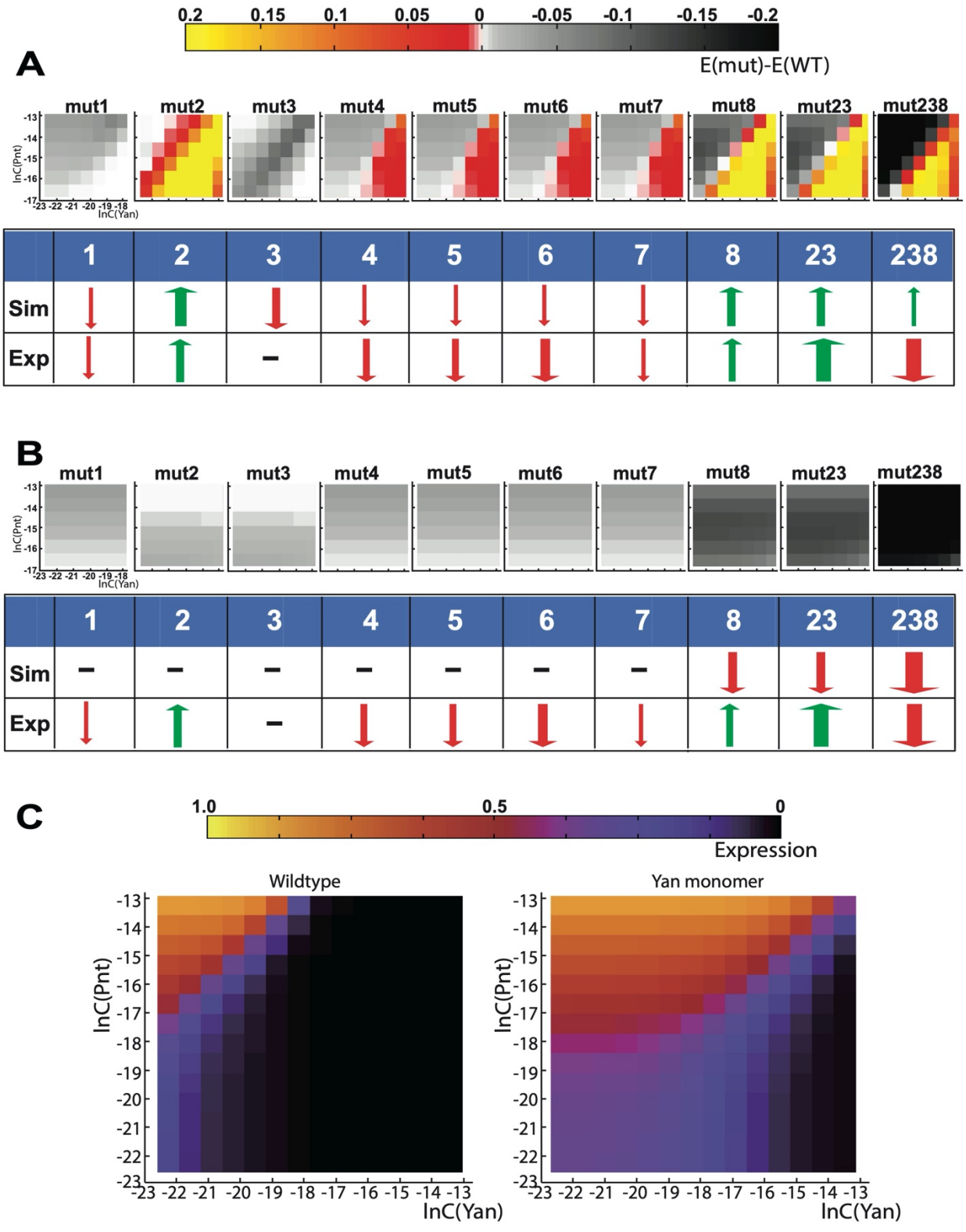


Fig 3.5 When only one molecule is permitted at the ETS23 pair or Yan monomer

(Fig 3.5 continued) A and B are both from the model that strong site ETS2, 3 and 8 were in close proximity with ETS8 being biased towards ETS2. A) When only one Yan/Pnt molecule was permitted at the ETS23 pair, the strong sites ETS2, ETS8 and ETS23 as pair still contributed to repression. Individual strong site ETS3 and weak site ETS1 contributed to activation. Meanwhile, ETS4, 5, 6, 7 contributed neither repression nor activation as the corresponding mutants showed minimal transcription changes. B) When Yan is a monomer mutant, all strong and weak sites except the individual ETS8 or ETS23 pair lost contribution to both repression and activation. Mut8 and mut23 showed reduced transcription, that should be test in my CRISPRed Yan monomer mutants. C) The expression of MHE wildtype reporter in wildtype vs. Yan monomer background. Yan monomer mutant showed a broader range of Yan and Pnt concentration enabling gene activation.

3.2.4 Additional inputs through other transcriptional activators binding and the threshold of total activators add another layer of regulation

3.2.4.1 Additional activating inputs from Mad/ Twist

Other transcription factors including Mad, Tin, Twi and dTCF have been identified to bind at MHE (Boisclair Lachance et al., 2018; Halfon et al., 2000). Because all four are activators, to simplify their incorporation into our model, we grouped them into a single category of extra activators. The binding sites for these extra activators are at the non-specific DNA sequence for ETS factors (Fig 3.1A). To simplify the calculation, our analysis only included the first element of the MHE comprised of the sequence from ETS1 to ETS7 (Fig 3.6A). We incorporated the DNA-protein interactions of the extra activators into the model by adding the term ω. ω is the free energy value of specific DNA binding by these additional activators. Although there was no direct measurements of the DNA-protein binding affinities of the extra activators, the known homeobox proteins, the category of which Tin belongs to, were found to have 1/10 fold of the DNA-protein affinity of TEL1 (De et al., 2014; Jung et al., 2018), and the SMADS proteins, the category of which Mad belongs to, have weak DNA-protein affinity (BabuRajendran et al., 2010; Hill, 2016). Therefore, we assigned θ the value to be 1/10 of ETS specific binding. Regarding protein-protein interactions, we were not aware of any direct interactions between Yan/Pnt and

the extra activators from previous experiments, and so did not include them in the model. To further translate the occupancy of extra activators to transcription, each bound activator interacted with BTC the same way as Pnt did, thereby adding one ω_{pnt} interaction with BTC for each bound activator.

Although the extra activators didn't affect the regulation from individual sites they disrupted the synergistic repressive functions that depend on long-range interactions. Adding extra activators didn't show significant change of gene activation (Fig 3.6B), suggesting that in the model MHE was predominantly controlled by Yan and Pnt binding at the ETS sites. Furthermore, as the independence of the binding of Yan/Pnt and extra activators, the individual ETS site' contribution to repression and activation also didn't change (Fig 3.6C, 3.3B). However, the double and triple mutants MHE^{mut23} and MHE^{mut238} showed unchanged and increased expression respectively, indicating the repression from the complex of ETS2, 3 and 8 was disrupted. This disruption might come from an indirect collaborative action for activation between the extra activators and Pnt. This simulation result was inconsistent with the experimental one that the expression of MHE^{mut23} significantly increased (Boisclair Lachance et al., 2018).

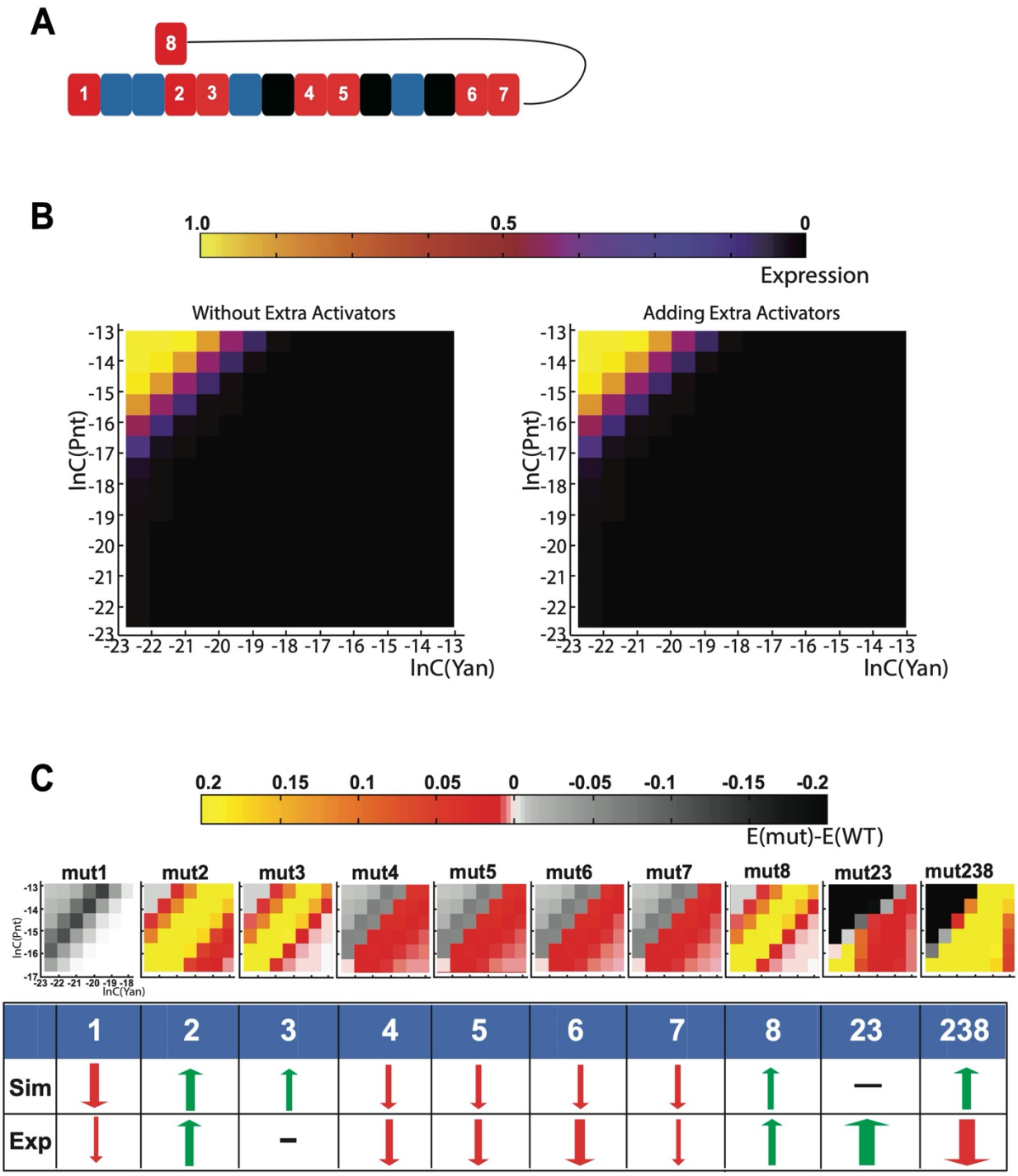


Fig 3.6 Adding extra activators disrupted the repression from the ETS23 pair

A) Conceptualized lattice representing MHE. Respective red and black squares represent ETS specific and nonspecific binding site as in Fig 3.2B. ETS8 is in close proximity to ETS23 pair. Blue squares represent the binding sites for the extra activators. B) The expression of MHE (Fig

3.6 continued) wildtype reporter in the background of without extra activators vs. adding extra activators. Adding extra activators didn't change gene activation. C) All images represent the relative expression change between MHE wildtype and mutant reporters. Adding extra activators didn't change the contribution of ETS2, 3, 8 to repression and ETS1, 4, 5, 6, and 7 to activation. However the repression from the ETS23 pair was lost.

3.2.4.2 Threshold of activator occupancy may switch MHE activation

To reconcile the conflict between the simulation and the experiments regarding the extra activators, we proposed an activator threshold model to attenuate the activation. The threshold model required that the activation occurred only when the number of activators including Pnt and the extra activators bound to the DNA was above a threshold. Threshold responses in transcription regulation was revealed in broad developmental context, particularly the *Drosophila* embryo patterning (Ashe et al., 2000; Courey and Huang, 1995; Kringstein et al., 1998; Stamataki et al., 2005; Struhl et al., 1989). For example, along the A-P axis of the *Drosophila* embryo, a smooth gradient of transcription activator Bicoid establishes precise target gene expression with sharp boundaries (Driever and Nüsslein-Volhard, 1988; Gao and Finkelstein, 1998). Directly illustrating the activation thresholds, uniform Bicoid expression at low, medium and high levels induced target gene expression profile representative of distinct positions along the A-P axis (Crauk and Dostatni, 2005). As MHE regulates *eve* transcription in *Drosophila* embryo, the threshold mechanism could also apply.

In additional to the extra activator in consistency, another discrepancy between the simulation and the experimental results about the MHE^{mut238} also motivated the threshold idea. MHE^{mut238} showed strong increased of simulated transcription as the three strong-affinity sites contributed to the formation of a strong repressive complex of Yan (Fig 3.4, 3.6), in contradiction to the strongly decreased transcription in experiments. We hypothesized that such decreased

transcription could be a result that transcription activators bound at MHE were below the activation threshold.

To test this idea in our model, we set the threshold of the number of activators, only above which the activators could facilitate the binding of BTC, and thus gene activation. We first tried setting five as the threshold considering that MHE^{mut238} had only five remaining ETS sites. When the threshold is five, as Fig 3.7 shown, all three strong-affinity sites strongly contributed to repression while transcription of mut238 was much reduced, and all weak-affinity sites strongly contributed activation, consistent with the experimental results. Additionally, adding the activation threshold into the model achieved a more precise range of Yan and Pnt concentrations for both repression and activation. That is consistent with the phenomenon that sharp boundaries of gene regulation are found throughout development.

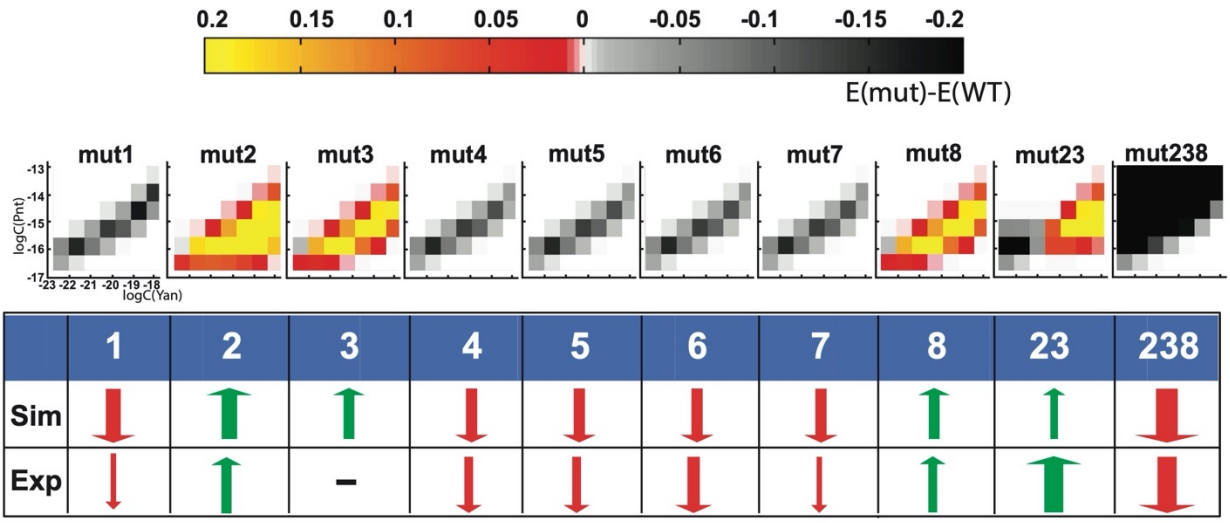


Fig 3.7 Activator occupancy of more than 9 sites can activate MHE transcription

In the model, strong site ETS2, 3 and 8 were in close proximity with ETS8 being biased towards ETS2, and extra activators were added. A) When a threshold of five activators were required for activation, strong sites ETS2, 3 and 8 contributed to repression strongly and at a precise concentration range. Meanwhile weak sites ETS1, 4, 5, 6 and 7 contributed to activation strongly and precisely. Losing three strong sites together also resulted in strong reduction of the transcription.

3.3 Discussion

3.3.1 Yan self-association organizes the 3D cooperation of the ETS sites for transcriptional repression

By developing a thermodynamic model focusing on the ETS-binding motifs within a conserved regulatory module, the *eve* MHE, we identified a cis-regulatory syntax to organize the repressive and active inputs for transcriptional regulation. Based on comparisons between the modeling and experiment results for multiple ETS mutants of the enhancer, we propose distinctive contributions of the strong- and weak-affinity sites to transcriptional repression and activation. In particular, we found the three strong-affinity sites collaborated to contribute strongly to the repression. Mechanistically, although one of the strong-affinity sites, ETS8, is distant from the other pair, our model suggested that Yan self-association facilitated their 3D interactions (Fig 3.8). The simulation showed that, the Yan monomer background lost the long-range interactions of the strong sites and thus the repressive regulation was disrupted.

Regarding the transcriptional activation contributed by the weak-affinity sites, the interactions among the three strong sites was surprisingly sufficient to control Pnt's preference to the weak sites despite the same DNA-protein binding affinities for Yan and Pnt. From a broad perspective, we speculate that the non-linear interactions among the TF binding sites, potentially facilitated by protein-protein interactions, are critical to organize transcriptional activation and repression distinctively.

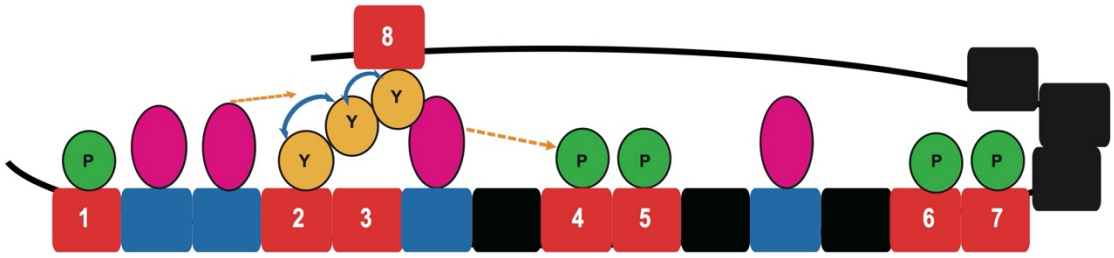


Fig 3.8 Illustration for MHE ETS2, 3 and 8 cooperativity and the interactions for all the factors involved

Green and orange circles represent Pnt and Yan molecules respectively. With the proximity of the three strong sites ETS2, 3 and 8, the self-association (blue double-sided arrows) of Yan facilitates the formation of a repressive complex. Although there was one Yan occupying the ETS2, 3 pair and one Yan occupying ETS8, there can be Yan molecule(s) in the middle of the two bound on DNA depending on the spatial arrangements of the ETS sites. Strong occupation of Yan at the strong-affinity sites made Pnt prefer weak-affinity sites ETS 1, 4, 5, 6 and 7. The pink ovals represent extra activators binding at their specific sites. The binding of the extra activators has impact on the occupancy of Yan and Pnt in an indirect way (yellow dash arrows).

Yan self-association may stabilize the interactions of ETS2, 3 and 8. Using ChIP-seq, it has been revealed that the genome-wide recruitment pattern was quite similar between Yan monomer and Yan wildtype (Webber et al., 2013a). However, it is worth noticing that the amplitude of binding was reduced in the Yan monomer background despite largely unchanged recruitment, suggesting the contribution of Yan self-association to stabilize occupancy (Webber et al., 2013a). This is consistent with our model that the occupancy of Yan at the ETS sites was mainly driven by the DNA-protein binding affinities whereas Yan self-association stabilized the binding of Yan molecule bound at the neighboring sites. In particular, for ETS 2, 3 and 8, which are not linear neighboring sites, Yan self-association may be key to maintain their spatial arrangements. To experimentally test the proximity between ETS23 and 8, current ChIP-seq/qPCR is not at the resolution to detect TF binding sites within an enhancer. Super resolution imaging for single

molecule may be able to help virtualize the Yan molecules bound at ETS23 and 8 to reveal their spatial relationship.

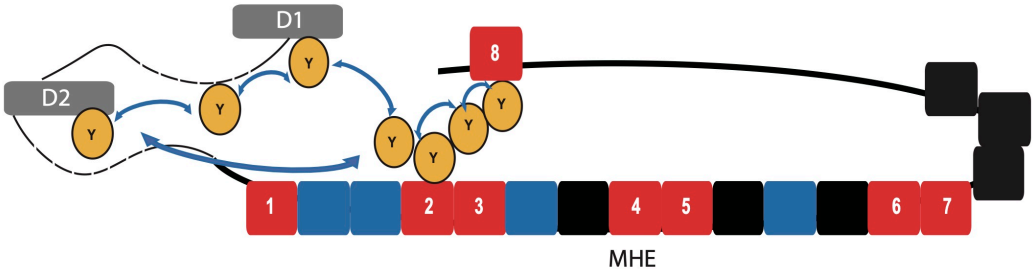


Fig 3.9 Illustration for collaborative repssion from D1, D2 and MHE

MHE, D1 and D2 are indicated. For simplicity, among the TFs, only Yan molecules are illustrated. Multiple Yan associates to form a repressive complex at the space of ETS23 and 8. The repressive complex further associates with Yan molecules bound at D1 and D2. Blue double-sided arrows represent Yan self-association.

The spatial arrangements of the three strong-affinity sites may determine the formation of the repressive complex. Among the three sites, ETS2 and 3 are a pair, which are too close to permit simultaneous occupancy of Yan (Boisclair Lachance et al., 2018). Our model indicated that such configuration of ETS2 and 3 reduced but not disrupted the repressive regulation if ETS8 still interacted with the pair (Fig 3.4A). This immediate juxtaposition of ETS2 and ETS3 may stabilize their interactions with ETS8. Speculating from the dynamics of Yan-DNA interactions, the presence of two overlapping strong-affinity sites could promote stable occupancy of at least one Yan molecule, which would associate with the Yan molecule bound at ETS8, by increasing the chance that a newly dissociated Yan would immediately rebind. There may be more than two Yan molecules within the space of ETS2, 3 and 8 depending on the dynamics of DNA bending (Fig 3.7). Supporting this DNA bending idea, both the experimental (Boisclair Lachance et al., 2018) and simulation results indicated that ETS8 biased towards interacting with ETS2. The bias

may be a result of the bending direction of the DNA, considering that the core consensus sequence of ETS2 and ETS3 are in antiparallel. The dynamic DNA 3D structure and state can modulate the strength of repression for spatiotemporal-precise transcription through allowing different numbers of Yan molecules forming the repressive complex at the space.

3.3.2 The local cooperativity within MHE and Yan self-association facilitate the long-range interactions of eve repressive CRMs

The strong self-associated Yan repressive complex organized by ETS2, 3 and 8 may facilitate the cooperativity among multiple repressive CRMs. Four CRMs bound by Yan had been identified: D1, D2, MHE and D4, and the long-range interactions among the first three CRMs organized Yan occupancy at the elements and the repressive function of Yan (Webber et al., 2013b). The mechanism of how the long-range interactions organize the Yan occupancy has not been fully investigated. Associating the reduced Yan occupancy across the genome with the collaborative repression among the three CRMs (Webber et al., 2013a), we speculate that Yan self-association may facilitate such long-range interaction for repressive function. Focusing on MHE, more speculatively, the ETS2, 3 and 8 interactions in addition to Yan self-association facilitate the long-range cooperativity (Fig 3.9). Supporting this idea, MHE^{mut23} showed reduced Yan occupancy at D1 and D2 (Boisclair Lachance et al., 2018). To further validate this speculation, I CRISPRed a MHE^{mut8} allele and Yan monomer mutant alleles. We hypothesized that like MHE mut23, MHE^{mut8} should also show reduced Yan occupancy at D1 and D2 and similar reduction should be seen in Yan monomer mutants.

3.3.3 Additional layer of regulation through the extra activators

Adding to the complexity of the regulation, additional activators interact with Yan and Pnt. Genetic interaction assays have shown synergistic regulation by ETS and other activator binding sites on MHE (Halfon et al., 2000). Losing ETS and other activator sites together resulted in synergistic reduction of MHE-*lacZ* expression, suggesting cooperative activation. The molecular mechanism of this cooperativity was unknown and there has not been any study to reveal direct protein-protein interactions between Pnt/Yan and these additional activators. Our model suggested an indirect interaction between Pnt and the extra activators that could disrupt the repression functions of the ETS sites even when the binding of the extra activators and Pnt/Yan was independent. This may be a result of the occupancy shift of Yan/Pnt at the ETS sites. The extra activators occupy the ETS nonspecific DNA, which constitutes the neighboring space to the ETS specific sites and thus if it is occupied by Yan could help stabilize Yan occupancy at the ETS specific sites through self-association. Therefore the extra activators at the nonspecific DNA indirectly destabilize Yan occupancy at the ETS sites. In turn, Pnt gains privilege in the competition with Yan for the ETS sites. This idea could be experimentally tested by measurements of the occupancy of Yan/Pnt at the MHE for different extra activators' mutants is worth testing for future experiments. We expected to see increased Yan occupancy and decreased Pnt occupancy in the extra activators' mutants. From a broad perspective, as clusters of ETS and the extra activators (including Mad, Twi, Tin and dTCF) were identified for other dorsal mesodermal enhancers (Halfon et al., 2002), such cooperativity could be implemented in regulating more CRMs and gene transcription.

Interaction between Yan/Pnt and the extra activators to integrate regulations from multiple signaling pathways, in combination with the threshold mechanism, could ensure more precise

transcription. Just as Yan and Pnt are effectors of the Ras signaling, the extra activators including Mad, Tin, Twi and dTCF, are respective effectors of Dpp and Wg signaling pathways (Azpiazu and Frasch, 1993; Baylies and Bate, 1996; Bodmer, 1993; Brunner et al., 1997; Sekelsky et al., 1995; van de Wetering et al., 1997). Through MHE, the effectors interact and thus the regulation from upstream signals are integrated. Specifically, the pleiotropic inductive Ras signaling, in combination with Dpp and Wg, which are more mesoderm restricted, can achieve spatiotemporal specific transcription (Halfon et al., 2000). Our simulation provided a theoretical basis for an indirect mechanism for signaling integration at molecular level. The spatiotemporal specificity also may come from the threshold mechanism. For example, the sharp transition for Eve+ from surrounding mesoderm, where the levels of some activators are high, and mesoderm may be the results of total number of activators below activating threshold. Overall the extra activators not only provide sufficiency for activation but also confer the specificity of gene transcription.

Further exploration of the extra activations may reveal differential interactions with distinct ETS sites. The in vivo experiments showed that the strengths of repression and activation contributed by distinct ETS sites were different, such as ETS6 conferred stronger activation whereas ETS7 conferred weaker activation. Our current model with Mad, Twi, Tin and dTCF all grouped into the same extra activator, didn't capture the strength differences from the in vivo results (Figure 3.5, 3.6 and 3.7). The inconsistent activation and repression strength between the simulation and the in vivo results may be a result of the allocation of different extra activators on MHE. For example, a cluster of Mad and Twi sites is in close proximity with ETS7 while a Tin site is close to ETS6. If the difference between MHE^{mut6} and MHE^{mut7} is contributed by the extra activators, *mad*, *twi* and *tin* mutants should tuned MHE^{mut6} and MHE^{mut7} reporter expression differentially.

How the interactions between extra activators and Yan/Pnt to fine tune transcriptional regulation is an exciting area for further studies.

3.4 Materials and methods

Calculation of occupancy at equilibrium

We represent the MHE as a lattice of potential binding sites for Yan and Pnt. A pattern of Yan and Pnt occupancy across MHE is referred to as a microstate. At constant temperature and pressure, to calculate the fractional occupancy of microstate k at equilibrium, we used

$$\text{fk}=\frac{\mathbf{exp}(-\Delta \mathbf{G}_{\boldsymbol{k}}/\mathbf{RT})[\boldsymbol{Yan}]^{\boldsymbol{a(k)}}[\boldsymbol{Pnt}]^{\boldsymbol{b(k)}}}{\boldsymbol{Z}}. \quad (1)$$

In this case $\Delta\mathbf{G}$ is the change in Gibbs free energy upon Yan and Pnt binding for microstate k, R is the gas constant ($\mathbf{1.987204118 \times 10^{3}}$ kcal/mol degree), T is the absolute temperature ($\mathbf{300}°$ K for all calculations), [Yan] and [Pnt] are the respective concentrations of free molecules of Yan and Pnt in the nucleus, a(k) and b(k) represent the respective numbers of bound Yan and Pnt molecules in microstate k. The numerator of Eq. 1 is a Boltzmann weight for a given microstate. The partition function, Z, is the sum of Boltzmann weights for all microstates considered, and is given by

$$Z = \sum_k \exp\left(\left(-\frac{\Delta G}{RT}\right)\right)[Yan]^{a(k)}[Pnt]^{b(k)}. \quad (2)$$

We considered two types of binding sites for Yan and Pnt, specific and nonspecific binding sites. αyan and αpnt represent the respective change in Gibbs free energy of Yan and Pnt specific binding to ETS sites, whereas βyan and βpnt represent the respective change in free energy of non-specific Yan and Pnt binding to nonspecific sites. γ represents the change in free energy of

association between two adjacent Yan molecules via their SAM domains. Reflecting Yan's ability to polymerize, each additional adjacent Yan molecule contributes an additional increment of γ. The sum of change in Gibbs free energy, in terms of αyan, αpnt, βyan, βpnt and γ, defines the ΔG_k of the microstate.

Calculation of transcription

We considered that the transcription is proportional to the binding of a basal transcription complex (BTC) at DNA and the binding of BTC is regulated by the transcription factors Yan and Pnt. The concept of BTC here is a necessary and rate limiting step for transcription but not related to any specific molecule. In addition to the binding sites for Yan and Pnt in MHE, we added another binding site for BTC. Then the fractional occupancy of BTC represent the relative transcription, which is

$$f=\frac{\boldsymbol{Z_2}}{\boldsymbol{Z}+\boldsymbol{Z_2}}. \quad (3)$$

In this case, $\mathbf{Z_2}$ is the sum of Boltzmann weights for all microstates containing a bound BTC, whereas Z is the sum for all microstates in which BTC is unbound. To calculate $\mathbf{Z_2}$, we used

$$Z2 = \sum_j \exp\left(\left(-\frac{\Delta G'}{RT}\right)\right)[Yan]^{a(j)}[Pnt]^{b(j)}. \quad (4)$$

All j microstates contain a bound BTC and the $\Delta G'$ equals to ΔG, which is determined by Yan and Pnt occupancy, adding the change in Gibbs free energy of BTC binding. As Pnt is an activator, for each molecule of Pnt binding at DNA, there is an ωpnt (<0) energy adding to the Gibbs free energy of BTC binding to strengthen its binding, whereas for each molecule of Yan binding at DNA, there is an ωyan (>0) energy adding the Gibbs free energy of BTC binding. Therefore, αyan, αpnt, βyan, βpnt,γ, ωyan and ωpnt defines $\Delta G'$.

MHE mutants

For MHE mutants, we changed the specific binding sites to nonspecific sites by switching the binding affinities from αyan, αpnt to βyan, βpnt respectively.

Parameter table

αyan	-9.955 kcal/mol
αpnt	-9.955 kcal/mol
βyan	-5.797 kcal/mol
βpnt	-5.797 kcal/mol
γ	-7.000 kcal/mol
ωyan	2 kcal/mol
ωpnt	-1 kcal/mol
Concentration (pnt/yan)	0.1 to 1000nM

3.5 References

Ashe, H. L., Mannervik, M. and Levine, M. (2000). Dpp signaling thresholds in the dorsal ectoderm of the Drosophila embryo. *Development* **127**, 3305 LP – 3312.

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

BabuRajendran, N., Palasingam, P., Narasimhan, K., Sun, W., Prabhakar, S., Jauch, R. and Kolatkar, P. R. (2010). Structure of Smad1 MH1/DNA complex reveals distinctive rearrangements of BMP and TGF-β effectors. *Nucleic Acids Res.* **38**, 3477–3488.

Baylies, M. K. and Bate, M. (1996). twist: A Myogenic Switch in Drosophila. *Science (80-.).* **272**, 1481 LP – 1484.

Biggin, M. D. (2011). Animal Transcription Networks as Highly Connected, Quantitative Continua. *Dev. Cell* **21**, 611–626.

Bodmer, R. (1993). The gene tinman is required for specification of the heart and visceral muscles in Drosophila. *Development* **118**, 719 LP – 729.

Boisclair Lachance, J.-F., Webber, J. L., Hong, L., Dinner, A. R. and Rebay, I. (2018). Cooperative recruitment of Yan via a high-affinity ETS supersite organizes repression to confer specificity and robustness to cardiac cell fate specification. *Genes Dev.* **32**, 389–401.

Brunner, E., Peter, O., Schweizer, L. and Basler, K. (1997). pangolinencodes a Lef-1 homologue that acts downstream of Armadillo to transduce the Wingless signal in Drosophila. *Nature* **385**, 829–833.

Carmena, A., Gisselbrecht, S., Harrison, J., Jiménez, F. and Michelson, A. M. (1998). Combinatorial signaling codes for the progressive determination of cell fates in the Drosophila embryonic mesoderm. *Genes Dev.* **12**, 3910–3922.

Carmena, A., Buff, E., Halfon, M. S., Gisselbrecht, S., Jiménez, F., Baylies, M. K. and Michelson, A. M. (2002). Reciprocal regulatory interactions between the Notch and Ras signaling pathways in the Drosophila embryonic mesoderm. *Dev. Biol.* **244**, 226–242.

Charlton-Perkins, M., Whitaker, S. L., Fei, Y., Xie, B., Li-Kroeger, D., Gebelein, B. and

Cook, T. (2011). Prospero and Pax2 combinatorially control neural cell fate decisions by modulating Ras- and Notch-dependent signaling. *Neural Dev.* **6**, 20.

Courey, A. J. and Huang, J.-D. (1995). The establishment and interpretation of transcription factor gradients in the Drosophila embryo. *Biochim. Biophys. Acta - Gene Struct. Expr.* **1261**, 1–18.

Crauk, O. and Dostatni, N. (2005). Bicoid Determines Sharp and Precise Target Gene Expression in the Drosophila Embryo. *Curr. Biol.* **15**, 1888–1898.

Crocker, J., Abe, N., Rinaldi, L., McGregor, A. P., Frankel, N., Wang, S., Alsawadi, A., Valenti, P., Plaza, S., Payre, F., et al. (2015). Low affinity binding site clusters confer HOX specificity and regulatory robustness. *Cell* **160**, 191–203.

De, S., Chan, A. C. K., Coyne, H. J., Bhachech, N., Hermsdorf, U., Okon, M., Murphy, M. E. P., Graves, B. J. and McIntosh, L. P. (2014). Steric Mechanism of Auto-Inhibitory Regulation of Specific and Non-Specific DNA Binding by the ETS Transcriptional Repressor ETV6. *J. Mol. Biol.* **426**, 1390–1406.

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

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 LP – 4193.

Green, S. M., Coyne, H. J., McIntosh, L. P. and Graves, B. J. (2010). DNA binding by the ETS protein TEL (ETV6) is regulated by autoinhibition and self-association. *J. Biol. Chem.* **285**, 18496–18504.

Gupta, M. and Liu, J. S. (2005). De novo cis-regulatory module elicitation for eukaryotic genomes. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 7079 LP – 7084.

Halfon, M. S., Carmena, A., Gisselbrecht, S., Sackerson, C. M., Jiménez, F., Baylies, M. K. and Michelson, A. M. (2000). Ras pathway specificity is determined by the integration of multiple signal-activated and tissue-restricted transcription factors. *Cell* **103**, 63–74.

Halfon, M. S., Grad, Y., Church, G. M. and Michelson, A. M. (2002). Computation-based discovery of related transcriptional regulatory modules and motifs using a experimentally validated combinatorial model. *Genome Res.* **12**, 1019–1028.

Hill, C. S. (2016). Transcriptional Control by the SMADs. *Cold Spring Harb. Perspect. Biol.* **8**, a022079.

Hope, C. M., Rebay, I. and Reinitz, J. (2017). DNA Occupancy of Polymerizing Transcription Factors: A Chemical Model of the ETS Family Factor Yan. *Biophys. J.* **112**, 180–192.

Jung, C., Bandilla, P., von Reutern, M., Schnepf, M., Rieder, S., Unnerstall, U. and Gaul, U. (2018). True equilibrium measurement of transcription factor-DNA binding affinities using automated polarization microscopy. *Nat. Commun.* **9**, 1605.

Knirr, S. and Frasch, M. (2001). Molecular integration of inductive and mesoderm-intrinsic inputs governs even-skipped enhancer activity in a subset of pericardial and dorsal muscle progenitors. *Dev. Biol.* **238**, 13–26.

Kribelbauer, J. F., Rastogi, C., Bussemaker, H. J. and Mann, R. S. (2019). Low-Affinity Binding Sites and the Transcription Factor Specificity Paradox in Eukaryotes. *Annu. Rev. Cell Dev. Biol.* **35**, 357–379.

Kringstein, A. M., Rossi, F. M. V, Hofmann, A. and Blau, H. M. (1998). Graded

transcriptional response to different concentrations of a single transactivator. *Proc. Natl. Acad. Sci.* **95**, 13670 LP – 13675.

Lambert, S. A., Jolma, A., Campitelli, L. F., Das, P. K., Yin, Y., Albu, M., Chen, X., Taipale, J., Hughes, T. R. and Weirauch, M. T. (2018). The Human Transcription Factors. *Cell* **172**, 650–665.

Le, D. D., Shimko, T. C., Aditham, A. K., Keys, A. M., Longwell, S. A., Orenstein, Y. and Fordyce, P. M. (2018). Comprehensive, high-resolution binding energy landscapes reveal context dependencies of transcription factor binding. *Proc. Natl. Acad. Sci. U. S. A.* **115**, E3702–E3711.

Mann, R. S., Lelli, K. M. and Joshi, R. (2009). Hox specificity unique roles for cofactors and collaborators. *Curr. Top. Dev. Biol.* **88**, 63–101.

Manning, G. S. (2006). The Persistence Length of DNA Is Reached from the Persistence Length of Its Null Isomer through an Internal Electrostatic Stretching Force. *Biophys. J.* **91**, 3607–3616.

Méthot, N. and Basler, K. (1999). Hedgehog Controls Limb Development by Regulating the Activities of Distinct Transcriptional Activator and Repressor Forms of Cubitus interruptus. *Cell* **96**, 819–831.

Nitta, K. R., Jolma, A., Yin, Y., Morgunova, E., Kivioja, T., Akhtar, J., Hens, K., Toivonen, J., Deplancke, B., Furlong, E. E. M., et al. (2015). Conservation of transcription factor binding specificities across 600 million years of bilateria evolution. *Elife* **4**, e04837.

O'Neill, E. M. D., Tjian, R. and Rubin, G. M. (1994). The Activities of Two Ets-Related Transcription Factors Required for Drosophila Eye Development Are Modulated by the Ras

/ MAPK Pathway. *Cell* **78**, 137–147.

Papatsenko, D. A., Makeev, V. J., Lifanov, A. P., Régnier, M., Nazina, A. G. and Desplan, C. (2002). Extraction of functional binding sites from unique regulatory regions: the Drosophila early developmental enhancers. *Genome Res.* **12**, 470–481.

Papatsenko, D., Goltsev, Y. and Levine, M. (2009). Organization of developmental enhancers in the Drosophila embryo. *Nucleic Acids Res.* **37**, 5665–5677.

Parker, D. S., White, M. A., Ramos, A. I., Cohen, B. A. and Barolo, S. (2011). The cis-Regulatory Logic of Hedgehog Gradient Responses: Key Roles for Gli Binding Affinity, Competition, and Cooperativity. *Sci. Signal.* **4**, ra38 LP-ra38.

Qiao, F., Song, H., Kim, C. A., Sawaya, M. R., Hunter, J. B., Gingery, M., Rebay, I., Courey, A. J. and Bowie, J. U. (2004). Derepression by Depolymerization. *Cell* **118**, 163–173.

Ramos, A. I. and Barolo, S. (2013). Low-affinity transcription factor binding sites shape morphogen responses and enhancer evolution. *Philos. Trans. R. Soc. B Biol. Sci.* **368**, 20130018–20130018.

Rebay, I. and Rubin, G. M. (1995). Yan functions as a general inhibitor of differentiation and is negatively regulated by activation of the Ras1/MAPK pathway. *Cell* **81**, 857–866.

Ruan, S., Swamidass, S. J. and Stormo, G. D. (2017). BEESEM: estimation of binding energy models using HT-SELEX data. *Bioinformatics* **33**, 2288–2295.

Sekelsky, J. J., Newfeld, S. J., Raftery, L. A., Chartoff, E. H. and Gelbart, W. M. (1995). Genetic characterization and cloning of mothers against dpp, a gene required for decapentaplegic function in Drosophila melanogaster. *Genetics* **139**, 1347 LP – 1358.

Stamataki, D., Ulloa, F., Tsoni, S. V, Mynett, A. and Briscoe, J. (2005). A gradient of Gli activity mediates graded Sonic Hedgehog signaling in the neural tube. *Genes Dev.* **19**, 626–641.

Stewart, A. J., Hannenhalli, S. and Plotkin, J. B. (2012). Why transcription factor binding sites are ten nucleotides long. *Genetics* **192**, 973–985.

Stormo, G. D., Zuo, Z. and Chang, Y. K. (2015). Spec-seq: determining protein-DNA-binding specificity by sequencing. *Brief. Funct. Genomics* **14**, 30–38.

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

van de Wetering, M., Cavallo, R., Dooijes, D., van Beest, M., van Es, J., Loureiro, J., Ypma, A., Hursh, D., Jones, T., Bejsovec, A., et al. (1997). Armadillo Coactivates Transcription Driven by the Product of the Drosophila Segment Polarity Gene dTCF. *Cell* **88**, 789–799.

Webber, J. L., Zhang, J., Cote, L., Vivekanand, P., Ni, X., Zhou, J., Nègre, N., Carthew, R. W., White, K. P. and Rebay, I. (2013a). The relationship between long-range chromatin occupancy and polymerization of the drosophila ets family transcriptional repressor yan. *Genetics* **193**, 633–649.

Webber, J. L., Zhang, J., Mitchell-Dick, A. and Rebay, I. (2013b). 3D chromatin interactions organize Yan chromatin occupancy and repression at the even-skipped locus. *Genes Dev.* **27**, 2293–2298.

Wei, G. H., Badis, G., Berger, M. F., Kivioja, T., Palin, K., Enge, M., Bonke, M., Jolma, A., Varjosalo, M., Gehrke, A. R., et al. (2010). Genome-wide analysis of ETS-family DNA-

binding in vitro and in vivo. *EMBO J.* **29**, 2147–2160.

Weirauch, M. T., Yang, A., Albu, M., Cote, A. G., Montenegro-Montero, A., Drewe, P., Najafabadi, H. S., Lambert, S. A., Mann, I., Cook, K., et al. (2014). Determination and inference of eukaryotic transcription factor sequence specificity. *Cell* **158**, 1431–1443.

Xu, C., Kauffmann, R. C., Zhang, J., Kladny, S. and Carthew, R. W. (2000). Overlapping Activators and Repressors Delimit Transcriptional Response to Receptor Tyrosine Kinase Signals in the Drosophila Eye. *Cell* **103**, 87–97.

Zhang, J., Graham, T. G. W., Vivekanand, P., Cote, L., Cetera, M. and Rebay, I. (2010). Sterile Alpha Motif Domain-Mediated Self-Association Plays an Essential Role in Modulating the Activity of the Drosophila ETS Family Transcriptional Repressor Yan. *Mol. Cell. Biol.* **30**, 1158–1170.

Chapter 4 Discussion and Future Directions

4.1 Exploration how cross-regulation between pntP2 and pntP3 fine-tunes Pnt output

The transcriptional cross-regulation between pntP2 and pntP3, in combination with their different transactivation potentials, may be critical to fine-tune the output of Pnt-activated transcriptional programs. The newly characterized PntP3 isoform has stronger transactivation potential than PntP2, and it acts redundantly with PntP2 to activate *pntP1* for the initiation of first-round photoreceptor fates. This redundancy confers robustness to cell fate specification against genetic and signaling variation. At the molecular level, the overall Pnt activity level in an individual cell reflects the overall concentration of PntP2, PntP3 and PntP1 proteins, the latter which is determined by the combined activity of PntP2 plus PntP3. Overall Pnt activity must precisely tuned, as much or too little Pnt both lead to abnormal activation of downstream transcriptional programs. Therefore, as the signaling-responsive initiators of cell fate transitions, PntP2 and PntP3 must be regulated, particularly when both are present redundantly.

Regarding their transcriptional regulation, my RT-qPCR and FISH results have showed the auto-repression of *pntP2* and *pntP3* and the cross-activation of *pntP2* by PntP3. Although I was unable to detect *pntP3* transcripts by FISH, the RT-qPCR showed a significant 12% increase in *pntP3* in pnt^{p2} mutant tissue, suggesting that PntP2 normally represses *pntP3*. Thus the auto- and cross-regulation between PntP2 and PntP3 may form a regulatory network for tuning the final overall PntP2 plus PntP3 in a context-dependent manner.

There are a few remaining questions. First, how transcription of *pntP3* is activated at the MF has not been revealed. As PntP3 is a stronger activator than PntP2, and can activate both *pntP1* and *pntP2*, exploration of the signaling inputs that activate *pntP3* expression is critical to understand

the initiation of the Pnt network. Second, the molecular mechanism behind the cross-regulatory interactions is unknown. In particular, my results did not reveal whether the regulations were direct or indirect. Assuming direct regulation, both the association with cofactors and the strength of RTK signaling to control the activities of PntP2 and PntP3 can be key determinants. Third, whether context-specific tuning of the Pnt network optimizes its activity level for each different cell fate specified needs to be explored. Depending on the developmental context, the existence and strength of cross-regulatory interactions may be adjusted. Below I suggest experimental approaches that could be pursued to address these three points.

Further exploration of *pntP3* transcriptional regulation can be achieved through constructing a *pntP3* transcriptional reporter. Although I could not detect *pntP3* transcript through FISH because of its short specific exon, a *pntP3* specific transcriptional reporter can be built by inserting a tag after its endogenous transcriptional starting site. To monitor the transcript instead of the protein, we can insert an RNA-level tag, either MS2 or PP7, which was developed for embryo live-imaging (Garcia et al., 2013). Upon transcription, the inserted MS2 or PP7 DNA sequence will lead to the formation of mRNA stem loops. These hairpins will be recognized by continuously expressing the coat proteins (MCP or PPCP) fused to a fluorescence tag like GFP; spatially localized GFP fluorescence will enable detection of *pntP3* transcript in live or fixed tissue.

Although the technology is sound, there are a few challenges to implement it for *pntP3*. First, as the stem loops will be inserted at the *pntP3* specific 5' end close to the start codon, the MS2-MCP complex may interfere with the initiation of translation. This would reduce PntP3 protein levels, which will weaken the negative autoregulation, and increase *pntP3* transcription. If this

happens, the MS2 system cannot monitor the transcript accurately. Therefore, the insertion position and number of stem loops need to be tested for minimal effect on the protein level by western blot. Because there is no specific PntP3 antibody, this can be assessed either with an antibody that recognizes all isoforms (Brunner et al., 1994) or by inserting the MS2 tag into the GFP-PntP3 BAC transgene or into an equivalent CRISPRed allele, which needs to be made first. Second, the protocol to detect the signal in the retina needs exploration. I am not aware of any publication using MS2-like system for live imaging in Drosophila retina. Retinal cells are smaller and more closely packed than those in the embryo so stronger fluorescence signal may be required to achieve single cell resolution. Staining with anti-GFP antibody should amplify the signal, but will only work in fixed tissue. Third, to quantify *pntP3* transcription and compare the results across genotypes and conditions, a normalization system is needed to monitor the binding efficiency of the coating protein to the MS2 loops. Similar MS2 reporters can be built for *pntP1* and *pntP2* for higher resolution study than is possible with FISH.

Once the technical challenges are resolved, *pntP3* transcription can be measured in different mutants to uncover the inputs that regulate its transcription. Relevant mutants can be divided into three categories: *pnt* isoforms, signaling pathway components and potential cofactors. Testing in pnt^{p1}, pnt^{p2}, pnt^{p3} mutants is critical to revel the Pnt network. Regarding signaling pathways, in addition to RTK signaling, Notch is a top candidate as it is an important contributor to photoreceptors fate specification, particularly R7 (Hayashi et al., 2008; Tomlinson and Struhl, 2001; Tomlinson et al., 2011). Additional signaling pathways to explore include Hh and Dpp. These two pathways play key roles in MF initiation(Heberlein and Mosest, 1995; Heberlein et al., 1993), and so are most active in the region where PntP3-GFP showed strong expression, putting them in the right place at the right time to contribute to the initial activation of *pntP3* and

the Pnt network. Last are the cofactors that may be involved in the transcription of *pntP3*, with Groucho and Atonal top candidates.

A differential equation-based model of the Pnt isoform network may then be built based on the quantification of *pntP1*, *pntP2* and *pntP3* transcript levels in wild type and isoform-specific *pnt* mutants made possible with the MS2-MCP system. The expectation is that the nature of the auto- and cross-regulation will lead to a feedback loop, through which the change in level of one isoform will not only adjust others but also tune its own. To understand the final Pnt output from the network (Fig 4.1A as one representative network topology), the modeling will incorporate the individual gene regulation for a single step change and then simulate for hundreds of thousands of steps until stable outputs, representative of the final states of the network initiated by a certain strength/frequency of EGFR signal, is achieved. After building the model, each nut and bolt of the system can be adjusted easily for testing different hypotheses. For example, I can modulate the strength of RTK signaling input quantitatively under different topology and regulation strength of the network. This may indicate which features of the network contribute to the robustness of Pnt output against the variations of the signaling input. In particular, negative feedback loops, which have been studied extensively to reduce variation and lead to stable output (Dublanche et al., 2006; Guinn and Balázsi, 2019; Singh and Hespanha, 2009), may be the focus of our future study. The strength of the individual regulation within the loop can be adjusted in the model, thereby modulating the variability of PntP2 and PntP3 levels, and in turn PntP1 expression.

Another interesting area to explore is the cell-to-cell variation in Pnt isoform protein and mRNA levels. My quantification for *pntP2* FISH showed that retinal cells at the same AP position,

which indicates the developmental state of cells, had significant variations of *pntP2* transcription and that such variations changed across the AP axis (Fig 4.1B). The variations of *pntP1* and *pntP2* can also be quantified by the MS2-reporter and FISH. Regarding the isoform protein levels, my preliminary quantification for GFP-PntP1, GFP-PntP2 and GFP-PntP3 all showed spikes of variations at photoreceptor fate transitions and such variations decreased across the AP axis. Studies have shown that variations in protein levels can have impact on cell fate decisions (Chang et al., 2008; Grün, 2020; Maamar et al., 2007). The source of the cell-to-cell variation in *pnt* isoform levels may be critical for further understanding the molecular mechanisms of the retinal cell specification. The source may include the production and dispersal of signaling ligands, signal transmission (e.g. kinetics of phosphorylation), and transcription of *pnt*. Incorporating randomness for different sources into the model to simulate the *pntP1, pntP2* and *pntP3* variation may reveal how different sources and the isoform network topology contribute and regulate the variations. Overall, the modeling will be a useful tool to carefully study each element of the Pnt network.

Regarding the mechanisms that confer cell-type specificity to Pnt network output, variation in signaling inputs and use of different network topologies may be critical. My work revealing different Pnt isoform expression patterns and regulatory interactions in the two rounds of photoreceptor specification provides a new entry point to explore this topic. Of particular interest is my discovery that PntP2 switches from auto-repressing its transcription in the first round of photoreceptor specification to auto-activating it in the second-round. The molecular mechanism that produces this switch is unknown. Two potential mechanisms to be investigated include the different levels of RTK signal that cells experience in the two rounds (discussed further in section 4.3) and the possibility that different cofactors associate with PntP2 at different stages of

ommatidial assembly. Cell-type specific cofactors include seven-up, which represses Prospero expression in R1, R6 cells, may be facilitate the *pntP2* regulatory switch. (Hayashi et al., 2008).

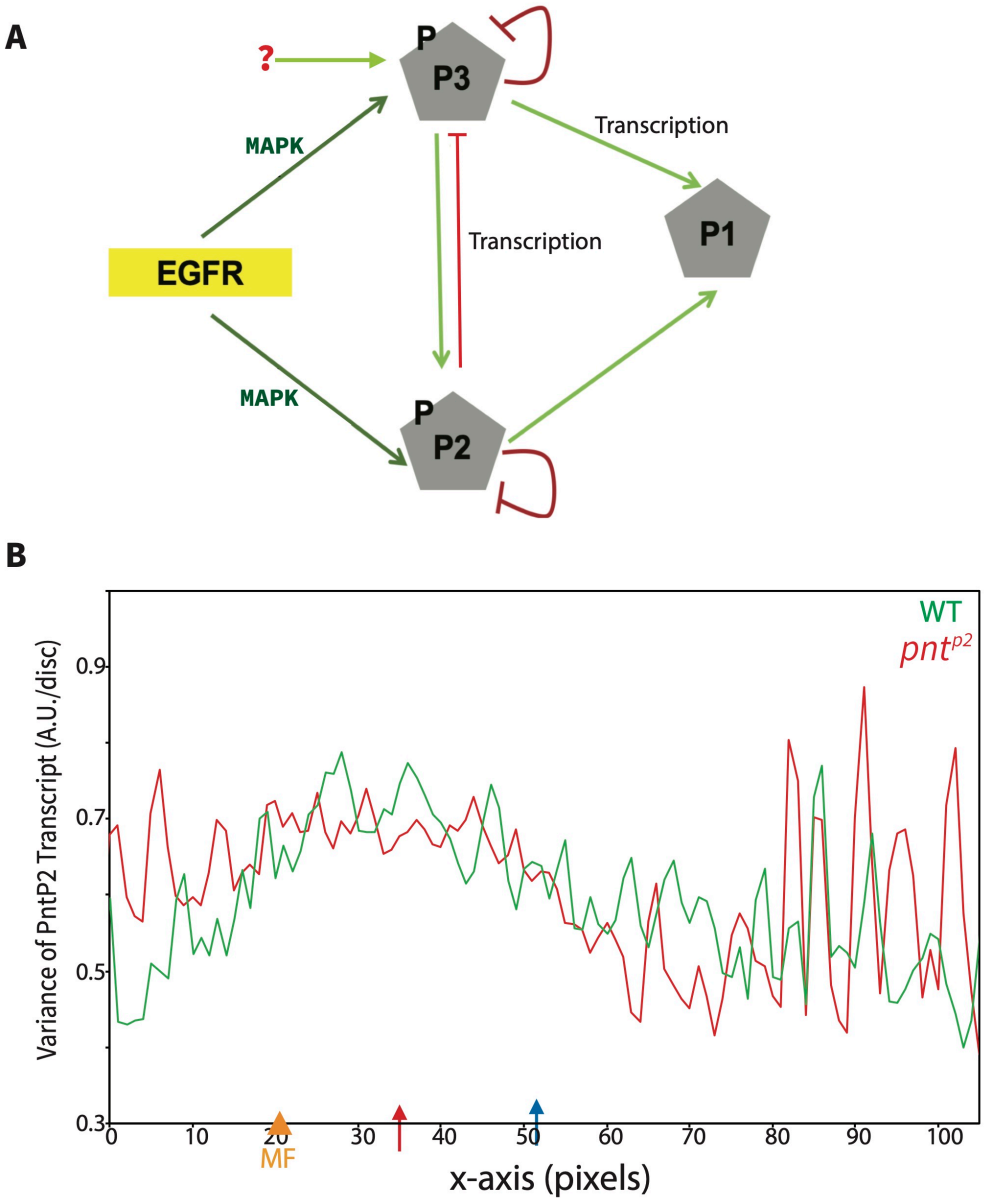


Fig 4.1 Potential Pnt network feedback topology and variance of pntP2 transcript across the AP axis in the eye disc

A) One potential feedback topology based on the current data: MAPK phosphorylation (dark green arrows) activates PntP2 and PntP3; PntP3 transcriptionally activates (light green arrows) *pntP2* while PntP2 transcriptionally represses *pntP3*; both PntP2 and PntP3 auto-repress their transcription and transcriptionally activate *pntP1*. Transcriptional activation of *pntP3* will be (Fig

4.1 continued) added based on the MS2-pntP3 reporter results. B) Variance of *pntP2* FISH in 6 wildtype (green) and 6 pnt^{p2} mutant (red) discs from maximum projections. In wild type, the variance rises anterior to the MF, peaks and then decreases to a steady state. Orange arrowheads mark the MF, red arrows mark the peak of *pntP2* expression and blue arrows mark the start of lower *pntP2* expression in the posterior half of the disc; the three can be mapped to correspondingly colored marks in Chapter 2. In pnt^{p2} discs, the variance levels are higher than normal anterior to the MF, then become similar to normal, but slightly lower in the posterior (right of blue arrow). The line connects the moving average of the coefficient of the variation of all foci within one-pixel windows along the x-axis.

4.2 Identification of transcriptional regulatory elements within the *pnt* locus

4.2.1 Discovery of potential regulatory elements for pnt isoforms

The distinct transcription patterns of *pnt* isoforms must reflect differential regulation through the multiple potential enhancer elements at *pnt* locus. Considering that the transcriptional starting sites of the isoforms are different, I suspect that the adjacent sequences may harbor isoform specific regulatory elements. *Pnt-GFP* ChIP-seq data from 3rd instar eye imaginal discs and stage 11 embryos (Webber et al., 2018 and unpublished by Jemma L. Webber) identified multiple Pnt bound peaks across the *pnt* locus, suggesting that the auto-regulation and cross-regulation of Pnt isoforms may be direct. Prominent peaks were identified near each of the isoform-specific transcription start sites, supporting my prediction. The published FAIRE-seq in 3rd instar eye discs (McKay and Lieb, 2013), indicating all potential regulatory regions, showed more peaks than *Pnt-GFP* ChIP-seq but included most of the Pnt-bound peaks. This makes that Pnt-bound regions within the *pnt* locus strong candidates for cis-regulatory elements responsible for control of *pnt* transcription.

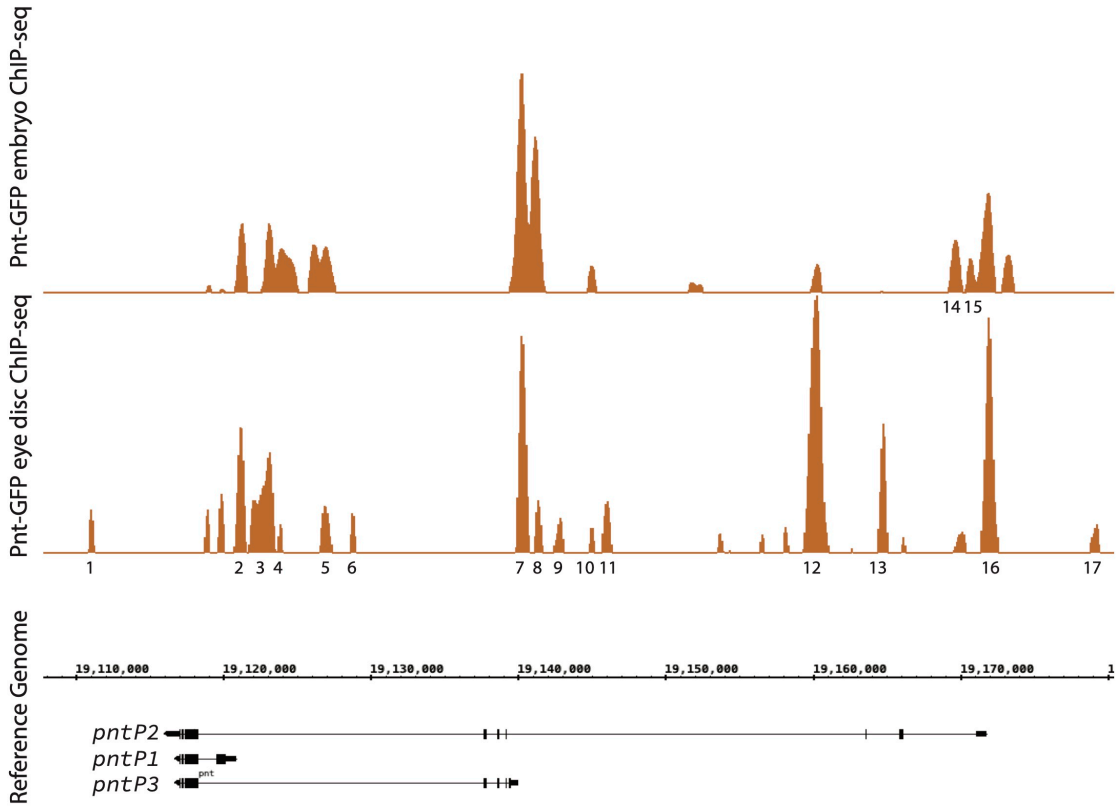


Fig 4.2 Potential cis-regulatory regions at the pnt locus

The top two panels are respective Pnt-bound peaks from Pnt-GFP embryo ChIP-seq and Pnt-GFP eye-antennal disc ChIP-seq; only the peaks above 80% of the highest peak value were picked (Webber et al., 2018 and unpublished by Jemma L. Webber). The bottom panel indicates the *pnt* locus. Peaks from the embryo and the eye ChIP-seq are largely overlapping. There are in total 17 peaks picked to construct reporters, as marked. The exact sequences of the 17 reporters are in appendix 5. Reporter 14 and 15 are identified based on the embryo ChIP-seq.

To validate whether any of the 17 Pnt-bound regions identified above the 80% peak height threshold (Fig 4.2) define a cis-regulatory module sufficient to drive isoform-specific patterns of expression in the 3rd instar eye imaginal disc, I cloned the regions and built transcriptional reporters for each one, and then integrated the transgenes into the 86Fb AttP landing site on the 3rd chromosome. Only the reporters showing expression in the 3rd instar retina can be retinal CRMs, and those regions should be prioritized for further research. Because CRMs can have

tissue-specific functions, examination of non-retinal tissues where Pnt is known to be expressed such as embryo, egg chamber and wing disc, will be important to mapping out the full regulatory complexity. My preliminary studies showed that reporter 7, 8 and 10 exhibited interesting expression patterns, that will be discussed in Appendix 4

4.2.2 In silico screen for TF binding sites in the pnt regulatory elements combined with genetic interaction assays to identify cell-type specific regulators of isoform-specific expression

CRMs frequently contain clusters of TF binding sites which organize interactions of multiple factors (Chapter 3). Thus the regulatory regions identified above may integrate Pnt isoforms and other factors for cell-type specific transcriptional regulation of *pnt*. Once the isoform specific CRMs are confirmed, we can search for other TF binding sites based on the known consensus sequences. I expect this will identify spatiotemporally restricted factors corresponding to the specific expression patterns of the isoforms. Then genetic assays can be implemented to validate the regulation. This *in silico*-based search may also suggest candidates for the cofactors that work with different Pnt isoforms to influence their DNA binding preferences; experiments to test these cofactors were discussed in section 4.1. From a broad perspective, the association between Pnt and the newly identified factors at specific CRMs may reveal cell-type specific cooperativity between RTK and other upstream signals. Although RTK/MAPK has long been identified as a generic inductive signal to activate cell fate transitions across development, the spatiotemporal specificity of the response is determined by its combination with other signals (Halfon et al., 2000; Halfon et al., 2002; Hayashi et al., 2008; Weber et al., 2008). The *pnt* CRMs can be useful tools to uncover the various combinatorial codes.

4.3 The modulation of Pnt transcriptional responses by signaling pathways

4.3.1 Control of the Pnt responses through RTK/MAPK activation

RTK/MAPK activation may control Pnt isoform network activity and the downstream transcriptional programs through modulating the pool of phosphorylated and unphosphorylated PntP2 and PntP3. My results suggested that PntP2 auto-repression and cross-repression of *pntP3* contributes regulation during the first round of photoreceptor specification but that in the second round, PntP2 switches to auto-activation. My hypothesis is that the different levels of RTK/MAPK activation in the two rounds determines the switch between repression and activation (Chapter 2). Briefly, RTK activation modulates the ratio of phosphorylated and unphosphorylated Pnt. For both PntP2 and PntP3, the unphosphorylated form has lower transactivation potential than the phosphorylated form. The predominance of unphosphorylated Pnt under conditions of low RTK activation will lead to overall low Pnt activity. If unphosphorylated Pnt binds an enhancer, but does not transactivate efficiently, the effect on gene expression will be repressive. To test this hypothesis, we can measure *pntP1*, *pntP2* and *pntP3* transcription with different adjustments of RTK signaling using loss- or gain-of-function alleles of RTK pathway genes such as *egfr, rl,* or *ras*. We can also use the identified *egfr* temperature sensitive mutant allele and adjust the signaling levels by shifting temperature. Using cell-type specific drivers to overexpress the dominant negative or dominant active forms of EGFR or Ras could modulate the signaling levels in a cell-type specific manner. If the hypothesis is correct, increasing RTK signaling levels may switch the auto- and cross-repression from PntP2 to activation at the anterior of the disc where first round fates are specified. On the other hand, reducing the RTK activation may switch the autoactivation of *pntP2* to autorepression at the posterior of the eye disc, where second round fates normally experience strong RTK activation

from two sources, the EGFR and Sev. Measuring *pntP2* transcript levels by FISH in a *sev* mutant background would test this.

The regulation from RTK activation may also rely on different enhancers having different preferences for the phosphorylated and unphosphorylated isoforms. For example, my finding that PntP2 activated *pntP1* while it auto-repressed and cross-repressed *pntP3* at the anterior, would suggest that the *pntP1* enhancer may preferentially recruit the phosphorylated PntP2 while the *pntP2* and *pntP3* primarily recruit unphosphorylated PntP2. How such preferences are determined is worth investigation.

The cis-regulatory syntax, as my research focused on in Chapter 3, is critical to organize different transcription factor inputs. To study the syntax organizing phosphorylated and unphosphorylated Pnt isoforms, the isoform specific CRMs must be identified first as I have discussed in section 4.2. As the phosphorylation site is next to the N-terminal SAM protein-protein interaction domain, the phosphorylation state may impact SAM-mediated interaction with other TFs and such cooperativity may facilitate the differential binding pattern of the phosphorylated- and unphosphorylated-Pnt isoforms. Therefore we should also mutate other TF sites to see if the cooperative interactions affect the binding preference. The binding preference for phosphorylated or unphosphorylated forms can be assessed by Phos-tag gels where a shift will be induced in electrophoretic mobility of phosphorylated proteins compared to their unphosphorylated counterparts (Nagy et al., 2018).

As another possibility, RTK activation may regulate the association of Pnt isoforms with specific cofactors. Previous work uncovered genetic interactions between Pnt and the transcriptional repressors Yan and Groucho (Webber et al., 2018). ChIP analyses showed that Pnt, Yan and

Groucho co-occupy significant number of enhancer regions and Pnt is required for stable occupancy of Yan and Groucho at these regions. Transcriptional repression of genes associated with co-occupied regulatory regions was reduced in *pnt* mutants, suggesting Pnt might mediate Yan and Groucho's repression. Therefore the auto- and cross-repression from PntP2 may be facilitated by these co-repressors. Supporting this idea, the Pnt bound peaks shown in Figure 4.2 overlapped with the peaks identified from the Yan ChIP-seq data in the eye imaginal discs (J. L.Webber, unpublished). As Yan and Groucho both respond to RTK signaling (Hasson et al., 2005; Rebay and Rubin, 1995), the Pnt-mediated repressive complex likely responds to the RTK signal. If the hypothesis about the cofactor association is correct, then mutants of Yan, Groucho or other potential co-repressors should disrupt the repressive function of PntP2. This could be assessed by measuring *pntP2* transcript levels in *yan;pnt*^{p2} or *pnt*^{p2}*,gro* double mutants.

4.3.2 Integration of RTK/MAPK and Notch signaling to regulate Pnt expression

The RTK/MAPK and Notch signaling pathways provide a great example to study how the integration of information from multiple signaling pathways regulates Pnt isoform-specific expression. The integration of RTK/MAPK and Notch inputs is required to specify R7 photoreceptors and cone cells (Cooper and Bray, 2000; Flores et al., 2000; Tomlinson and Struhl, 2001). For both *pros*, the R7 and cone cell specific gene, and *D-Pax2*, the cone cell specific gene, RTK activation is mediated by Pnt for proper expression (Charlton-Perkins et al., 2011; Flores et al., 2000; Hayashi et al., 2008; Tomlinson et al., 2011). I can think of two possibilities for Notch to integrate with RTK to regulate the target gene expression. Primarily, the downstream effectors of Notch, including Su(H), interact with Pnt through co-occupation at the same enhancers (Flores et al., 2000; Hayashi et al., 2008). Secondly, Notch can regulate the

expression of RTK transcription effectors, evidenced by the expression change of Yan and Pnt in *notch* mutant comparing to the wildtype (Bernasek et al., 2018). Interestingly, the change occurred at both anterior and posterior, where there are differential expression patterns of Pnt isoforms, suggesting Notch may regulate more than one isoform. According to the isoform specific expression patterns, Notch may regulate *pntP1* and/or *pntP3* at the anterior when R8 cells are specified, *pntP2* in R7 photoreceptors, and *pntP2* and/or *pntP3* at the posterior in cone cells differentially. To test these ideas, we can measure isoform specific expressions by FISH or transcriptional reporters in Notch mutant. Thus, in addition to RTK cascade, Notch signaling adds another layer of transcriptional regulation on Pnt isoforms.

4.4 Differential Pnt transcriptional responses in the embryo and wing

4.4.1 The role of MAPK activation and Pnt during embryogenesis

While the proposed ideas above focused mainly on the retina tissue, similar issues should be addressed in other developmental contexts, including the wing and embryo where RTK/MAPK/Pnt activation plays critical roles. Such studies would provide a broader picture of how isoform-specific Pnt expression and activity is regulated in response to different signaling conditions. It was in the embryo where *pntP1* and *pntP2* were originally identified, and their differential transcript distributions were illustrated (Klambt, 1993). Then following studies uncovered major roles of Pnt during ventral nerve cord development (Estes et al., 2008; Giesen et al., 1997; Granderath and Klämbt, 1999; Granderath et al., 2000; Klaes et al., 1994; Scholz et al., 1997), and cardiac cell specification at stage 11 embryo (Boisclair Lachance et al., 2018; Carmena et al., 1998; Carmena et al., 2002; Halfon et al., 2000; Knirr and Frasch, 2001; Webber et al., 2013). The functions of Pnt isoforms, particularly that of the newly identified PntP3,

during embryo development is worth exploration. Just like *pntP1* and *pntP2*, GFP-PntP3 showed strong expression in the embryo across developmental stages (Jean-François Boisclair Lachance, unpublished), suggesting its potential functions. Considering that the homozygous pnt^{p3} mutant survived normally, the function of PntP3 may be redundant with either PntP2 or PntP1. The survival rate of pnt^{p2} mutant animals is higher than that of pnt^{p2p3} mutants, suggesting that PntP3 may be redundant with PntP2 (Chapter 2). However more detailed analysis for the embryonic development for the mutants will be essential to confirm the idea. Furthermore, based on my research in the retina, modulating RTK signaling levels may reveal PntP3's role if PntP3 mainly contributes to the robustness of embryogenesis. For example, I supplied repeated 18°C to 31°C temperature shifts at the 3rd instar larvae stage in a sensitized *egfr/rl* background to investigate photoreceptor development defects. To study embryogenesis, similar treatment could be supplied for embryo developmental stages.

If we want to focus on a specific gene regulation in the embryo, *eve* at stage 11will be a good choice. Its expression regulation by RTK/Pnt signaling has been well characterized and the CRMs have been identified. GFP-PntP3 had expression in the stage 11 mesoderm (Jean-François Boisclair Lachance, unpublished), where the Eve positive cells originate, and thus PntP3 has the potential to interact with the CRMs and regulate *eve* expression. After confirming expression of the isoforms in this context, Genetic interaction assays can be carried out between *eve* alleles carrying deletions of specific CRMs and my *pnt* isoform-specific crispr mutants. Then using *eve* expression as a read out, we might find synergistic reduction of Eve, indicating a particular isoform's preference for the CRM. Studying how differences in regulatory syntax organizes/influences the isoforms at the MHE is also a good experiment. I expect that introducing Pnt isoform specific mutants will result in different changes in expression of the

MHE ETS site mutant reporters (see Chapter 3 and Boisclair-Lachance et al 2018) if PntP1, PntP2 and PntP3 have different preferences for the ETS binding sites.

4.4.2 The role of MAPK activation and Pnt during wing development

The wing is another context where multiple roles of RTK/MAPK/Pnt activation have been revealed. Detailed analysis and discussion will be in Appendix 2.

4.5 References

Bernasek, S. M., Lachance, J.-F. B., Peláez, N., Bakker, R., Navarro, H. T., Amaral, L. A. N., Bagheri, N., Rebay, I. and Carthew, R. W. (2018). Ratio-based sensing of two transcription factors regulates the transit to differentiation. *bioRxiv* 430744.

Boisclair Lachance, J.-F., Webber, J. L., Hong, L., Dinner, A. R. and Rebay, I. (2018). Cooperative recruitment of Yan via a high-affinity ETS supersite organizes repression to confer specificity and robustness to cardiac cell fate specification. *Genes Dev.* **32**, 389–401.

Brunner, D., Dücker, K., Oellers, N., Hafen, E., Scholz, H. and Klämbt, C. (1994). The ETS domain protein pointed-P2 is a target of MAP kinase in the sevenless signal transduction pathway. *Nature* **370**, 386–389.

Carmena, A., Gisselbrecht, S., Harrison, J., Jiménez, F. and Michelson, A. M. (1998). Combinatorial signaling codes for the progressive determination of cell fates in the Drosophila embryonic mesoderm. *Genes Dev.* **12**, 3910–3922.

Carmena, A., Buff, E., Halfon, M. S., Gisselbrecht, S., Jiménez, F., Baylies, M. K. and Michelson, A. M. (2002). Reciprocal regulatory interactions between the Notch and Ras signaling pathways in the Drosophila embryonic mesoderm. *Dev. Biol.* **244**, 226–242.

Chang, H. H., Hemberg, M., Barahona, M., Ingber, D. E. and Huang, S. (2008). Transcriptome-wide noise controls lineage choice in mammalian progenitor cells. *Nature* **453**, 544–547.

Charlton-Perkins, M., Whitaker, S. L., Fei, Y., Xie, B., Li-Kroeger, D., Gebelein, B. and Cook, T. (2011). Prospero and Pax2 combinatorially control neural cell fate decisions by modulating Ras- and Notch-dependent signaling. *Neural Dev.* **6**, 20.

Cooper, M. T. D. and Bray, S. J. (2000). R7 photoreceptor specification requires Notch activity. *Curr. Biol.* **10**, 1507–1510.

Dublanche, Y., Michalodimitrakis, K., Kümmerer, N., Foglierini, M. and Serrano, L. (2006). Noise in transcription negative feedback loops: simulation and experimental analysis. *Mol. Syst. Biol.* **2**, 41.

Estes, P., Fulkerson, E. and Zhang, Y. (2008). Identification of motifs that are conserved in 12 Drosophila species and regulate midline glia vs. neuron expression. *Genetics* **178**, 787–799.

Flores, G. V., Duan, H., Yan, H., Nagaraj, R., Fu, W., Zou, Y., Noll, M. and Banerjee, U. (2000). Combinatorial signaling in the specification of unique cell fates. *Cell* **103**, 75–85.

Garcia, H. G., Tikhonov, M., Lin, A. and Gregor, T. (2013). Quantitative Imaging of Transcription in Living Drosophila Embryos Links Polymerase Activity to Patterning. *Curr. Biol.* **23**, 2140–2145.

Giesen, K., Hummel, T., Stollewerk, A., Harrison, S., Travers, A. and Klambt, C. (1997). Glial development in the Drosophila CNS requires concomitant activation of glial and repression of neuronal differentiation genes. *Development* **124**, 2307 LP – 2316.

Granderath, S. and Klämbt, C. (1999). Glia development in the embryonic CNS of

Drosophila. *Curr. Opin. Neurobiol.* **9**, 531—536.

Granderath, S., Bunse, I. and Klämbt, C. (2000). gcm and pointed synergistically control glial transcription of the Drosophila gene loco. *Mech. Dev.* **91**, 197–208.

Grün, D. (2020). Revealing dynamics of gene expression variability in cell state space. *Nat. Methods* **17**, 45–49.

Guinn, M. T. and Balázsi, G. (2019). Noise-reducing optogenetic negative-feedback gene circuits in human cells. *Nucleic Acids Res.* **47**, 7703–7714.

Halfon, M. S., Carmena, A., Gisselbrecht, S., Sackerson, C. M., Jiménez, F., Baylies, M. K. and Michelson, A. M. (2000). Ras pathway specificity is determined by the integration of multiple signal-activated and tissue-restricted transcription factors. *Cell* **103**, 63–74.

Halfon, M. S., Grad, Y., Church, G. M. and Michelson, A. M. (2002). Computation-based discovery of related transcriptional regulatory modules and motifs using a experimentally validated combinatorial model. *Genome Res.* **12**, 1019–1028.

Hasson, P., Egoz, N., Winkler, C., Volohonsky, G., Jia, S., Dinur, T., Volk, T., Courey, A. J. and Paroush, Z. (2005). EGFR signaling attenuates Groucho-dependent repression to antagonize Notch transcriptional output. *Nat. Genet.* **37**, 101–105.

Hayashi, T., Xu, C. and Carthew, R. W. (2008). Cell-type-specific transcription of prospero is controlled by combinatorial signaling in the Drosophila eye. *Development* **135**, 2787–2796.

Heberlein, U. and Mosest, K. (1995). Mechanisms of drosophila retinal morphogenesis: The virtues of being progressive. *Cell* **81**, 987–990.

Heberlein, U., Wolff, T. and Rubin, G. M. (1993). The TGFβ homolog dpp and the segment polarity gene hedgehog are required for

propagation of a morphogenetic wave in the Drosophila retina. *Cell* **75**, 913–926.

Klaes, A., Menne, T., Stollewerk, A., Scholz, H. and Klämbt, C. (1994). The ets transcription factor encoded by the drosophila gene pointed direct glial cell differentiation in the embryonnic CNS. *Cell* **78**, 149–160.

Klambt, C. (1993). The Drosophila gene pointed encodes two ETS-like proteins which are involved in the development of the midline glial cells. *Development* **117**, 163–176.

Knirr, S. and Frasch, M. (2001). Molecular integration of inductive and mesoderm-intrinsic inputs governs even-skipped enhancer activity in a subset of pericardial and dorsal muscle progenitors. *Dev. Biol.* **238**, 13–26.

Maamar, H., Raj, A. and Dubnau, D. (2007). Noise in Gene Expression Determines Cell Fate in Bacillus subtilis *Science (80-.).* **317**, 526 LP – 529.

McKay, D. J. and Lieb, J. D. (2013). A common set of DNA regulatory elements shapes Drosophila appendages. *Dev. Cell* **27**, 306–318.

Nagy, Z., Comer, S. and Smolenski, A. (2018). Analysis of Protein Phosphorylation Using Phos-Tag Gels. *Curr. Protoc. Protein Sci.* **93**, e64.

Rebay, I. and Rubin, G. M. (1995). Yan functions as a general inhibitor of differentiation and is negatively regulated by activation of the Ras1/MAPK pathway. *Cell* **81**, 857–866.

Scholz, H., Sadlowski, E., Klaes, A. and Klämbt, C. (1997). Control of midline glia development in the embryonic Drosophila CNS. *Mech. Dev.* **62**, 79–91.

Singh, A. and Hespanha, J. P. (2009). Optimal feedback strength for noise suppression in autoregulatory gene networks. *Biophys. J.* **96**, 4013–4023.

Tomlinson, A. and Struhl, G. (2001). Delta/Notch and Boss/Sevenless Signals Act

Combinatorially to Specify the Drosophila R7 Photoreceptor. *Mol. Cell* **7**, 487–495.

Tomlinson, A., Mavromatakis, Y. E. and Struhl, G. (2011). Three distinct roles for Notch in Drosophila R7 photoreceptor specification. *PLoS Biol.* **9**,.

Webber, J. L., Zhang, J., Mitchell-Dick, A. and Rebay, I. (2013). 3D chromatin interactions organize Yan chromatin occupancy and repression at the even-skipped locus. *Genes Dev.* **27**, 2293–2298.

Webber, J. L., Zhang, J., Massey, A., Sanchez-Luege, N. and Rebay, I. (2018). Collaborative repressive action of the antagonistic ETS transcription factors Pointed and Yan fine-tunes gene expression to confer robustness in Drosophila. *Dev.* **145**,.

Weber, U., Pataki, C., Mihaly, J. and Mlodzik, M. (2008). Combinatorial signaling by the Frizzled/PCP and Egfr pathways during planar cell polarity establishment in the Drosophila eye. *Dev. Biol.* **316**, 110–123.

APPENDIX1: Quantification of GFP-PntP1, GFP-PntP2, and GFP-PntP3 during photoreceptor specification in 3rd instar eye disc

A1.1 Results and Discussion

The expression dynamics of Pnt isoforms may enable the transition to distinct photoreceptor fates. Our model demonstrated that context-dependent Pnt networks determine the first and the second round photoreceptor fates (Chapter 2). If Pnt isoforms also function differently to determine the cell fates within the same round, such as R2R5 vs. R3R4, then differences in expression dynamics might be evident. Interestingly, differential expression dynamics of Yan and Pnt (all isoforms together) have been revealed in distinct R1-R8 photoreceptor cells (Bernasek et al., 2018; Peláez et al., 2015). Thus, to understand Pnt isoform in individual photoreceptor types, I studied the expression dynamics of Pnt isoforms in each photoreceptor type.

Using genomic BAC transgenes in which GFP tags had been added at the unique N-terminus, I showed that the three Pnt isoforms have differential expression dynamics during photoreceptor fate transitions. I used DAPI to label all eye cell nuclei, and then based on the anterior-posterior and apical-basal position, and morphology I identified all the individual cell types including progenitors, photoreceptors and cone cells (details in A1.2 Materials and methods, (Peláez et al., 2015; Tomlinson, 1985; Tomlinson and Ready, 1987; Wolff and Ready, 1993)). The cells were also mapped to a point of developmental time considering that the MF moves from posterior to anterior to initiate one array of R8 cells every two hours (Basler and Hafen, 1991). Then GFP-PntP1, -PntP2 or -PntP3 fluorescence levels within individual nuclei was normalized to DAPI to control for measurement variation. The normalized values represent the expression of GFP-

PntP1, GFP-PntP2 or GFP-PntP3 within individual cells. Focusing on specific cell types, these data provide a dynamic view of Pnt isoform expression through cell fate transitioning processes. GFP-PntP1, -PntP2 and -PntP3 expression initiated in two groups of progenitor cells: PntP1 and PntP3 reached peak expression in the earlier group, which is close to MF, while PntP2 reached peak expression in the later group (Fig. A1.1). For photoreceptors, consistent with my characterization of the expression of *pntP2* and PntP3 in Chapter 2, GFP-PntP3 expressed in the first round of photoreceptors, R2R5 pairs and R3R4 pairs, while GFP-PntP2 expressed in both the first and the second round photoreceptor fates. GFP-PntP1 expressed in all photoreceptor cell types. All isoform expression in the photoreceptors showed similar trends across developmental time such that expression levels were highest at the beginning of fate transitions, sharply decreased over time until reaching a low and steady level at the latest developmental time points. However, there were differences between distinct cell types. For GFP-PntP3, it reached the highest expression in R8 cells, lower in R2R5 pairs and lowest in R3R4 pairs (Fig. A1.1C); for GFP-PntP2, the expression was the highest in R7 cells, lower in R3R4 and R1R6 pairs and lowest in R2R5 pairs and R8 cells (Fig. A1.1B); for GFP-PntP1, the expression was the highest in R2R5 pairs, lower in R8 cells, even lower in R3R4 pairs, then R7 cells and the lowest in R1R6 pairs (Fig. A1.1A). The differences among the photoreceptors suggest that on top of the Pnt networks for the two rounds of photoreceptor fates, there may be distinct Pnt isoform mediated regulation for individual photoreceptor fates.

Based on the differential expression dynamics, further exploration of the contribution of the dynamics to distinct photoreceptor fates will be important. A primary question is to understand if the detected PntP2 and PntP3 directly activate the transcription of photoreceptor specification genes or whether they only transcriptionally activate ***PntP1***. Favoring the latter mechanism, the

sequential model that PntP2 and PntP3 transcriptionally activate PntP1 from Chapter 2 predicts that the photoreceptor fate decisions directly associate with PntP1 dynamics.

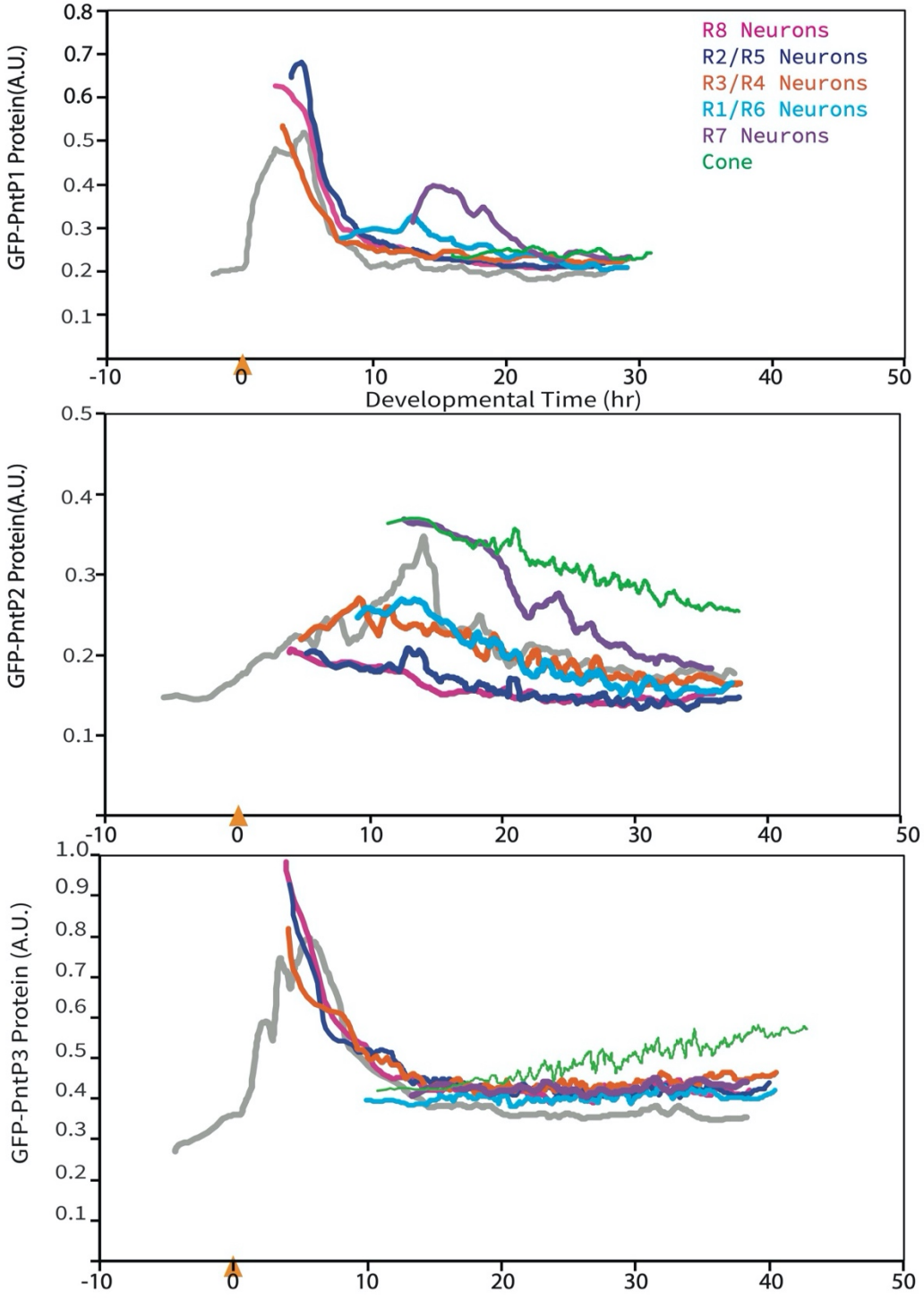


Fig A 1.1 Quantification of GFP-PntP1, GFP-PntP2 and GFP-PntP3 in multipotent cells, R1-R8 and cone cells

(Fig A 1.1 continued) Moving averages of GFP-PntP1 **(A)** and GFP-PntP3 **(C)** levels highlight the peaks of expression at the MF (yellow arrowhead) region in the multipotent cells that will be specified into the first five photoreceptors recruited to each ommatidium, R8, R2, R5, R3 and R4. GFP-PntP2 **(B)** levels reach the peak of expression in the multipotent cells that will be specified into the second round of photoreceptors, R1, R6 and R7. GFP-PntP1 expresses in R1-R8, with relative lower expression in R1, R6 and R7. GFP-PntP2 expresses in R1-R8, with relative higher expression in R1, R6 and R7. GFP-PntP3 expresses in the first five photoreceptors but its expression in photoreceptors, R1, R6, and R7 was not above baseline. Only GFP-PntP2 and GFP-PntP3 express in cone cells but with reverse trends along the developmental time. Data plotted are from two discs of each genotype from independent experiments.

Therefore, within the first round of specification, R2R5 pairs may require higher levels of PntP1 compared to R3R4 pairs. This higher PntP1 may be transcriptionally activated by the higher level of PntP3 in R2R5 pairs. Interestingly, the transcriptional repressor Yan showed lower expression in R2R5 pairs than that in R3R4 pairs (Peláez et al., 2015). Adding the information of expression dynamics to the conventional model that High Pnt/Low Yan specify photoreceptor fates, R2R5 pairs may require a higher PntP1/Yan ratio to be specified. If the detected PntP2 and PntP3, together with PntP1 directly activate the transcription of photoreceptor fate genes, the sum of the transcriptional activities of three isoforms determines the fates. PntP1 plus PntP3 was still higher in R2R5 pairs compared to R3R4 but PntP2 showed lower expression in R2R5 pairs. One important caveat to be considered was that unlike GFP-PntP1 and GFP-PntP3, the GFP-PntP2 BAC transgene did not fully complement pnt^{null} flies, suggesting the GFP tag may interfere with the dynamics of PntP2. The fact that GFP-PntP2 peaks were much lower than the peaks of GFP-PntP1 and GFP-PntP3 might be a reflection of interfered protein dynamics. Other fluorescent tags or insertion locations may be introduced to monitor endogenous PntP2 dynamics. Once PntP2 expression dynamics is more accurately measured, the association between the total Pnt transcriptional activity and the photoreceptor fates can be better proposed.

The expression dynamics of Pnt isoforms in cone cells are different from that in all photoreceptor cells. Consistent with the characterization from Chapter2, GFP-PntP2 and GFP-PntP3 both expressed in cone cells. However, unlike the decrease of PntP2 along the developmental time axis, GFP-PntP3 increased in cone cells (Fig. A1.1BC). There was no expression of GFP-PntP1 detected in cone cells (Fig. A1.1A). The opposite expression dynamics of PntP2 and PntP3 in cone cells suggests another working model, in addition to the redundancy, between them. As I have shown that pnt^{p2} mutants lost some cone cells while *pntp3* mutants showed normally specified cone cells (Chapter 2), my hypothesis is that PntP2 is the major determinant for cone cell fates while PntP3 is only partially redundant with PntP2 for cone cell fate specification but plays an additional later role for cone cell functionality.

A1.2 Materials and methods

Immunohistochemistry and microscopy

Third instar eye-antennal imaginal discs were dissected in cold S2 cell medium, fixed in 4% PFA/PBS for 30 minutes at room temperature, washed 3X in PBST (1X PBS, 0.1% Triton), permeabilized in PBST (1X PBS, 0.5% Triton) for 20 min, washed 3X in PBS, incubated with DAPI (0.1ug/ml DAPI in PBST (0.1% Triton)) for 20 min at room temperature, washed 1X in PBS and mounted immediately. Imaging was performed within the same day of fixation with a Zeiss LSM 880 confocal microscope, using 0.8 μm steps for a stack of ~60 optical sections.

Quantification of expression levels

Expression dynamics were inferred from confocal image stacks using an updated version of an existing segmentation and annotation pipeline (Bernasek et al., 2018; Peláez et al., 2015). The new pipeline includes FlyEye Silhouette (Nicelio Sanchez-Luege et al. unpublished); an open-source package for macOS that integrates our image segmentation algorithm with a GUI for cell

type annotation. Cell segmentation was performed with DAPI signals for identification of cell nuclei and the mean GFP fluorescence signal was normalized to the mean DAPI signal of each nucleus. Cell positions along the anterior-posterior axis were mapped to developmental time according to the previous description (Bernasek et al., 2018; Peláez et al., 2015).

A1.3 References

Basler, K. and Hafen, E. (1991). Specification of Cell Fate in the Developing Eye of Drosophila. *Bioessays* **13**, 621–631.

Bernasek, S. M., Lachance, J.-F. B., Peláez, N., Bakker, R., Navarro, H. T., Amaral, L. A. N., Bagheri, N., Rebay, I. and Carthew, R. W. (2018). Ratio-based sensing of two transcription factors regulates the transit to differentiation. *bioRxiv* 430744.

Jean-François Boisclair Lachance, Nicolás Peláez, Justin J. Cassidy, J. L. and Webber, Ilaria Rebay, and R. W. C. (2014). A comparative study of Pointed and Yan expression reveals new complexity to the transcriptional networks downstream of receptor tyrosine kinase signaling. *Dev Biol.* **385**, 263–278.

Mavromatakis, Y. E. and Tomlinson, A. (2013). Switching cell fates in the developing Drosophila eye. *Development* **140**, 4353–61.

Pallavi, S. K. (2003). Egfr/Ras pathway mediates interactions between peripodial and disc proper cells in Drosophila wing discs. *Development* **130**, 4931–4941.

Paul, L., Wang, S., Manivannan, S. N., Bonanno, L., Lewis, S., Austin, C. L. and Simcox, A. (2013). Dpp-induced Egfr signaling triggers postembryonic wing development in Drosophila. *Proc. Natl. Acad. Sci.* **110**, 5058–63.

Peláez, N., Gavalda-Miralles, A., Wang, B., Navarro, H. T., Gudjonson, H., Rebay, I., Dinner, A. R., Katsaggelos, A. K., Amaral, L. A. N. and Carthew, R. W. (2015). Dynamics and heterogeneity of a fate determinant during transition towards cell differentiation. *Elife* **4**,.

Scholz, H., Deatrick, J., Klaes, A. and Klambt, C. (1993). Genetic dissection of pointed, a Drosophila gene encoding two ETS-related proteins. *Genetics* **135**, 455–468.

Tomlinson, a (1985). The cellular dynamics of pattern formation in the eye of Drosophila. *J. Embryol. Exp. Morphol.* **89**, 313–31.

Tomlinson, A. and Ready, D. F. (1987). Neuronal differentiation in the Drosophila ommatidium. *Dev. Biol.* **120**, 366–376.

Wang, S. H., Simcox, A. and Campbell, G. (2000). Dual role for Drosophila epidermal growth factor receptor signaling in early wing disc development. *Genes Dev.* **14**, 2271–2276.

Wolff, T. and Ready, D. F. (1993). Pattern formation in the Drosophila retina. *Cold Spring Harb. Lab. Press* **2**, 1277–1325.

Zecca, M. and Struhl, G. (2002). Control of growth and patterning of the Drosophila wing imaginal disc by EGFR-mediated signaling. *Development* **129**, 1369–1376.

APPENDIX2: PntP2 and PntP3 in 3rd instar wing disc

A2.1 Results and Discussion

In addition to the eye and the embryo, the wing is another context where multiple roles of RTK/MAPK/Pnt activation have been revealed. PntP2 has been the major isoform implicated in EGFR-mediated regulation of wing disc patterning (Paul et al., 2013; Scholz et al., 1993). The EGFR signaling initially specifies the dorsal compartment and then as the disc grows, continued strong signaling specifies the notum, restricting development of the wing proper to the rest of the epithelium (Pallavi, 2003; Paul et al., 2013; Wang et al., 2000; Zecca and Struhl, 2002). Reflecting these roles, partially duplicated wing pouches were reported in hypomorphic pntP2-specific allelic combinations (Scholz et al., 1993). As predicted by this prior work, my homozygous pnt^{p2} null mutant adult escapers had duplicated and malformed wings (Fig A2.1 B, green line marks the wing tissue); examination of 3rd instar wing discs revealed duplication of the wing pouch and reduction of notum tissue (Fig A2.1 A, C, yellow arrow). In contrast, the pnt^{p3} null mutant wings and wing discs were normally patterned (Fig A2.1 D and not shown). This together with the results from my Chapter 2 that *egfr/rl*; pnt^{p3} flies showed wing margin defects suggest that the role of PntP3 during wing development is redundant with PntP2 or PntP1.

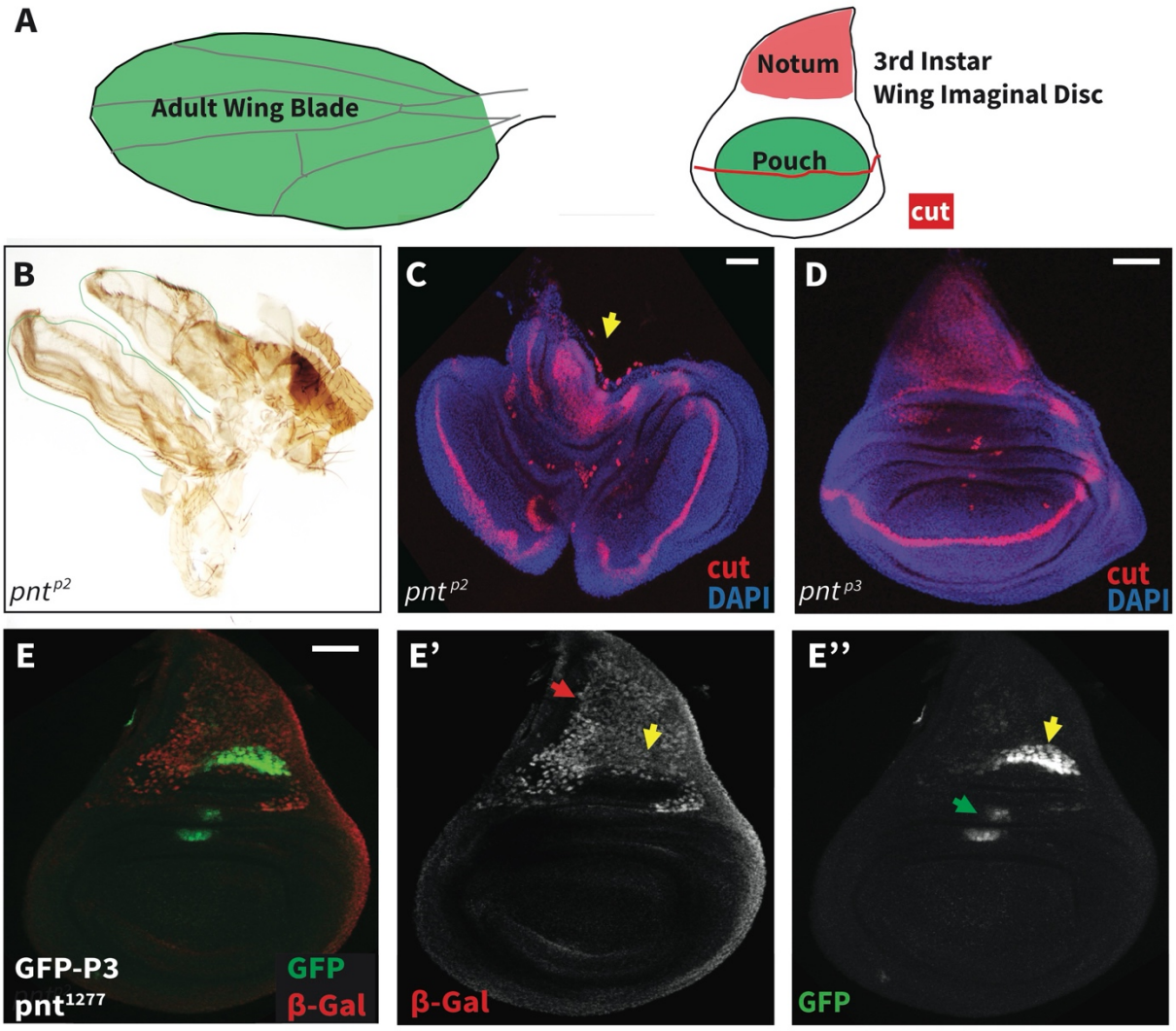


Fig A2.1 Expression and functions of PntP2 and PntP3 during wing patterning

Adult wings are oriented anterior up and distal left. Maximum projections of 3rd instar wing imaginal discs are oriented dorsal up. **(A)** A schematic depicting the adult wing blade that develops from the pouch region (green) of the wing imaginal disc. Cut expression (red) marks the prospective wing margin. **(B-C)** Loss of *pnt*^{p2} results in duplication of wing material. (B) A duplicated and malformed wing dissected from an adult (B) or 3rd instar (C) *pnt*^{p2} escaper. Cut expression (red) highlights the wing pouch duplication and reduced notum (yellow arrow) of *pnt*^{p2} discs (C) and the wild type patterning of *pnt*^{p3} discs (D). DAPI (blue) marks all nuclei. Scale bar: 50 µm.**(E)** β-gal expression from the *pnt*^{1277} (red) is seen throughout the notum (red arrow) and overlaps GFP-PntP3 (green) in the posterior compartment (yellow arrow). PntP3 is also expressed in the dorsal hinge region (green arrow). No expression of either isoform was detected in the pouch. Scale bar: 50 µm.

To further study the role of the isoforms in the wing, their expression patterns should be understood first. I compared the expression patterns of PntP2 and PntP3 in the 3rd instar wing disc. As in the eye, both overlapping and non-overlapping expression domains were detected (Fig A2.1 E). Thus both isoforms were expressed in a cluster of cells in the posterior compartment of the notum (Fig A2.1 E', E'', yellow arrow), cells elsewhere in the notum expressed primarily *pntP2* (FigA 2.1 E', red arrow), cell clusters in the dorsal hinge region expressed primarily GFP-PntP3 (FigA 2.1 E'', green arrow). The expression in the notum and hinge is consistent with the expression of Pnt-GFP, for which the GFP tagged the C-terminus of all three Pnt isoforms. This expression pattern suggested redundant functions of PntP2 and PntP3 for notum and hinge development, that can be further explored. However neither PntP2 nor PntP3 showed strong expression in the wing pouch although Pnt-GFP had cross-like pattern of stripes of expression in the wing pouch (Boisclair Lachance et al., 2014). Therefore it seems that PntP1 is the only isoform left that can contribute to the wing pouch expression of Pnt-GFP, but my preliminary results showed that GFP-PntP1 is largely overlapping with GFP-PntP3, expressing in the notum and hinge (data not shown). To solve this confusion, we should be aware of the caveat that the GFP-PntP2 and the *pntP2-lacZ* reporters are both hypomorphic *pntP2* alleles, so that they might not capture the wildtype endogenous PntP2 expression. New PntP2 protein and transcription reporters need to be built. Once a wildtype PntP2 reporter is constructed, we should consider expand the expression analysis to the 2nd instar larvae wing disc, considering that EGFR initiated the dorsal development during the 2nd instar.

Functionally, the synergic effect between PntP2 and PntP3 is worthy of further exploration. I showed that *egfr/rl; pnt^{p3}* and *egfr/rl; pnt^{p2}/pnt^{p3}* exhibited much higher penetrance of wing margin defects than the penetrance in the single heterozygotes (Chapter 2), suggesting strong

synergy between *pntP2* and *pntP3*. The synergistic function should occur during the development of pouch, which will grow into the wing blade. Given I didn't see co-expression of PntP2 and PntP3 at the 3rd instar, further research should focus on earlier time points. This synergy effect may be manifested in the cross-regulatory network of the isoforms, which can exploit the mathematical modeling from discussion session 4.1 for better understanding.

A2.2 Materials and methods

Immunohistochemistry and microscopy

Third instar wing discs were dissected in cold S2 cell medium, fixed for 10 min in 4% PFA with 0.1% Triton X-100, washed 3X in PBT (1X PBS, 0.1% Triton), blocked in PNT (1X PBS, 0.1% Triton, 1% normal goat serum), stained with primary antibodies in PNT overnight at 4° C, washed 3X in PBT, and stained with secondary antibodies in PNT overnight at 4° C. Primary antibodies used were: rabbit α-GFP (1:2000, Molecular Probes); mouse cut (1:50, DSHB, 2B10); mouse α-β-galactosidase (1:1000, Promega). Secondary antibodies were from Jackson ImmunoResearch: donkey α-rabbit-Cy3 (1:2000), donkey α-rabbit-488 (1:2000), donkey α-mouse-Cy3 (1:2000. DAPI (1:2000, Invitrogen) was used to detect DNA. Imaging was performed with a Zeiss LSM 880 confocal microscope, using 1.2-1.5 μm steps. Adult wing was imaged with a Canon EOS Rebel camera fitted to a Leica stereo microscope. Individual slices were merged using iSolution-Lite software (IMT-Digital).

A2.3 References

Basler, K. and Hafen, E. (1991). Specification of Cell Fate in the Developing Eye of Drosophila. *Bioessays* **13**, 621–631.

Bernasek, S. M., Lachance, J.-F. B., Peláez, N., Bakker, R., Navarro, H. T., Amaral, L. A. N., Bagheri, N., Rebay, I. and Carthew, R. W. (2018). Ratio-based sensing of two transcription factors regulates the transit to differentiation. *bioRxiv* 430744.

Jean-François Boisclair Lachance, Nicolás Peláez, Justin J. Cassidy, J. L. and Webber, Ilaria Rebay, and R. W. C. (2014). A comparative study of Pointed and Yan expression reveals new complexity to the transcriptional networks downstream of receptor tyrosine kinase signaling. *Dev Biol.* **385**, 263–278.

Mavromatakis, Y. E. and Tomlinson, A. (2013). Switching cell fates in the developing Drosophila eye. *Development* **140**, 4353–61.

Pallavi, S. K. (2003). Egfr/Ras pathway mediates interactions between peripodial and disc

proper cells in Drosophila wing discs. *Development* **130**, 4931–4941.

Paul, L., Wang, S., Manivannan, S. N., Bonanno, L., Lewis, S., Austin, C. L. and Simcox, A. (2013). Dpp-induced Egfr signaling triggers postembryonic wing development in Drosophila. *Proc. Natl. Acad. Sci.* **110**, 5058–63.

Peláez, N., Gavalda-Miralles, A., Wang, B., Navarro, H. T., Gudjonson, H., Rebay, I., Dinner, A. R., Katsaggelos, A. K., Amaral, L. A. N. and Carthew, R. W. (2015). Dynamics and heterogeneity of a fate determinant during transition towards cell differentiation. *Elife* **4**,.

Scholz, H., Deatrick, J., Klaes, A. and Klambt, C. (1993). Genetic dissection of pointed, a Drosophila gene encoding two ETS-related proteins. *Genetics* **135**, 455–468.

Tomlinson, a (1985). The cellular dynamics of pattern formation in the eye of Drosophila. *J. Embryol. Exp. Morphol.* **89**, 313–31.

Tomlinson, A. and Ready, D. F. (1987). Neuronal differentiation in the Drosophila ommatidium. *Dev. Biol.* **120**, 366–376.

Wang, S. H., Simcox, A. and Campbell, G. (2000). Dual role for Drosophila epidermal growth factor receptor signaling in early wing disc development. *Genes Dev.* **14**, 2271–2276.

Wolff, T. and Ready, D. F. (1993). Pattern formation in the Drosophila retina. *Cold Spring Harb. Lab. Press* **2**, 1277–1325.

Zecca, M. and Struhl, G. (2002). Control of growth and patterning of the Drosophila wing imaginal disc by EGFR-mediated signaling. *Development* **129**, 1369–1376.

APPENDIX3: Fate switched experiments: GFP-PntP2 expression associated with R7 fates

A3.1 Results and discussion

I further confirmed that GFP-PntP2 expression associated with R7 and cone cell fates by the fate switched experiments (Mavromatakis and Tomlinson, 2013). I found that GFP-PntP2 expressed in ectopically induced R7 and cone cells, which were Pros positive.

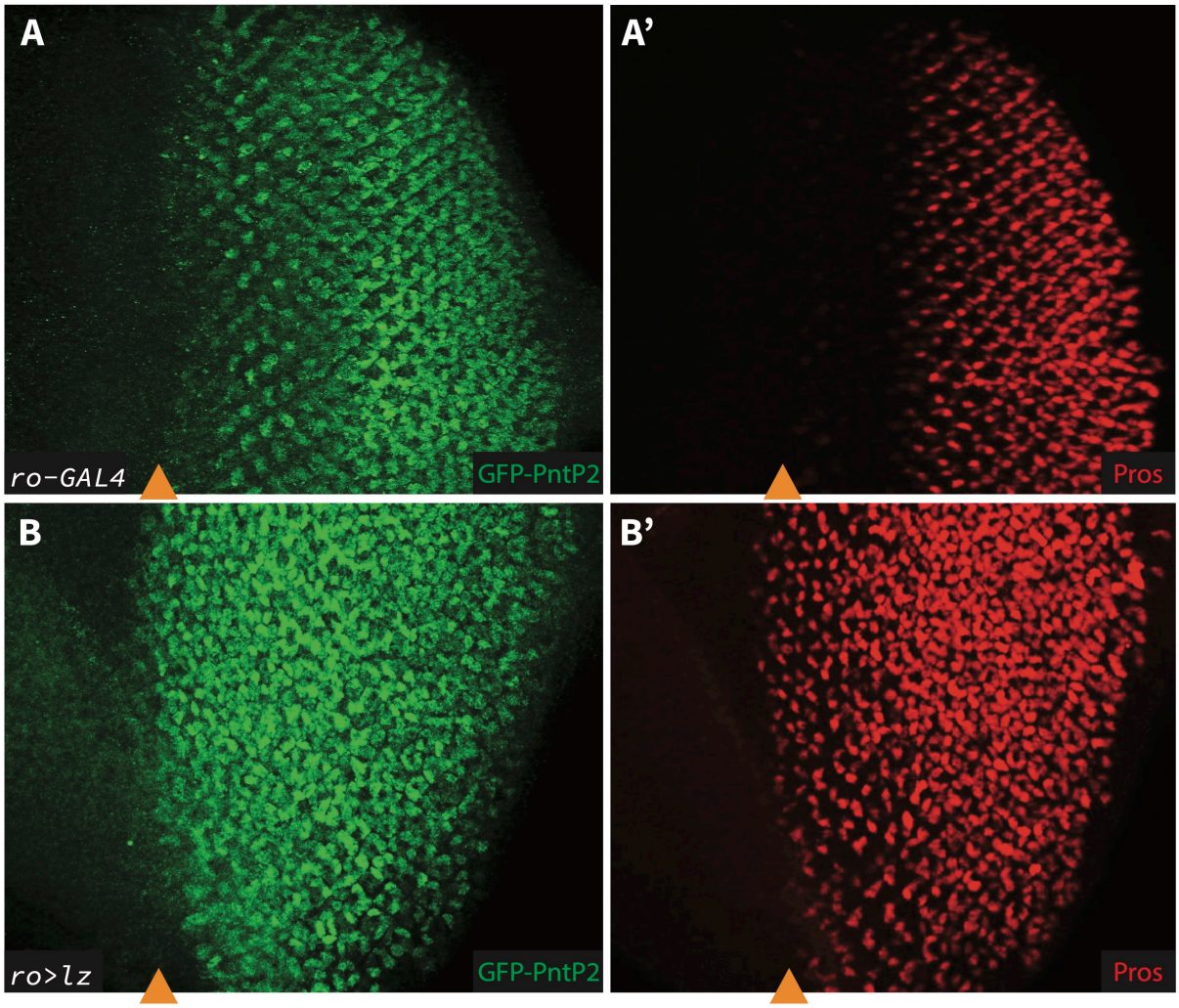


Fig A3.1 Expression GFP-PntP2 in fate-switched R7 cells

(Fig A3.1 continued) All images show maximum projections of optical confocal sections of representative 3rd instar eye imaginal discs, oriented anterior left. Scale bars: 10 μm. Pros (red) marks R7 photoreceptors and cone cells. **(B)** Overexpression of *UAS-lozenge (lz)* driven by *ro-Gal4*. Overexpression of *lozenge* induced ectopic Pros-positive cells. GFP-PntP2 expression is associated with the ectopic Pros-positive cells, thereby strongly increasing at the anterior of the disc.

A3.2 References

Mavromatakis, Y. E. and Tomlinson, A. (2013). Switching cell fates in the developing Drosophila eye. *Development* **140**, 4353–61.

APPENDIX4: Transcriptional reporters for *pnt* locus

A4.1 Results and Discussion

In my preliminary studies in the retina, I found reporter 7, 8 and 10 showed expression in the first round photoreceptor R3/R4 pairs and the second round cone cells (Fig A4.1A, B). This expression pattern is interesting because it is a subset of that of *pntP3*, and the positions of these three peaks are close to the 5'UTR of *pntP3* (Fig 4.2 in Chapter 4). Therefore these regions have the potential to regulate *pntP3* transcription. To confirm if these are essential regulatory regions for *pntP3* specifically, we can CRISPR deletions of the regions and see the change of *pntP3* expression. I expect reduction of *pntP3* expression in the individual deletions, although larger deletions may be needed to reveal function if these three bound regions comprise a single enhancer element. Once confirmed, these transcriptional reporters can be used to explore transcriptional regulation of *pntP3* under different genetic and environmental conditions as discussed in discussion section 4.1.

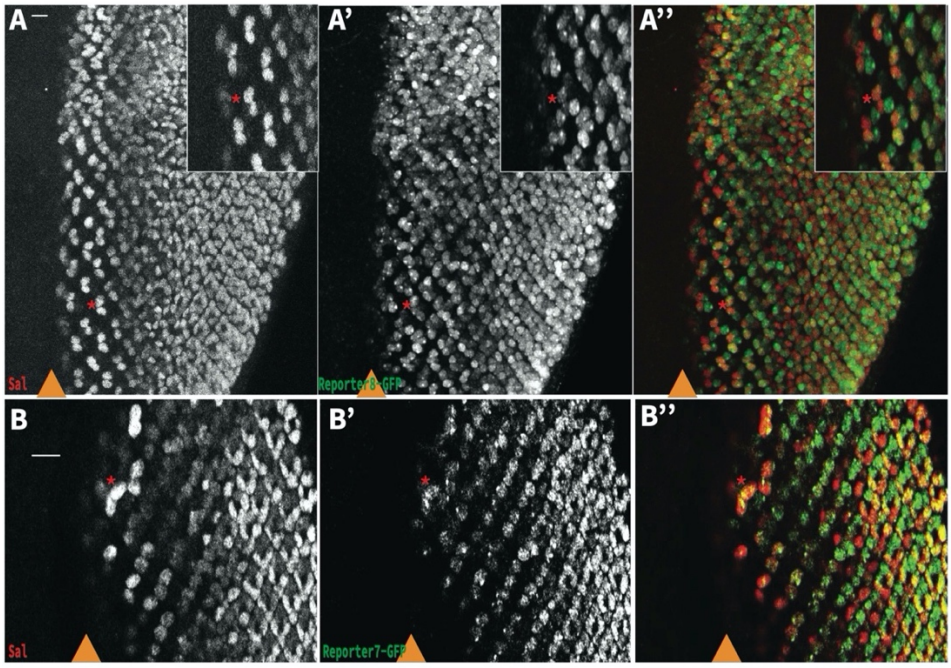


Fig A4.1 The expression of Reporter 8 and Reporter 7 in R3R4 pairs and in cone cells

Maximum projection of representative 3rd instar eye discs, oriented anterior left. R3R4 pairs (red star) at the anterior, and cone cells at the posterior are marked by Sal expression (red, anti-Sal). MF is marked by orange arrowheads. **(A)** Reporter 8-GFP showed expression in R3R4 pairs and in the cone cells, overlapping with Sal. **(B)** Reporter 8-GFP showed expression in R3R4 pairs and in the cone cells, overlapping with Sal. Scale bars: 10 μm.

A2.2 Materials and methods

Immunohistochemistry and microscopy

Details are the same in session A2.2.

Flies

The reporter regions were cloned (sequences in Appendix 5) and transcriptional reporters were constructed for each one. Then the transgenes were integrated into the 86Fb AttP landing site on the 3rd chromosome

APPENDIX5: Sequences of the transcriptional reporters

Reporter 1

CATTGGGAACGGCGCTCTTTGGCAAACCGGATGAGGAGCTAACGGCGAATCGTGTG
GGTAACTTTAGTGCGGATGCAGATGAGATGAATCCTGAGGAACTGGGCAGCTATCT
GGAGGGTGATATGCTGGTGCCGCAAACGGATCTGATTATGAAAAATGGACTGCCCA
CTCAGTCCTCCCGTTGGCCAAATGGTGTGGTGCCATATGAAATCCGAGGAAACTTCA
ATGCCAGAGATATGGCCACAATTGAAAATGCCATTGGTGAGTACCATCGTCGGACG
TGCATTCGGTTTGTGAAGCGCAGCTCTGAGCGGGACTACATCTCCATTCGCGGCGAC
AACTCCGGCTGTTGGTCCTCCGTTGGTCGAGTGGGTGGCAAGCAGGAGGTGAACCTA
CAGTCTCCCGGCTGTCTCAGCCGTCCAGGAACCGCCATGCACGAGCTGATGCACGCC
CTCGGCTTCCTTCATGAACAGAACCGTATGGAGCGCGATGGTTATGTGGCCATCCAG
TACAACAATGTCCAGTCCTCGGCGATGAACAACTTCGAAAAGGCCGCCCGCACGGA
AGCCTTTGGGGTGCCCTACGACTATGGCAGCGTGATGCACTACTCCAAGAACGCCTT
CTCCATCAACGGACAGCCCACCATACTGGCCATGGTGAGTACAATGGTACCAAAAG
TCTCCTGGGTTTAACTTAACACTCTTCAATTTCTGCATTC

Reporter 2

TCAGATACACGTAATTTGCAGCATTCGCAGTTCACACACCGTCGCGAATTCTTAGCA
CTTTTGCGAATTAATGGCAATTAATTGTGAGCACTTCACGATCGTCGCTCTGGCACG
GGCGATAAATTAATGCGCATATCGTAACTATATGTACATTGTTTTCATTCCCCGCGC
GACAATCAGTATGTTTGGTGTTGTTGTTTTTCTTCTTTCGTTCCTTAAATTCACATTTA
TTTCTATTTTTTTTTTTTTTATCGTTATTGTGTATTTTGCGCGCCTTTTCCCCGATAAAA
AAAACAGAACAGAACGCGTTGTATGAGGGACCCAAGATCAAGTTGGGACCACTAGA
GTATGGTACTATATAGATGGTGGGATACTATATATACACAATGAGCATGCGACGTCA
GCTCTGCCACCTTCGACATCCGTTCAAGAACTGAAGTTGCTCTGAAGTTGCACTCGG
ATTGTTGGCTGCTCGTCGGCCAGCTCGGCATTCGCTCTCTCTTTCTGGCCGGGTTGCC
TCTCGCTCGCACTCAAACGACCGTGTGTGGGAGAGAGCGCTTCGTGTGTACTTCGGC
TCGACGACTTCGACTCGCATTCCGAGATACCGAGATACAGATACGAATGTGCCAATG
GCAAATAATAATAATCGAGTCCGAGAATACTTTCACAGGGTGTTTCTCACACGCTCT
CGCTCAATCAGCCATCCATCGACTGCTCTCTTAGTTAGTGGCAGTGACTAAAAAAAA
AAGCCACAGAGACTAAAATTAAATGGTACGCCGCTGCGCATTTAAGGCAAACGTTT
GCTTAAAGAGTTTCGTAGTTTCTACTTGTTGAATGGGAACTCTAAACGCAGCTTAAA
TTATCGGAATATTTGAATTTGCTTAATAT

Reporter 3

ACTTGGCCCCACTTTCAGATACACATATGTATTATATAAATTATTTTTTGTTTTGGCT
AATGAACAGCAAAAAAAAAAAAAAAATGGTGGCCATAGCTCTGGAAACGTGTTTTGG
GGCATTTTTTCCGCTGTCCGTTAGCCAATGACAATTTGCTTTGGATTTTCGGGCTAAA
TACGAGTCCGAAACGTGAAGTCCAGAATGACACGTTCCCACTTCGAAAAGATAGAT
CCATGGCTGAGTGGTTAGAGATATTGAACGCCATAAAGAAGCCATAGGTGAGGTGC
ATTTTTCGCTTATCAGGTTTTTGGGGGCAGCTGCTCTTATCGGCGGGGGAATCGCTTG
TTGATTCTCAACCCGCTGATACAAATTGCAATGGCCTCAAGTGGTCGTCTTTTCTTGG
AAATCCCCAATGCACCTCGAGCTGATAAAGCACAAAAGTTATGCGCTGGGGTTATTA

CACGAAACAAAGTTGCCAATTTAACGGCATACATAATAAATACGCATAAGGAAGCC
CCAAAACCAAACAGCTGTTGACCATGTTTGGTTGTGGCCACAAaaggatggattgttttttttttttttt
ttttggccacttaagaattcttgcagcacaatcatatttATAATGtaaaatatatttaaaaaaatatttttCGTTAGCTTCTTGT
TTTTCCTTCTTTGATTTGCTGCGCAAAGTGCTGGAGAACAATAGAGAGACACAAAAA
ACACAAGCGGAAATACAAAAAAAACCCAGTGAAAGAAGCAGAGCTACCTTAATGC
GTGCAAGGGAGAGGGGCTACCGAGAGCGAGATAGATGATGCGCCTCGCTTACGACC
ATTTCCGGTTGAAGCACGTGCGCACATTTTGAAGAGGAGCGTAACCATCGAACCCCA
GCTACTCCTCCCATTTCGTCTCCTCCCAAAGACTGCACATGAAAAGCCACCACCAAA
GCAGAAAACCAATCACAATCTCATAAAACCGGCGACACTTCCACCAAAAAAGGGAT
AAAGAAAACAAAACAAGCCCCGAGTCTTCGTTTTCGTCTTCAGGGGAGTAAATGTA
GATCCTACTTGCCGAAAACGCACACGTTCGCAGCCAAGGGTTAACATAATAATTGTC
GTTGTCTTGAGTAAAAGCACACACAAAAAAAAAATCAGATCAAGCTAACTTAAATA
TATAATAACTATAAACATAAAGCGGAAAATCAAGAAACCTCAAAGTTAAAGCTACA
AAAGCAATTAATCCTTAAGCTTATACTAATATTTTCTAGCCTTAAATTGCATTAATCC
AAAACG

Reporter 4
CAATTAAAACTCGTTCTGTTCGCACAAAAGGATTACACATTTAATGCAATAAGGATT
GTCAGGGAGGAGGAGTTGCAGGAAGCGCCTACGGGAAGAACAAAGGCGGATGGGC
AAGGTACTCCCGAATCTAGGGATTCCTTCGGATGGGCCTGCCAACAACAGCTGTCGA
CTGACACTCGAGATCATCAGCCCATTCAATTGATATCAAATTGGAGTAGGATCCTTG
CCAGAATCTCGCTCGATCCTGTGCACGTGTACAAATTGCACTctctgagctctgagctctgagctctg
TATCCTATAATATGGTCTGAGCCCTATGCTCTGTCTTTTGTACCATATAGCACTGCAA
GCACTTCCGGCCCCATCCCGCTAATGTAAATCGGATGTCAATGCATTGCGCGTGAGG
CATATCAATTAAAAAAGACGAGGCGAGCGAGAGAGCAAGTGAGTGCGAGGCGAAA
ATGCCAGAGAGAACCTGTTAGAAGACAAGCCGCATAAATAAATAAACTGTTAGGCT
GGAAAACGGAAGGAAATCGTAAACACAAGCCAGTAGTTGTTGCTGATGTTTTGACA
CGCATACTATCATAATTACAGCTCCCAGTAATATCCATTTCCC

Reporter 5
CAGATTGTGTGAGCCTACGCACACAGATGCAGGTATCGCCCACACAATAACGATGA
GAGTGTGAGAAACTGCAGGCAGCCTGTGTTCCTCCTACGAACGCTTCAGCTCCCTTT
CCATGATCCTGCCTCCTGCCCTTTTCCTGGTTCCTTCTTCCTTGTTCCTTATTCCTTATT
TATTCCACTTTCTGTCCCCCAACTCCTCCACTTTCCTGTGTGTCTACTGTGTGCCTACT
TCGGGCGAACAAGATGCAGTCAAGCGATATAGGGATATTAGAATGCCCGACATGCG
GAGCCAGGGGATCTATGGATCTCAAGATCCGCAGGGCCGAAGGGCAGAGGAGGGA
ATCTACCTGCCGGGCCACTTTGGGCTTGCCTAATTTACTTGTGCCCTTCTTGCCGCAT
TGTTCTTGCAGTTTTCATTCATAAAATCTGAAATCTCTACCCTGGCGAGGCATTGAAC
TGGTCTGTGGATAGGATCAGAATAAAGGAATCTCAAGGTGCATCGATTCGGAACTA
AAGGAAGTAAATTACCGGATGTTGTCTACTCATTTGCCTGTTGTTCTTTTCGTATTCC
CTTTGCCCTTTTGCTTTTGATTCATCGAAATGTCCTGCGGGCAGTCCTTTGTGGACGA
CCTTGCTGTGCTGCTTGTTGCTCATTTTTCACTTGAAACCGAAACCAAAGCGAACAA
AACGAGCTGAATGCAAGCCAATAAAACATTTTGCCAATCCCTGGGATACACCTTTTT
ATTCTTGTTGTTGTTGCGTCGGCGGTTAGTGGGTATCGAATGGCACGATATCCTGGC
AGCCCTTTTAATTGCACACCTCGTATGCAATGA

Reporter 6

GCAGGCAGGTCTCAAACTGATTAAGATCAGAAAATACGGAAAATAGTTAAAGCCGA
AAGAGCAACATTAACGGAGACTGAGAAATGTTCAAGTTTTGCACAGAATTTACGAG
TTCGCTTTTGAGCCCGGCCAAATTGCATTAGCACCGGCCACAGGGGTAAAACGGGTT
GGCTGGGGAAATGAGAATTTTCCCACGTTCCGACTTTGGAAACTCAGGTGTTCGAGT
TGGACAAATTATTGCCGCAAATGTAGGCAATCCACACGAACTACGCGGCCAGTCAG
CCAGCCGAACAACCAGCACAAGGAGACGCCAACCCGAAGGACGCATTCACACACA
GGTCCTGTGCACGGTGGTTGCATCCACTGATAGGCATCGCACCACCAGCACCACTTG
CACCACTTGCACCCCCTATCATTCGAACCTGCACTTG

Reporter 7

CATATATCACAGCACCATTTAGCCATTAACAGATTGTCAGCATTAGCAAACGGGTTT
ACAGAACACACTGCAGCACTCAAACACACACACGTACTCCCTTCTGGAGATTTATTC
AAATTTGTTCTTATTTTTAATGAGTTCGATAAAAACAACACACGGAACGGAAAAATA
TATTTTGTAATTCGCTCGTCGTTCGTTCTGTGTATCTATCGGATAGCTGGATTTGTGA
AACCAAACTGAAACGAAACGATCGTATCCGTAGCTATAGCTGCATATCCTAAGCAA
AGCGAGCGAAGCGAAGAGAGAGAGCGCCAGAGAGCGCAACAAGTGGGAACGGAA
CTCGTAAATCTGGAAACCCTCTTGCGGAACGAAATGGAATTCATTCATAAAACGATG
CGAGTGCGTGTGTGTGCGTGCGCTCTCTCGGCTGCGTCTGTGTGTTGTGTGCTTGGAA
CGGAAACGGCCATCATCAGAGCAACAACAAGAGAAACAACAACTACAGCTAGTAA
AACAACAACGCCAGGTTGTTTGTTTAGCCGAGAGTAGAAGGCAAAAGGCAAGGGaaa
aaaaaaacaaaaaataaaacaaaacgcacagccacacgaaaaaattaaagcaaaaaactgcaagaaagagaacaaacacaaaTAAA
ACAGCCAGCGCTTGTTTTTTAGGCAACAGGTGACCAAATGTGCGATTATACAGGGTG
TAACAGGTGCACAAAATAGAAGTGTTTTAAAATGTCAGCATTACAAAAAACTATAC
ATCTTGGTGAGTTAAGTAATTGCTAAAAAGACCTGGGACTCAAATTAG

Reporter 8

TCCCCCTGCAATGACTCGAGAAATGCCCAGGCACACAGGCAATTTATGAATGAATTT
CTTGTGTTGCCTCACACGCGCATTCTTCTCAAAAGAAACTCCAGACAACGAACACAA
CGAAAACAACAACAACTAAAAAGAAGAAGACACCGATATTGTAGTAAGGGAACGC
TTCGTATTTCGAAGATACAATACCTCCAGGCGATAGAAGCAACAACAACAATCGAT
CGAGACGTCGACGTCGACAGAGGCCGCTTCCGCTGCTTCCGTTCGATGTATTGATAT
AGGGTGCAAAAGAGATGGCTAGATGCAGAACGGCCAGCGTGGAAAAGAGAAGATA
GTGGCCAGGGCAAAGGAGAGGAGGAGGATAAGATAAAGCTATTATTATGATGACTA
CGCGACCAGGTAAATTTATGAATGGGATACGAAACGAAACGAGACGCGAAACATTG
ACTCCGCGGCTTAACGTCTATGAATGAATTGAAATTACGTGCGACGGCGACGGAGA
CGTAAGGGAGATCATTTGGTCATTATCACAAAACCGCATAATTGGCGACAAATGCA
GGGGAAACCCTTTGGCTTTATCTTCGCCACGACGTTTACCTGGACGGCACCGGAAAA
AATAACCAATCCGCTTGCTAGGTATACCAAAGTAAGCTATTTCATTAGTCACAATCT
TTAAGAACCAAAAGCCATATTGTATCTTTCAGTTAACTTGCCTCTCTATTTTCATAGC
TTTTCTTATCACATCAAATGAACTTCTTCTCTCAGTGTATGTTTTAGTCCCAAGGTGG
GCTGCGTGTTAAAGAAATTACTTCTCATAGCCATGACTATGGGCCCTGTACATATTT
G

Reporter 9

GATTGTTTTTAGATCTGCAGCCGTTACTCAGGCTCAACTGCGATATTGAAGTCACGT
AGTCACACAAATGGCAAGCGATTGCATTGGATGTTACGATCTTTTCAATGCTAGACT
TTATGGCTTCAAAATGGCAATCTATGCGAGGCCCTTTCGTACCAGGCACCCGCTTTG
GTCTTATTTAATCGGGCACTAAGTAAATGGAGATTGGAAACTGTGTTCGTGTTTCGT
GGGTGGACTTGTGCCAATCTTAAAGTGTGAGTAATTGCGCAATTAGCACACGAGTCT
GAAGAAATCTTATCGCAATTCCCAGAAATGGGACTCGAGCGAAGTACTAGCTGAAA
TTTATTAGTGCAGGTTATTTTTGAACCCAATACATCTTATCAGCTGAAGACCTGGCTT
CCGATAACTTGGCCGAGTGTCGAGTGTTCATTCATAAAAAGTCGGAAAACGATCGG
CGTATTTCAGCGCGTAAGCGTGGACAATGAAATTCTATTTCCGTTTCCTGTTGTGATC
AGATCGTTATCGAGCGATAACCGGAGAAATTTCAAATTTAATGGAGTGTCTGATTGT
CGA

Reporter 10

ACGTATGCATGACAGGAGGCATGGAGTTGGAGTTGGAATCTTGAAAAGGGGGAGTG
AGCgtgtgtgtgtgtgtgtatcgtgtgtgtgCGGGGAGTTTTCGATATAAAAATAATTGCAGGAATT
GTATGTTGTTGACGACGTTCGAAAAAGAAAAAATACAAAAAAATAAAAGGCCATAT
AGGAAAAAATCAGCGAGACGATCGGCTTCTTCTTGTTGCACTTTTGCAGCTGAAACA
GCAAAAGTAGAAAGAGACGGGGCGAAGCAAGTGTAAAAGAGAAGTGCGCTGGCTA
ACAAGAAGAAGAGGGAGCTGTGCGACCCTGGTGAAAAAATAAAAAAAAAATGGGG
GGGGGGGGAAAAGAAAGAACCGAGGAGTTGTCGTACGAAACTTACTACACCGGAA
ATATAAAATTACACACTCAACACTTAAATGGCAtgtgtgtgtgtgtgtgtgtgCGAGTGAGTTG
CATGCTGCAGTTGAGATTCTTCGAGATACGTGGAGATGCAGTGCGTCACGTCTACAG
GGTGCAAAACGATCGAACCCAATTTGACCGATTGTTCCGTCGGCAAAGGGCAGGGA
AACGCAAATTAATTGCCCAATTTGAAAGGTAGGTAGTAGCCCCCATTCGACAATCAA
TGGACATGAAAAGTGTTGCACCTCAGGTGTGTTCCATAATTATGGGAATATTTATTT
AGCCAACTGCACGAGAGAATGATTTACGCACGAAGTGCGCATAAAGCGGATTATAT
TTATGGCATTTGCTGGGGATTAAGGGTGCAAATACGGTGTAGAAAATCCCAATATCA
AGAGAGTGAAATTTTGAGAAGGATATGAAAGTGAAGGGTCCTCAATTGGAAGGGTC
CAGCAATAATATCAATTAAAAGATGATGCAT

Reporter 11

TCAAATTACACAGATACTCACTCGCACACAAGCACAAGCGAAGGCGAGGACTGCCG
AAAGATTAACAACACACAATATGTCTGTGTCCTGCAGCTTTTACCTGCTAATGATCC
TAATGAAAAGATCGCATACCCTCTCTTTTGGCCGCGGTCATGTCCGACTTGTTGAGT
CCTTTGGCGGCTTTTGACCCCTCTCCACTCCCCCAATCTGCGGACGATCAAGGACTCT
GCTCGCTTGGCGTCTGTGGCGTGGGAAAAATCGGTCCAAAAAATCGCTCCTTCTGTT
CGGTTTCTCTTTTTCCCCCATTTAATATTTCTTTCTTTTTTTTTTTGGCAAAAAAAAAT
ATTTCTCAAGCCGCGCCGAGATCTTGTCTCCCCCTTCAATAATTTCTGCACTGATTTT
TCGAAGGGGACTCAGTCTGGATTTTTCGAACTGGACTGAGTCACACTGCAACAGGTA
GTGAAATTCAATGGCTGATTACAAAGGCAACCCGCTGTCTAGTAGCCCTCAAAAGTC
CAGCAGACAGCTTGACCCTGCGATAGTTATGCATACATATGTACATATATGCAATAT
ATACTCACTAACAGTAGATAGGCCATGGTTGCTGGCAAACATTCTTGACTCATTGCC
GATAAAGTCCGGAATTCTAGGAAACATTTCGAGCTGA

Reporter 12

ACAAAATAGTAGATTGGCAATTAGATTTTTCGAAAAACAAAGGCTTTTGTTCAATAC
GAATTTGGTTTAAACATTCTTAAAAGTATAGTTAAGGTTAGTTTTCAAGATTTTTTAG
AATTTCAGCGACTTGGTTGCAAAATCCCAACTAAACCATAATACCCTACACTACACT
GCAGGGTATAAAAAGGCAACGACGACTGCAGTTGGATTTCGAGCCAAAGTCGACGC
GCCCACGGGCAGCGACagaagaagaagaagatgagaagctggagcagcgataaaagaggagccccaaagacaaagcc
ggagaacaagaagaagagtaagagccggtttgcaggaacaagcctagaagaagaggaagaaggaagaggagCACTGGTGG
AGTGGAGGAGGGGAGGGGGCGACGGGGCAATGCAGGCACACGCAAAAAAGGAAG
GAAACAAACATCAAGCCCCAATGACAGAGGAAGCAACCACAACAACAACAAATGG
CCGTGGAAGAAGGGGAGACAGCGGCAACAAGATCCACTCGTCGCTGGCTCTTATCC
CCACTCCCATTCCCAATTCCATTTCCATTCCCCCATTCATCAACTGCATTTTAATTCTT
TTTTGGTTGTCTTTGCGTTTTGTTTTTTTTTCTTCGTTTTATTATTCGAGAACCAACTA
GTTTTGTTGTTTTTTGCTGCCTGGCGAGTATGAATGAAAACGAACGTGTTTCTTGGAG
AGCGGGAGCCCCAATTTGAGAGCGGGAGAGCGCTGGGGAGCAGCGGCGCCGGCAG
CGACGTCGACAGCGCGGGGATGAATGAAATGAGCAGCAAAATGAATGAAAGTTGTG
GGGCTCGCGCAGTCAACGTCGCTGCCAGCGTCAACGTTAGGCTCTCCCGATCTCTCA
CTCTCCCACTCTCCTTGTGAAGAGAATTTGTATATTTTTTCTACTTTTTATACAAATCA
TAGCGCACATATTTAATAAGCGAAAATATAATTTGCTTTATTTATTgtttcgtttggttcatttcgttt
cgttttgtttcTTGTTATTTTCAAAGTTTATGATTAATGAGAGATGATTACGGCGTGTGAGTG
TGCGTGCGTGAGTGTACTTTTGTATCTGTGTGCGCGTATAATTAAAATAAGGGCAGC
CTTTTGCATACACTTGCAAAAAAAGGATTTTTTTTATTTCTGCTAAGAAAAAATTGTT
TTTGCACTGTTTTAGTCATTCCAATTGCAGGAGTCCGAAATGGTTTTAAACGTACAAT
ATGATTACGTTATTATATTTTTAATAAACGATTTCCCAAGGTTTTATGCATTTTAAGT
ACATTTAGATTTCAGGGCGACCCTTATGCCAAGTGTATCT

Reporter 13

TTTGACTTGCTCAAACGAGGGAACTTTGTCTGCGGGAATGCCATTAATGATAATGGA
GGTCAGAGTAGTCAAACGAAATGCTCCGTAAATATCGTAATCAATTGCTGAGCGAA
TACCAGCTGAATTCACATACTCAATTGGAATCCAATTTTTGGAGCTCTACGATCCAC
AAGGCTTCAATTACTGTTCAGACTGATGGAAACCCAAAACTTGGCTCCTTCTAAAGG
CCAGACACTTCGGATTGGATTACCTGGCTGCATTTTCCAAGCGACTGCGCATGCGCA
TGTCCTTTTTCTTTTTTTTTGGGGGGGGGGGAGTGGTCACTGGCAGGGCAGCGGGGT
AAGGATCAGCTGGGGGTCAGGCTGTTGGCCCCGAAAGACCGGATGAAAGTGGTCAA
CGGCTCGGCGGGCCAACGGATATGCTTACCGCAGGCGATGTGTCAAATGCCTTCACT
CGGTCGTCAATTTCTTGGGGACACAAAGAATCCAAGGGAGTGGGAAAATAAGAGAA
TGACTGCTTTGCATAGCCGAAGATTTGGATACTCTTGTTATTTGAGGAGGCAGCTCT
ATTAAGACAAAGTTGGTTGGGCATACAATAACTGATCCTACTTAAGACGCAAAAAA
CTTAAATTACTTCAAAATAATCTAT

Reporter 14

GCAGATGCTTAAAGAACAGATTGTAAACGAATATGATCTTTTGATATGGTGCGCACT
CCTCGGCCAGacgacaacgacagcaacaactaaaacaccagcagcggcaacaaATAGGGGTGCAATTTTTCG
ATGTGTACGTAACGGTAGAATACACATTTAGTCGCTGCTCTTTTCTTTCCCAGCGCCT
GGCCCTTTTTTCCTCGCGCTGAATGTAAAGAAGGGATCTCCGATCCGCTGATGATCC
TCGGGATGCCTGCCTGTGTACACATGTGTTGTTCTATCTGTCGCAGAACAGTCCGAA
CTGCATGTGTGTGCGGAATATCTATGTGTGCGGAGTATCTGTGTGTGTGGCATATCT
GTGCGAGTGTGTCGTCTTGTCTTGTCATGGGATCGATCTTGATCTCGGCCTGTCTGCA

GCAGGTTGATACTTTTAGCATCGGTCCGGGCACAGAATCCGAGTTTAACGTGGCATG
GGGCAACGAACCAGCCCGTCGAACATCCACTGAACTGTAATTGTCTGTGGGTATTTA
TGGAAAAGGATCTTGAACTGGTCTTTCATTTCGGACCAGACGCTCGAGGAGGGAAA
GGGAGAGGAGCGGTTGGCGGTAGCTGGGCTCAATTATGATCCGACGATCGTCACAC
ATCAAATTATCGATTATGGCAATAATTTCTTATCATTCCCTTGTGTATTTTGCTCGCA
AATTGGTTTTTTTTAAATAGCCCAAGGGATATTAAACATGTCTTTAGAGTAGTTCAAT
GATCGTAACTCCTTAAATTATCGTTTTTTGGCTTTAAGACTGCGATAAGGACGTAAA
TCAAGTTTTATCATAGAAATAATAACAATTGATCTAAGCAAAATCTTATTGTAAAAT
TAATCATTTTAATGGTAACATTTTTATAAAATATGATATCGGTTAAGCCTTTAACTAT
CCTTTGATCTTTTTTGTATTAATAATTACTTTAATTTACATCCCAAATTGTAAGATCCT
TTTAAAAACCCTTTCAGGTAACCCAAAGTGCAGAACA

Reporter 15

TATAAAATATGATATCGGTTAAGCCTTTAACTATCCTTTGATCTTTTTTGTATTAATA
ATTACTTTAATTTACATCCCAAATTGTAAGATCCTTTTAAAAACCCTTTCAGGTAACC
CAAAGTGCAGAACAGGAAAGCCCCAATTATACATACTTTAAAACCGAAAGGTCAAA
AGAGTAGCACTCCGGTATCAGTTTTATTGATACTTAATTGTCTAATGCGATCGCTGTG
CTAGATTCTTGACCCTAACCAAAAACCGTTCACACATTCTCGATTACCAAACCCCTT
GGATTGATGGCCCGAAACGCGATCGTGGGGCAGGCGAAGCTCATTTGCCGCACCGG
AAGCTGTAAAAGGAACAGTAGTAATTTCGAGCGTGTTTCCCACAAGGGTTGTGGAA
GTGCCCCACACACACACACTCACGCACACTGGCGGCTAACTAGCTTGGCTATGATGC
CGAGTAGAAAACCAATTCAAATTCTGTTGGCCTGGCCAACAACGAGTTTATAGAGG
CAGCAATCCTGAATCGAGGCGAGATCGCGAGATCGGCGGATCTACCGAGCGGTAGC
CAAACCAGTTTTCTCCTAGATTCATTCAAGAAAACAGGCAACACCAACACAAAAAG
GGTTTACATGCACTCAAAAGAAACTCTCTAAATTTCATATTTTTAAGATAAATATGA
TTTAAAAGTTATATTTCAAAACTATATATCAACAGCAATTAAGAGATATATGGTTTA
TTATTTGTTTTCAATTGTTTTGTCTTTTAATTTTCGCTGGATTTTTGTTTCGAGTGTAG
GGAAAATAACAAAAACTAATACTTGAACTACTGT

Reporter 16

ACTCCGCAGCACCGATGACTATCTCTTGCTCGTTCTCTCTCTCTATCACGCTGTATAG
ATGTGTATGTTTTAACCCTCCGGTGGGATCAGCAGGTTGATTAGGGGGAGAGAGAG
AGGGATTACGTGCCGGGCACAAAGACAAACGAAAAACCGAAAGGAACAACTCGAG
AAGAAATCGGGGATCGGTTTAGCTCGGCTTATAAGACGCGAAAAGAAACAAACCGA
AAAAAGGCCGACGGATCGCTGGCAAACACACACACGCGCGCGCACACTCACACTCA
GGCACTCTCACGCGATGTTTGAGGAGATTTTTGTGGTTTTATGATGAATTTCGTATAT
AGAAATTGGTTTCCTGTGTTTCTGAACCGCGTTTTTTTTTCTGTTGCTGTTTGCCTTCC
TTTTTTTGATGTATATTGCGCGCGCAGATCTTTAGATGTTTTGCTATTGCACTCATTC
ACTTACACACTGGGTTTTTGCTTGGCTTGCAAAGTGTTTTGCATTTGTTATTTATTTGT
TATTTGATTCGTATAATTCATTCAAATCGAATTTACCGTAGCGCAAGTGCACGCTTCC
CGAAAGAAAGACAAAAACAGACTGACCGACTGGCAACAGGCGGCAGCGACGCCGA
CGTTGGCGCAGACGCAGCGTTGACGCCAACGCCGACGCACCTctctctccgtttctctctctctttctc
tcCGCTGCTAAGATTAATATTTTTTTGTGTTGCTCTGGCAAGTGATAACAAAAATGGC
AAATAATATGAAAATCGAGAGTTATATCAATCTGACTTATAAACATTATTCGCTATT
TTACTTGATCGCTGGAATGAAAACGAAAAGGCGTGGGTGATCCTCAATCTCGCTATC

TCTTTCGCACTCGCACTAGAGCGTGCATCATGAGCGCAAGAAAAAAATGGAAATCA
GAATTACATATTTCGAGCAGCAAAAACAAAGTGTTTGCAGCGACAAAAAAATAGCA
ACAATTTATAAAAAAAATCTGCTTGCTCAGAACTCAGTGTAAGTGCGAATATGCGAG
AGAgtgtgtgggtgcgt

Reporter 17
CCTTCAAACGGCGCTGTCGGCGCGTAAACTCAGCCTGGCCATGCCGATCCTGCTGGA
GTTTCCCCAGGCTGAGATTCATATCTCCGATGACGCTAGTTTGCTTTCTGGCCTTGAT
TGGCAGCGACAGGGCTCCAGGGCAGTGATACGCCACTTTCTGAATCTGAACAATGTT
CTTGATTTGATGTTGGATCAGTGTCGCTACTTCCATGTGCCCATTGGCAATATGCCGC
CGGATACTGTGCTTTTCGGAGCGGATCTTTTCTTCGCTCGCTTGCTGCAGCGGCATAA
CTTTGTGCTGTGGTGGTCGGCGAGTACCAGACCAGATTTGGGTGGCCGGGAGGCGG
ACGACAGCCGGCTGTTGGCGGAATTCGAGGAGAGCATTAGTGTGGTGCAAAACAAG
GCCGGTTTCTATCCGGATGTTTGCGTGGAGCTGGCTCTGGATAGCCTGGCGGTGAGT
GCCCTGCTCCAATCGACTAGGATTCAGGAAATGGAAGGCGCCTCATCTGCCATTACG
TTCGATGTGATGCCGCAGGTCTCGCTGGAGGAGATGATTGGCACTGTTCCGGCGGCC
ACCTTGCCGAGTTATGATGAAACGGCCCTCTGTTCCGCCGCCTTCCGCGTTATGCGC
TCCATGGTGAATGGTTGGTTGCGAGAGGTATCCATCAATAGGAACATCT