THE UNIVERSITY OF CHICAGO

QUANTITATIVE INSIGHTS INTO THE ROLE OF RTK/MAPK SIGNALING PATHWAY ON DEVELOPMENTAL CELL STATE TRANSITIONS

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 DEVELOPMENT, REGENERATION, AND STEM CELL BIOLOGY

BY

SOO JI HUR

CHICAGO, ILLINOIS

AUGUST 2023

TABLE OF CONTENTS

LIST OF FIGURES v

ACKNOWLEDGMENTS vii

ABSTRACT viii

1 INTRODUCTION 1
1.1 Overview 1
1.2 How Cells "Find Themselves": Understanding Cell Fate Transitions 2
1.2.1 Waddington's Intuition: Early Framework Linking Genes to Cellular Phenotypes 2
1.2.2 From Flies to Insights: Identification of Developmental Genes 4
1.2.3 High-throughput RNA Sequencing Methods Reveal Cellular Heterogeneity in Gene Expression 5
1.2.4 Using Dynamical Systems to Define Cell States and the Transitions Between Them 6
1.2.5 Discoveries in Cell Fate Transitions: What's Next? 7
1.3 Regulation by Transcription Factors: Moving Beyond "On or Off" Description to Quantifying Input-Output Functions 9
1.3.1 The Role of Transcription Factors in the Flow of Information 9
1.3.2 Common Transcription Factors and Where to Find Their Specificity . 11
1.3.3 Dynamic Regulation of Genes by Transcription Factors 12
1.4 A Model of Precision: Cell Fate Transitions in the Developing Drosophila Eye 14
1.4.1 Building A Neurocrystalline Lattice: Drosophila Compound Eye Development 14
1.4.2 Signaling the Way: The Wave of Differentiation in the Developing Eye Requires the Interplay of Multiple Signaling Pathways 14
1.4.3 Reiterative Use of RTK/MAPK Signaling Pathway Directs Progenitor Cell Transitions to Photoreceptor Neurons 16
1.4.4 Transcriptional Regulation by the Pnt and Yan Network 19
1.4.5 How Do Antagonistic Transcription Factors Pnt and Yan Control Progenitor to Photoreceptor Cell Fate Transitions? 21
1.5 Scope of the Thesis 22
1.6 References 23

2 RATIOMETRIC SENSING OF PNT AND YAN TRANSCRIPTION FACTOR LEVELS REGULATE PHOTORECEPTOR FATE TRANSITIONS IN DROSOPHILA 40
2.1 Statement of Contributions 40
2.2 Abstract 41
2.3 Introduction 42
2.4 Results 44

2.4.1 Pnt and Yan positively regulate each other's expression 47
2.4.2 An increase in the Pnt/Yan ratio accompanies cell state transitions . 51
2.4.3 Perturbation of the Pnt/Yan ratio biases cells towards or against R-cell fates 52
2.4.4 A biophysical model explains how small changes in stoichiometric ratio can have a large effect on Yan and Pnt DNA occupancy 56
2.5 Discussion 61
2.6 Materials and Methods 65
2.6.1 Genetics 65
2.6.2 Immunohistochemistry 66
2.6.3 Quantification of expression levels 67
2.6.4 Conversion of distance to time 69
2.6.5 Alignment of expression data 70
2.6.6 Analysis of yan clones 70
2.6.7 Simple competitive binding model 71
2.6.8 Competitive binding model with cooperativity 72
2.7 Supplemental Figures 81
2.8 References 81

3 UNDERSTANDING HOW PNT AND YAN EXPRESSION LEVELS DETERMINE THE MRNA OUTPUT OF TARGET GENES 90
3.1 Statement of Contributions 90
3.2 Abstract 90
3.3 Introduction 91
3.4 Results 94
3.4.1 Developing an experimental approach to visualize TF protein and mRNA at single nuclei resolution in the developing eye disc 94
3.4.2 Developing an image processing pipeline to optimize signal-to-noise ratio from Drosophila eye disc smFISH images 97
3.4.3 Quantifying the effect of nuclear Pnt and Yan protein levels on photoreceptor fate-determining gene transcription 100
3.5 Discussion 103
3.6 Materials and Methods 109
3.6.1 Fly Genetics 109
3.6.2 Generation of smFISH probes 110
3.6.3 Combined Immunostaining and smFISH 110
3.6.4 Confocal Imaging 111
3.6.5 smFISH Channel Processing 111
3.6.6 Linking Expression Levels to Individual Segmented Nuclei 112
3.7 References 112

4 CONCLUSIONS AND DISCUSSION 115
4.1 Summary and Implications 115
4.2 Future Directions 116

4.2.1 Pnt isoform/ phosphorylation state-dependent regulation 116
4.2.2 Yan oligomerization and Cell State Transitions 117
4.2.3 Dissecting Transcriptional Kinetics Downstream of Pnt and Yan . . . 119
4.3 References 120

A APPENDIX:THE ROLE OF PNT AND YAN IN OMMATIDIAL PATTERNING: VIEW FROM E-CADHERIN SUBCELLULAR LOCALIZATION 123
A.1 Introduction 123
A.2 Rationale 125
A.3 Results and Discussion 126
A.4 References 126

B APPENDIX: THE EFFECT OF INCREASED YAN SELF-ASSOCIATION AFFINITY ON EMBRYONIC MESODERM SPECIFICATION 129
B.1 Introduction 129
B.2 Rationale 130
B.3 Results and Discussion 130
B.4 References 131

C APPENDIX: SEQUENCES FOR SMFISH PROBES 134

LIST OF FIGURES

1.1 **Waddington's Epigenetic Landscape** . . . 3
1.2 **Transcription Factor Activities Represented by a Gene Regulatory Network Diagram** . . . 10
1.3 **Complex interplay of Hh, RTK, Dpp, Wg, and Notch signaling in the Drosophila eye imaginal disc** . . . 17
1.4 **Pnt and Yan are downstream effectors of the RTK signaling pathway** 21

2.1 **Pnt and Yan expression defines two waves of cell state transitions** . . . 48
2.2 **Pnt and Yan activate each other's expression** . . . 50
2.3 **A two-fold shift in the Pnt/Yan ratio accompanies R cell fate transitions** 53
2.4 **An artificial increase in the Pnt/Yan ratio biases progenitor cells towards an R fate** . . . 55
2.5 **An artificial decrease in the Pnt/Yan ratio biases progenitor cells against an R fate** . . . 57
2.6 **Cooperative binding by Yan greatly sensitizes DNA occupancy to the Pnt/Yan ratio** . . . 60
2 S1 **Identification of cell nuclei in eye discs** . . . 74
2 S2 **Pnt-GFP and Yan expression in progenitor cells** . . . 75
2 S3 **Characterization of Pnt-GFP** . . . 75
2 S4 **Expression dynamics for various eye cell types** . . . 76
2 S5 **Heterogeneity of Yan and Pnt expression** . . . 77
2 S6 **Sev-Gal4 drives UAS transgene expression in subsets of progenitor cell** 78
2 S7 **Fate specification errors in sev>RasV12 and sev>YanACT** . . . 78
2 S8 **YanACT has no effect on the Pnt/Yan ratio in progenitor cells but does alter Yan levels in R7 cells** . . . 79
2 S9 **YanACT has a weak effect on the Pnt/Yan ratio in R3/R4 cells** . . . 80
2 S10 **A simple equilibrium model of Yan / Pnt competition for DNA binding sites** . . . 81
2 S11 **A simple equilibrium model of Yan / Pnt competition for DNA binding sites** . . . 82
2 S12 **Accuracy of cell-type identification** . . . 83

3.1 **Pnt and Yan protein levels and Cell State** . . . 93
3.2 **Pnt and Yan protein levels and Transcriptional Output at Target Loci** 93
3.3 **Overview of the Combined Immunostaining and smFISH Protocol** . . . 94
3.4 **Pnt and Yan protein expression patterns are preserved in combined smFISH and immunostaining protocol** . . . 96
3.5 **4-channel imaging allows simultaneous visualization of Pnt and Yan proteins and nascent RNA of target gene *Pros*** . . . 97
3.6 **Overview of smFISH spot detection and quantification pipeline** . . . 99
3.7 **Photoreceptor gene transcription during P vs. R7 cell fate decision** . . . 101

3.8 **Salm transcription as a function of Pnt levels, Yan levels, and Pnt/Yan ratio** . . . 102
3.9 **Prospero transcription as a function of Pnt levels, Yan levels, and Pnt/Yan ratio** . . . 104
3.10 **Combined effect of Pnt and Yan levels on Salm nascent RNA expression** 105
3.11 **Modeling potential switch and rheostat mechanism of Pnt and Yan regulation** . . . 106

A.1 **Ommatidial patterning involves line-arc-five cell cluster transition.** . . . 124
A.2 **Effects of perturbing Pnt and Yan on E-cad Remodelling During Ommatidia Morphogenesis** . . . 127

B.1 **Increasing Yan SAM affinity produces no significant effect on the number of cardiac muscle progenitor cells in stage 11 embryonic mesoderm** 132

C.1 **Pros smFISH probe set** . . . 135
C.2 **Salm smFISH probe set** . . . 136
C.3 **Senseless smFISH probe set** . . . 137

ACKNOWLEDGMENTS

I am indebted to my PI Ilaria Rebay for her generosity, trust, and guidance through my doctoral journey. I also thank my past and current colleagues in the Rebay lab - Catherine Wu, Xiao Sun, Julio Miranda-Alban, Jacob Decker, Juana Delao, Saman Tabatabaee, Amanda Hill, and Mischa Ludwig for their support. They have challenged me to grow as an independent thinker and a rigorous scientist. I would like to thank my thesis committee, Chip Ferguson, Edwin Munro, and Michael Glotzer. Their advice, constructive feedback, and wisdom have greatly enhanced the quality of this research. Thank you all your guidance as I learnt how to execute my research ideas.

I extend my heartfelt gratitude to my family - my mother, Minji, Michael, Andrew, and uncle Park - for their unwavering support and encouragement as I embarked on this journey far from my home in Australia. Your belief in me has been the driving force behind my work, and I cannot thank you enough for being my rock-solid supporters.

To my dear friends in Sydney and Chicago, especially Melissa and Juan, thank you for not only celebrating every milestone in my graduate school career but also standing by me like siblings during the tough times. Your presence made all the difference, and I cherish our bond dearly.

A special acknowledgment goes to my best friend and partner, Carlos Cortez. Your unwavering support throughout this journey has been my guiding light. You bring warmth and joy to my life, and I am forever grateful.

Lastly, this work is dedicated to my nephew Julian and niece Serena. I have no doubt that your wonder and curiosity will lead you to remarkable discoveries that fill your hearts and minds with joy and fulfillment. Keep exploring, questioning, and dreaming big, for the world is yours to discover.

ABSTRACT

Cell state transitions play a pivotal role in the development of multicellular organisms, orchestrating the precise arrangement and abundance of distinct cell types within the organism. This critical task is carried out by gene regulatory networks, complex systems composed of molecular regulators including transcriptional activators, repressors, and their target genes. Unraveling the mechanisms by which these regulators drive such a precise and reliable process despite intrinsic random fluctuations is an intriguing question in developmental biology. In this thesis, I focus on the gene regulatory network of the highly-conserved Receptor Tyrosine Kinase (RTK)/ Mitogen-activated Protein Kinase (MAPK) pathway as a case study. Owing to decades of research, main components of the Drosophila RTK/MAPK signaling network are well-established. Here, I leverage the wealth of qualitative knowledge to gain quantitative insights on how the RTK/MAPK pathway effectors regulate cell state transitions in the developing Drosophila compound eye.

The findings are consolidated into two results chapters. Chapter 2 focuses on elucidating how expression dynamics of the activator Pointed (Pnt) and repressor Yan - effectors of the Drosophila RTK/MAPK signaling pathway - are modulated to inform whether a progenitor cell remains in the multipotent state or differentiate into a photoreceptor neuron. Here I show that the expression ratio of Pnt to Yan in the nucleus is stabilized in the progenitor cell population, and that this activator-to-repressor ratio, not the absolute level of either transcription factor, determines the progenitor vs. photoreceptor neuron state.

In Chapter 3, I delve deeper into how Pnt and Yan expression dynamics are used to fine-tune the transcriptional output at target loci. Specifically, I explore the regulation of two well-characterized photoreceptor-fate determining genes, *spalt major* and *prospero* at progenitor vs. photoreceptor fate decision windows. To this end, I developed a quantitative imaging and analysis pipeline that can simultaneously measure the abundance of proteins and nascent RNA at single nuclei resolution *in vivo* developing Drosophila eye. Results

suggest that the nuclear concentration of Yan may be used to determine whether the target gene promoter is in a transcriptionally active ("On") or inactive ("Off") state. Meanwhile, Pnt concentration is weakly correlated to the level of nascent transcripts produced from active promoters.

In Chapter 4, I summarize the main findings presented in this thesis and discuss avenues for further exploration. Overall, my work explored how the interplay of antagonistic transcription factors Pnt and Yan contributes to the accuracy and robustness of cell fate decisions during complex tissue development. There are many examples in living cells where co-expressed transcription factors exert opposing regulatory inputs on a common set of target genes. The findings and the conceptual framework presented in this thesis may be relevant to other such systems. Finally,the experimental and computational analysis methods I developed may prove useful to the broader community using the developing Drosophila eye as a model system.

CHAPTER 1

INTRODUCTION

1.1 Overview

All multicellular organisms, be it delicate flowers, soaring birds, or sentient humans, begin their life as a single cell. During development, this initial cell ultimately gives rise to a myriad of specialized cell types, each of which serves a unique function within the organism. This process of cellular differentiation has long captured the curiosity of mathematicians, biologists, chemists, and physicists alike. Of particular interest is the remarkable accuracy and reliability with which cells in developing tissues are programmed towards a particular cellular identity (e.g. a muscle cell, or a neuron, or a blood cell) while rejecting possible alternative fates to form correctly proportioned tissues.

Over the past several decades, significant progress has been made to understand the molecular drivers of cell fate decision-making. We now understand that cell intrinsic and extrinsic factors converge onto gene regulatory networks to determine the state of overall gene expression within a cell. Although numerous gene regulatory networks that drive developmental decisions have been identified, many questions remain. For instance, how is precision achieved from the interactions of regulatory molecules whose expression levels fluctuate in space and time? How do cells produce a sufficient amount of gene products required for a particular cell type? How is the timing of gene expression regulated? These are the questions that motivated my research. We are currently in an exciting era of research where the integration of quantitative data and theoretical frameworks is paving the way towards a more predictive understanding of cellular differentiation.

1.2 How Cells "Find Themselves": Understanding Cell Fate Transitions

1.2.1 Waddington's Intuition: Early Framework Linking Genes to Cellular Phenotypes

A widely used metaphor to understand developmental cell fate decisions is the "epigenetic landscape" introduced by C.H. Waddington in his book *The Strategy of Genes* (Waddington, 1957). Waddington likened the process of cell differentiation to a ball rolling down a landscape, where its final destination is a probabilistic event influenced by numerous ridges and valleys. These valleys were illustrated to have a high degree of canalization, constraining the possible paths the ball can take. Beneath this landscape, interconnected cables attached to pegs determine the steepness of the valleys, symbolizing the intricate interactions between genes and their products (Figure 1.1).

The epigenetic landscape and the concept of canalization were developed to explain the robustness of phenotypes observed in nature against genetic and environmental variations (Waddington, 1957). He proposed that cell fate decisions during tissue development are highly reliable, yet probabilistic outcomes. The complex interplay between genes and the environment was predicted to determine the probabilities associated with these outcomes. Although Waddington's landscape lacked specific examples of developmental genes (pegs) and much of the nature of gene products driving these interactions (cables), it laid the groundwork for understanding cell fate decisions by bridging the activities of genes and cellular phenotypes. Furthermore, this early framework encompassed both qualitative and quantitative aspects of cell fate determination.

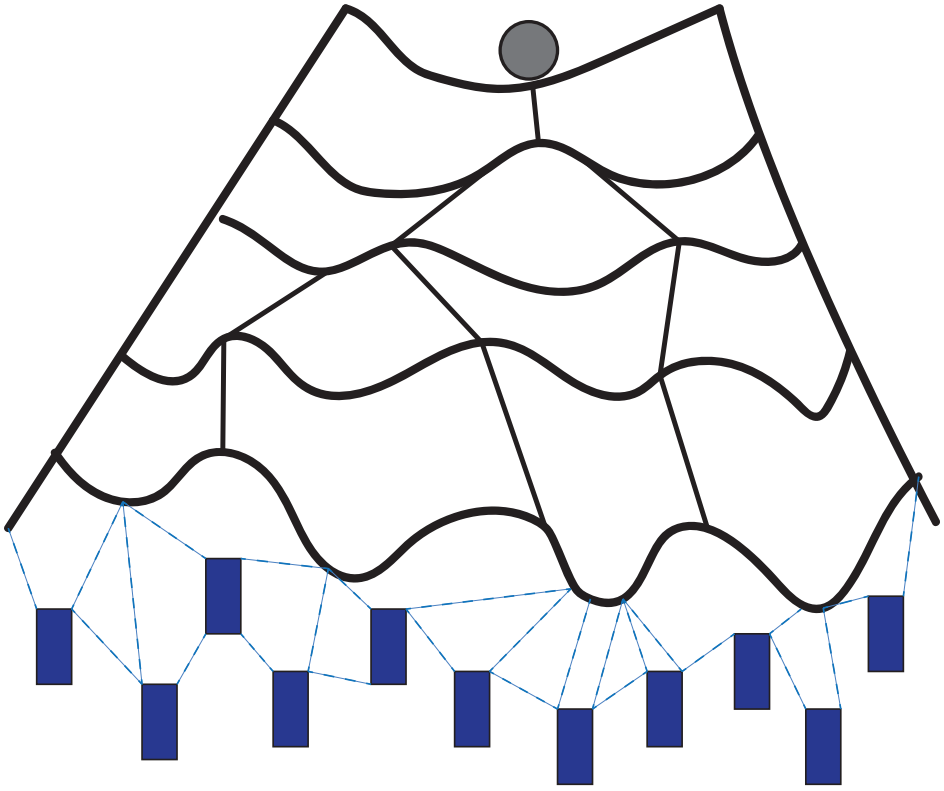

Figure 1.1: Waddington's Epigenetic Landscape

Adapted from *The Strategy of the Genes* by C.H. Waddington (1957). The ball depicts a developing cell during differentiation. The probability of its trajectory down the landscape is determined by the ridges and valleys, created by networks of connected pegs, a metaphor for genes.

1.2.2 From Flies to Insights: Identification of Developmental Genes

The idea that genes contributed to heritable traits existed since Mendel's early work in the mid-19th century (Mendel, 1866). However, no systematic experiments existed that could test the role of individual genes on specific developmental pathways. In 1980s, Nusslein-Volhard and Weischahus conducted influential genetic studies in the fruit fly Drosophila Melanogaster known as the Heidelberg Screen (Wieschaus and Nüsslein-Volhard, 2016). This work was inspired by the success of systematic mutagenesis screenings performed in bacteria (Jacob and Monod, 1961) and fungi (Beadle and Tatum, 1941; Hartwell et al., 1970) in dissecting the components of biochemical pathways. Nusslein-Volhard and Weischaus reasoned that a similar approach could be used to identify the genes responsible for embryonic development. They chose pattern formation during the early stage of Drosophila development as a model system. Through genetic screens and mutagenesis experiments followed by genetic mapping, they identified numerous genes involved in early embryonic development (Nüsslein-Volhard and Wieschaus, 1980; Nüsslein-Volhard et al., 1984). This spurred further developmental genetics research, culminating in the identification of hundreds of developmentally critical genes in other contexts of Drosophila development as well as other study organisms such as worms and mice.

With the advent of molecular cloning and other approaches, the nature of protein products encoded by these genes were unveiled. In fact, many of the early-acting genes identified in the Heidelberg screen encoded highly-conserved transcription factors such as a family of homeobox DNA-binding domain transcription factors including *Ultrabithorax* (McGinnis et al., 1984) and *fushi tarazu* (Laughon and Scott, 1984). Other lethal alleles identified were components of major signaling pathways including the Receptor Tyrosine Kinase (RTK) pathway (Bier et al., 1990; Mayer and Nüsslein-Volhard, 1988), Wingless/ Wingless Integrated (Wnt) (Peifer and Wieschaus, 1990; Riggleman et al., 1990; Nusse and Varmus, 1992; Sanson et al., 1999), Hedgehog (Hooper and Scott, 1989; Ingham and McMahon, 2001;

Nakano et al., 1989), and the Decapentaplegic (Dpp)/ Bone Morphogenetic Protein (BMP) pathway (Ferguson and Anderson, 1992; Penton et al., 1994). Further genetic and molecular investigation resulted in a comprehensive mapping of these major signaling pathways, leading to a surprising discovery that only a handful of highly-conserved signaling pathways orchestrate numerous developmental fate decisions across animal species including humans.

1.2.3 High-throughput RNA Sequencing Methods Reveal Cellular Heterogeneity in Gene Expression

The definition of a "cell state" has evolved considerably over time, since the term's first usage by Virchow in the late nineteenth century to describe cells of the same functional group in the human body as a state or a society ("Zellenstaat") (Virchow, 2020). In the decades preceding the 2000s, cell type classification relied on traits such as morphology, location, functional characteristics, and the expression of cell-type specific marker genes observed through antibody staining (Mulas et al., 2021; Raff et al., 1979; Hockfield and McKay, 1985; Hickey et al., 2021).

Over the last two decades, an explosion of high-throughput techniques such as RNA sequencing (RNA-seq) has enabled an unbiased, global profiling of transcripts to categorize cell states. Initially, gene expression profiling could only be performed on bulk cell populations, such as cultured cells or harvested animal tissues at various developmental stages (Marioni et al., 2008; Wang et al., 2009). However, the recent development of single cell RNA-seq (scRNA-seq) has allowed investigators to profile gene expression in individual cells (Trapnell et al., 2014; Trapnell, 2015; Cao et al., 2017).

Recent studies using approaches such as scRNA-seq have uncovered a striking heterogeneity among individual cells, even within those traditionally classified under the same cell type using earlier approaches (Trapnell et al., 2014; Trapnell, 2015; Cao et al., 2017; Deng et al., 2014). Gene expression heterogeneity among cells in developing tissues has been con-

sistently observed across metazoa from worms (Cao et al., 2017; Packer et al., 2019) and flies (Bollepogu Raja et al., 2023; Li, 2021) to human fetal tissues (Olaniru et al., 2023). This observation poses a new question: how do we use global gene expression profiles with high cell-to-cell variability to define cell states and the transition between them?

1.2.4 Using Dynamical Systems to Define Cell States and the Transitions Between Them

Dynamical systems theory provides a framework for understanding the qualitative behavior of complex systems. While Waddington drew inspiration from dynamical systems to develop his landscape metaphor, it was Stuart Kauffman who directly applied its theory to describe cellular homeostasis and differentiation (Kauffman, 1969; Moris et al., 2016). This approach has since been employed by numerous studies to theorize cell state transitions, as well as to guide the interpretation of experimental data (Glass and Kauffman, 1973; Hsu, 1982; Huang et al., 2005; Casey et al., 2020).

Through the dynamical systems lens, one can interpret the observed heterogeneity in single cell data to mean that each cell occupies a point in a high-dimensional gene expression space (Trapnell, 2015). That is, if there are 20,000 genes in the genome, each cell is a point in a 20,000-dimensional space. The cell's state is defined by this point in the total gene expression state space. Mapping cells onto a point in this geometrical state space consistently demonstrated that certain regions in the gene expression state space are more likely to be occupied by cells compared to other regions (Moris et al., 2016). These regions are referred to as "attractors" (Huang et al., 2005). Cells that are found within the same attractor space can be categorized as one group (or fate in a developmental context).

Attractors, akin to the valleys in Waddington's landscape, are produced by the activities of gene regulatory networks that promote certain transcriptional outputs (Moris et al., 2016). Since gene regulatory networks have evolved feedback mechanisms to stabilize their activities

(Britten and Davidson, 1969; Kauffman, 1971, 1969), cells within attractors are considered to be in steady states of gene expression (Hsu, 1982; Moris et al., 2016). Transitions between different states have been proposed to arise from changes in the cell's environment leading to changes in the expression of gene regulatory network components (Britten and Davidson, 1969). This view is now supported by extensive experimental evidence pointing to the role of transcription factors that change their expression levels and/or activities to mediate the effects of various signals (will be discussed in later sections of this chapter). Such changes are thought to trigger a cell to exit the previous steady state and enter a new steady state through one or more intermediate transition states (Moris et al., 2016).

Concepts from dynamical systems theory can therefore be used to understand how cell fates are established, and how cells transit between multiple states on the course of its differentiation. Recently, a study that interpreted qunatitative data obtained from programmed differentiation of pluripotent stem cells *in vitro* using concepts from dynamical systems theory revealed two distinct transition paths that cells take to transition to either neural or mesoderm fates. Strikingly, this model informed by experimental data and theory recapitulated the cell fate outcomes of pluripotent stem cells exposed to different concentrations of signaling factors (Sáez et al., 2022). It would be challenging to apply the same approach to *in vivo* developing tissues due to the difficulties associated with controlling and measuring multiple signaling inputs and tracking cell fate transitions with sufficiently high spatial and temporal resolution. Nevertheless, such studies promise to produce predictive insights for deciphering the geometries of cell fate transitions and are currently being attempted (J Briscoe, personal communication).

1.2.5 Discoveries in Cell Fate Transitions: What's Next?

Together, conceptual and experimental progress in developmental biology influenced and complemented each other in several ways. Waddington's epigenetic landscape provided an

early conceptual framework for understanding cell fate decisions, emphasizing the interplay between genes. Developmental genetics studies such as that of Nusslein-Volhard and Weischaus produced the earliest evidence for developmental genes, and mapped them onto regulatory networks. Recent studies using unbiased sequencing technologies have expanded our understanding of cell types by providing a large-scale characterization of gene expression at single cell resolution. The integration of dynamical systems theory further expanded the understanding of cell fate transitions by providing useful framework to think about cellular heterogeneity (gene expression state space) and cell fates (stable attractor states).

While high-throughput approaches, including but not limited to scRNA-seq, have gained enormous popularity over the past decade for studying cell fate decisions, these techniques often overlook multiple layers of cell fate regulation. The limitations of high-throughput sequencing approaches in investigating cell fate transitions are significant. They include the absence of experimentally-derived temporal and spatial information, the inability to detect low abundance transcripts (Wheat et al., 2020), and the oversight of transcription-independent modes of regulation, such as post-translational modifications and protein-protein interactions in sub-cellular compartments.

Recent advances in quantitative microscopy techniques have emerged to address some of these limitations. These include the advent of single-molecule RNA fluorescent *in situ* hybridization (smFISH) (Femino et al., 1998; Raj et al., 2008) and its more high-throughput versions, in which barcoded smFISH probes are used to measure the expression of hundreds of mRNA species in a single fixed tissue (Zhu et al., 2018; Chen et al., 2015). These tools provide an unprecedented opportunity to investigate the regulation of developmental genes within cells in their endogenous environments, which is critical to capture the full biological complexities of cell fate decision-making, and for better understanding emerging concepts such as tissue self-organization that may underlie cell fate precision and robustness in developing organs (Kim et al., 2018).

Additionally, advances in live imaging of tagged proteins and mRNA expression *in vivo* tissues such as the Drosophila embryo have provided novel aspects of gene expression regulation (these studies will be discussed in the later section "Dynamic Regulation of Genes by Transcription Factors"). Such studies as well as the study by Sáez et al. (2022) mark a new era in studying the dynamic control of gene expression and cell fate transitions.

1.3 Regulation by Transcription Factors: Moving Beyond "On or Off" Description to Quantifying Input-Output Functions

1.3.1 The Role of Transcription Factors in the Flow of Information

Transcription factors mediate the flow of genetic information (Figure 1.2) that underlie almost all cellular functions we see in biology. They do so by activating or repressing the transcription of target genes, including their own genes and those encoding other transcription factors. The complex interplay of transcription factors are commonly represented in the form of gene regulatory network diagrams (Figure 1.2). In these diagrams, the conventional arrow (with a pointy end) is used to depict the activating input of a transcription factor on a target gene, while another type of arrow (with a flat end) denotes the repressive input (Figure 1.2). From this representation, one can conclude that transcription factor activities are binary: they are either activating or repressive. While this representation is useful in understanding their general mechanism of action as well as their place within a larger network of interacting components, this view has a limited potential to explain how the diversity in gene expression arises under different cellular contexts through the activities of a few common transcription factors.

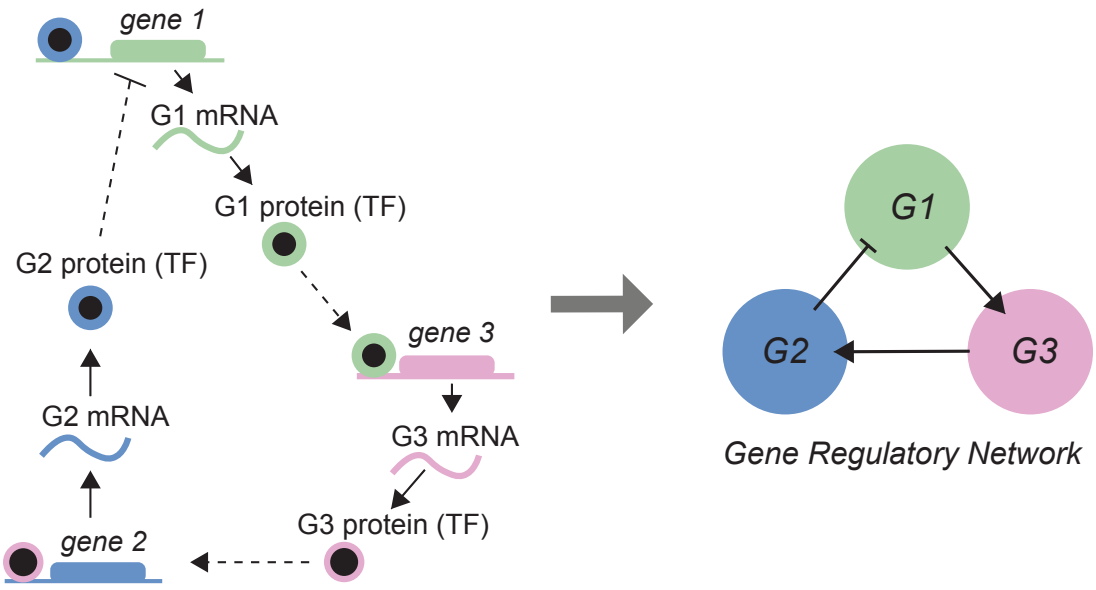

Figure 1.2: Transcription Factor Activities Represented by a Gene Regulatory Network Diagram

Left: the role of transcription factors (TF) in the flow of information from gene to mRNA to protein. Colored circles with a black center denote TF proteins that bind to regulatory regions of target genes. Right: a schematic of a gene regulatory network composed of the three transcription factors encoded by genes 1-3 shown on the left. This network diagram describes the hypothetical interactions between the three transcription factors, G1-G3, involving two positive regulations (G1 to G3; G3 to G2) and a negative regulation (G2 to G1). Adapted from Huynh-Thu and Sanguinetti (2019)

1.3.2 Common Transcription Factors and Where to Find Their Specificity

Context-specific activities of transcription factors arise from the interplay of their subcellular localization, protein-protein interactions, protein-DNA interactions, and their local concentrations. For example, changing the subcellular localization of a transcription factor could influence its local concentration without changing its total concentration in the cell. Protein-protein interactions can affect protein-DNA interactions by the dimerization of transcription factors that stabilize its interaction with the DNA (Lachance et al., 2018), or by the binding of co-factors that recruit transcription factors to new target loci (Webber et al., 2018).

Additionally, all of these features have been shown to be regulated by extracellular signals during cell state transitions. A common mode of regulation by which extracellular signals affect the behavior of transcription factors is through post-translational modifications such as phosphorylation or dephosphorylation. These modifications can directly alter the localization of transcription factors. An example is the Nuclear Factor of Activated T cells (NF-AT) family of transcription factors in vertebrates. The localization of these transcription factors are regulated by calcium signaling: the serine residues on its N-terminus is dephosphorylated by calcinuerin in the presence of high intracellular calcium leading to the accessibility of its nuclear localization signal (NLS) and nuclear import (Zhu et al., 1998), while its phosphorylation in the nucleus by kinases leads to the accessibility of its nuclear export signal (NES) and subsequent nuclear export (Beals et al., 1997). Furthermore, phosphorylation/ dephosphorylation of transcription factors can affect their activities by affecting protein-protein interaction with itself (Lluís et al., 2005; Braunstein et al., 2003) and other co-regulators (Janknecht and Hunter, 1996; Parker et al., 1998; Chawla et al., 1998), targeting or protecting them from proteolytic machinery (Maniatis, 1999; Pahl and Baeurle, 1996), and changing their binding strength to DNA (Dérijard et al., 1994; Sakamoto et al., 1997; Kasibhatla et al., 1999). These effects can be attributed to the ability of these post-translational modifications to induce protein conformational changes as well as altering charges on residues on DNA

binding or multimerization interfaces.

Another layer of complexity stems from the cis-regulatory syntax of gene enhancers, which deals with the number, affinity, position, spacing, and orientation of transcription factor binding sites (Siggers and Gordân, 2014). The interplay between enhancer-specific syntax and available transcription factors in the nucleus has been shown to control where and when genes are expressed in numerous instances (Bothma et al., 2015; Farley et al., 2016; Parker et al., 1998; Crocker and Ilsley, 2017; Lachance et al., 2018).

Collectively, these regulatory mechanisms can be used to tune the space, time, and amount of target gene transcription. These mechanisms highlight the complex and context-dependent nature of transcription factor activities that are not captured by gene regulatory network diagrams. As such, we are currently heading towards the use of increasingly quantitative approaches where different regulatory parameters that tune transcription factor activities are measured and/or perturbed while also measuring their effects on output transcription dynamics and cell fate.

1.3.3 Dynamic Regulation of Genes by Transcription Factors

Recent advances in protein labelling techniques and imaging modalities have provided an unprecedented opportunity to study transcription factor dynamics with high spatial and temporal resolution. These technologies have enabled the measurement of transcription factor concentrations, diffusion time, and degradation time - all of which are crucial for developing an intuitive understanding of transcription factor activities.

In addition, we can now measure mRNA expression as a direct read out of transcription factor activities at particular loci. The first tool to make this a possibility is single molecule fluorescent in-situ hybridization (smFISH). smFISH is performed on fixed tissues, which allows the measurement of both cytoplasmic mRNA as well as nascent transcripts from DNA loci at single molecule resolution *in vivo* developing tissues. smFISH data is used to

measure transcriptional activities at both binary and continuous scales. The presence or absence of a bright transcription spot in the nucleus is used to assess whether or not a locus is engaged in transcription. In the meantime, fluorescence intensities scales with increasing number of nascent RNA molecules at active transcription sites, and therefore provides an estimate of the number of bound RNA Polymerase II complexes that are present at the moment of tissue fixation.

Another powerful technique to measure gene expression is the MS2-MCP live imaging system. This tool involves tagging the end of a protein coding sequence with repeats of virally derived MS2 stem-loop sequences which are recognized by the viral protein MCP. As the tagged mRNA of interest is transcribed, multiple MCP proteins bind to the nascent transcripts allowing researchers to visualize the dynamics of transcription in real time.

Together, these techniques have resulted in a significant discovery that transcription is a dynamic process characterized by transient bursts of mRNA production interspersed by periods of quiescence (Golding et al., 2005; Paré et al., 2009). Transcriptional bursting can be characterized by its constituent features, the rate of transcription and the frequency of bursting, both of which influence total synthesized mRNA (Lammers et al., 2020). Emerging studies in syncytial Drosophila embryos revealed that transcriptional dynamics of zygotic genes are fine-tuned by the activity of maternally deposited transcription factors (Jaeger et al., 2004; Gregor et al., 2007; Liu et al., 2013; Crauk and Dostatni, 2005). These studies have provided a quantitative framework to characterize transcriptional bursting and provided insight on its regulation by transcription factors (reviewed in (Gregor et al., 2014) and (Garcia et al., 2020)). However, they are limited to pre-cellularized embryos in which all nuclei share a common cytoplasm, long before cells commit to a particular fate.

1.4 A Model of Precision: Cell Fate Transitions in the Developing Drosophila Eye

1.4.1 Building A Neurocrystalline Lattice: Drosophila Compound Eye Development

In a 1976 paper that carefully studies the structure of the Drosophila compound eye (Ready et al., 1976), Ready and his colleagues states that the eye is akin to a crystal lattice. The compound eye is a highly ordered array of 750-800 optical units called ommatidia. In each ommatidium, there is a precise complement of eight photoreceptor neuron subtypes R1-R8, along with four lens-secreting cone cells and pigment cells (Ready et al., 1976; Wolff and Ready, 1991). The development of the Drosophila eye has emerged as a popular model system for studying the genetic mechanisms underlying developmental decisions. Several factors contribute to its popularity. Firstly, any disruptions to cell fate determination result in a distinct "rough eye" phenotype, easily distinguishable from the highly organized wild-type eye. Secondly, genetic perturbations specific to the eye do not compromise the overall viability of the fly, enabling the consequences of each perturbation to be characterized at various developmental stages of the tissue. Lastly, the acquisition of distinct cell types within the eye relies on the intricate interplay of multiple signaling pathways and gene regulatory networks.

1.4.2 Signaling the Way: The Wave of Differentiation in the Developing Eye Requires the Interplay of Multiple Signaling Pathways

Development of the compound eye is initiated at the larval stage in a specialized epithelial tissue called the eye imaginal disc (Tomlinson and Ready, 1987). Initially, the eye epithelium is a disorganized monolayer of asynchronously dividing progenitor cells. At the third

instar larval stage, the morphogenetic furrow (MF), a pro-differentiation wave marked by a dorsoventral indentation of the epithelium, emerges at the posterior margin of the tissue. The MF moves at a constant velocity towards the anterior of the epithelium, inducing progenitor to photoreceptor transitions upon its wake.

The initiation and propagation of the MF are intricately regulated by several signaling pathways. Positive control is exerted by Hedgehog (Hh) signaling (Ma et al., 1993), Epidermal Growth Factor Receptor (EGFR) signaling, and Decapentaplegic (Dpp) signaling (Heberlein et al., 1993a), while negative control is influenced by Wingless (Wg) signaling (Treisman and Rubin, 1995). Notch signaling also contributes to the initiation of the MF (Kumar and Moses, 2001). The Hh signal originating from the MF instructs cells at and just anterior to the MF to differentiate into photoreceptor neurons by activating the expression of Atonal. The Atonal-expressing photoreceptor neurons, known as R8 cells, release the EGFR ligand Spitz, which activates the Receptor Tyrosine Kinase (RTK) signaling pathway in neighboring cells. This leads to the activation of the RTK transcription activator Pnt, which in turn promotes the expression of *hh* in these cells, creating a feedback loop of Hh signaling that drives the anterior propagation of the MF (Roignant and Treisman, 2009).

Furthermore, the pace and spacing of the advancing MF are controlled by both short and long-range interactions. Hh is also involved in short-range signaling of Dpp (Heberlein et al., 1993b; Burke and Basler, 1996), which in turn activates long-range Dpp signaling, leading to the repression of Homeothorax (Hth) expression (Bessa et al., 2002). The inhibition of Hth allows for sustained Hh signaling and contributes to the progression of the MF. Additionally, Wg signaling promotes the expression of Hth from the anterior side, further influencing the dynamics of the MF (Pichaud and Casares, 2000; Tsai et al., 2007). The complex interplay between these signaling pathways, as described above, is summarized in Figure 1.2A, illustrating the intricate regulatory network governing the initiation and propagation of the morphogenetic furrow in the Drosophila eye disc.

The R8 cells are the founder cells, recruiting the rest of the photoreceptor fates to a growing ommatidium. The R8 cells are specified via the expression of Atonal, which is broadly expressed in the proneural region anterior to the MF, but becomes restricted to one R8 cell per ommatidium due to lateral inhibition by Notch (Figure 1.2B) (Chen and Chien, 1999; Baker and Zitron, 1995; Cagan and Ready, 1989).

After the specification of an R8 cell in each ommatidium, the remaining photoreceptor fates are recruited in a temporally ordered fashion: R2/R5, then R3/R4, then R1/R6 and finally, the R7 (Tomlinson and Ready, 1987). This temporal order is spatially captured within the eye disc epithelium, due to the progressive movement of the MF from the posterior to the anterior boundary of the tissue (Wolff and Ready, 1991; Wolff, 1993). Additionally, the nuclei of progenitor cells as well as photoreceptor cells each exhibit a highly stereotypical position and morphology within the tissue. As such, a single fixed eye disc offers a unique opportunity to examine the temporal progression of transcription factors and their target gene expression during progenitor to various photoreceptor fate transitions (Peláez et al., 2015; Bernasek et al., 2023). This is discussed in further detail in the introduction of Chapter 2.

1.4.3 Reiterative Use of RTK/MAPK Signaling Pathway Directs Progenitor Cell Transitions to Photoreceptor Neurons

Extensive communication between cells in the developing eye disc is mediated by inductive RTK/MAPK signaling via the EGF and Sevenless (Sev) receptors. As such, RTK signaling is used repeatedly to form all photoreceptor fates (R1-R7) except the R8 founder cell (Freeman, 1996). The activation of epidermal growth factor receptor (EGFR), a prototypical RTK, allows progenitor cells to transition to specified R cells (Baker and Rubin, 1989; Xu and Rubin, 1993; Freeman, 1996). Activation of EGFR is achieved by the secreted ligand Spitz from the neighboring R8 cell and potentially other specified R cells (Freeman, 1994; Tio

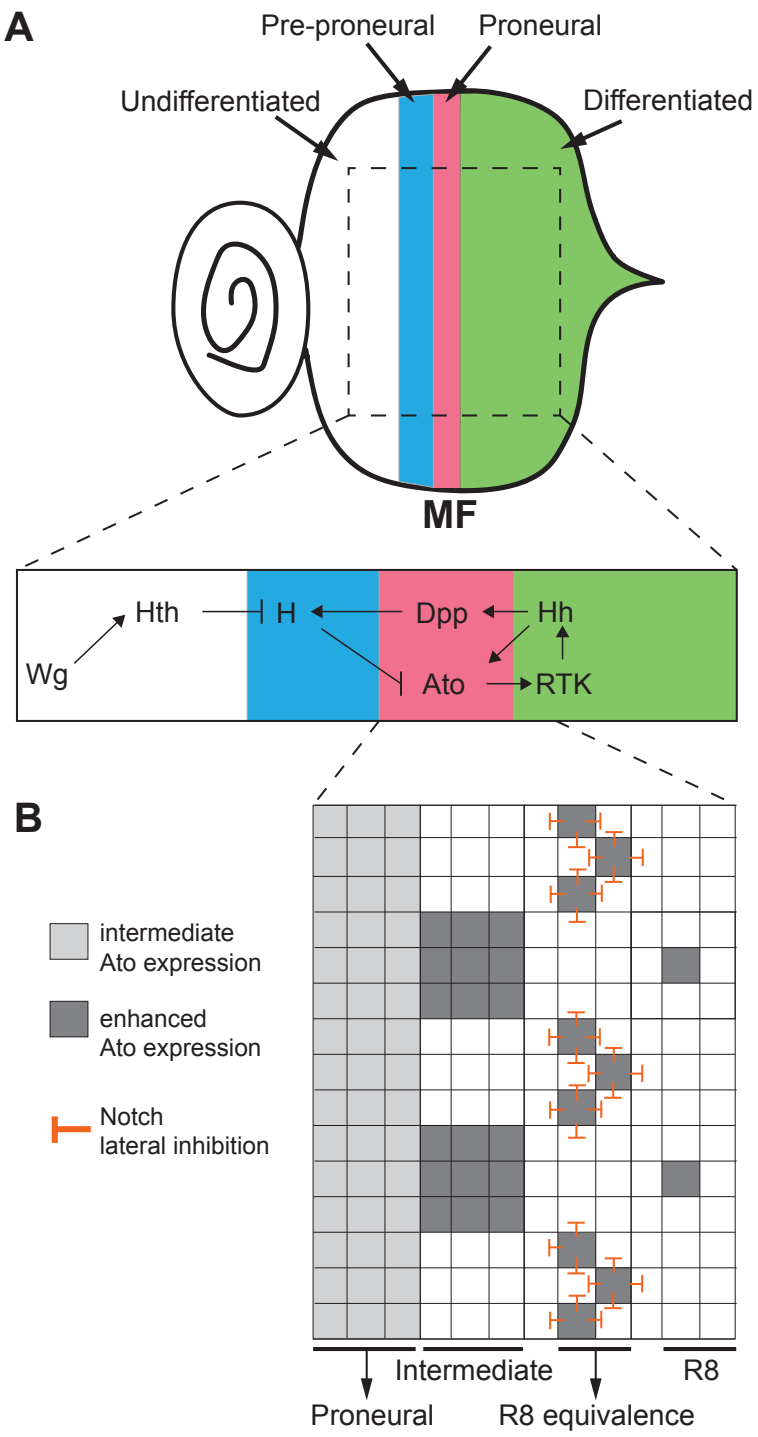

Figure 1.3: Complex interplay of Hh, RTK, Dpp, Wg, and Notch signaling in the Drosophila eye imaginal disc

A. Summary of major signaling pathways that set up distinct developmental stages across the eye disc and move the MF forward (adapted from (Greenwood and Struhl, 1999); B. Regular spacing of the founder R8 cells that initiate the assembly of ommatidia is mediated by Atonal expression and Notch signaling.

et al., 1994; Tio and Moses, 1997). Spitz is a short-range (1-2 cell diameter) acting ligand produced by R8 as well as subsequently specified photoreceptor cells. The R7 cell requires an additional Sev receptor (Campos-Ortega and Hofbauer, 1977), which is activated by another R8 secreted RTK ligand, Bride of Sevenless (Boss) (Reinke and Zipursky, 1988).

EGFR/Sev activation triggers the kinase cascade in which phosphorylated MAPK enters the nucleus, and activate/ repress downstream transcription factors Pointed (Pnt) (O'Neill et al., 1994) and Yan (also known as Anterior Open, or Aop) (Rebay and Rubin, 1995) (Figure 1.4). Regulation of photoreceptor fates by these transcription factors is the main focus of my thesis research. I will synthesize current knowledge on these transcription factors in the subsequent section.

Reiterative usage of the same RTK/MAPK signaling prompts the question of how the distinct photoreceptor fates are specified in the developing eye. It was initially thought that RTK/MAPK signaling acts as a mere survival cue permitting progenitor cells to receive further instructive cues that specify them into specific photoreceptor fates. Indeed, RTK/MAPK signaling prevents apoptosis through mechanisms such as downregulating the pro-apototic gene *hid* (Bergmann et al., 1998; Kurada and White, 1998). Nevertheless, the observation that progenitor cells cannot transition to any of the photoreceptor fates except R8 in Egfr null clones even when apoptosis is prevented (p35-expression background) (Baonza et al., 2001; Baker and Yu, 1997) suggested that signaling through the EGF and Sev receptors provides some level of instruction in addition to acting as a pro-survival cue.

The current understanding is that a combination of temporally and spatially controlled signals including RTK/MAPK and Notch signaling creates unique combinations of expressed transcription factors in subsets of progenitor cells that instruct them to adopt a particular photoreceptor fate over others. The combinatorial code that instructs each photoreceptor fate is extensively reviewed in Voas and Rebay (2004). Furthermore, recent scRNA-seq data from the Mardon lab has produced an unbiased, global profiling of genes expressed in indi-

vidual cells of the developing eye disc (Bollepogu Raja et al., 2023). This dataset contains previously identified genes underlying each photoreceptor fate's combinatorial code as well as novel genes that are uniquely expressed in each photoreceptor fate. I provide an updated summary in Table 1.1.

Cell Fate	RTK signal	Notch signal	Cell Fate Determinant
R8	Inactive	Active	Atonal, Salm
R2/R5	Active (EGFR)	Inactive	Rough, Liprin-γ
R3/R4	Active (EGFR)	Inactive	Svp, Salm, Prickle, DIP-δ (R3), CG4341 (R4)
R1/R6	Active (EGFR)	Inactive	Lozenge, Svp, BarH1,2
R7	Active (EGFR, Sev)	Active	Lozenge, Prospero, Salm

Table 1.1: Combinatorial Code for the Specification of Photoreceptor Fates Updated from Voas and Rebay (2004). Citations **R8**:Dokucu et al. (1996); Domingos et al. (2004a); **R2/R5**: Kimmel et al. (1990); Heberlein et al. (1991); Bollepogu Raja et al. (2023); **R3/R4**: Domingos et al. (2004b); Mlodzik et al. (1990); Rawls and Wolff (2003); Bollepogu Raja et al. (2023); **R1/R6**: Daga et al. (1996); Flores et al. (1998, 2000); **R7**: Lai et al. (1996); Domingos et al. (2004a); Xu et al. (2000)

1.4.4 Transcriptional Regulation by the Pnt and Yan Network

Owing to the numerous developmental decisions they regulate in Drosophila, *pnt* and *yan* (also known as *anterior open, aop*) genes were also identified by the Heidelberg screen (Wieschaus and Nüsslein-Volhard, 2016). *pnt* was named because of its hypomorphic phenotype in the embryo which produces a "pointy" head skeleton, while *yan/aop* was named because of the dorsal anterior deletion in the embryo (Wieschaus and Nüsslein-Volhard, 2016). It is now known that proteins encoded by the *pnt* and *yan* (*aop*) genes are members of the E26 Transformation Specific (ETS) family of transcription factors. Pnt and Yan bind DNA via their shared ETS binding domain, a conserved helix-turn-helix DNA binding domain of ∼85 amino acids in length (Wei et al., 2010; Yordy and Muise-Helmericks, 2000) that recognizes

a core 'GGAA/T' consensus sequence (Wei et al., 2010).

There are three functionally characterized Pnt protein isoforms: PntP1, PntP2, and PntP3. They share an identical C-terminal region which includes the ETS DNA binding domain but differ in their N-termini. While PntP1 is constitutively active when expressed, PntP1 and P2 requires activation by MAPK phosphorylation in their unique N-termini (Brunner et al., 1994; O'Neill et al., 1994; Shwartz et al., 2013; Wu et al., 2020). Nevertheless, they have the ability to bind to the same target genes, and all act as general transcriptional activators. Meanwhile, Yan acts as a general repressor (Rebay and Rubin, 1995). Two protein domains are highly relevant for its function: the ETS DNA binding domain and the SAM domain. The SAM domain facilitates multimerization of Yan monomers leading its self-association, which is essential for its repressive ability at target genes (Qiao et al., 2004; Zhang et al., 2010).

Pnt and Yan act as nuclear effectors of the RTK/MAPK signaling pathway and govern numerous developmental decisions across different stages and tissues, including but not limited to, the differentiation and maintenance of neuroectodermal cells during embryonic head development (Dumstrei et al., 1998), follicle cell migration during oogenesis (Morimoto et al., 1996), branching of cells during tracheal development (Ohshiro et al., 2002), specification of cardiac and muscle cell fates in the embryonic mesoderm (Halfon et al., 2000), as well as the specification of photoreceptor fates in the developing eye (O'Neill et al., 1994; Rebay and Rubin, 1995).

Yan-mediated repression is relieved through RTK signaling. In the absence of RTK stimulation, Yan is localized in the nucleus, exerting transcriptional repression on its target genes. However, when RTK is activated, Yan undergoes phosphorylation by the MAP kinase Rolled, which reduces its repressive activity (Rebay and Rubin, 1995). Phosphorylation of Yan by MAPK is facilitated by another protein called Modulator of the Activity of ETS (Mae) (Baker et al., 2001). Mae binds Yan monomers with around 1000-fold higher affinity

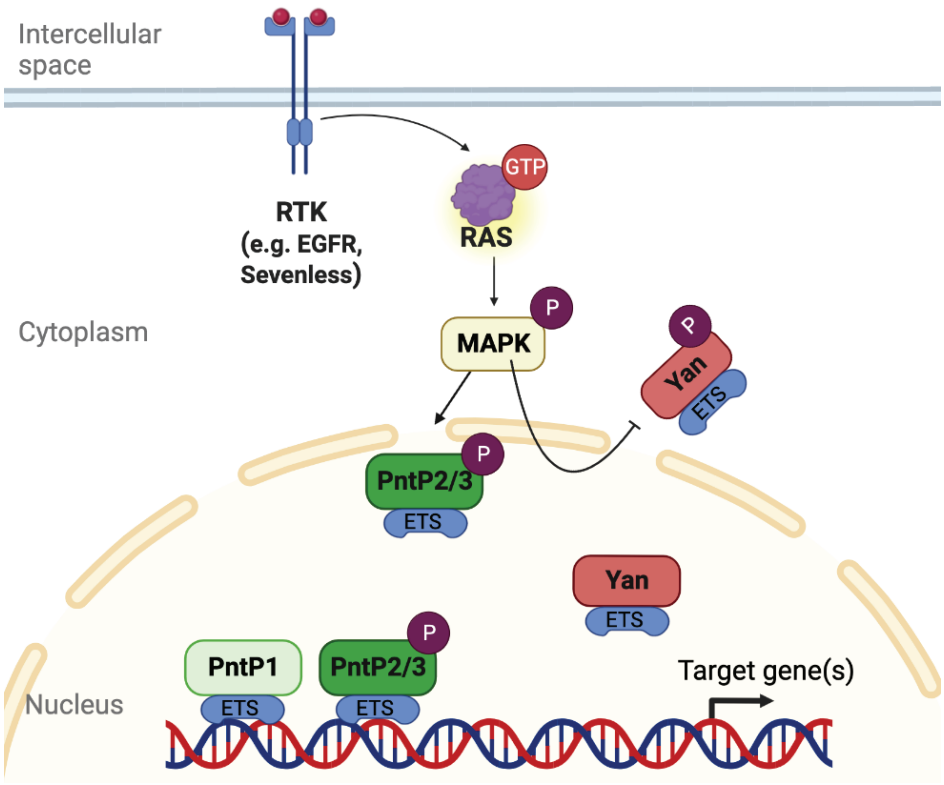

Figure 1.4: Pnt and Yan are downstream effectors of the RTK signaling pathway
Schematic of the RTK signaling pathway

via its SAM domain (Qiao et al., 2004) leading to Yan complex disassembly. Phosphorylation of Yan leads to its export from the nucleus, mediated by an exportin protein called CRM1 (Rebay and Rubin, 1995; Tootle et al., 2003). It is thought that the removal of Yan from the nucleus allows Pnt to occupy the sites previously occupied by Yan, resulting in transcriptional activation (Gabay et al., 1996; O'Neill et al., 1994; Klämbt, 1993).

1.4.5 How Do Antagonistic Transcription Factors Pnt and Yan Control Progenitor to Photoreceptor Cell Fate Transitions?

Previously, it was thought that Pnt and Yan display mutually exclusive expression patterns to regulate cell fate transitions (Graham et al., 2010). However, qualitative characterization of their expression patterns across multiple developing Drosophila tissues by the Rebay lab

has led to the surprising discovery that Pnt and Yan are co-expressed in cells during developmental transitions (Boisclair Lachance et al., 2014). This finding was further supported by a follow-up study (Peláez et al., 2015) that characterized Yan expression dynamics by measuring Yan levels in individual progenitor and photoreceptor cells across the eye disc.

How Pnt levels change during developmental fate transitions alongside changes in Yan levels has not been previously examined. Furthermore, the observation that Pnt and Yan are both present in the nuclei of progenitor and photoreceptor cells at the time of photoreceptor transitions prompt the question of how opposing inputs from co-expressed Pnt and Yan levels are integrated to drive distinct transcriptional and cell fate outcomes.

1.5 Scope of the Thesis

The central aim of my thesis research is to understand how opposing inputs from co-expressed transcriptional activator Pnt and repressor Yan are integrated to inform precise and robust transitions of multipotent progenitor cells to specified photoreceptor fates during Drosophila eye development. Specifically, I examine how nuclear levels of Pnt and Yan proteins are integrated at the cellular level to control the progenitor vs. photoreceptor fate decision (Chapter 2), and at the level of individual DNA loci to control the probability and/or the amount of transcription of photoreceptor-fate determining genes.

Work presented in this thesis combines genetics, quantitative microscopy, computational analysis, and computational modeling approaches to uncover novel insights into how the nuclear effectors of the highly-conserved RTK/MAPK signaling pathway regulate cell fate decisions. While my research is primarily focused on Drosophila eye development, the regulatory strategies unveiled in this study may hold broader relevance to other systems, including but not limited to other RTK/MAPK signaling contexts, in which accurate cell fate transitions rely on inputs from multiple antagonistic transcription factors.

1.6 References

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(6835):330–334.

Baker, N. E. and Rubin, G. M. (1989). Effect on eye development of dominant mutations in drosophila homologue of the egf receptor. *Nature*, 340(6229):150–153.

Baker, N. E. and Yu, S.-Y. (1997). Proneural function of neurogenic genes in the developing drosophila eye. *Current Biology*, 7(2):122–132.

Baker, N. E. and Zitron, A. E. (1995). Drosophila eye development: Notch and delta amplify a neurogenic pattern conferred on the morphogenetic furrow by scabrous. *Mechanisms of development*, 49(3):173–189.

Baonza, A., Casci, T., and Freeman, M. (2001). A primary role for the epidermal growth factor receptor in ommatidial spacing in the drosophila eye. *Current Biology*, 11(6):396–404.

Beadle, G. W. and Tatum, E. L. (1941). Genetic control of biochemical reactions in neurospora. *Proceedings of the National Academy of Sciences*, 27(11):499–506.

Beals, C. R., Sheridan, C. M., Turck, C. W., Gardner, P., and Crabtree, G. R. (1997). Nuclear export of nf-atc enhanced by glycogen synthase kinase-3. *Science*, 275(5308):1930–1933.

Bergmann, A., Agapite, J., McCall, K., and Steller, H. (1998). The drosophila gene hid is a direct molecular target of ras-dependent survival signaling. *Cell*, 95(3):331–341.

Bernasek, S. M., Hur, S. S., Peláez-Restrepo, N., Lachance, J.-F. B., Bakker, R., Navarro, H. T., Sanchez-Luege, N., Amaral, L. A., Bagheri, N., Rebay, I., et al. (2023). Ratiometric

sensing of pnt and yan transcription factor levels confers ultrasensitivity to photoreceptor fate transitions in drosophila. *Development*, pages dev–201467.

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 & development*, 16(18):2415–2427.

Bier, E., Jan, L. Y., and Jan, Y. N. (1990). rhomboid, a gene required for dorsoventral axis establishment and peripheral nervous system development in drosophila melanogaster. *Genes & development*, 4(2):190–203.

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. *Developmental Biology*, 385(2):263–278.

Bollepogu Raja, K. K., Yeung, K., Shim, Y. K., Li, Y., Chen, R., and Mardon, G. (2023). A single cell genomics atlas of the drosophila larval eye reveals distinct developmental timelines and novel markers for all photoreceptor subtypes. *bioRxiv*, pages 2023–02.

Bothma, J. P., Garcia, H. G., Ng, S., Perry, M. W., Gregor, T., and Levine, M. (2015). Enhancer additivity and non-additivity are determined by enhancer strength in the *Drosophila* embryo. *eLife*, 4:e07956. Publisher: eLife Sciences Publications, Ltd.

Braunstein, J., Brutsaert, S., Olson, R., and Schindler, C. (2003). Stats dimerize in the absence of phosphorylation. *Journal of Biological Chemistry*, 278(36):34133–34140.

Britten, R. J. and Davidson, E. H. (1969). Gene regulation for higher cells: A theory: New facts regarding the organization of the genome provide clues to the nature of gene regulation. *Science*, 165(3891):349–357.

Brunner, D., Dücker, K., Oellers, N., Hafen, E., Scholzi, H., and Klambt, C. (1994). The ets domain protein pointed-p2 is a target of map kinase in the sevenless signal transduction pathway. *Nature*, 370(6488):386–389.

Burke, R. and Basler, K. (1996). Hedgehog-dependent patterning in thedrosophilaeye can occur in the absence of dpp signaling. *Developmental biology*, 179(2):360–368.

Cagan, R. L. and Ready, D. F. (1989). Notch is required for successive cell decisions in the developing drosophila retina. *Genes & development*, 3(8):1099–1112.

Campos-Ortega, J. and Hofbauer, A. (1977). Cell clones and pattern formation: On the lineage of photoreceptor cells in the compound eye of drosophila. *Wilhelm Roux's archives of developmental biology*, 181:227–245.

Cao, J., Packer, J. S., Ramani, V., Cusanovich, D. A., Huynh, C., Daza, R., Qiu, X., Lee, C., Furlan, S. N., Steemers, F. J., et al. (2017). Comprehensive single-cell transcriptional profiling of a multicellular organism. *Science*, 357(6352):661–667.

Casey, M. J., Stumpf, P. S., and MacArthur, B. D. (2020). Theory of cell fate. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*, 12(2):e1471.

Chawla, S., Hardingham, G. E., Quinn, D. R., and Bading, H. (1998). Cbp: a signal-regulated transcriptional coactivator controlled by nuclear calcium and cam kinase iv. *Science*, 281(5382):1505–1509.

Chen, C.-K. and Chien, C.-T. (1999). Negative regulation of atonal in proneural cluster formation of drosophila r8 photoreceptors. *Proceedings of the National Academy of Sciences*, 96(9):5055–5060.

Chen, K. H., Boettiger, A. N., Moffitt, J. R., Wang, S., and Zhuang, X. (2015). Spatially resolved, highly multiplexed rna profiling in single cells. *Science*, 348(6233):aaa6090.

Crauk, O. and Dostatni, N. (2005). Bicoid Determines Sharp and Precise Target Gene Expression in the Drosophila Embryo. *Current Biology*, 15(21):1888–1898.

Crocker, J. and Ilsley, G. R. (2017). Using synthetic biology to study gene regulatory evolution. *Current opinion in genetics & development*, 47:91–101.

Daga, A., Karlovich, C. A., Dumstrei, K., and Banerjee, U. (1996). Patterning of cells in the drosophila eye by lozenge, which shares homologous domains with aml1. *Genes & development*, 10(10):1194–1205.

Deng, Q., Ramsköld, D., Reinius, B., and Sandberg, R. (2014). Single-cell rna-seq reveals dynamic, random monoallelic gene expression in mammalian cells. *Science*, 343(6167):193–196.

Dérijard, B., Hibi, M., Wu, I.-H., Barrett, T., Su, B., Deng, T., Karin, M., and Davis, R. J. (1994). Jnk1: a protein kinase stimulated by uv light and ha-ras that binds and phosphorylates the c-jun activation domain. *Cell*, 76(6):1025–1037.

Dokucu, M. E., Zipursky, S. L., and Cagan, R. L. (1996). Atonal, rough and the resolution of proneural clusters in the developing drosophila retina. *Development*, 122(12):4139–4147.

Domingos, P. M., Brown, S., Barrio, R., Ratnakumar, K., Frankfort, B. J., Mardon, G., Steller, H., and Mollereau, B. (2004a). Regulation of r7 and r8 differentiation by the spalt genes. *Developmental biology*, 273(1):121–133.

Domingos, P. M., Mlodzik, M., Mendes, C. S., Brown, S., Steller, H., and Mollereau, B. (2004b). Spalt transcription factors are required for r3/r4 specification and establishment of planar cell polarity in the drosophila eye.

Dumstrei, K., Nassif, C., Abboud, G., Aryai, A., Aryai, A., and Hartenstein, V. (1998). Egfr signaling is required for the differentiation and maintenance of neural progenitors along the dorsal midline of the drosophila embryonic head. *Development*, 125(17):3417–3426.

Farley, E. K., Olson, K. M., Zhang, W., Rokhsar, D. S., and Levine, M. S. (2016). Syntax compensates for poor binding sites to encode tissue specificity of developmental enhancers. *Proceedings of the National Academy of Sciences*, 113(23):6508–6513.

Femino, A. M., Fay, F. S., Fogarty, K., and Singer, R. H. (1998). Visualization of Single RNA Transcripts in Situ. *Science*, 280(5363):585–590. Publisher: American Association for the Advancement of Science Section: Report.

Ferguson, E. L. and Anderson, K. V. (1992). Decapentaplegic acts as a morphogen to organize dorsal-ventral pattern in the drosophila embryo. *Cell*, 71(3):451–461.

Flores, G. V., Daga, A., Kalhor, H. R., and Banerjee, U. (1998). Lozenge is expressed in pluripotent precursor cells and patterns multiple cell types in the drosophila eye through the control of cell-specific transcription factors. *Development*, 125(18):3681–3687.

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(1):75–85.

Freeman, M. (1994). The spitz gene is required for photoreceptor determination in the drosophila eye where it interacts with the egf receptor. *Mechanisms of development*, 48(1):25–33.

Freeman, M. (1996). Reiterative use of the egf receptor triggers differentiation of all cell types in the drosophila eye. *Cell*, 87(4):651–660.

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(11):3355–3362.

Garcia, H. G., Berrocal, A., Kim, Y. J., Martini, G., and Zhao, J. (2020). Lighting up the central dogma for predictive developmental biology. In *Current Topics in Developmental Biology*, volume 137, pages 1–35. Elsevier.

Glass, L. and Kauffman, S. A. (1973). The logical analysis of continuous, non-linear biochemical control networks. *Journal of theoretical Biology*, 39(1):103–129.

Golding, I., Paulsson, J., Zawilski, S. M., and Cox, E. C. (2005). Real-Time Kinetics of Gene Activity in Individual Bacteria. *Cell*, 123(6):1025–1036.

Graham, T. G. W., Tabei, S. M. A., Dinner, A. R., and Rebay, I. (2010). Modeling bistable cell-fate choices in the Drosophila eye: qualitative and quantitative perspectives. *Development*, 137(14):2265–2278.

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(24):5795–5808.

Gregor, T., Garcia, H. G., and Little, S. C. (2014). The embryo as a laboratory: quantifying transcription in Drosophila. *Trends in Genetics*, 30(8):364–375.

Gregor, T., Tank, D. W., Wieschaus, E. F., and Bialek, W. (2007). Probing the Limits to Positional Information. *Cell*, 130(1):153–164.

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(1):63–74.

Hartwell, L. H., Culotti, J., and Reid, B. (1970). Genetic control of the cell-division cycle in yeast, i. detection of mutants. *Proceedings of the National Academy of Sciences*, 66(2):352–359.

Heberlein, U., Mlodzik, M., and Rubin, G. M. (1991). Cell-fate determination in the developing drosophila eye: role of the rough gene. *Development*, 112(3):703–712.

Heberlein, U., Wolff, T., and Rubin, G. M. (1993a). The tgfβ homolog dpp and the segment polarity gene hedgehog are required for propagation of a morphogenetic wave in the drosophila retina. *Cell*, 75(5):913–926.

Heberlein, U., Wolff, T., and Rubin, G. M. (1993b). The tgfβ homolog dpp and the segment polarity gene hedgehog are required for propagation of a morphogenetic wave in the drosophila retina. *Cell*, 75(5):913–926.

Hickey, J. W., Tan, Y., Nolan, G. P., and Goltsev, Y. (2021). Strategies for accurate cell type identification in codex multiplexed imaging data. *Frontiers in Immunology*, 12:727626.

Hockfield, S. and McKay, R. (1985). Identification of major cell classes in the developing mammalian nervous system. *Journal of Neuroscience*, 5(12):3310–3328.

Hooper, J. E. and Scott, M. P. (1989). The drosophila patched gene encodes a putative membrane protein required for segmental patterning. *Cell*, 59(4):751–765.

Hsu, C. (1982). A probabilistic theory of nonlinear dynamical systems based on the cell state space concept.

Huang, S., Eichler, G., Bar-Yam, Y., and Ingber, D. E. (2005). Cell fates as high-dimensional attractor states of a complex gene regulatory network. *Physical review letters*, 94(12):128701.

Huynh-Thu, V. A. and Sanguinetti, G. (2019). Gene regulatory network inference: an introductory survey. *Gene Regulatory Networks: Methods and Protocols*, pages 1–23.

Ingham, P. W. and McMahon, A. P. (2001). Hedgehog signaling in animal development: paradigms and principles. *Genes & development*, 15(23):3059–3087.

Jacob, F. and Monod, J. (1961). Genetic regulatory mechanisms in the synthesis of proteins. *Journal of molecular biology*, 3(3):318–356.

Jaeger, J., Surkova, S., Blagov, M., Janssens, H., Kosman, D., Kozlov, K. N., Manu, Myasnikova, E., Vanario-Alonso, C. E., Samsonova, M., Sharp, D. H., and Reinitz, J. (2004). Dynamic control of positional information in the early Drosophila embryo. *Nature*, 430(6997):368–371.

Janknecht, R. and Hunter, T. (1996). Versatile molecular glue. transcriptional control. *Current biology: CB*, 6(8):951–954.

Kasibhatla, S., Tailor, P., Bonefoy-Berard, N., Mustelin, T., Altman, A., and Fotedar, A. (1999). Jun kinase phosphorylates and regulates the dna binding activity of an octamer binding protein, t-cell factor β1. *Molecular and Cellular Biology*, 19(3):2021–2031.

Kauffman, S. (1969). Homeostasis and differentiation in random genetic control networks. *Nature*, 224(5215):177–178.

Kauffman, S. (1971). Gene regulation networks: A theory for their global structure and behaviors. In *Current topics in developmental biology*, volume 6, pages 145–182. Elsevier.

Kim, E. J. Y., Korotkevich, E., and Hiiragi, T. (2018). Coordination of cell polarity, mechanics and fate in tissue self-organization. *Trends in cell biology*, 28(7):541–550.

Kimmel, B. E., Heberlein, U., and Rubin, G. M. (1990). The homeo domain protein rough is expressed in a subset of cells in the developing drosophila eye where it can specify photoreceptor cell subtype. *Genes & Development*, 4(5):712–727.

Klämbt, C. (1993). The drosophila gene pointed encodes two ets-like proteins which are involved in the development of the midline glial cells. *Development*, 117(1):163–176.

Kumar, J. P. and Moses, K. (2001). The egf receptor and notch signaling pathways control the initiation of the morphogenetic furrow during drosophila eye development.

Kurada, P. and White, K. (1998). Ras promotes cell survival in drosophila by downregulating hid expression. *Cell*, 95(3):319–329.

Lachance, J.-F. B., 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 & Development*, 32(5-6):389–401.

Lai, Z.-C., Harrison, S. D., Karim, F., Li, Y., and Rubin, G. M. (1996). Loss of tramtrack gene activity results in ectopic r7 cell formation, even in a sina mutant background. *Proceedings of the National Academy of Sciences*, 93(10):5025–5030.

Lammers, N. C., Kim, Y. J., Zhao, J., and Garcia, H. G. (2020). A matter of time: Using dynamics and theory to uncover mechanisms of transcriptional bursting. *Current Opinion in Cell Biology*, 67:147–157.

Laughon, A. and Scott, M. P. (1984). Sequence of a drosophila segmentation gene: protein structure homology with dna-binding proteins. *Nature*, 310(5972):25–31.

Li, H. (2021). Single-cell rna sequencing in drosophila: Technologies and applications. *Wiley Interdisciplinary Reviews: Developmental Biology*, 10(5):e396.

Liu, F., Morrison, A. H., and Gregor, T. (2013). Dynamic interpretation of maternal inputs by the Drosophila segmentation gene network. *Proceedings of the National Academy of Sciences*, 110(17):6724–6729.

Lluís, F., Ballestar, E., Suelves, M., Esteller, M., and Muñoz-Cánoves, P. (2005). E47 phosphorylation by p38 mapk promotes myod/e47 association and muscle-specific gene transcription. *The EMBO journal*, 24(5):974–984.

Ma, C., Zhou, Y., Beachy, P. A., and Moses, K. (1993). The segment polarity gene hedgehog is required for progression of the morphogenetic furrow in the developing drosophila eye. *Cell*, 75(5):927–938.

Maniatis, T. (1999). A ubiquitin ligase complex essential for the nf-κb, wnt/wingless, and hedgehog signaling pathways. *Genes & development*, 13(5):505–510.

Marioni, J. C., Mason, C. E., Mane, S. M., Stephens, M., and Gilad, Y. (2008). Rna-seq: an assessment of technical reproducibility and comparison with gene expression arrays. *Genome research*, 18(9):1509–1517.

Mayer, U. and Nüsslein-Volhard, C. (1988). A group of genes required for pattern formation in the ventral ectoderm of the drosophila embryo. *Genes & development*, 2(11):1496–1511.

McGinnis, W., Garber, R. L., Wirz, J., Kuroiwa, A., and Gehring, W. J. (1984). A homologous protein-coding sequence in drosophila homeotic genes and its conservation in other metazoans. *Cell*, 37(2):403–408.

Mlodzik, M., Hiromi, Y., Weber, U., Goodman, C. S., and Rubin, G. M. (1990). The drosophila seven-up gene, a member of the steroid receptor gene superfamily, controls photoreceptor cell fates. *Cell*, 60(2):211–224.

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(12):3745–3754.

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

Mulas, C., Chaigne, A., Smith, A., and Chalut, K. J. (2021). Cell state transitions: definitions and challenges. *Development*, 148(20):dev199950.

Nakano, Y., Guerrero, I., Hidalgo, A., Taylor, A., Whittle, J., and Ingham, P. (1989). A protein with several possible membrane-spanning domains encoded by the drosophila segment polarity gene patched. *Nature*, 341(6242):508–513.

Nusse, R. and Varmus, H. E. (1992). Wnt genes. *Cell*, 69(7):1073–1087.

Nüsslein-Volhard, C. and Wieschaus, E. (1980). Mutations affecting segment number and polarity in drosophila. *Nature*, 287(5785):795–801.

Nüsslein-Volhard, C., Wieschaus, E., and Kluding, H. (1984). Mutations affecting the pattern of the larval cuticle in drosophila melanogaster: I. zygotic loci on the second chromosome. *Wilhelm Roux's archives of developmental biology*, 193:267–282.

Ohshiro, T., Emori, Y., and Saigo, K. (2002). Ligand-dependent activation of breathless fgf receptor gene in drosophila developing trachea. *Mechanisms of development*, 114(1-2):3–11.

Olaniru, O. E., Kadolsky, U., Kannambath, S., Vaikkinen, H., Fung, K., Dhami, P., and Persaud, S. J. (2023). Single-cell transcriptomic and spatial landscapes of the developing human pancreas. *Cell Metabolism*, 35(1):184–199.

O'Neill, E. M., Rebay, I., 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(1):137–147.

Packer, J. S., Zhu, Q., Huynh, C., Sivaramakrishnan, P., Preston, E., Dueck, H., Stefanik, D., Tan, K., Trapnell, C., Kim, J., et al. (2019). A lineage-resolved molecular atlas of c. elegans embryogenesis at single-cell resolution. *Science*, 365(6459):eaax1971.

Pahl, H. L. and Baeurle, P. A. (1996). Control of gene expression by proteolysis. *Current opinion in cell biology*, 8(3):340–347.

Parker, D., Jhala, U. S., Radhakrishnan, I., Yaffe, M. B., Reyes, C., Shulman, A. I., Cantley, L. C., Wright, P. E., and Montminy, M. (1998). Analysis of an activator: coactivator complex reveals an essential role for secondary structure in transcriptional activation. *Molecular cell*, 2(3):353–359.

Paré, A., Lemons, D., Kosman, D., Beaver, W., Freund, Y., and McGinnis, W. (2009). Visualization of Individual Scr mRNAs during Drosophila Embryogenesis Yields Evidence for Transcriptional Bursting. *Current Biology*, 19(23):2037–2042.

Peifer, M. and Wieschaus, E. (1990). Mutations in the drosophila gene extradenticle affect the way specific homeo domain proteins regulate segmental identity. *Genes & development*, 4(7):1209–1223.

Peláez, N., Gavalda-Miralles, A., Wang, B., Navarro, H. T., Gudjonson, H., Rebay, I., Dinner, A. R., Katsaggelos, A. K., Amaral, L. A., and Carthew, R. W. (2015). Dynamics and heterogeneity of a fate determinant during transition towards cell differentiation. *eLife*, 4:e08924. Publisher: eLife Sciences Publications, Ltd.

Penton, A., Chen, Y., Staehling-Hampton, K., Wrana, J. L., Attisano, L., Szidonya, J., Cassill, J. A., Massagué, J., and Hoffmann, F. M. (1994). Identification of two bone morphogenetic protein type i receptors in drosophila and evidence that brk25d is a decapentaplegic receptor. *Cell*, 78(2):239–250.

Pichaud, F. and Casares, F. (2000). homothorax and iroquois-c genes are required for the establishment of territories within the developing eye disc. *Mechanisms of development*, 96(1):15–25.

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: structural insights into the regulation of yan by mae. *Cell*, 118(2):163–173.

Raff, M. C., Fields, K. L., Hakomori, S.-I., Mirsky, R., Pruss, R. M., and Winter, J. (1979). Cell-type-specific markers for distinguishing and studying neurons and the major classes of glial cells in culture. *Brain research*, 174(2):283–308.

Raj, A., van den Bogaard, P., Rifkin, S. A., van Oudenaarden, A., and Tyagi, S. (2008). Imaging individual mRNA molecules using multiple singly labeled probes. *Nature Methods*, 5(10):877–879. Number: 10 Publisher: Nature Publishing Group.

Rawls, A. S. and Wolff, T. (2003). Strabismus requires flamingo and prickle function to regulate tissue polarity in the drosophila eye.

Ready, D. F., Hanson, T. E., and Benzer, S. (1976). Development of the drosophila retina, a neurocrystalline lattice. *Developmental biology*, 53(2):217–240.

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(6):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(2):321–330.

Riggleman, B., Schedl, P., and Wieschaus, E. (1990). Spatial expression of the drosophila segment polarity gene armadillo is posttranscriptionally regulated by wingless. *Cell*, 63(3):549–560.

Roignant, J.-Y. and Treisman, J. E. (2009). Pattern formation in the drosophila eye disc. *The International journal of developmental biology*, 53(5-6):795.

Sáez, M., Blassberg, R., Camacho-Aguilar, E., Siggia, E. D., Rand, D. A., and Briscoe, J. (2022). Statistically derived geometrical landscapes capture principles of decision-making dynamics during cell fate transitions. *Cell Systems*, 13(1):12–28.

Sakamoto, Y., Yoshida, M., Semba, K., and Hunter, T. (1997). Inhibition of the dna-binding and transcriptional repression activity of the wilms' tumor gene product, wt1, by camp-dependent protein kinase-mediated phosphorylation of ser-365 and ser-393 in the zinc finger domain. *Oncogene*, 15(17):2001–2012.

Sanson, B., Alexandre, C., Fascetti, N., and Vincent, J.-P. (1999). Engrailed and hedgehog make the range of wingless asymmetric in drosophila embryos. *Cell*, 98(2):207–216.

Shwartz, A., Yogev, S., Schejter, E. D., and Shilo, B.-Z. (2013). Sequential activation of ets proteins provides a sustained transcriptional response to egfr signaling. *Development*, 140(13):2746–2754.

Siggers, T. and Gordân, R. (2014). Protein–dna binding: complexities and multi-protein codes. *Nucleic acids research*, 42(4):2099–2111.

Tio, M., Ma, C., and Moses, K. (1994). spitz, a drosophila homolog of transforming growth factor-α, is required in the founding photoreceptor cells of the compound eye facets. *Mechanisms of development*, 48(1):13–23.

Tio, M. and Moses, K. (1997). The drosophila tgf alpha homolog spitz acts in photoreceptor recruitment in the developing retina. *Development*, 124(2):343–351.

Tomlinson, A. and Ready, D. F. (1987). Cell fate in the Drosophila ommatidium. *Developmental Biology*, 123(1):264–275.

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.

Trapnell, C. (2015). Defining cell types and states with single-cell genomics. *Genome research*, 25(10):1491–1498.

Trapnell, C., Cacchiarelli, D., Grimsby, J., Pokharel, P., Li, S., Morse, M., Lennon, N. J., Livak, K. J., Mikkelsen, T. S., and Rinn, J. L. (2014). The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nature biotechnology*, 32(4):381–386.

Treisman, J. E. and Rubin, G. M. (1995). wingless inhibits morphogenetic furrow movement in the drosophila eye disc. *Development*, 121(11):3519–3527.

Tsai, Y.-C., Yao, J.-G., Chen, P.-H., Posakony, J. W., Barolo, S., Kim, J., and Sun, Y. H. (2007). Upd/jak/stat signaling represses wg transcription to allow initiation of morphogenetic furrow in drosophila eye development. *Developmental biology*, 306(2):760–771.

Virchow, R. (2020). *Die Cellularpathologie in ihrer Begründung auf physiologische und pathologische Gewebelehre*. Good Press.

Voas, M. G. and Rebay, I. (2004). Signal integration during development: insights from the drosophila eye. *Developmental dynamics: an official publication of the American Association of Anatomists*, 229(1):162–175.

Waddington, C. H. (1957). *The strategy of the genes*. Routledge.

Wang, Z., Gerstein, M., and Snyder, M. (2009). Rna-seq: a revolutionary tool for transcriptomics. *Nature reviews genetics*, 10(1):57–63.

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. *Development*, 145(dev165985).

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. *The EMBO journal*, 29(13):2147–2160.

Wheat, J. C., Sella, Y., Willcockson, M., Skoultchi, A. I., Bergman, A., Singer, R. H., and Steidl, U. (2020). Single-molecule imaging of transcription dynamics in somatic stem cells. *Nature*, 583(7816):431–436.

Wieschaus, E. and Nüsslein-Volhard, C. (2016). The heidelberg screen for pattern mutants of drosophila: a personal account. *Annual review of cell and developmental biology*, 32:1–46.

Wolff, T. (1993). Pattern formation in the drosophila retina. *The development of Drosophila melanogaster*, pages 1277–1325.

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(3):841–850.

Wu, C., Boisclair Lachance, J.-F., Ludwig, M. Z., and Rebay, I. (2020). A context-dependent bifurcation in the pointed transcriptional effector network contributes specificity and robustness to retinal cell fate acquisition. *PLoS genetics*, 16(11):e1009216.

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(1):87–97.

Xu, T. and Rubin, G. M. (1993). Analysis of genetic mosaics in developing and adult drosophila tissues. *Development*, 117(4):1223–1237.

Yordy, J. S. and Muise-Helmericks, R. C. (2000). Signal transduction and the ets family of transcription factors. *Oncogene*, 19(55):6503–6513.

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. *Molecular and Cellular Biology*, 30(5):1158–1170.

Zhu, J., Shibasaki, F., Price, R., Guillemot, J.-C., Yano, T., Dötsch, V., Wagner, G., Ferrara, P., and McKeon, F. (1998). Intramolecular masking of nuclear import signal on nf-at4 by casein kinase i and mekk1. *Cell*, 93(5):851–861.

Zhu, Q., Shah, S., Dries, R., Cai, L., and Yuan, G.-C. (2018). Identification of spatially associated subpopulations by combining scrnaseq and sequential fluorescence in situ hybridization data. *Nature biotechnology*, 36(12):1183–1190.

CHAPTER 2

RATIOMETRIC SENSING OF PNT AND YAN TRANSCRIPTION FACTOR LEVELS REGULATE PHOTORECEPTOR FATE TRANSITIONS IN DROSOPHILA

Sebastian M. Bernasek*, Suzy S. J. Hur*, Nicolás Peláez-Restrepo*, Jean-François Boisclair Lachance, Rachael Bakker, Heliodoro Tejedor Navarro, Nicelio Sanchez-Luege, Luís A. N. Amaral, Neda Bagheri, Ilaria Rebay, and Richard W. Carthew

*Joint first-authors

This work is now published and can be accessed at: https://doi.org/10.1242/dev.201467 (Bernasek et al., 2023)

2.1 Statement of Contributions

This work is the result of a long-standing collaboration between members of the Rebay lab, Carthew lab, and the Amaral lab. This project took on a new conceptual/ experimental direction when I formally entered the collaboration, but the original idea to measure Pnt/Yan ratio dynamics was conceived by NP under the guidance of IR and RC. SB developed the computational tools including the FlyEye Software, analyzed the data, and performed mathematical modeling. NP, SB, and I worked very closely throughout this project. I performed the experiments in Figures 1,3,4,5 and 6 (+ supplemental figures) under technical guidance from NP, and analyzed data with extensive support from SB. I contributed to the re-writing of the manuscript in light of our new results. A full list of contributions as shown in (Bernasek et al., 2023) is provided below.

Conceptualization: S.M.B., N.P.-R., L.A.N.A., N.B., I.R., R.W.C.; Methodology: S.M.B., S.S.J.H., N.P.-R., H.T.N.; Software: S.M.B., H.T.N., N.S.-L.; Formal analysis: S.M.B.,

S.S.J.H., N.P.-R., J.-F.B.L., I.R., R.W.C.; Investigation: S.S.J.H., N.P.-R., J.-F.B.L., R.B.; Data curation: S.M.B.; Writing - original draft: R.W.C.; Writing - review and editing: S.M.B., S.S.J.H., N.P.-R., J.-F.B.L., R.B., H.T.N., N.S.-L., L.A.N.A., N.B., I.R.; Visualization: S.S.J.H., N.P.-R.; Supervision: L.A.N.A., N.B., I.R., R.W.C.; Project administration: L.A.N.A., N.B., I.R., R.W.C.; Funding acquisition: I.R., R.W.C.

2.2 Abstract

Cell state transitions are often triggered by large changes in the concentrations of transcription factors and therefore large differences in their stoichiometric ratios. Whether cells can elicit transitions using modest changes in the ratios of co-expressed factors is unclear. Here we investigate how cells in the Drosophila eye resolve state transitions by quantifying the expression dynamics of the ETS transcription factors Pnt and Yan. Eye progenitor cells maintain a relatively constant ratio of Pnt/Yan protein despite expressing both proteins with pulsatile dynamics. A rapid and sustained two-fold increase in the Pnt/Yan ratio accompanies transitions to photoreceptor fates. Genetic perturbations that modestly disrupt the Pnt/Yan ratio produce fate transition defects consistent with the hypothesis that transitions are normally driven by a two-fold shift in the ratio. A biophysical model based on cooperative Yan-DNA binding coupled with non-cooperative Pnt-DNA binding illustrates how two-fold ratio changes could generate ultrasensitive changes in target gene transcription to drive fate transitions. Thus, coupling cell state transitions to the Pnt/Yan ratio sensitizes the system to modest fold-changes, conferring robustness and ultrasensitivity to the developmental program.

2.3 Introduction

Organismal development proceeds through a sequence of cell transitions that yield increasingly restricted cellular states or fates. A common feature of cells that undergo fate transitions is the apparent irreversibility of the transitions, even when triggered by transient stimuli. Dynamical modeling of these transitions often assumes multistability, in which cells can be resting in one of two or more stable states (Rand et al., 2021; Sáez et al., 2022). These stable states are typically a progenitor state and one or more derived states with lesser developmental potential.

Transcription factors are frequently utilized to define cell states, and transitions to a new state are often driven by varying their absolute or relative protein abundance (Laslo et al., 2006; Park et al., 2012; Yao et al., 2008). A considerable number of developmental transitions rely on mutually exclusive expression of two antagonistic transcription factors. Each factor programs a distinct cell state such that cells adopt one of two possible fates depending on whether they express high levels of one transcription factor or the other. Transitions are triggered when transient stimuli alter the balance between the two factors, resulting in expression of a new dominant factor. First documented in the early embryo of Drosophila (Jaeger et al., 2004), such regulatory switches built on mutual repression function in mouse embryonic germ-layer restriction, hematopoiesis, neural tube patterning, mesoderm development, and pancreatic lineage restriction (Acloque et al., 2011; Briscoe and Ericson, 2001; Schaffer et al., 2010; Schrode et al., 2014). A hallmark of such systems is the mutually exclusive expression of the two opposing transcription factor regulators, which is frequently controlled by a bistable mechanism and double negative feedback loops connecting the two factors.

In Drosophila, two ETS-domain transcription factors, Pointed (Pnt) and Yan (also known as Aop, Anterior open), act downstream of signals mediated by receptor tyrosine kinases (RTKs) to regulate cell fate transitions in a wide variety of tissues across the body and across

the life cycle (Flores et al., 2000; Gabay et al., 1996; Halfon et al., 2000; Lachance et al., 2014; Melen et al., 2005; O'Neill et al., 1994; Rebay and Rubin, 1995; Rogge et al., 1995; Shwartz et al., 2013; Xu et al., 2000). Consistently across many different developmental transitions, loss of yan results in too many progenitors transitioning to a particular fate, while loss of pnt prevents fate transitions (Brunner et al., 1994; Flores et al., 2000; Gabay et al., 1996; Halfon et al., 2000; O'Neill et al., 1994; Xu et al., 2000). Since both Pnt and Yan bind to a common GGA(A/T)-containing DNA sequence motif (Nitta et al., 2015; Wei et al., 2010; Zhu et al., 2011), competition for DNA occupancy and antagonistic regulation of common target genes is thought to orchestrate the transcriptional changes that drive particular transitions (Flores et al., 2000; Halfon et al., 2000; Hollenhorst et al., 2011; Lachance et al., 2018; Nitta et al., 2015; Webber et al., 2018, 2013; Wei et al., 2010; Xu et al., 2000; Zhu et al., 2011).

The opposing genetic phenotypes and transcriptional activities of Pnt and Yan, together with their mutually exclusive expression patterns in certain tissues, inspired development of a bistable model for cell fate specification based on mutual repression of one another's expression (Gabay et al., 1996; Golembo et al., 1996; Graham et al., 2010; Melen et al., 2005). For example, in the embryonic ventral ectoderm, ventral-most cells express Pnt while ventro-lateral cells express Yan (Gabay et al., 1996). This pattern is established by secretion of a ligand for the EGF Receptor (EGFR) from the ventral midline (Golembo et al., 1996). EGFR activation in nearby cells induces the Ras pathway to express Pnt and degrade Yan (Gabay et al., 1996; Melen et al., 2005). Cells with insufficient EGFR activation express Yan, which represses Pnt. Dynamical modeling has described this as a bistable system in which cells assume different fates based on whether they are in a low Pnt/high Yan state vs. a high Pnt/low Yan state (Graham et al., 2010; Melen et al., 2005).

While the bistable model readily explains transitions in which the different states are marked by mutually exclusive Pnt and Yan expression, an unresolved paradox is that many tissues have extensive co-expression of Yan and Pnt (Lachance et al., 2014). The third-instar

larval eye disc is one such tissue. Qualitative imaging has shown that rather than adopting mutually exclusive expression states, eye progenitor cells co-express high levels of Pnt and Yan for prolonged periods. Further, when progenitors transit to photoreceptor (R cell) fates, they continue to co-express both proteins, transiently increasing Pnt levels before initiating a parallel decay in both Pnt and Yan (Lachance et al., 2014). These observations are at odds with the long-standing assumption that cross-repression within a bistable network creates mutually exclusive Yan and Pnt expression states during R cell fate specification (Graham et al., 2010; Lai and Rubin, 1992; Rebay and Rubin, 1995).

Here, we have explored how Yan and Pnt expression dynamics in the third instar eye disc impact progenitor to R-cell fate transitions. Quantitative fluorescence-based microscopy was coupled with automated high-throughput image analysis to measure Yan and Pnt protein levels at single cell resolution in thousands of eye disc cells. We find that despite variation in Pnt and Yan levels, cells in the progenitor state maintain a fairly constant ratio of Pnt/Yan for up to 50 hours of development. When specific progenitors transit to R-cell fates, the Pnt/Yan ratio rapidly increases, but only by at most two-fold. Experimental perturbation of the ratio results in progenitor cells either prematurely transiting or failing to transit to R-cell fates, implicating the Pnt/Yan ratio as the driver of cell fate transitions. Using a biophysical model, we demonstrate how even a modest change in the Pnt/Yan ratio will cause a large change in the DNA occupancy of the two proteins on their common target genes. We conclude that the transition to photoreceptor fates is sensitive to the stoichiometric ratio of Pnt and Yan and propose that ratiometric sensing mechanisms could confer ultrasensitivity to biological transitions regulated by multiple factors.

2.4 Results

During the final 50 hours of the larval phase of the Drosophila life cycle, a morphogenetic furrow (MF) moves across the eye disc epithelium at a constant velocity from posterior to

anterior (Figure 1A). As the MF moves, every 2 hours a column of periodic R8 photoreceptor cells is formed with near simultaneity immediately posterior to the MF. Each R8 cell in the column secretes ligands that activate the RTK – Ras – MAPK pathway in seven neighboring progenitor cells and causes them to transit from progenitor to R-cell fates (Freeman, 1996). These transitions occur stepwise every ~2 hours, first generating R2 and R5 cells, then R3 and R4 cells, then R1 and R6 cells, and finally the R7 cell (Figure S1A; (Wolff and Ready, 1993)). Since the MF repeatedly induces new columns of R-cell clusters or ommatidia, it ultimately forms a lattice pattern of ~750 ommatidia in each eye. Although ~6,000 progenitor cells adopt R-cell fates, a pool of ~14,000 progenitors remain uncommitted, and later adopt other retinal cell fates (Wolff and Ready, 1991).

Progenitor cells posterior to the MF co-express Pnt and Yan and are poised to adopt R-cell fates (Figure 1B). When progenitors transition to R-cell fates, RTK – Ras – MAPK signals regulate the opposing transcriptional inputs from Pnt and Yan on their shared target genes. The signals concomitantly downregulate Yan activity and upregulate Pnt activity, which is thought to shift the state of the progenitor cell towards an R-cell state by reprogramming target gene expression.

To measure Pnt protein abundance in individual eye cells, we used a genomic transgene of the pnt locus in which fast-fold green fluorescent protein (GFP) was inserted in-frame at the carboxy-terminus of the coding sequence, tagging all Pnt isoforms. The Pnt-GFP transgene fully complemented null mutations in the endogenous pnt gene and completely restored normal eye development (Lachance et al., 2014), demonstrating that the GFP tag does not compromise Pnt function and that all essential genomic regulatory sequences are included. To measure Yan protein abundance, we used a monoclonal antibody that elicits immunofluorescence linearly proportional to native Yan protein concentration (Peláez et al., 2015). This enabled us to measure nuclear Pnt-GFP and Yan protein abundance simultaneously in individual cells. Using an established computational pipeline (Peláez et al., 2015)

we segmented all eye cell nuclei and then calculated Pnt-GFP and Yan fluorescence levels and exact 3D positions in each eye disc (Figure S1B-E). A unique feature of eye development allowed us to map each cell's spatial position in the plane of the eye disc onto developmental time. Because the MF moves at a fixed velocity, forming one column of R8 cells every 2 hours (Basler and Hafen, 1989; Campos-Ortega and Hofbauer, 1977; Gallagher et al., 2022), the distance between the MF and any cell posterior to it is linearly proportional to the time elapsed since the MF passed. This mapping was applied to the thousands of segmented nuclei in each eye disc and repeated for multiple eye discs.

A second feature of eye development enabled us to annotate cell states without relying on marker genes but instead on reproducible changes in cell morphology and apical-basal nuclear position. These stereotyped morphological changes are observable before the earliest known marker genes begin to be expressed, allowing accurate annotation of progenitor and R-cell states (Figure S1A,F) (Peláez et al., 2015; Tomlinson et al., 1987; Wolff, 1993). Cross-referencing with cell-type-specific markers confirmed that we could make cell state assigments with >97% accuracy using only morphology (Table S1). Combining these state assignments with the measurements of Pnt-GFP expression, Yan expression, and developmental time for many thousands of cells generated a composite picture of Yan and Pnt expression dynamics (Figure 1D). Although this approach cannot track the behavior of any individual cell over time, it nevertheless provides a dynamic view of eye cells as they develop. From this composite picture, individual cell behaviors can be inferred.

Focusing first on progenitor cells, we found that both Pnt-GFP and Yan expression dramatically increased coincident with passage of the MF (Figure 1C,D). This rapid increase culminated in two successive peaks of Pnt-GFP expression and a single peak of Yan expression at a time period in between the two Pnt-GFP peaks (Figure S2). The peaks of Pnt-GFP in progenitor cells coincided with the time periods when they transition to R-cell fates (Figure 1C-F). Transitions to R8, R2/R5, and R3/R4 fates occurred during the first

peak of expression (Figure 1E), while transitions to R1/R6, and R7 fates occurred during the second peak (Figure 1F). After their peaks, both Pnt-GFP and Yan decayed to a basal level in progenitors (Figure 1D). Despite the alternating peaks of Pnt and Yan expression, the overall trend of induction and decay was highly similar for the two proteins.

Focusing next on Pnt-GFP and Yan dynamics in cells that had transitioned to R-cell fates, we found that these transitions coincided with a rapid increase in Pnt-GFP and rapid decrease in Yan (Figure 1D-F). The Pnt-GFP increase was so rapid that the youngest cells we could classify as R-cells already had a 25-50% higher level of Pnt-GFP than progenitor cells at comparable time points (Figure 1D). Thereafter, the R-cells maintained this elevated Pnt-GFP for 2 – 4 hours before down-regulating Pnt-GFP to a basal level of expression (Figure 1E,F). Similarly, the decrease in Yan in transitioning cells was so rapid that the youngest cells classified as R-cells already had a 25-50% lower level of Yan than progenitors at comparable time points (Figure 1D). Thereafter, R-cells also maintained this diminished level of Yan for 2 – 4 hours before further down-regulating its expression (Figure 1E,F).

Two key conclusions emerged from this analysis. First, Pnt and Yan expression appear positively correlated in progenitor cells, confirming prior qualitative observations (Lachance et al., 2014). Second, transitions to R-cell fates coincide with concomitant but modest up- and down-regulation of Pnt and Yan expression, respectively. This suggests that cell fate specification in the eye does not rely on a classic bistable mechanism in which mutual repression between Pnt and Yan produces large changes in expression to distinguish the two states (Graham et al., 2010; Melen et al., 2005b).

2.4.1 Pnt and Yan positively regulate each other's expression

The positive correlation in expression of Yan and Pnt in progenitor cells led us to hypothesize that they might positively regulate each another's expression. To test this possibility, we first manipulated Pnt protein levels by varying the copy number of the pnt gene. If Pnt

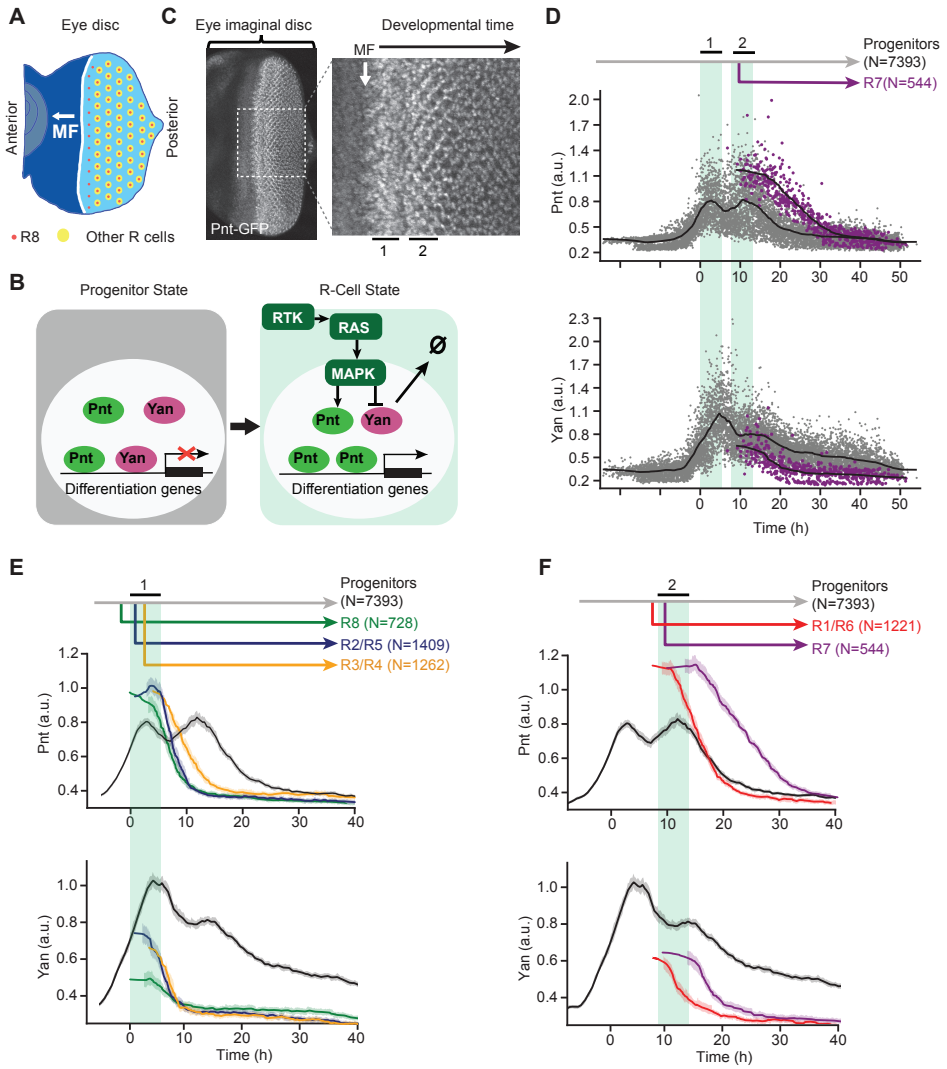

Figure 2.1: Pnt and Yan expression defines two waves of cell state transitions

(A) Schematic of a developing eye disc ~100 hours post fertilization. Passage of the MF (white arrow) initiates R cell fate specification. Specification of regularly spaced R8 cells (red dots) marks the start of ommatidial assembly and is followed by recruitment of additional R-cell types (yellow dots). Adapted from (Peláez et al., 2015).(B) Schematic summary of RTK-mediated regulation of progenitor to R-cell fate transitions. Progenitor cells co-express Pnt and Yan (left). In response to RTK signals, cells increase Pnt and decrease Yan (right).(C) Maximum intensity projection of Pnt-GFP fluorescence in an eye disc oriented anterior left and dorsal up. Magnified view shows the two zones where progenitor cells undergo fate transitions coincide with peaks of Pnt-GFP (black bars, labeled 1 and 2).(D) Pnt-GFP and Yan levels in individual progenitor (grey) and R7 cells (purple) over developmental time. N, number of cells analyzed in each group. Progenitor cells are present across all time (grey arrow) while R7 cells arise later in time (purple arrow). Solid black lines are smoothed moving averages across 250 and 75 individual nuclei for progenitor and R7 cells, respectively. Black bars labeled 1 and 2 denote the two peaks of Pnt-GFP in progenitor cells. Shaded vertical stripes highlight these two regions, which coincide with transition of progenitor to various R fates. Although time appears continuous in these plots (see experimental procedures), temporal resolution is about two hours, i.e. the time between specification of each column of R8 cells.(E,F) Line averages, with 95% confidence intervals shaded, for Pnt-GFP and Yan levels in progenitor (grey), R8 (green), R2/R5 (blue), R3/R4 (orange), R1/R6 (red), and R7 (purple) cells over developmental time. N, number of cells analyzed in each group

positively regulates Yan expression, then we predicted Yan levels should change in parallel to altered Pnt levels. Cells with two copies of the pnt gene expressed ~50% more Pnt protein than cells with one copy (Figure 2A,B). This difference was observed both in progenitor cells and R cells. In spite of the different levels of Pnt protein in cells with one or two copies of the pnt gene, progenitors transited to R-cell fates at similar times (Figure 2A,B), and produced adult eyes with wildtype morphologies (Figure S3). When we measured Yan, we found that Yan levels also differed between the two genotypes, despite both having two copies of the yan gene (Figure 2A,B). Yan output was 25-50% greater in progenitor cells with two copies of the pnt gene than in cells with one copy, consistent with a role for Pnt in positively regulating Yan expression.

To ask if this positive regulatory relationship might be mutual, we next assessed the effect of Yan levels on Pnt-GFP expression by varying the copy number of the yan gene. We conducted this experiment by generating clones of yan homozygous mutant cells in an eye where other cells have one or two copies of the yan gene. Qualitatively, cells with different yan gene copy numbers exhibited Pnt-GFP levels that positively correlated with yan gene dosage (Figure 2C). Using a published pipeline to segment and measure cells within eye disc clones (Bernasek et al., 2020), we quantified and compared Pnt-GFP protein levels in progenitor cells with zero vs. one wildtype copy of the yan gene (Figure 2D). Pnt-GFP expression was ~30% greater in progenitor cells with one copy of yan than in cells with zero copies of yan (Figure 2D). This result suggests that Pnt expression is positively dependent on Yan.

The parallel behaviors of Pnt and Yan in these experiments prompted us to quantitatively assess the extent to which Pnt and Yan levels are coordinated within individual cells. To do this, we calculated the Pnt/Yan ratio in progenitor cells containing one or two copies of the pnt gene. The Pnt/Yan ratio was invariant to the pnt gene copy number in progenitor cells that were sampled at discrete time points over their developmental trajectory (Figure 2E).

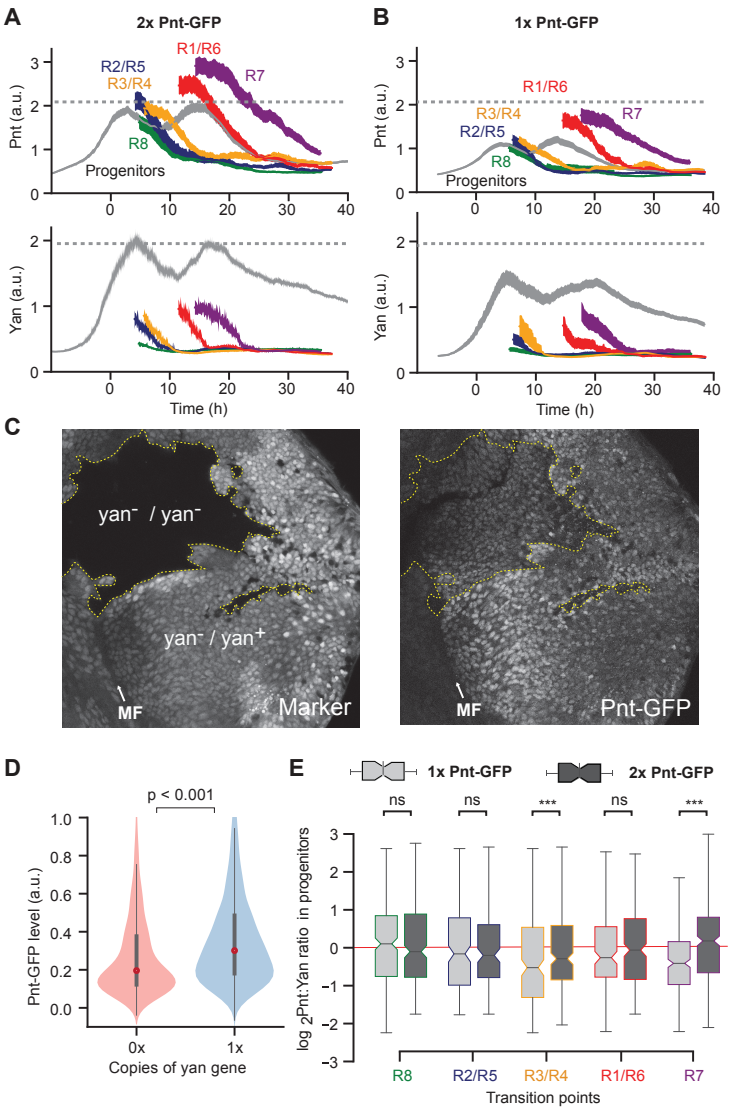

Figure 2.2: Pnt and Yan activate each other's expression

(A,B) Line averages, with 95% confidence intervals shaded, for Pnt-GFP and Yan levels in progenitor and R cells containing two (A) or one copy (B) of the pnt-GFP gene. The number of cells analyzed for each cell-class ranged from 312 (R7 cells) to 3,127 (progenitors) in (A) and 271 (R7 cells) to 2,834 (progenitors) in (B).(C) Confocal section showing progenitor cell nuclei in an eye disc containing several patches (clones) of homozygous yan mutant cells, marked by lack of RFP and outlined by dotted yellow lines. Right panel, Pnt-GFP fluorescence signal is lower in the yan mutant clones.(D) Pnt-GFP abundance is significantly lower in homozygous yan mutant (0x) versus heterozygous (1x) progenitor cells. Violin plots contain red dots that denote the median of each distribution and grey lines that denote the interquartile range. $p<0.001$, Mann-Whitney U test.(E) Comparison of Pnt/Yan ratios in progenitor cells with one copy (light grey) versus two copies (dark grey) of the pnt-GFP gene. Cells were differentially sampled across five time-windows of cell fate transitions. These time windows were defined in each disc by the time spanned by the first ten identifiable R cells of a given class. Progenitor cells were sampled from these time windows. Median ratios, boxed by the quartile ratios, are shown. The ratio in progenitor cells is statistically indistinguishable between cells bearing different gene dosages during the R8, R2/R5, and R1/R6 cell fate transitions ($P>0.1$, KS 2-sample test).

It was striking that the ratio measured in the progenitors was fairly constant between the sampled time points. To examine this more closely, we calculated the ratio for all progenitor cells over the entire 50 hours of eye development (Figure 3A). The average Pnt/Yan ratio in progenitor cells remained dynamically stable about a fixed value over time (Figure 3A). In conclusion, these results demonstrate that Yan and Pnt expression are tightly and positively coordinated in progenitor cells.

2.4.2 An increase in the Pnt/Yan ratio accompanies cell state transitions

We next determined the Pnt/Yan ratio for R cells. We first measured the ratio in R7 cells along their developmental trajectory. Transitions to the R7 fate coincided with a rapid increase in the ratio (Figure 3A). The increase was so rapid that the youngest cells we could classify as R7 cells already had a two-fold higher ratio than progenitor cells at comparable time points. Thereafter, the ratio rose and then gradually fell over time, but always remained at higher values than that of progenitor cells. Examination of the Pnt/Yan ratio in all other R-cell types revealed a similar and consistent rapid increase as they transitioned from a progenitor state, followed by a gradual decline to a new steady value that was higher than that of progenitor cells (Figures 3B and S4).

The rapid shift in the Pnt/Yan ratio during cell state transitions suggested that the progenitor ratio becomes destabilized at the onset of these transitions. If so, then cell-to-cell heterogeneity of the ratio should rise during the time-windows of state transitions, and then subside as cell states are resolved. Indeed, greater cell-to-cell variation in the ratio coincided with the times when Pnt and Yan levels peaked (Figure 3A), corresponding to the two waves of R-cell differentiation. To further characterize this instability, we used an established method (Peláez et al., 2015) to quantify the cell-to-cell heterogeneity of Pnt, Yan, and the Pnt/Yan ratio over time (Figure S5). Progenitor cells exhibited maximal ratio heterogeneity during the two time periods that coincide with cell state transitions.

Thereafter, heterogeneity diminished in the remaining progenitor cells. R cells exhibited an abrupt drop in heterogeneity immediately after their fate specification, and heterogeneity descended further thereafter. These results support the idea that the Pnt/Yan ratio becomes specifically destabilized when progenitor cells are poised to undergo state transitions.

If cells undergoing state transitions experience a rise in the Pnt/Yan ratio, we should observe cells at transition points that cannot be morphologically classified as R cells but that possess an elevated ratio. This prediction was verified since there was an overlap in ratio values between the youngest cells we could classify as R cells and a subset of other cells sampled at comparable time points (Figure 3C-G). Nevertheless, the median ratio values were significantly different between the R cells and all the other cells (Figure 3H, $P<0.001$, KS 2-sample test). Together these measurements suggest that a subset of progenitor cells elevate their Pnt/Yan ratio by approximately two-fold at transition points, and that this modest increase is sufficient to confer molecular competency for transitioning to a photoreceptor state.

2.4.3 Perturbation of the Pnt/Yan ratio biases cells towards or against R-cell fates

The previous model for eye development suggested that Yan and Pnt antagonize one another to maintain cells in a progenitor state defined by high Yan and to move cells towards a R-cell state defined by high Pnt (Graham et al., 2010). Our results demonstrate that the cell states are not distinguished by large differences in the Pnt/Yan ratio. Instead, we hypothesize that a modest two-fold change in the ratio is sufficient to trigger R-cell differentiation. We tested this by experimentally altering the ratio in progenitor cells and then asking whether this perturbation altered the normal course of R-cell fate transitions.

Ras-dependent MAPK phosphorylation of Yan and Pnt had been previously found to stimulate turnover of Yan and synthesis of Pnt (Brunner et al., 1994; Gabay et al., 1996a;

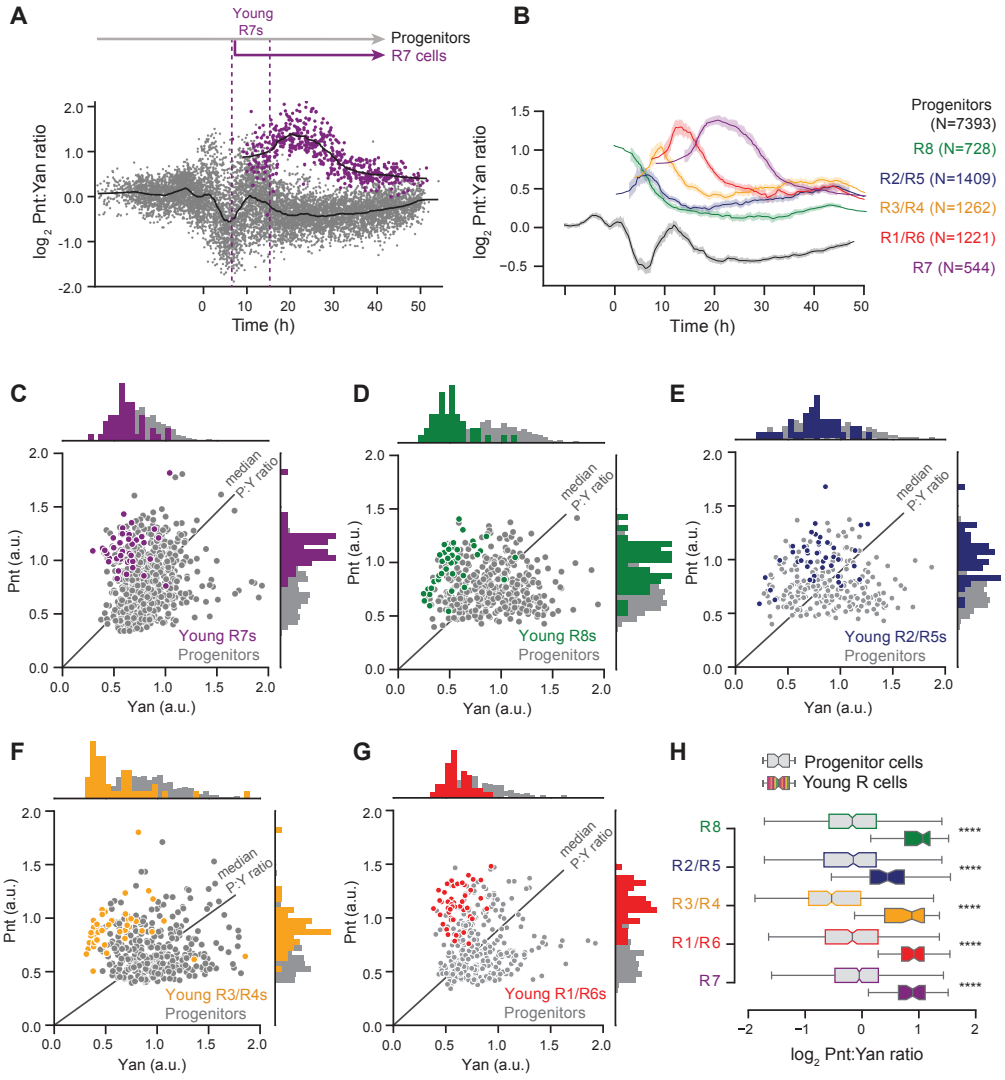

Figure 2.3: A two-fold shift in the Pnt/Yan ratio accompanies R cell fate transitions

(A) The log2-transformed Pnt/Yan ratio in individual progenitor (grey, 7393 cells) and R7 (purple, 544 cells) cells and their moving line averages (solid black lines). Cells marked as "Young R7" are the initial cohort of R7 cells that are present in the timespan of the first ten identifiable R7 cells ("young" R7s become "mature" Elav-positive R7s in about two hours).(B) Line averages, with 95% confidence intervals shaded, of the log2-transformed Pnt/Yan ratio for all annotated cell types over time. N, number of cells sampled. (C-G) Joint distributions of Pnt-GFP and Yan levels in Young R7 (C), Young R8 (D), Young R2/R5 (E), Young R3/R4 (F), and Young R1/R6 (G) cells. Progenitor cells (grey) present in the timespan of the young R cells were analyzed for comparison. Black lines denote median Pnt/Yan ratios among progenitor cells. Histograms of the data at the top and side of each scatterplot highlight the overlap between R cells and some of the concurrent progenitors. (H) Comparison of the Pnt/Yan ratios in young R cells (color boxes and whiskers) versus progenitor cells (grey boxes and whiskers) sampled over comparable time-periods. 40 cells were analyzed for each group of young R cells, and 153-487 cells were analyzed for each group of progenitors. Median ratios boxed by the quartile ratios are shown. ***, $p<0.001$, one-tailed Mann-Whitney U test.

O'Neill et al., 1994; Rebay and Rubin, 1995). Therefore, we expressed a mutant allele of the Ras1 gene encoding a constitutively active version of the Ras GTPase, reasoning that it would artificially increase the Pnt/Yan ratio (Figure 4A). RasV12 was transiently expressed in subsets of progenitor cells during the R3/R4 and R7 transition points by driving it with the sevenless (sev) promoter (Figure S6; (Fortini et al., 1992)). The promoter also drove stable expression in R3/R4 and R7 cells (Figure S6). As predicted, sev>RasV12 progenitor cells had an elevated second wave of Pnt expression, which coincides with the second time-window of cell state transitions (Figure 4B). Thereafter, progenitors maintained abnormally high levels of Pnt. Yan levels were weakly reduced in progenitor cells during the second transition period (Figure 4C). The overall effect was a modest two-fold elevation in the Pnt/Yan ratio in progenitor cells at the second transition period and thereafter (Figure 4D). When we looked at R-cell fate specification, we observed supernumerary R cells in the sev>RasV12 eye discs (Figure 4E). These ectopic R cells were typically located beside R7 cells (Figure S7A,B) and had comparable Pnt/Yan ratios to R7 cells (Figure 4E). Thus, just as a two-fold ratio increase accompanies wildtype R7 fate acquisition, artificially inducing a two-fold elevation in the Pnt/Yan ratio causes additional progenitor cells to inappropriately transit to R-cell fates.

We next asked whether inducing modest reductions in the ratio would prevent R-cell state transitions. It was previously shown that MAPK phosphorylation of Yan strongly stimulates Yan turnover (Rebay and Rubin, 1995). If the serines and threonines targeted by MAPK are mutated in Yan, then signaling-induced Yan turnover is blocked (Figure 5A). We transiently expressed this mutant YanACT protein in progenitor cells and R3/R4/R7 cells using the sev promoter. The Pnt/Yan ratio in progenitor cells was not detectably altered in sev>YanACT eye discs (Figure 5B-D and Figure S8A-C). This was not surprising since RTK – Ras – MAPK pathway is low in progenitors, and therefore YanACT protein is not abnormally stable in these cells. Nor was the ratio altered in R8/R2/R5/R1/R6 cells of

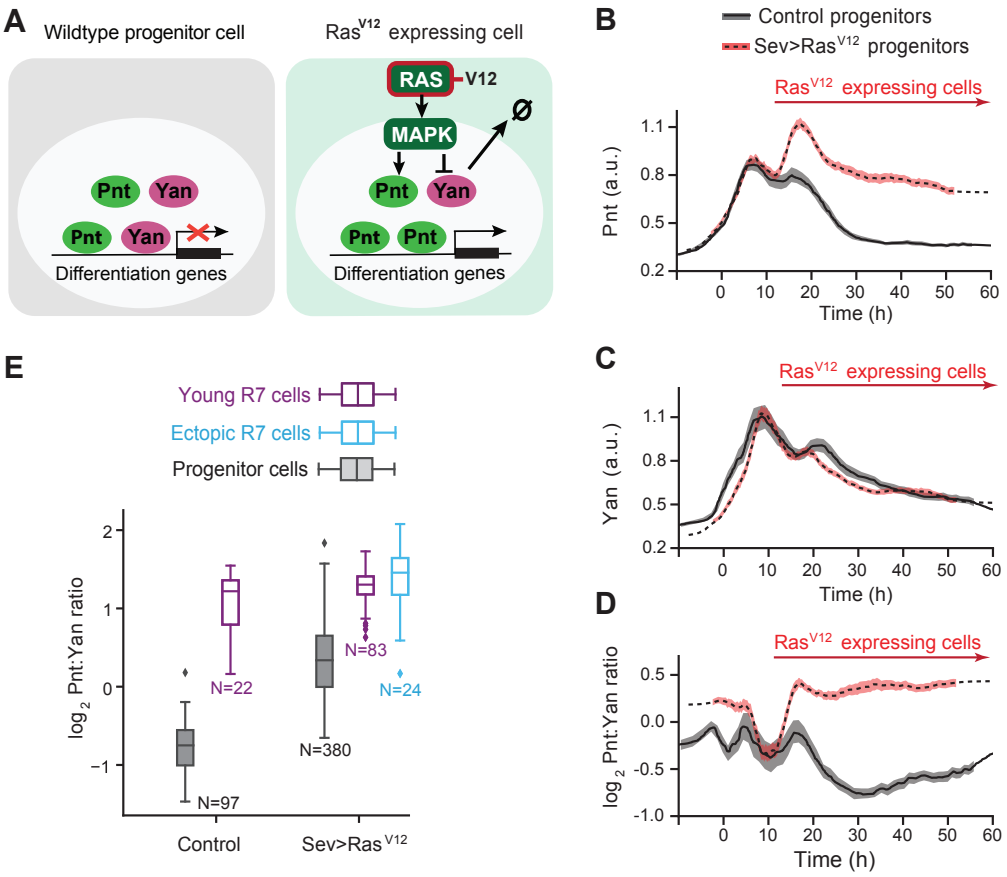

Figure 2.4: An artificial increase in the Pnt/Yan ratio biases progenitor cells towards an R fate

(A) Schematic showing the effect of RasV12 on signal transduction and Yan and Pnt expression in eye cells.(B-D) Line averages, with 95% confidence intervals shaded, of Pnt-GFP (B), Yan (C) and log2- transformed ratio of Pnt/Yan ratio (D) in wildtype (black) and sev>RasV12 (red) progenitor cells over developmental time. Red arrows mark time during which cells express RasV12. We previously reported a modest increase in the duration of Yan-YFP expression in sev>RasV12 progenitor cells (Peláez et al. 2015), but this difference was not detected using the Yan antibody (B). 1,381 (wildtype) and 4,038 (sev>RasV12) were analyzed. (E) The log2-transformed ratio of Pnt/Yan in wildtype and sev>RasV12 young R7 cells (purple). Progenitor cells (grey) present in the same timespan were also analyzed. In sev>RasV12 eye discs, there were ectopic R7 cells (blue). Boxplots are median and quartiles. N, number of cells analyzed in each group.

sev>YanACT eye discs since the sev promoter does not induce YanACT expression in these cells (Figure S9A-D,G,H). However, in R3/R4/R7 cells, which both receive a strong RTK signal and express YanACT, there was a reduction in the Pnt/Yan ratio. The reduction was weak in R3/R4 cells (Figures S9E,F) but was larger in R7 cells, which had a ratio value similar to progenitor cells sampled at comparable times (Figure 5B-D). Consistent with the experimental design, the modest reduction in the Pnt/Yan ratio in R cells was primarily driven by higher levels of Yan protein (Figure S8D).

Changing the Pnt/Yan ratio to make it Yan-biased in young R7 cells induced a remarkable reduction in the total number of R7 cells that developed (Figure 5B,C and Figure S7A,C). Only a few R7 cells could be identified shortly after fate specification, and more mature R7 cells could not be found. Overall, the progressive fate specification of the R7 cells was reduced seven-fold in the sev>YanACT eye discs (Figure 5E). In contrast, R3/R4 cell fate specification was not impaired even though the Pnt/Yan ratio was slightly lower than normal (Figure 5E). This suggests that cell state transitions can withstand weak or transient variation in the ratio, but not sustained larger changes.

In conclusion, two independent approaches to artificially alter the Pnt/Yan ratio resulted in cells making errors in state transitions. Modestly increasing the ratio caused cells to undergo ectopic state transitions while decreasing the ratio caused cells to fail to undergo transitions. The results suggest that progenitor cells can sense a two-fold change in the Pnt/Yan stoichiometric ratio and respond by transitioning to R-cell fates.

2.4.4 A biophysical model explains how small changes in stoichiometric ratio can have a large effect on Yan and Pnt DNA occupancy

The effect of transcription factors on their target genes is highly correlated with their occupancy on DNA elements within these genes. When transcription factors do not compete for binding sites, occupancy is driven by total nuclear concentration of the transcription factor

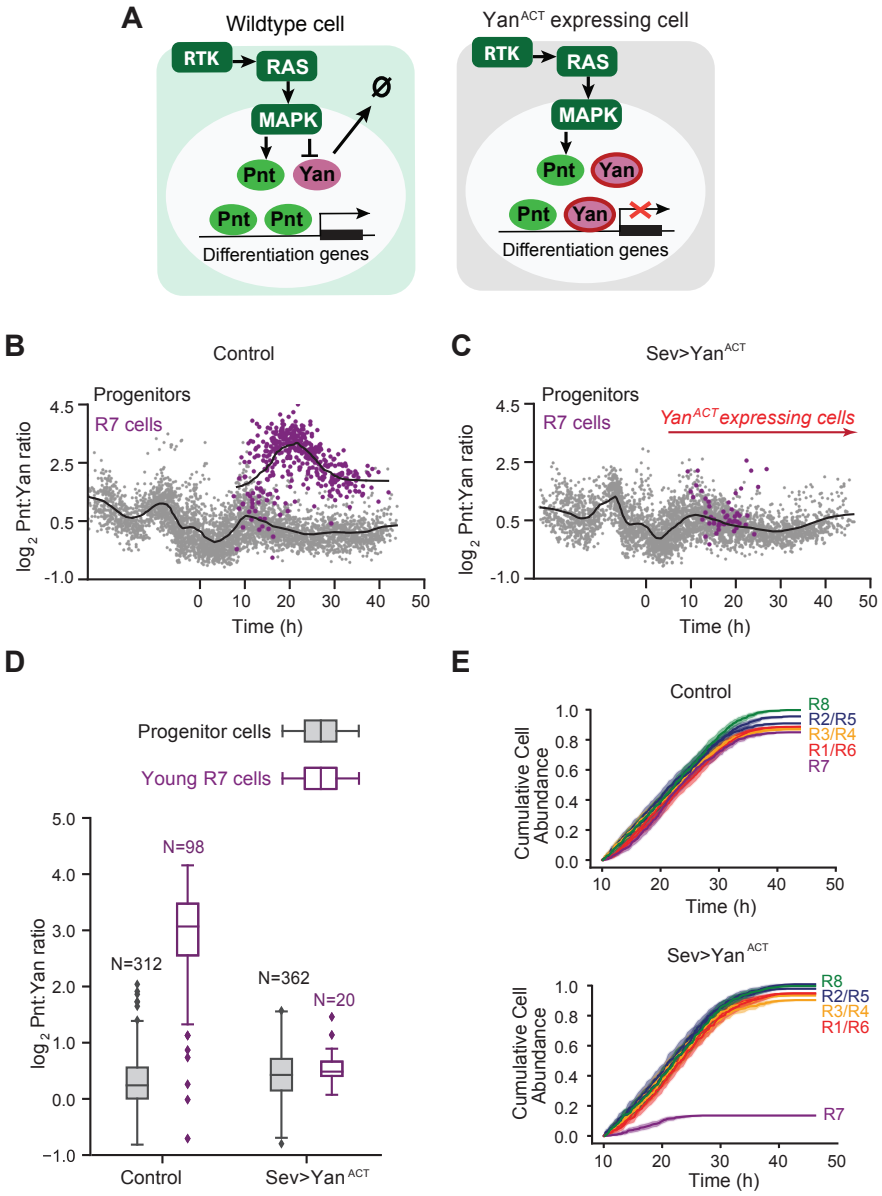

Figure 2.5: An artificial decrease in the Pnt/Yan ratio biases progenitor cells against an R fate

(A) Schematic showing the effect of YanACT on signal transduction and Yan and Pnt expression in eye cells.(B) The log2-transformed Pnt/Yan ratio in wildtype control progenitor (grey, 3796 cells) and R7 (purple, 413 cells) cells and their moving line averages (solid black lines).(C) The log2-transformed Pnt/Yan ratio in sev>YanACT progenitor (grey) and R7 (purple) cells and their moving line averages (solid black lines). Red arrow marks time during which cells express YanACT.(D) The log2-transformed ratio of Pnt/Yan in wildtype-control and sev>YanACT young R7 cells (purple). Progenitor cells (grey) that were present in the timespan of these young R7 cells were also analyzed. Boxplots are median and quartiles. N, number of cells analyzed in each group.(E) Cumulative abundance of each R cell type as a function of developmental time in wildtype (top) and sev>YanACT (bottom) eye discs. Counts are normalized by the total number of R8 cells detected across the time-course, and R2, R5, R3, R4, R1, and R6 cells are each tallied separately. Shading denotes 95% confidence intervals.

and its binding affinity. In systems with transcription factors that mutually repress each other's expression, the stoichiometry of the factors can be very different in cells. Thus, DNA occupancy by antagonistic transcription factors reflects the large difference in their relative abundance. Since Pnt and Yan have overlapping sequence specificity for DNA binding both in vitro and in vivo, they are thought to compete for common binding sites in target genes (Flores et al., 2000; Halfon et al., 2000; Lachance et al., 2018; Nitta et al., 2015; Webber et al., 2018; Wei et al., 2010; Xu et al., 2000; Zhu et al., 2011). However, the differences in Pnt/Yan stoichiometry that we measured in eye cells are quite modest – two-fold. How could such a modest difference in stoichiometry profoundly affect relative DNA occupancy by Yan and Pnt to alter transcription of target genes?

Measuring DNA occupancy at single cell resolution is not yet experimentally feasible, so we instead investigated this question in silico. We first created a simple biophysical model for equilibrium binding of two transcription factors that compete for common DNA binding sites. If the proteins have a similar binding affinity for DNA and if the binding sites are saturated, then at equilibrium, occupancy will be sensitive to the relative concentration of each factor across a broad range of absolute concentrations (Figure S10).

However, the situation with Pnt and Yan is more complex. Yan binds to common binding sites with higher affinity than Pnt (Xu et al., 2000). Moreover, recent experiments suggest Yan and Pnt differentially interpret the structural syntax of enhancers and promoters (Lachance et al., 2018). This complexity is a consequence of cooperative interactions between DNA-bound Yan molecules. Yan monomers are able to self-associate via their sterile α motifs (SAM), enabling DNA-bound Yan monomers to stabilize the recruitment of additional Yan monomers to adjacent sites (Hope et al., 2017; Lachance et al., 2018; Qiao et al., 2004). Since Pnt does not self-associate (Mackereth et al., 2004; Meruelo and Bowie, 2009), this Yan-Yan cooperativity will bias its competitiveness for common binding sites.

Therefore, we adapted a modeling framework developed to probe the effects of cis-

regulatory syntax on Yan binding site occupancy (Hope et al., 2017). The model considers an ensemble of microstates, each defined by a unique configuration of vacant or Yan-bound sites. Each microstate is assigned a thermodynamic potential based on the cumulative influence of strong sequence-specific binding to consensus ETS sites, weak non-specific DNA binding, and SAM-SAM polymerization. We generalized the earlier model by incorporating Pnt as a second transcription factor that competes for occupancy of the same binding sites but does not self-associate (Figure S11A,B). Using this model, we explored how Yan-Yan cooperativity affects Pnt binding site occupancy.

We first considered a scenario in which neither Yan nor Pnt exhibit cooperativity (Figure 6A). In the absence of stabilizing SAM-SAM interactions, the landscape of overall binding site occupancy was identical to that obtained with the simple equilibrium model described earlier (Figures 6A, S10, and S11). Increasing the Pnt/Yan ratio resulted in a gradual increase in Pnt occupancy for all individual binding sites, producing a titration contour closely resembling a Michaelian curve (Figures 6C and S11C).

We then introduced stabilizing SAM-SAM interactions for Yan, which dramatically sharpened the differences between Yan and Pnt occupancy (Figure 6B). The model output resembles a phase diagram in which there is a sharp boundary that separates the state in which all sites are bound by Yan from the state in which all sites are bound by Pnt. This phase boundary lies along a region in which Pnt and Yan are in 1:1 stoichiometry, regardless of total protein concentration. Weighting the energetic contributions of binding strength and polymerization by the statistical frequency of each microstate revealed that the transition was driven by an abrupt change in the dominant binding mechanism. Polymerization-associated cooperativity effects dominate binding site occupancy when the Pnt/Yan ratio is low, while binding strength dominates when the ratio is high (Figure S11D).

Increasing the Pnt/Yan ratio revealed nonlinear transitions from low to high Pnt occupancy for each individual binding site (Figures 6C and S11E). These transitions resembled

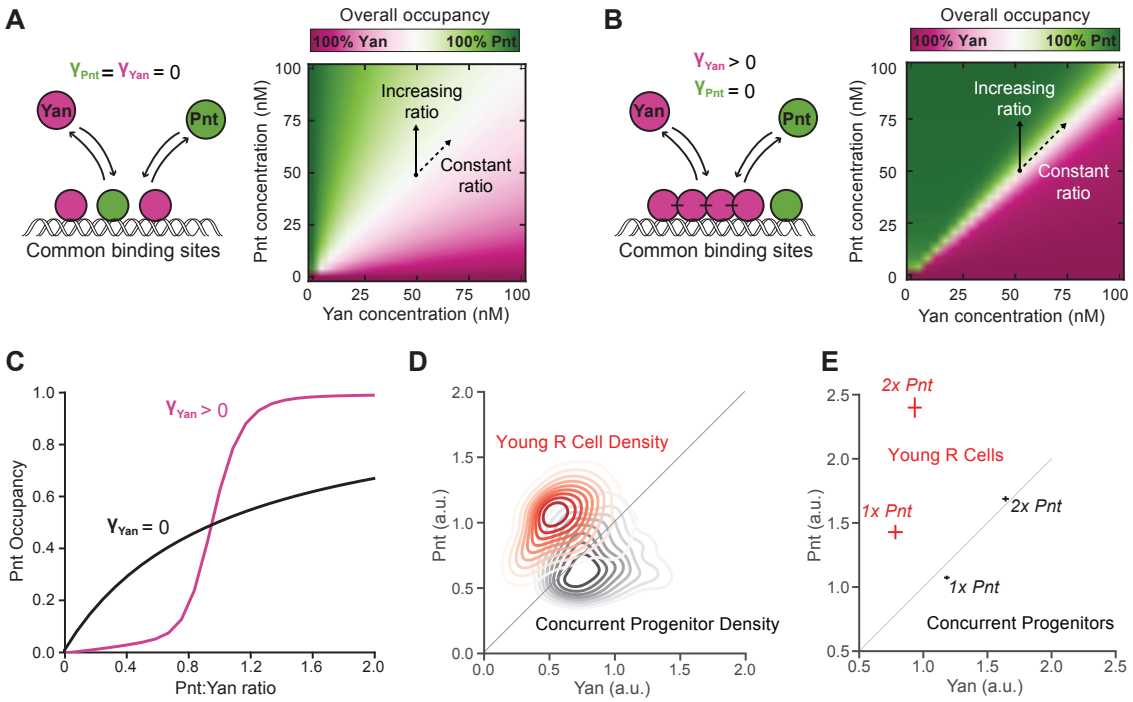

Figure 2.6: Cooperative binding by Yan greatly sensitizes DNA occupancy to the Pnt/Yan ratio

(A) Left – Schematic showing Pnt and Yan compete for occupancy of mutual DNA binding sites when Yan is unable to self-polymerize. Right – Overall binding site occupancy calculated as a function of transcription factor abundance. We use a diverging scale because all sites are fully saturated at total transcription factor concentrations above 1 nM. Under the range of conditions shown, this implies that Yan occupies all sites left vacant by Pnt. Simultaneous proportional increases in absolute abundance of both species have minimal impact on Pnt occupancy (dashed arrow), while varying ratio confers gradual change (solid arrow).(B) Left – Schematic of Pnt and Yan competition for occupancy of mutual DNA binding sites when Yan can self-polymerize. Right – Overall binding site occupancy calculated as a function of transcription factor abundance. The sharpness of the transition between all-Pnt occupancy and all-Yan occupancy is increased compared to (A).(C) Average Pnt occupancy across all binding sites as a function of the Pnt/Yan ratio when Yan is unable to self-polymerize (black line) versus when Yan can self-polymerize (pink line). Contours correspond to vertical paths traversed across the phase-plots in (A) and (B) at a fixed Yan concentration of 50 nM.(D) Probability density function of experimentally measured Pnt-GFP and Yan levels for the combined populations of all young R cells (red) and concurrent progenitor cells (black). Contours decrease in opacity with decreasing density and are not shown for values below 0.1. Density function is estimated using Gaussian kernels with a bandwidth selected using Scott's rule. Thin diagonal line shows a constant Pnt/Yan ratio to aid interpretation.(E) Experimentally measured average Pnt-GFP and Yan levels for the combined populations of all young R cells (red) and concurrent progenitor cells (black), sampled from eye discs containing either one (1x Pnt) or two (2x Pnt) copies of the pnt-GFP gene. Horizontal and vertical lines about each point denote the standard errors of the mean. Thin diagonal line shows a constant Pnt/Yan ratio to aid interpretation.

Hill function forms, with sharp thresholds delimiting distinct outputs, suggesting that transitions from low to high Pnt occupancy are ultra-sensitive to changes in the ratio. At low Pnt/Yan ratios, Yan occupancy dominated, while at some critical ratio, Pnt-bound molecules prevented Yan-Yan interactions, ultimately allowing Pnt to outcompete Yan as the ratio increased further. If transcription factor levels are sufficient to saturate binding sites, then relative occupancy by Pnt and Yan was agnostic to changes in the absolute abundance of either factor, as long as the ratio remained constant.

This model suggests a molecular mechanism for how even a two-fold change in the Pnt/Yan ratio could shift DNA occupancy and therefore regulation of common target genes from an all-Yan to an all-Pnt regime. Therefore, the model provides guidance on how to interpret the experimental data. Although our fluorescence measurements do not provide us with absolute concentration estimates for Yan and Pnt, nevertheless the Pnt/Yan ratio measured in progenitor and R cells can be viewed as a parallel representation to the phase spaces calculated in the model (Figure 6B,D). In this way, the small changes in the Pnt/Yan ratio we measured could elicit large changes in DNA binding site occupancy, while significant changes in absolute Pnt and Yan concentration would have little effect so long as the relative ratio remains constant (Figure 6E). Together this ratio-based mechanism provides both ultrasensitivity and robustness to retinal cell state transitions.

2.5 Discussion

The Drosophila retina has an extraordinary track record as a tractable genetic system for elucidating the conserved molecular mechanisms that direct cell state transitions. By introducing single-cell resolution quantitative analysis of Yan and Pnt expression dynamics, and then computationally modeling how they might impact DNA binding complexes and hence transcriptional output, our study has uncovered important and unanticipated mechanistic features of this well-studied system. We find that eye cells in the progenitor pool maintain a

dynamically stable Pnt/Yan ratio predicted to favor the Yan-dominated enhancer occupancy necessary to sustain the progenitor state. In transitioning R cells, RTK signaling disrupts the progenitor ratio set-point and produces a sustained increase in the ratio to now favor Pnt DNA occupancy. In this way, transcriptional programs associated with the progenitor state are shifted to those appropriate to the R-cell state. We suggest that sensitivity to small but sustained relative changes makes this ratiometric system robust to fluctuations in absolute Pnt and Yan protein levels while enabling rapid ultrasensitive responses to inductive signaling.

For a transition to occur, the Pnt/Yan ratio set point must be shifted. However, prior to undergoing a state transition, the progenitor's ratio set-point is dynamically stable over time even while cells vary in Yan and Pnt expression over a four-fold range. This suggests a stabilizing mechanism that monitors the ratio of Pnt and Yan and takes corrective action when the ratio deviates from a specified reference value. The results of experiments in which we manipulated pnt or yan gene copy number suggest the stabilizing mechanism includes a weak positive feedback loop between Pnt and Yan. As a result of this positive feedback, any fluctuation in the level of one factor will bring about a compensatory change in the other factor, maintaining dynamic but approximately constant stoichiometry of the two transcription factors. Overall, this allows cell fate transitions to occur reliably over a wide range of absolute protein concentrations and prevents transient fluctuations in either Yan or Pnt from inappropriately triggering or blocking differentiation.

Our results show that experimental manipulation of Ras and MAPK activity can change the Pnt/Yan ratio, most likely by regulating the synthesis and degradation of Pnt and Yan (Graham et al., 2010). Presumably, the ratio shift that naturally occurs at state transitions is due to inductive signals coming from R8 cells, since the signals are received by RTKs and transduced by the Ras – MAPK pathway. The ratio adopts a new set-point in R cells, twice the value of the old set-point, and is sustained thereafter. If the positive feedback loop

between Yan and Pnt regulates the set-point, then the strength of their mutual interactions must be differentially altered in order for the expression ratio to maintain a new value. Alternatively, feedback might remain the same but new regulatory inputs could tune Yan and Pnt expression to the new ratio in R cells. Investigating the contributions of the multiple regulatory interactions that have been documented within the Pnt/Yan network (Graham et al., 2010; Rohrbaugh et al., 2002; Webber et al., 2013b; Webber et al., 2018; Wu et al., 2020) to maintenance and modulation of the ratio will be an interesting future direction.

Previous studies have shown how cells can sense relative concentration changes of individual molecular species using fold-change detection (Frick et al., 2017; Adler et al., 2017; Alon et al., 1999; Barkai and Leibler, 1997; Goentoro and Kirschner, 2009; Lee et al., 2014; Mesibov et al., 1973; Shoval et al., 2010). To compute fold-change of a single protein, regulatory circuits first "measure" and establish a molecular memory of the state of one variable, and later compare it to its background level or to a new measurement of the same variable (Frick et al., 2017; Lyashenko et al., 2020). In contrast, rather than measuring the individual fold-change in Pnt or Yan over time, eye cells likely rely on concurrent comparison of Pnt and Yan levels, or more speculatively as suggested by the biophysical model, by concurrent comparison of the transcriptional outcome of Yan and Pnt competition for DNA binding sites.

Progenitor cells exhibit pulsatile expression of Yan and Pnt in which their expression levels decay to a basal state by 50 hours of developmental time. If cells transit to a R-cell fate, this decay in expression is accelerated. Nevertheless, the Pnt/Yan ratio remains stable about a fixed set-point specific for cells in either progenitor or R-cell states. This is in contrast to many developmental systems in which the ratio of antagonistic transcription factors remains stable because expression of the factors is maintained constant when cells exist in a particular developmental state (Laslo et al., 2008). A ratiometric system with modest variation may provide several advantages not just for the eye, but potentially to

other systems.

First, such a system can permit fast response times. In the eye, a cascade of four inductive events occurs over a brief 8-hour window of time to recruit R1 – R7 cells. If a classical mechanism of mutual transcriptional repression existed between Yan and Pnt expression, then the time to resolve the ratio change would likely be too slow. By using a system in which a two-fold shift in the ratio triggers cells to switch states, transient RTK inputs are sufficient to change the relative levels of the two proteins, thereby robustly and rapidly achieving the objective.

Second, a ratio-sensing system can define precise temporal windows of competence for differentiation. In the eye, competence of cells to respond to RTK signals is not only dependent on the transduction machinery but also on the presence of Yan and Pnt, which regulate target gene transcription downstream of RTKs. Eye progenitor cells anterior to the MF do not express Yan and Pnt, likely to prevent these cells from inappropriately responding to RTK activity. Over a 50-hour time window as the MF moves across the eye disc, progenitor cells undergo the first phase of cell fate specification during which R cells and cone cells are recruited into ommatidia. Since fate specification occurs as a wave across the eye, posterior ommatidia complete the first specification program ~50 hours before anterior ommatidia do. Yet posterior progenitor cells do not carry on with further fate transitions until the anterior ommatidia complete the initial induction of R cells and cone cells (Cagan and Ready, 1989). It is only then, that progenitor cells resume undergoing RTK-induced fate transitions – forming the pigment cells (Freeman, 1996; Miller and Cagan, 1998). We suggest that Yan and Pnt are shut off after the R-cell specification phase in order to create a multi-hour gap in progenitor competence that separates the two phases of fate specification.

Finally, a ratio-based mechanism can confer ultrasensitivity to modest changes in molecular stoichiometry when it is coupled to additional molecular mechanisms, for example, competition and cooperativity in enhancer occupancy. Although the ratio of Pnt and Yan

shifts only by two-fold when cell state transitions occur, this change is sufficient to produce robust regulation of the transit to differentiation. Our biophysical model shows how the composition of transcription factor complexes on Yan/Pnt target genes could dramatically change when a two-fold ratio change occurs, going from a repressive to an activating composition. Stabilizing interactions between Yan molecules bound to adjacent DNA sites sensitize enhancer occupancy to the relative abundance of the competing transcription factors. Indeed, prior work has shown that DNA-bound Yan monomers enhance recruitment of Yan to adjacent binding sites through cooperative protein-protein interactions (Lachance et al., 2018; Qiao et al., 2004). Given that cooperativity and competition are features of many biological systems, ratiometric control mechanisms analogous to the one we have described in the Drosophila eye could prove to be a broadly used regulatory strategy.

2.6 Materials and Methods

2.6.1 Genetics

Unless otherwise noted, all experiments were performed with animals reared at 25°C on standard cornmeal-molasses Drosophila culture medium. Fly stocks from the Bloomington Stock Center were as follows: pnt-GFP BAC transgene, BL42680; pnt2, BL2222; H2Av-mRFP BL23650, pntΔ88 (O'Neill et al., 1994); sev>RasV12 (Fortini et al., 1992); sev>YanACT (Rebay and Rubin, 1995); and sev>Gal4 (Jemc and Rebay, 2006). Measurements of wild-type dynamics of Pnt-GFP were made in eye discs from w1118; pnt-GFP; pntΔ88/pnt2, H2Av-mRFP. The effect of pnt gene dose on Pnt/Yan ratio was measured in w1118; pnt-gfp/+ ; pntΔ88/pnt2 (1× pnt) and w1118; pnt-gfp; pntΔ88/pnt2 (2× pnt) discs. In experiments with RasV12 or YanACT, Pnt-GFP and Yan levels were measured in eye discs from animals of the following genotypes: w1118; pnt-gfp/sev>RasV12; pntΔ88/+ relative to w1118; pnt-gfp/+; pntΔ88/+; and w1118; pnt-gfp/sev>YanACT; pntΔ88/+ relative to

w1118; pnt-gfp/+; pntΔ88/+. Yan mutant eye clones were generated using the yan833 null allele, ey>FLP and the FRT40 crossover point. Yan+ cells were labeled using the clonal marker Ubi>mRFPNLS (Bloomington Stock 34500). Developing eyes were dissected from white prepupae carrying w, ey>FLP ; pnt-gfp, yan833, FRT40A/ pnt-gfp, Ubi>mRFPNLS, FRT40A.

2.6.2 Immunohistochemistry

All eye-antennal discs were dissected from white prepupae. For experiments in which $4x'$,6-diamidino-2-phenylindole (DAPI) was used as the nuclear marker, samples were fixed in 4% paraformaldehyde (w/v)/PBS-Triton X-100 0.1% (v/v) for 25 min at room temperature. After fixation, eye discs were blocked in PBS-Triton X-100 0.1% (v/v) and 1% (v/v) normal goat serum for 30 min at room temperature. Primary and secondary antibodies were incubated each for 2 h at room temperature with DAPI. Including DAPI during primary and secondary antibody staining was important for its even penetration. Antibodies used were as follows: mouse anti-Yan 8B12 (DHSB, 1:200) and goat anti-mouse Cy3 (1:2000, Jackson Immunoresearch). Discs were washed for 20 min in PBS-Triton X-100 0.1% (v/v), and mounted in VectaShield Antifade mounting medium (Vector Laboratories). To avoid rapid changes in cell volume, which would affect relative fluorescence measurements, the substitution of PBS-Triton X-100 with VectaShield was carried out gradually using increments of ∼33% in VectaShield concentration. For experiments in which H2Av-mRFP was used as the nuclear marker, discs were fixed in 4% (w/v) paraformaldehyde/PBS for ∼45 min and goat anti-mouse Pacific Blue antibody (Life Technologies, 1:200 dilution) was used as the secondary antibody.

Samples were imaged within 18-24h after fixation. 1024×1024 16-bit images were captured using either a Zeiss LSM880 or a Leica SP5 confocal microscope equipped with 40× oil objectives. Discs were oriented with the equator (dorsal-ventral boundary) parallel to

the x-axis of the image. Optical slices were set at 0.8μm (45-60 optical slices per disc) with an additional digital zoom of 1.2-1.4. Images were recorded a region of at least six rows of ommatidia on each side of the equator. All discs for a given perturbation were fixed, mounted and imaged in parallel with control discs to reduce measurement error. Sample preparation, imaging and analysis were not performed under blind conditions, nor was sex of the animals noted at the time of dissection.

2.6.3 Quantification of expression levels

A minimum of three experimental replicate eye discs were imaged for each experimental condition, with the exception of the sev>RasV12 condition for which only one control disc was used. For each wild-type eye disc, ∼160 R8, ∼320 R2/5, ∼300 R3/4, ∼260 R1/6, ∼220 R7 and >1000 progenitor cells were manually selected from a set of computationally 2D-segmented nuclei using Silhouette. The fluorescence intensity of each nucleus was measured by our established procedure (Peláez et al., 2015). Our new pipeline includes Silhouette, an open-source package for macOS that integrates our image segmentation algorithm with a GUI (graphical user interface) for cell type annotation (available at: https://www.silhouette.amaral.northwestern.edu/). Subsequent analysis and visualization procedures were carried out using the accompanying FlyEye python package, which is freely available on GitHub. The resulting expression dynamics were inferred from confocal image stacks using an updated version of the segmentation and annotation pipeline described previously (Peláez et al., 2015). Empirically, the replicate number of three eye discs provided reproducible composite dynamics. No outlier eye discs were found for any experimental condition, so all samples were pooled, temporally aligned and included in the analysis. The sample size was not pre-computed before data collection.

Nuclear segmentation was performed using either H2Av-mRFP (Figs 1, 3, 4, Figs S1-S2, S4 and S5) or DAPI (Figs 2, 5 and Figs S7-S9) as a reference channel to identify individual nu-

clear boundaries. Nuclei expressing H2Av-mRFP were segmented using the Silhouette app's optimized implementation of the GraphCut algorithm (Qi, 2013), while DAPI-stained discs were segmented using an open-source implementation of the watershed algorithm designed to mitigate the effect of bright spots caused by DAPI accumulation in nucleoli (Bernasek et al., 2020). Each layer of the confocal stack was segmented in 2D independently in both cases. For each nucleus measured, a single contour was manually selected and assigned a cell type (cell state) using Silhouette, with care taken to avoid annotating nucleolar contours rather than complete nuclear contours in DAPI-stained discs. As each nucleus spans multiple optical sections, contours were always selected in the middle of the nucleus. The population of progenitor cells was measured in multiple layers to include cells whose nuclei were at different positions in the apical-ventral axis of the eye disc. In all cases, each nucleus was measured only in one layer to avoid measurement pseudoreplication.

Cells were identified and annotated by nuclear position in xyz space and shape using the His2Av-mRFP channel with the other channels turned off and the marker protein channel turned off. To validate the accuracy of this method of cell identification/annotation, cells were also immunostained for various cell-type-specific marker proteins, as noted in Table S1. Annotations made with tie His2Av-mRFP channel were then compared with the cell-type specific markers visualized in other channels. For each annotated nucleus contour, expression measurements were obtained by normalizing the mean fluorescence of the Pnt-GFP and Yan antibody channels by the mean fluorescence of the reference channel used for segmentation (DAPI or His2Av-mRFP). This normalization mitigates variability due to potentially uneven sample illumination, segment area, nuclear volume and, in the case of His-RFP, differences in protein expression capacity between cells.

2.6.4 Conversion of distance to time

Cell positions along the anterior-posterior axis were mapped to developmental time as described previously (Peláez et al., 2015). This is predicted based on two assumptions: the furrow proceeds at a constant rate of one column of R8 neurons per 2 h; and minimal cell migration occurs. For each disc, Delaunay triangulations were used to estimate the median distance between adjacent columns of R8 neurons (Fortune, 1992). We used the median rather than the mean distance because it minimized the influence of non-adjacent R8s that were falsely identified by the triangulation (Peláez et al., 2015). Dividing the furrow velocity of 2 h per column by this median distance yields a single conversion factor from position along the anterior-posterior axis to developmental time. This factor was applied to all cell measurements within the corresponding disc, yielding expression time series. Notably, these are not single cell dynamics, but rather aggregate dynamics across the developmental time course of a spatially organized cell population.

Computation of moving averages and confidence intervals Moving averages were computed by first-order Savitzky-Golay filtration (Savitzky and Golay, 1964). This method augments the simple windowing approach used by Peláez et al. (2015) by enabling visualization of expression trends at early time-points that are otherwise obscured by large window sizes. A secondary first-order filtration with one-fifth the original window size was applied to smooth lines for visualization purposes. None of our conclusions is sensitive to our choice of filtration or smoothing method. Primary window sizes of 250 and 75 cells were used for reporting the expression of multipotent and differentiated cells, unless noted otherwise. Confidence intervals for the moving average were inferred from the 2.5th and 97.5th percentile of ten-thousand-point estimates of the mean within each window. Point estimates were generated by bootstrap resampling with replacement of the expression levels within each window.

2.6.5 Alignment of expression data

To align multiple eye disc samples, cells of each sample were aligned with a reference population by shifting them in time. The magnitude of this shift was determined by maximizing the cross-correlation of progenitor Pnt-GFP expression Y(t) with the corresponding reference time series X(t). Rather than raw measurements, moving averages within a window of ten cells were used to improve robustness against noise.

For each experimental treatment, a disc was randomly chosen and shifted in time, such that time zero corresponds to the first annotated R8 neuron. This disc then served as the reference population for the alignment of all subsequent biological replicates within the treatment. Similarly, different experimental treatments (e.g. control and perturbation) were aligned by first aligning the discs within each treatment, then aggregating all cells within each treatment and repeating the procedure, with the first treatment serving as the reference.

This approach differs from the previous implementation of our pipeline in which discs were manually aligned by the inflection point of their Yan-YFP expression profiles (Peláez et al., 2015). Manual alignment entails arbitrary prioritization of certain dynamic features over others. Our revised protocol yields consistent, reproducible alignment of expression time series that equally weighs the entire time course. The automated approach is more principled but less robust than the manual approach. Specifically, it fails when dynamic forms qualitatively differ between experimental treatments. In these instances, we revert to manual alignment using the inflection point of Pnt-GFP induction as a reference.

2.6.6 Analysis of yan clones

We used ey>FLP and FRT40A to generate yan833 null clones within 23 eye discs carrying the Pnt-GFP transgene (see Genetics section). The chromosome carrying the wild-type yan allele was marked with a Ubi>mRFPNLS transgene. Discs were dissected, fixed and co-stained with DAPI before confocal imaging of RFP and GFP. Images of 36 unique vertical

cross-sections spanning non-overlapping cells were collected in total. Images were analyzed using our previously published method called Fly-QMA that quantitatively analyzes mosaic imaginal discs (Bernasek et al., 2020). Briefly, cell nuclei were identified by watershed segmentation of the DAPI channel. For each segment, Ubi-mRFPNLS and Pnt-GFP fluorescence was quantified by normalizing the average intensity of all pixels within the respective fluorescence channel by the average DAPI fluorescence. Fluorescence bleed-through between the GFP and RFP channels was corrected as described previously (Bernasek et al., 2020). Bleed-through correction successfully eliminated any detectable difference in Pnt-GFP expression in mosaic eye discs in which all cells were wild type for yan. The same correction procedure was therefore applied to all measurements of mosaic eye discs containing yan mutant clones.

Mitotic recombination between the mutant and wild-type chromosomes yields cell populations exhibiting low, medium and high levels of RFP fluorescence, which correspond to cells with 0, 1 and 2 copies of the wild-type allele, respectively (Bernasek et al., 2020). Segmented nuclei were assigned to one of three groups using a k-means classifier. The procedure was validated through comparison with manual annotation of 2500 cells. The overall classification rate was 95%, Cells residing on the border of clones were excluded from all analyses to mitigate edge effects. The remaining measurements were aggregated across all eye discs for comparison of cells with zero versus one copy of the wild-type yan gene.

2.6.7 Simple competitive binding model

Fig. S10 presents results for an equilibrium model of two species, Yan (Y) and Pnt (P), competing for a finite pool of shared binding sites, S: formula where KD,Yan and KD,Pnt are equilibrium association constants and SY and SP denote the bound species. Applying a mass balance to the total protein and binding site (S0) abundances: formula yields an analytically tractable system of nonlinear equations (Wang, 1995). For each pair of absolute

protein abundances (Y0, P0), the Pnt binding site occupancy is simply SP/S0.

2.6.8 Competitive binding model with cooperativity

The model presented in Fig. 6 and Fig. S11 expands upon the work of Hope et al. (2017). The model is based on a single cis-regulatory element consisting of n adjacent binding sites, each of which may be designated as ETS or non-ETS. Each binding site may exist in one of three binding states: bound by a single copy of Yan, bound by a single copy of Pnt or unbound. Thermodynamic potentials were assigned to each binding state using two parameters for each transcription factor. The parameter X defines the free energy of transcription factor X binding to an ETS site, while X defines the free energy of binding to a non-ETS site. A unique configuration of binding states for all n binding sites constitutes a single microstate, k. The thermodynamic potential of each microstate was taken to be the sum of thermodynamic potentials for each of its constituent binding sites. For each microstate, the stabilizing effect of polymerization was incorporated via a third parameter, X, that defines the free energy of SAM-SAM binding between a pair of similar transcription factors bound to adjacent sites. The net result is a total thermodynamic potential, Gk, for each microstate. An example enumeration of all possible microstates for an element consisting of one ETS site preceding two non-ETS sites is provided in Fig. S11B. The statistical frequencies of each microstate were evaluated by constructing a canonical ensemble: formula in which pk is the statistical frequency of microstate k, [P] is the Pnt concentration, [Y] is the Yan concentration, aP(k) and aY(k) are functions representing the number of bound molecules of P and Y within microstate k, T is a fixed temperature set to 300K, and R is the gas constant. Fractional occupancies for each binding site correspond to the cumulative statistical frequency of all microstates in which the site is occupied by a given transcription factor. Overall fractional occupancies are similarly evaluated across all sites within the element. We considered regulatory elements comprising 12 binding sites in which only the first site carries the ETS

designation. We retained the same parameterization of Yan binding used by Hope et al. (2017): Y=9.955 kcal/mol, Y=5.837 kcal/mol and Y=7.043 kcal/mol. We parameterized Pnt-binding thermodynamics to provide balanced competition between Pnt and Yan in the absence of any SAM-mediated polymerization of Pnt. That is, we set Pnt-binding affinities such that the transition from Pnt to Yan occupancy occurs when Pnt and Yan concentrations are approximately equal. The model used to generate Fig. 6B assumes that Pnt binds individual sites with elevated affinities P=0.96(Y+Y) and P=0.96(Y+Y). The model used to generate Fig. 6A uses these same elevated binding affinities for Yan, while setting Y=0 kcal/mol. Qualitatively, our results are not sensitive to this parameterization.

Statistical frequencies of all microstates were enumerated using a recursive graph-traversal algorithm implemented in python, where each node in the graph represents an individual binding site and each edge reflects occupancy by a specific transcription factor. An open-source python implementation of this modeling framework has been published on GitHub (https://doi.org/10.5281/zenodo.5520493).

Acknowledgements

We thank members of the Amaral, Carthew and Rebay labs for helpful discussions during the course of this project, and C. LaBonne and M. Glotzer for helpful comments on the manuscript. We thank Kevin White's lab and the MODEncode Consortium for recombineering the Pnt-GFP transgene. the Bloomington Drosophila Stock Center for flies, Laura Nilson for the use of computational resources, the Developmental Studies Hybridoma Bank for antibody reagents and the Northwestern Biological Imaging Facility (BIF) for technical imaging support.

Figure S1

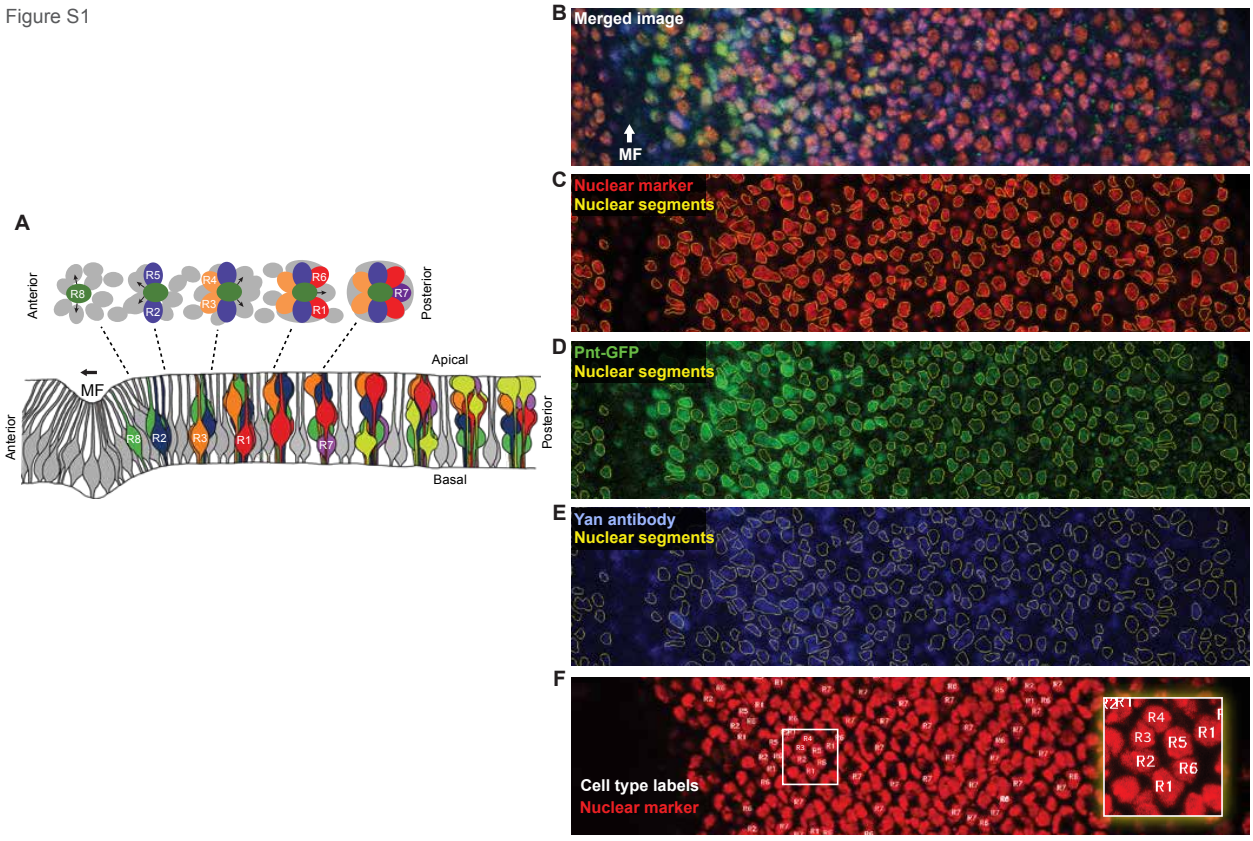

Figure 2 S1: Identification of cell nuclei in eye discs

(A) Schematic cross-sectional view along the anteroposterior axis of an eye disc, showing the epithelial invagination that marks the MF and the assembling clusters of differentiated cells forming ommatidia. Note the dynamic and characteristic apical-basal positions of nuclei from progenitor cells (grey), R cells (various colors), and cone cells (light green). Shown above is a apical-basal view of five ommatidia sampled at different locations along the anteroposterior axis. The relative positions of progenitor and R cell nuclei are highly stereotyped in this plane. Arrows denote signals transmitted from the R8 to neighboring progenitor cells and transduced via the RTK – Ras – MAPK pathway. Adapted from (Peláez et al., 2015).(B-E) Quantitative analysis pipeline showing microscopy and segmentation of one representative eye disc. Shown is a partial cross-section of a confocal z-stack with the merged channels (B) and separate fluorescence channels for Histone- RFP (C), Pnt-GFP (D), and Yan antibody (E). MF marks the location of the MF. Yellow lines depict nuclear contours segmented by FlyEye Silhouette.(F) Manual annotation of R-cell types. Image shows a projection of His-RFP fluorescence across the Z-sections spanning R cell nuclei. The region imaged is identical to that shown in (B-E). The R-cell types are documented and are labelled as shown. The inset shows an ommatidium magnified to see nuclear positions.

Figure S2

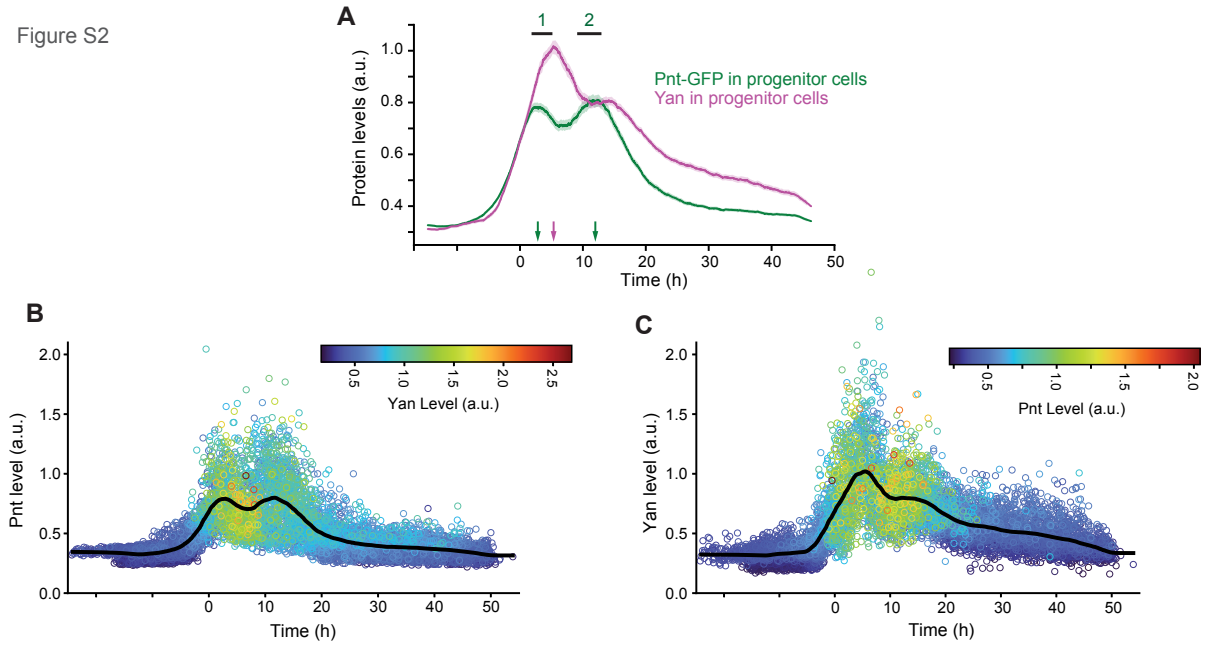

Figure 2 S2: Pnt-GFP and Yan expression in progenitor cells

(A) Line averages of Pnt-GFP (green) and Yan (magenta) levels in progenitor cells. Lines are smoothed moving averages across 500 sequential progenitors, shaded regions are 95% confidence intervals. Arrows indicate the times at which local maxima occur. Note the peaks of Pnt and Yan expression take place sequentially in progenitor cells.(B, C) Expression levels of Pnt-GFP (B) and Yan (C) in individual progenitor cells, color-coded in accordance with the expression level of the complementary transcription factor, as shown at the right of each panel. Datapoints are arranged such that those with higher levels of the complementary transcription factor are displayed in front. Lines are smoothed moving averages across 500 sequential progenitor cells.

Figure S3

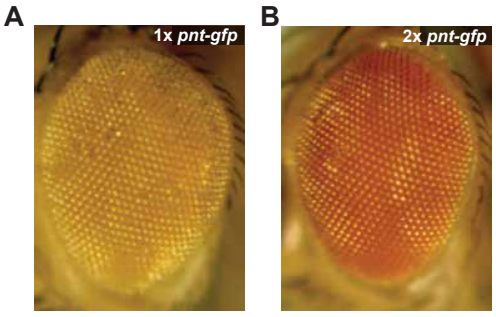

Figure 2 S3: Characterization of Pnt-GFP

Adult eyes of flies carrying (A) one or (B) two copies of the pnt-gfp transgene in a pnt null mutant background.

Figure S4

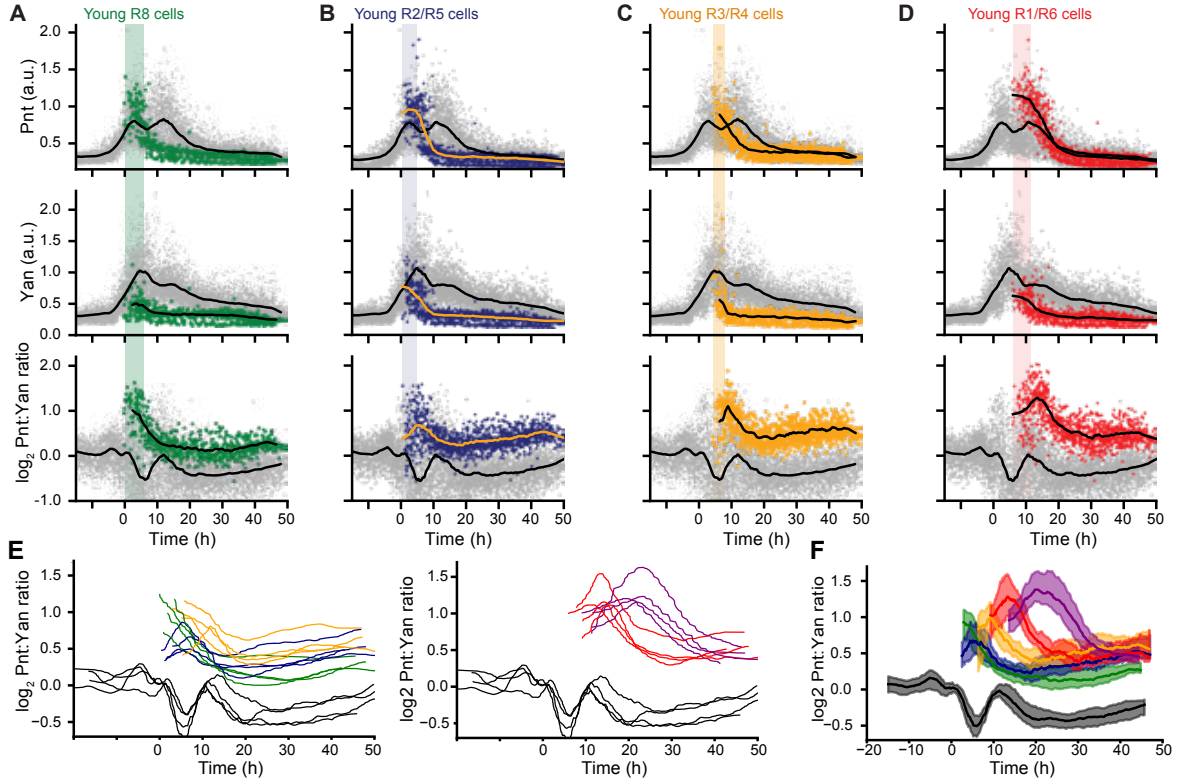

Figure 2 S4: Expression dynamics for various eye cell types

(A-D) Pnt-GFP and Yan protein levels, and the log2-transformed Pnt/Yan ratio for individual progenitor (grey), and R8 cells (A), R2/R5 cells (B), R3/R4 cells (C), and R1/R6 cells (D) over developmental time. Progenitor cells are present across all time (grey arrow) while R cells arise later in time (colored arrows). Solid black lines are smoothed moving averages across 250 and 75 individual nuclei for progenitor and R cells, respectively. Black bars labeled 1 and 2 denote the two peaks of Pnt-GFP in progenitor cells. Shaded vertical stripes highlight these two regions, which coincide with transition of progenitor to various R fates.(E) Moving averages for the log2-transformed Pnt/Yan ratio for progenitor (grey), R8 cells (green), R2/R5 cells (blue), R3/R4 cells (orange), R1/R6 cells (red), and R7 cells (purple) over developmental time from individual WT replicate discs. (F) Moving averages for the log2-transformed Pnt/Yan ratio for progenitor (grey), R8 cells (green), R2/R5 cells (blue), R3/R4 cells (orange), R1/R6 cells (red), and R7 cells (purple) from three replicate WT discs, along with 95% confidence interval of the mean for each cell type obtained from two-stage hierarchical bootstrapping approach that first subsamples the independent discs, then subsamples the cell measurements within them.

Figure S5

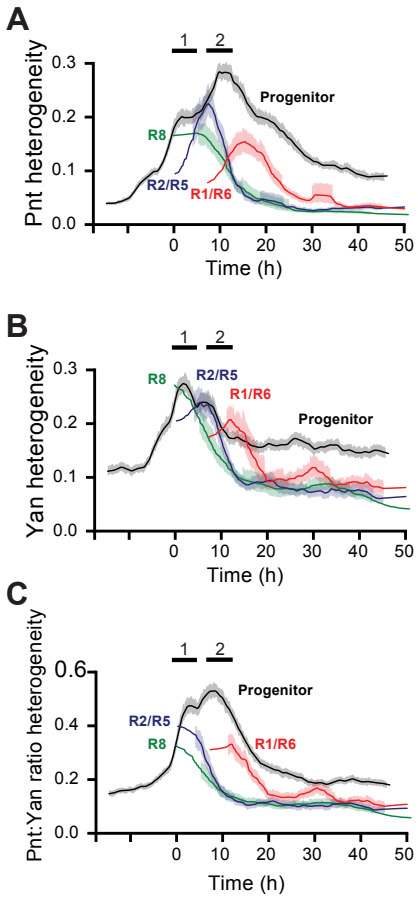

Figure 2 S5: Heterogeneity of Yan and Pnt expression

Cell-to-cell expression heterogeneity of Pnt-GFP (A), Yan (B), and the log2-transformed Pnt/Yan ratio (C). Heterogeneity was estimated by de-trending fluctuations about a moving average of 250 sequential cells as described (Paleaz 2015). Lines shown are moving averages of 250 sequential fluctuations. Shaded regions are 95% confidence intervals. Colors denote cell type.

Figure S6

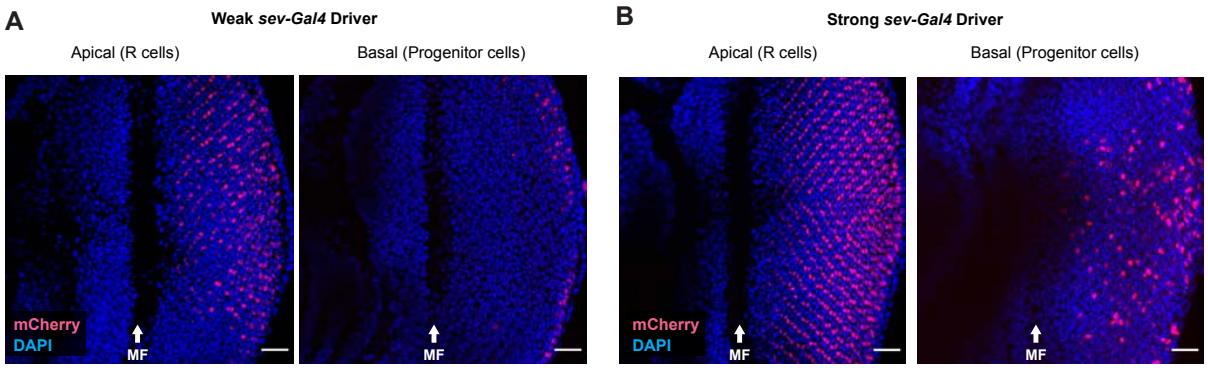

Figure 2 S6: Sev-Gal4 drives UAS transgene expression in subsets of progenitor cell

Projection of sev-Gal4>UAS-NLS-mCherry protein fluorescence (red) and DAPI (blue) in apical z-slices containing R-cell nuclei and basal z- slices containing progenitor cell nuclei. Anterior is left and MF is morphogenetic furrow. Scale bars denote 20 μm. (A) Expression from a weak sev-Gal4 driver. (B) Expression from a strong sev-Gal4 driver. When expressed from the weak driver, mCherry is predominantly limited to R7, R3, and R4 cells. When expressed from the strong driver, mCherry is also detected in some progenitor cells.

Figure S7

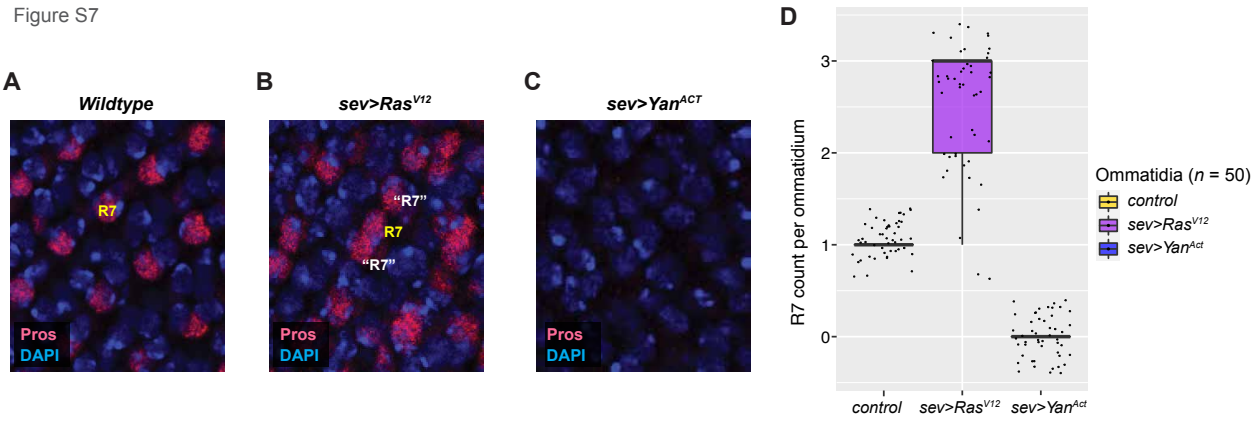

Figure 2 S7: Fate specification errors in sev>RasV12 and sev>YanACT

Prospero (Pros) protein fluorescence in a region of the eye disc where ommatidia have recruited cells to form R7 cells. Pros is specifically expressed in R7 cells. (A) Ommatidia in wildtype eye disc, which have one Pros-positive cell per ommatidium. (B) Some ommatidia in a sev>RasV12 eye disc have ectopic R7 cells labelled "R" (white font) that are Pros-positive and located beside the endogenous R7 cells (labeled with yellow font). (C) Ommatidia from a sev>YanACT eye disc have no Pros-positive cells, indicating that they lack R7 cells. S7D. Number of R7 cells positive for the R7 marker Prospero in individual mature ommatidia of wildtype, sev>RasV12 and sev>YanACT discs (n = 50 ommatidia analyzed for each group across 3 replicate discs). Boxplots are median and quartiles.

Figure S8

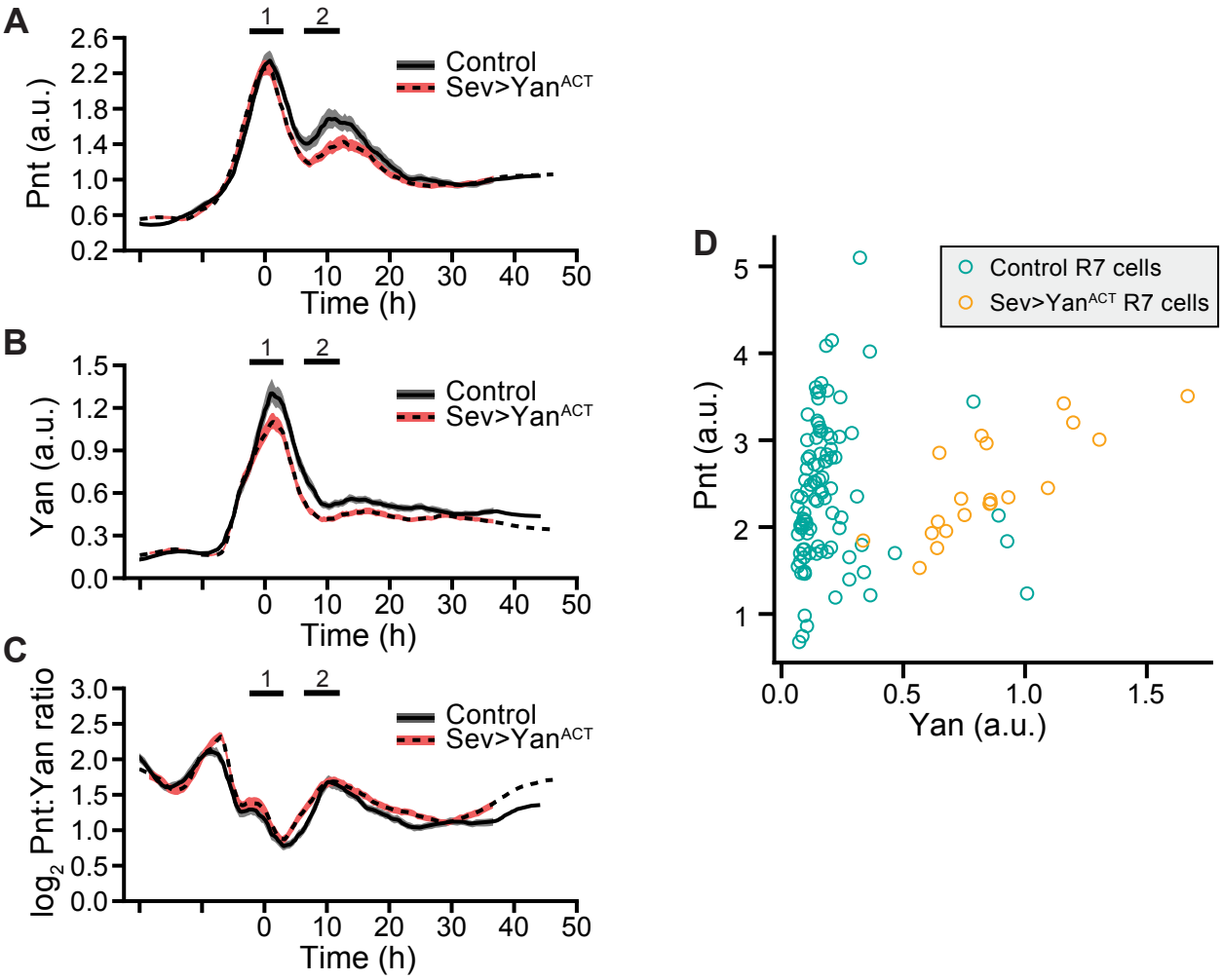

Figure 2 S8: YanACT has no effect on the Pnt/Yan ratio in progenitor cells but does alter Yan levels in R7 cells

(A-C) Effect of sev>YanACT expression on Pnt-GFP (A), Yan (B), and the log2-transformed Pnt/Yan ratio (C) in progenitor cells over time. Lines are moving averages across 250 sequential cells. Shaded regions are 95% confidence intervals. Black lines denote wildtype cells. Red lines denote sev>YanACT cells. (D) Joint distribution of Pnt and Yan levels among R7 cells sampled from the region spanning 15 to 20 h of developmental time. The sev>YanACT R7 cells show higher Yan levels ($P<0.001$, one-tailed Mann-Whitney U test), while Pnt levels are indistinguishable from wildtype ($P=0.4$, KS 2-sample test).

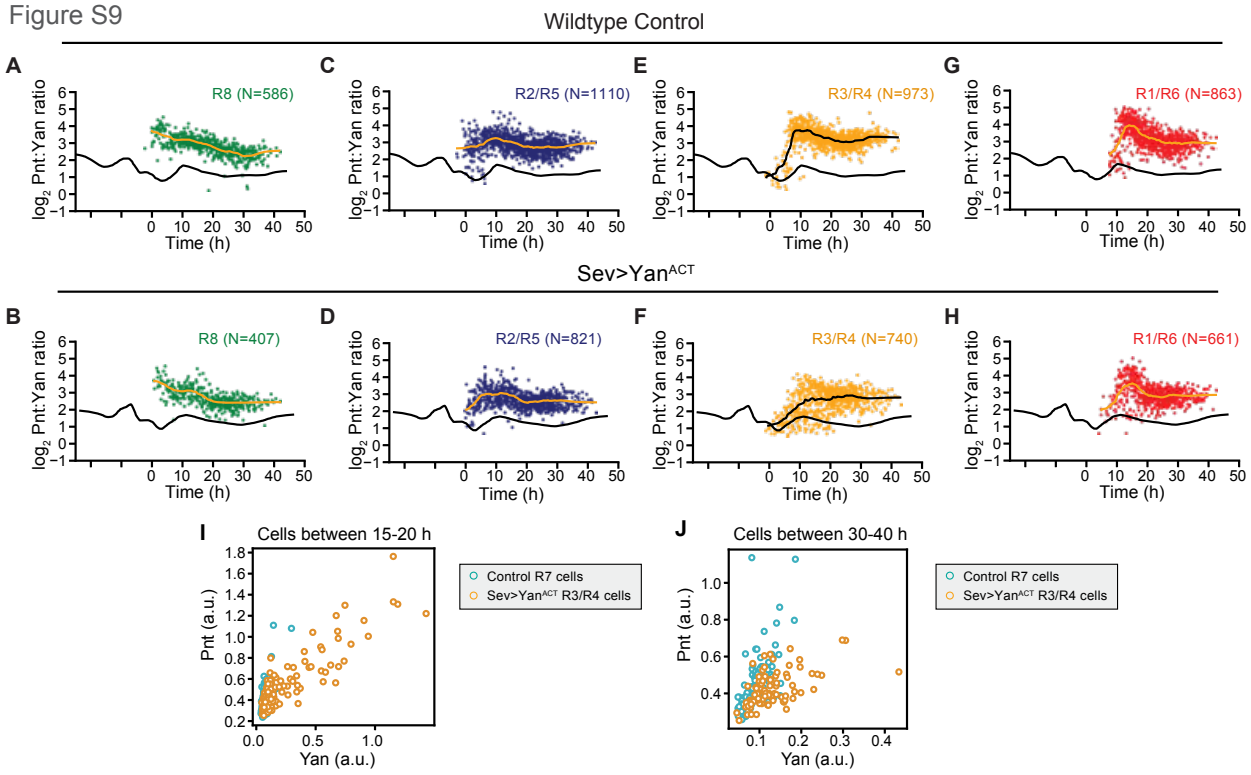

Figure 2 S9: YanACT has a weak effect on the Pnt/Yan ratio in R3/R4 cells
The log2-transformed Pnt/Yan ratio for individual R cells (dots) and their moving line averages (black lines) over developmental time in wildtype (A,C,E,G) and sev>YanACT eyes (B,D,F,H). The moving line average for progenitor cells (black lines) is also shown in each plot. Black bars labeled 1 and 2 denote the two peaks of Pnt-GFP in progenitor cells, which coincide with transition of progenitor to various R fates. Various R cell types are shown, R8 (A,B), R2/R5 (C,D), R3/R4 (E,F), and R1/R6 (G,H). S9 (I, J) Joint distribution of Pnt and Yan levels among R3/R4 cells sampled from the region spanning 15 to 20 h (I) and 30 to 40 h (J) of developmental time. In both cases, a KS 2-sample test on the log2 Pnt:Yan ratio yields $p < 0.00001$.

Figure S10

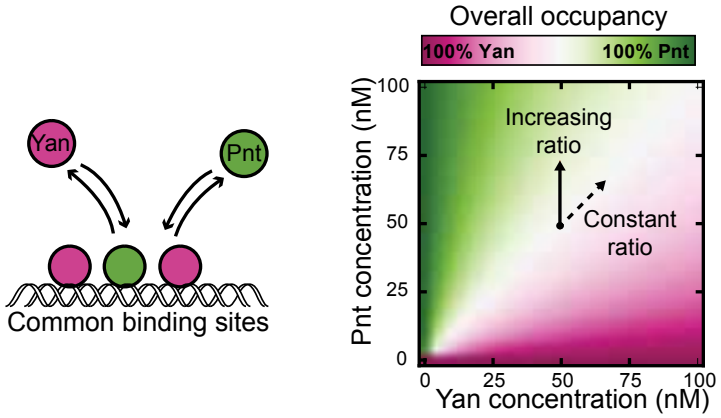

Figure 2 S10: A simple equilibrium model of Yan / Pnt competition for DNA binding sites

Schematic of a simple two-species competitive binding model (left), alongside theoretical Pnt site occupancy as a function of transcription factor abundance (right). Equivalent binding affinities are used for illustrative purposes. Color scale reflects overall Pnt site occupancy. A diverging scale was used because all sites are fully saturated at total transcription factor concentrations above 1 nM. Under the range of conditions shown, this implies that Yan occupies all sites left vacant by Pnt. Simultaneous proportional increases in absolute abundance of both species have minimal impact on Pnt occupancy (dashed arrow), while varying ratio confers gradual change in occupancy (solid arrow).

2.7 Supplemental Figures

2.8 References

Acloque, H., Ocana, O. H., Matheu, A., Rizzoti, K., Wise, C., Lovell-Badge, R., and Nieto, M. A. (2011). Reciprocal repression between sox3 and snail transcription factors defines embryonic territories at gastrulation. *Developmental cell*, 21(3):546–558.

Adler, M., Szekely, P., Mayo, A., and Alon, U. (2017). Optimal regulatory circuit topologies for fold-change detection. *Cell systems*, 4(2):171–181.

Alon, U., Surette, M. G., Barkai, N., and Leibler, S. (1999). Robustness in bacterial chemotaxis. *Nature*, 397(6715):168–171.

Figure S11

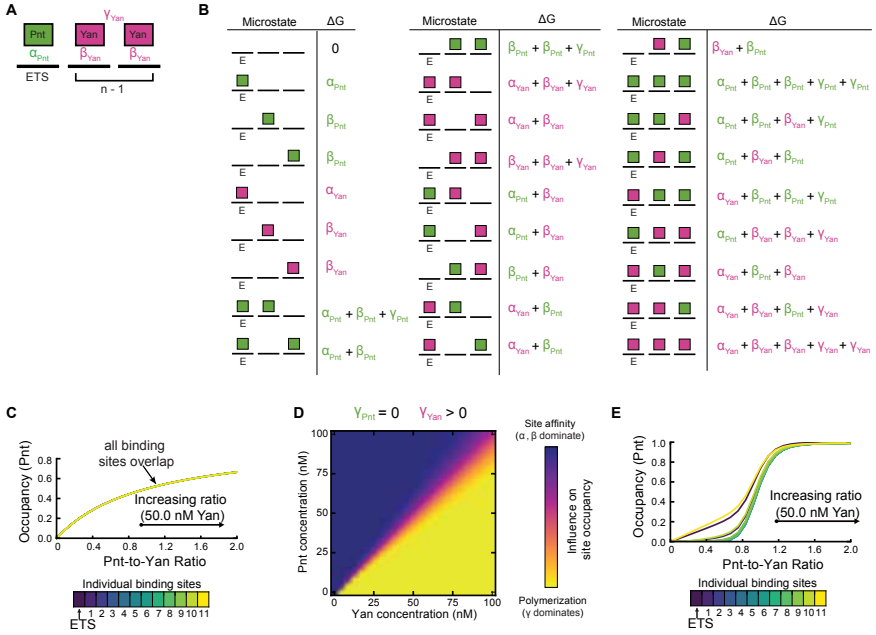

Figure 2 S11: A simple equilibrium model of Yan / Pnt competition for DNA binding sites

(A)Summary of thermodynamic interactions within one microstate of a cis-regulatory element containing one ETS site and two non- ETS sites. Solid black lines represent individual binding sites. Green and magenta rectangles denote Pnt and Yan molecules. Example thermodynamic potentials of strong ETS-binding, weak non-ETS binding, and polymerization interactions are denoted by αPnt, βYan, and γYan, respectively. For this microstate, αP(k)=1 and αY(k)=2. (B) Enumeration of all possible microstates for a cis-regulatory element of length 3 in which only the first site carries the ETS designation. Solid black lines denote binding sites, green and magenta rectangles denote bound Pnt and Yan molecules. The cumulative thermodynamic potentials of each microstate, δG k, are listed beside each graphical depiction. (C) Pnt occupancy of individual binding sites as a function of Pnt/Yan ratio in the absence of Yan polymerization. Contours correspond to a vertical path traversed across Figure 6A at a fixed Yan concentration of 50 nM. All binding sites behave similarly. (D) Relative thermodynamic contributions of binding site affinity versus polymerization to microstate statistical frequencies as a function of Pnt and Yan concentration. For each point in the plane, the influence of site affinity was calculated by weighting the sum of all ETS and non-ETS thermodynamic potentials for each microstate by the statistical frequency of the corresponding microstate. The influence of polymerization was similarly determined. The color scale reflects the relative magnitude of these two summations, normalized by limits of zero and complete polymerization. (E) Pnt occupancy of individual binding sites as a function of Pnt/Yan ratio when Yan polymerizes. Contours correspond to a vertical path traversed across Figure 6B at a fixed Yan concentration of 50 nM. Line colors reflect binding site positions within the cis- regulatory element. Sites at intermediate distances from the strong ETS site (green lines) transition at higher ratios than those nearest and furthest from the strong ETS site (blue and yellow lines).

Table S1. Accuracy of cell-type identification

Cell Type Scored	Cell Type Marker	Sample Size (*n*)	% Correct Type Assignment
Progenitors	Elav (-)	462	99.8
All mature Rs	Elav (+)	258	99.2
R8	Sens	106	97.2
R2/R5	Rough	162	99.4
R3/R4	Rough	142	100.0
R7	Pros	104	99.0

Figure 2 S12: Accuracy of cell-type identification

Barkai, N. and Leibler, S. (1997). Robustness in simple biochemical networks. *Nature*, 387(6636):913–917.

Basler, K. and Hafen, E. (1989). Dynamics of drosophila eye development and temporal requirements of sevenless expression. *Development*, 107(4):723–731.

Bernasek, S. M., Hur, S. S., Peláez-Restrepo, N., Lachance, J.-F. B., Bakker, R., Navarro, H. T., Sanchez-Luege, N., Amaral, L. A., Bagheri, N., Rebay, I., et al. (2023). Ratiometric sensing of pnt and yan transcription factor levels confers ultrasensitivity to photoreceptor fate transitions in drosophila. *Development*, pages dev–201467.

Briscoe, J. and Ericson, J. (2001). Specification of neuronal fates in the ventral neural tube. *Current opinion in neurobiology*, 11(1):43–49.

Brunner, D., Dücker, K., Oellers, N., Hafen, E., Scholzi, H., and Klambt, C. (1994). The ETS domain protein Pointed-P2 is a target of MAP kinase in the Sevenless signal transduction pathway. *Nature*, 370(6488):386–389. Number: 6488 Publisher: Nature Publishing Group.

Campos-Ortega, J. and Hofbauer, A. (1977). Cell clones and pattern formation: On the

lineage of photoreceptor cells in the compound eye of drosophila. *Wilhelm Roux's archives of developmental biology*, 181:227–245.

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(1):75–85.

Fortini, M. E., Simon, M. A., and Rubin, G. M. (1992). Signalling by the sevenless protein tyrosine kinase is mimicked by rasl activation. *Nature*, 355(6360):559–561.

Freeman, M. (1996). Reiterative use of the egf receptor triggers differentiation of all cell types in the drosophila eye. *Cell*, 87(4):651–660.

Frick, C. L., Yarka, C., Nunns, H., and Goentoro, L. (2017). Sensing relative signal in the tgf-β/smad pathway. *Proceedings of the National Academy of Sciences*, 114(14):E2975–E2982.

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(11):3355–3362.

Gallagher, K. D., Mani, M., and Carthew, R. W. (2022). Emergence of a geometric pattern of cell fates from tissue-scale mechanics in the drosophila eye. *Elife*, 11:e72806.

Goentoro, L. and Kirschner, M. W. (2009). Evidence that fold-change, and not absolute level, of β-catenin dictates wnt signaling. *Molecular cell*, 36(5):872–884.

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(1):223–230.

Graham, T. G. W., Tabei, S. M. A., Dinner, A. R., and Rebay, I. (2010). Modeling bistable

cell-fate choices in the Drosophila eye: qualitative and quantitative perspectives. *Development*, 137(14):2265–2278.

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(1):63–74.

Hollenhorst, P. C., McIntosh, L. P., and Graves, B. J. (2011). Genomic and biochemical insights into the specificity of ets transcription factors. *Annual review of biochemistry*, 80:437–471.

Hope, C. M., Rebay, I., and Reinitz, J. (2017). DNA Occupancy of Polymerizing Transcription Factors: A Chemical Model of the ETS Family Factor Yan. *Biophysical Journal*, 112(1):180–192.

Jaeger, J., Surkova, S., Blagov, M., Janssens, H., Kosman, D., Kozlov, K. N., Manu, Myasnikova, E., Vanario-Alonso, C. E., Samsonova, M., Sharp, D. H., and Reinitz, J. (2004). Dynamic control of positional information in the early Drosophila embryo. *Nature*, 430(6997):368–371.

Jemc, J. and Rebay, I. (2006). Characterization of the split ends-like gene spenito reveals functional antagonism between spoc family members during drosophila eye development. *Genetics*, 173(1):279–286.

Lachance, J.-F. B., 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. *Developmental biology*, 385(2):263–278.

Lachance, J.-F. B., 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 & Development*, 32(5-6):389–401.

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(4):609–620.

Laslo, P., Pongubala, J. M., Lancki, D. W., and Singh, H. (2008). Gene regulatory networks directing myeloid and lymphoid cell fates within the immune system. In *Seminars in immunology*, volume 20, pages 228–235. Elsevier.

Laslo, P., Spooner, C. J., Warmflash, A., Lancki, D. W., Lee, H.-J., Sciammas, R., Gantner, B. N., Dinner, A. R., and Singh, H. (2006). Multilineage transcriptional priming and determination of alternate hematopoietic cell fates. *Cell*, 126(4):755–766.

Lee, R. E., Walker, S. R., Savery, K., Frank, D. A., and Gaudet, S. (2014). Fold change of nuclear nf-κb determines tnf-induced transcription in single cells. *Molecular cell*, 53(6):867–879.

Lyashenko, E., Niepel, M., Dixit, P. D., Lim, S. K., Sorger, P. K., and Vitkup, D. (2020). Receptor-based mechanism of relative sensing and cell memory in mammalian signaling networks. *Elife*, 9:e50342.

Mackereth, C. D., Schärpf, M., Gentile, L. N., MacIntosh, S. E., Slupsky, C. M., and McIntosh, L. P. (2004). Diversity in structure and function of the ets family pnt domains. *Journal of molecular biology*, 342(4):1249–1264.

Melen, G. J., Levy, S., Barkai, N., and Shilo, B.-Z. (2005). Threshold responses to morphogen gradients by zero-order ultrasensitivity. *Molecular systems biology*, 1(1):2005–0028.

Meruelo, A. D. and Bowie, J. U. (2009). Identifying polymer-forming sam domains. *Proteins: Structure, Function, and Bioinformatics*, 74(1):1–5.

Mesibov, R., Ordal, G. W., and Adler, J. (1973). The range of attractant concentrations for bacterial chemotaxis and the threshold and size of response over this range: Weber law and related phenomena. *The Journal of general physiology*, 62(2):203–223.

Miller, D. T. and Cagan, R. L. (1998). Local induction of patterning and programmed cell death in the developing drosophila retina. *Development*, 125(12):2327–2335.

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

O'Neill, E. M., Rebay, I., 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(1):137–147.

Park, B., Ahrends, R., and Teruel, M. (2012). Consecutive Positive Feedback Loops Create a Bistable Switch that Controls Preadipocyte-to-Adipocyte Conversion. *Cell Reports*, 2(4):976–990.

Peláez, N., Gavalda-Miralles, A., Wang, B., Navarro, H. T., Gudjonson, H., Rebay, I., Dinner, A. R., Katsaggelos, A. K., Amaral, L. A., and Carthew, R. W. (2015). Dynamics and heterogeneity of a fate determinant during transition towards cell differentiation. *eLife*, 4:e08924. Publisher: eLife Sciences Publications, Ltd.

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: structural insights into the regulation of yan by mae. *Cell*, 118(2):163–173.

Rand, D. A., Raju, A., Sáez, M., Corson, F., and Siggia, E. D. (2021). Geometry of gene regulatory dynamics. *Proceedings of the National Academy of Sciences*, 118(38):e2109729118.

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(6):857–866.

Rogge, R., Green, P. J., Urano, J., Horn-Saban, S., Mlodzik, M., Shilo, B.-Z., Hartenstein, V., and Banerjee, U. (1995). The role of yan in mediating the choice between cell division and differentiation. *Development*, 121(12):3947–3958.

Sáez, M., Blassberg, R., Camacho-Aguilar, E., Siggia, E. D., Rand, D. A., and Briscoe, J. (2022). Statistically derived geometrical landscapes capture principles of decision-making dynamics during cell fate transitions. *Cell Systems*, 13(1):12–28.

Savitzky, A. and Golay, M. J. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Analytical chemistry*, 36(8):1627–1639.

Schaffer, A. E., Freude, K. K., Nelson, S. B., and Sander, M. (2010). Nkx6 transcription factors and ptf1a function as antagonistic lineage determinants in multipotent pancreatic progenitors. *Developmental cell*, 18(6):1022–1029.

Schrode, N., Saiz, N., Di Talia, S., and Hadjantonakis, A.-K. (2014). Gata6 levels modulate primitive endoderm cell fate choice and timing in the mouse blastocyst. *Developmental cell*, 29(4):454–467.

Shoval, O., Goentoro, L., Hart, Y., Mayo, A., Sontag, E., and Alon, U. (2010). Fold-change detection and scalar symmetry of sensory input fields. *Proceedings of the National Academy of Sciences*, 107(36):15995–16000.

Shwartz, A., Yogev, S., Schejter, E. D., and Shilo, B.-Z. (2013). Sequential activation of ets proteins provides a sustained transcriptional response to egfr signaling. *Development*, 140(13):2746–2754.

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(1):143–150.

Webber, J. L., Zhang, J., Cote, L., Vivekanand, P., Ni, X., Zhou, J., Nègre, N., Carthew, R. W., White, K. P., and Rebay, I. (2013). The relationship between long-range chromatin occupancy and polymerization of the drosophila ets family transcriptional repressor yan. *Genetics*, 193(2):633–649.

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. *Development*, 145(dev165985).

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. *The EMBO journal*, 29(13):2147–2160.

Wolff, T. (1993). Pattern formation in the drosophila retina. *The development of Drosophila melanogaster*, pages 1277–1325.

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(1):87–97.

Yao, G., Lee, T. J., Mori, S., Nevins, J. R., and You, L. (2008). A bistable Rb–E2F switch underlies the restriction point. *Nature Cell Biology*, 10(4):476–482. Number: 4 Publisher: Nature Publishing Group.

Zhu, L. J., Christensen, R. G., Kazemian, M., Hull, C. J., Enuameh, M. S., Basciotta, M. D., Brasefield, J. A., Zhu, C., Asriyan, Y., Lapointe, D. S., et al. (2011). Flyfactorsurvey: a database of drosophila transcription factor binding specificities determined using the bacterial one-hybrid system. *Nucleic acids research*, 39(suppl_1):D111–D117.

CHAPTER 3

UNDERSTANDING HOW PNT AND YAN EXPRESSION LEVELS DETERMINE THE MRNA OUTPUT OF TARGET GENES

3.1 Statement of Contributions

Suzy S.J. Hur: conceptualization, methodology, investigation, data collection, visualization; Andrea Herman: software, validation, formal analysis, visualization; Hernan G. Garcia: supervision, methodology, funding acquisition; Ilaria Rebay: supervision, conceptualization, funding acquisition

3.2 Abstract

During tissue development, cells adopt increasingly tissue-specific states of gene expression. Producing the right amount of gene products at the right time is critical for cells to transition to a newly-specified state. In the developing Drosophila retina, multipotent progenitor cells face a binary decision – whether to differentiate into a photoreceptor neuron or remain as progenitor cells that differentiate into alternative, subsequent fates. This decision is mediated by RTK/MAPK signaling that modulates the concentration of the transcription activator Pointed and repressor Yan in the nucleus. Here, we investigate how Pointed and Yan protein abundance in the nucleus regulates the mRNA expression of key photoreceptor genes during progenitor to photoreceptor neuron transitions by combining immunostaining, single-molecule fluorescence in situ hybridization (smFISH), and quantitative image analysis. Our results provide insights into how the concentration of input TFs affects downstream transcription of fate-determining genes and ultimately influence cell fate choice.

3.3 Introduction

The transition of progenitor cells into specific cell fates at the right time in the right place is essential for normal tissue development, as any disruptions in this process can lead to developmental abnormalities. During multicellular tissue development, the acquisition of distinct cellular identities by progenitor cells is governed by transcription factors that regulate the expression of cell fate-specific genes. At critical cell fate decision points, cells in various developmental contexts encounter antagonistic transcription factors that either promote or inhibit a particular fate transition. Resolving these conflicting inputs is necessary for cells to establish the correct gene expression program. However, the underlying mechanisms by which antagonistic transcription factors elicit precise responses at individual fate-determining gene loci, coordinating accurate and reliable cell fate decisions, remain poorly understood.

Antagonistic transcription factors play a critical role in regulating cell fate decisions across various developmental contexts. Notable examples include PU.1 and GATA2, which govern the choice between erythroid and myeloid fates (Walsh et al., 2002), Nanog and OTX2, which control rostral and caudal fate in neural patterning (Acampora et al., 2017), Esg and Scute, which influence progenitor and enteroendocrine fate decisions (Li et al., 2017), and the Nkx6 and Ptf1a pair, which regulate the duct and acinar fates in pancreas development (Schaffer et al., 2010). Currently, the prevailing mechanism for regulating binary fate decisions by antagonistic transcription factors involves the reciprocal repression of each other's expression while promoting their own expression. This mutually exclusive expression pattern governs the adoption of specific gene expression programs by progenitor cells. However, it is noteworthy that in certain instances, antagonistic transcription factors are co-expressed in transitioning cells, suggesting the existence of additional mechanisms that enable accurate decision-making in developing tissues.

Pnt and Yan are a pair of antagonistic transcription factors that regulate the binary progenitor vs. photoreceptor fate decision in the Drosophila larval eye imaginal disc. Pnt and

Yan hypomorphic mutants lead to loss and gain of photoreceptor cell counts respectively. Pnt and Yan compete for common binding sites, and interestingly, they are co-expressed in progenitors as well as differentiating photoreceptor cells. In Chapter 2, I presented our published study showing that the ratio of Pnt/Yan, not the absolute level of either transcription factor, determines whether progenitor cells transition to photoreceptor fates (R1-R7) or remain multipotent ((Bernasek et al., 2023); Figure 3.1).

At spatial windows of the developing eye disc that correspond to progenitor vs. photoreceptor decision points, there is an increased heterogeneity in the Pnt/Yan ratio among the progenitor cell population (Chapter 2). Among these cells, ones that stably elevate the ratio of up to two-fold transition to photoreceptor cells. When the ratio was perturbed, this led to premature or failed transitions of progenitor cells to photoreceptor cells. This suggests that the expression of photoreceptor genes are regulated by the relative levels of Pnt and Yan.

In Chapter 2, equilibrium binding model was performed to explore DNA occupancy patterns at varying Pnt and Yan concentrations. This predicted an ultrasensitive response in binding occupancy to Pnt/Yan ratio changes, when Pnt and Yan competition for binding and Yan monomer's ability for cooperative recruitment were considered. The model's findings suggest that the binary switch in gene expression is primarily driven by the dominant binding of either the activator or the repressor molecules at individual gene loci, rather than being determined by a mutually exclusive expression pattern in the nucleus.

Single-molecule fluorescence in situ hybridization (smFISH) has emerged as a powerful approach to study the transcriptional output from desired loci at single molecule resolution (Raj et al., 2008). This technique has been used in syncitial Drosophila embryos to reveal mechanisms by which transcriptional activities of gap genes downstream of maternal protein gradients are regulated (Gregor et al., 2007; Little et al., 2013; Zoller et al., 2018). However, how input transcription factor levels determine output mRNA expression levels of fate-determining genes in individual cells, and how these interactions drive robust cell fate

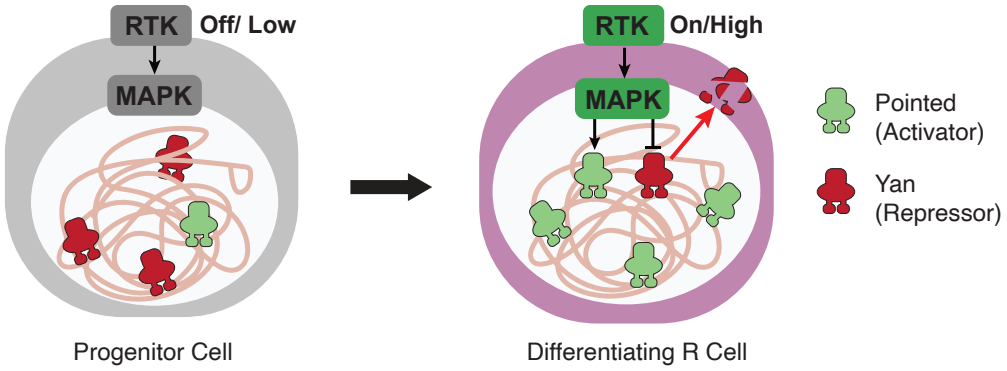

Figure 3.1: Pnt and Yan protein levels and Cell State

(Left) Low Pnt/Yan ratio is associated with the undifferentiated progenitor state. (Right) Differentiating R cells elevate their Pnt/Yan ratio. Ratio dynamics are influenced by RTK/MAPK signaling activity - high RTK signaling promotes rapid degradation of Yan as well as activation of Pnt isoforoms.

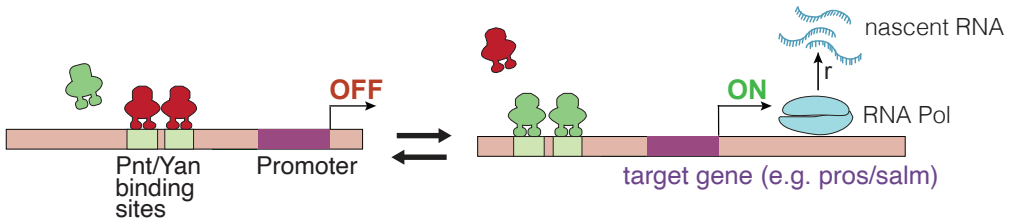

Figure 3.2: Pnt and Yan protein levels and Transcriptional Output at Target Loci

transitions remain unknown. The imaginal eye disc provides a remarkable opportunity to link the expression levels of Pnt and Yan to their activities at photoreceptor-fate determining genes in progenitor cells vs. transitioning photoreceptor cells.

In this Chapter, I combine smFISH, protein immunostaining, and quantitative analysis to link input Pnt and Yan protein levels to output nascent RNA levels at single nuclei resolution in the imaginal eye disc. I use two well-characterized genes that are critical for the acquisition of the R7 cell fate. I hypothesized that a combinatorial effect of co-expressed Pnt and Yan control whether these fate-determining genes are turned "On" or "Off" and/or the rate of mRNA production from these loci (Figure 3.2).

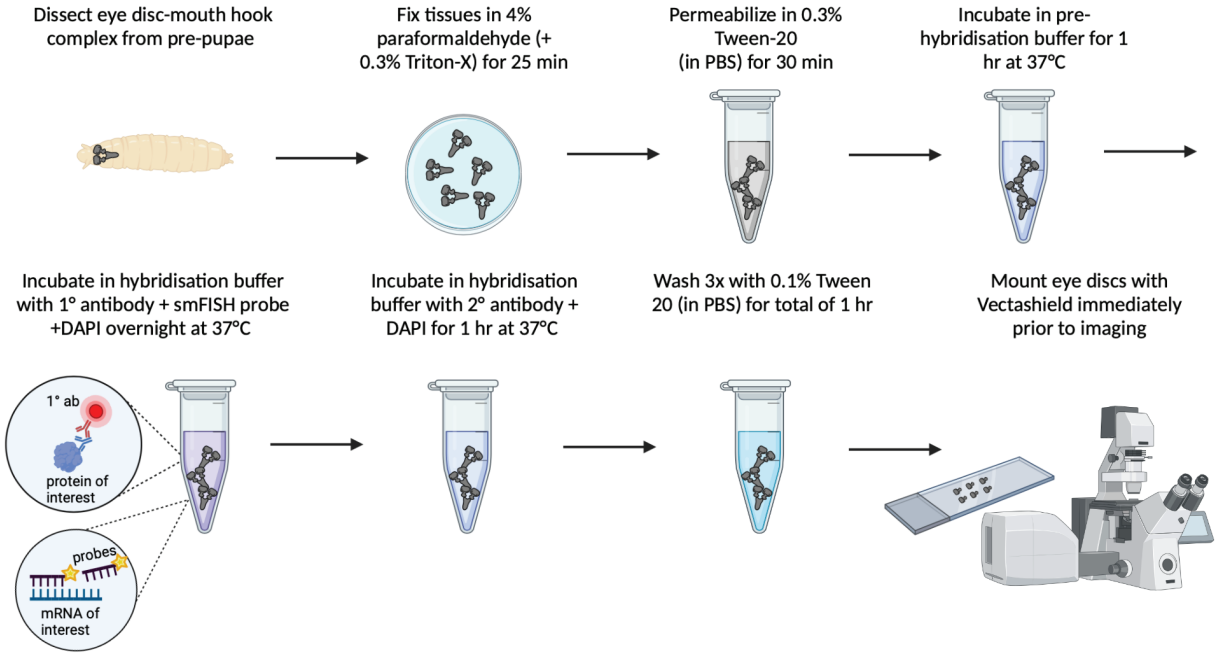

Figure 3.3: Overview of the Combined Immunostaining and smFISH Protocol Drosophila embryo smFISH protocol (Little and Gregor, 2018) was modified to allow preservation of GFP signal and combined usage with protein immunostaining in the eye imaginal disc.

3.4 Results

3.4.1 Developing an experimental approach to visualize TF protein and mRNA at single nuclei resolution in the developing eye disc

Simultaneously measuring mRNA and protein abundance in fixed tissues has been a challenge because the standard smFISH method involves a heating step which denatures the epitopes of antibodies and GFP-tagged proteins(Bakker et al., 2020). To address this, I developed a modified smFISH protocol with a lower smFISH probe annealing temperature and a longer incubation time (Figure 3.3). This protocol preserves GFP siganls, and when combined immunostaining using monoclonal antibodies, allows us to visualize and measure endogenous protein expression.

To bridge the expression levels of Pnt and Yan proteins and the transcriptional activity

of their target genes during P vs. R fate decision, I chose two well-studied Pnt/Yan target genes, *spalt major* (*salm*) and *prospero* (*pros*), both of which are required for the specification of R cell subtypes. Salm encodes a zinc-finger TF, and is responsible for the specification of R3/R4 (Domingos et al., 2004b) as well as R7 and R8 (Domingos et al., 2004a). Pros encodes a homeodomain TF, responsible for the specification of R7 cells as well as cone cells (Doe et al., 1991; Xu et al., 2000; Cook et al., 2003; Hayashi et al., 2008). Pnt and Yan protein levels are measured using GFP-tagged Pnt and monoclonal antibody against Yan respectively as described in Chapter 2.

Performing the newly developed smFISH and immunostaining protocol produced comparable Pnt and Yan expression dynamics as performing immunostaining alone. This was first qualitatively confirmed by the presence of one Yan expression peak and two Pnt-GFP peaks posterior to the morphogenetic furrow (Figure 3.4 A,C), and quantitatively by comparing the moving averages of Pnt-GFP and Yan expression between discs that were treated with the new protocol or the standard immunostaining protocol (Figure 3.4 B, D).

As expected, smFISH produces bright spots inside cell nuclei that are indicative of nascent transcription sites of multiple primary mRNA molecules that emanate from their genetic loci. Additionally, the presence of both Salm and Pros nascent transcription spots match the previously characterized expression pattern of Salm and Pros proteins in the developing eye disc.I qualitatively examined whether Salm and Pros transcribing nuclei also contained Pnt protein, Yan protein, or both. In hundreds of nuclei examined, all Salm and Pros transcription spot containing nuclei were also positive for Pnt expression, and some of these nuclei also displayed substantial Yan expression. Qualitatively, this suggests that while the presence of the activator Pnt may be required for Pros/Salm transcription, exclusion of the repressor Yan is not necessary to turn on these target genes.

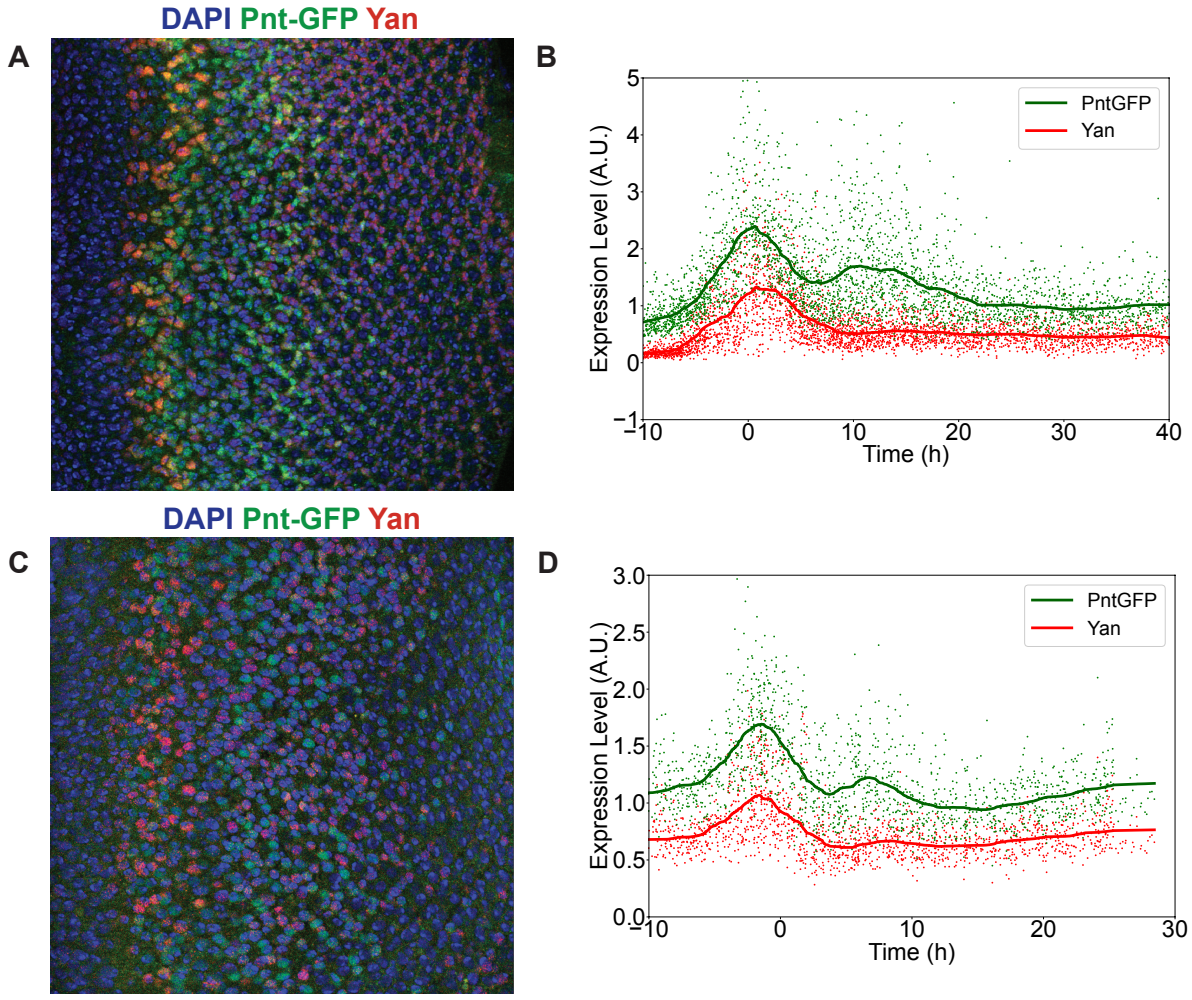

Figure 3.4: Pnt and Yan protein expression patterns are preserved in combined smFISH and immunostaining protocol

A. Single basal z-slice of the wildtype eye disc obtained from the standard immunostaning protocol. B. Moving average of Pnt-GFP and Yan expression in the progenitor cell population from discs standard immunostaining protocol (n = 3). C. Single basal z-slice of the wildtype eye disc obtained from the new combined smFISH and immunostaning protocol. D. Moving average of Pnt-GFP and Yan expression in the progenitor cell population from discs from the combined protocol (n = 2).

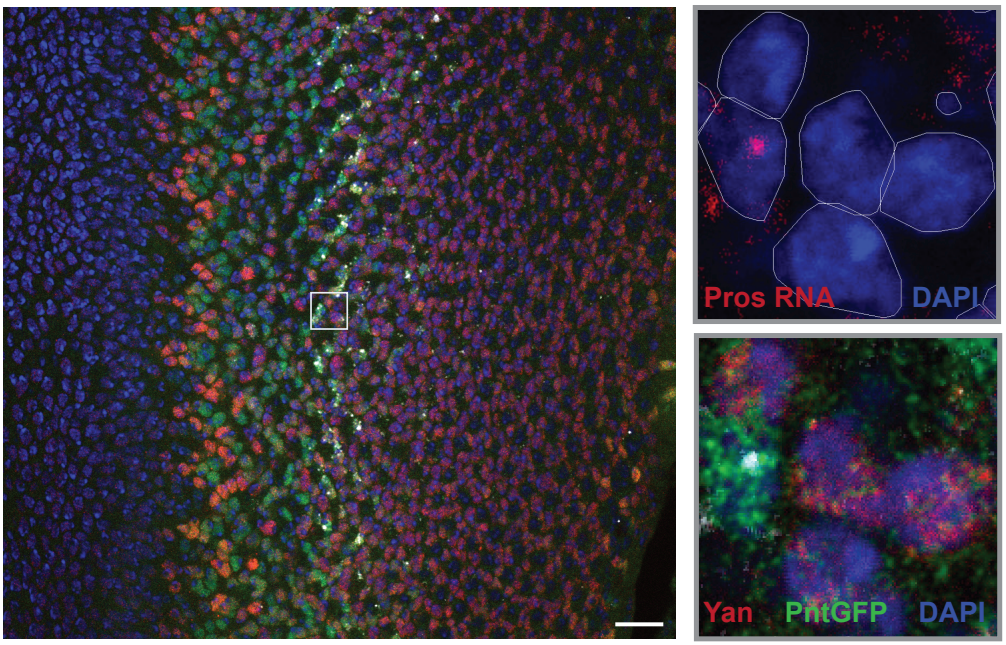

Figure 3.5: 4-channel imaging allows simultaneous visualization of Pnt and Yan proteins and nascent RNA of target gene *Pros*
(Left) Eye imaginal disc single optical section. Scale bar = 20μM (Right) Zoom-in of a single ommatidial cluster showing a transitioning R7 cell with an active Pros transcription site surrounded by neighboring progenitor cells. Top and bottom panels show the same cluster with different channels .

3.4.2 Developing an image processing pipeline to optimize signal-to-noise ratio from Drosophila eye disc smFISH images

In this study, our aim was to quantify how the output transcriptional activities at the *salm* and *pros* loci change as a function of Pnt and Yan concentrations in the nuclei. To measure nascent transcript levels, we employed a method of summing the intensities of all pixels within the smFISH channel for each individual nucleus. Theoretically, this approach would yield expression levels for nuclei containing either one or two primary transcription spots corresponding to one or two actively transcribing loci respectively, while no expression level would be detected for inactive loci that do not produce any transcription spot. In practice, substantial noise in smFISH imaging data can lead to false positive expression levels. Thus, it is critical to distinguish the pixel intensity features between noise and true signals, and use these features to eliminate noise in the imaging data.

There are existing smFISH data filtering methods for cultured cells and *in vivo* models, such as early stage Drosophila embryos and the wing imaginal disc. However these methods cannot be directly applied to our smFISH data sets due to two main reasons. First, our imaging setup involving the use of four distinct lasers to obtain good-quality signals for all channels (DAPI, Pnt-GFP, Yan-Cy3, and mRNA-ATTO633) while keeping the imaging duration to a reasonable time frame. In doing so, the resolution of smFISH data was compromised such that mRNA cannot be captured at single molecule resolution. Secondly, the three-dimensional tissue morphology of the eye disc, characterized by densely packed nuclei along the z-axis, creates a unique signal-to-noise profile across the 3D tissue space. To address these challenges, we developed a custom image processing pipeline specifically tailored for the eye imaginal disc, while incorporating commonly used filtering techniques for processing smFISH data. As shown in Figure 3.4, our smFISH processing pipeline has four main components. (1) Difference of Gaussian (D.o.G) Filtering; (2) Intensity Peak Detection; (3) Shadow Particle Filtering; and (4) Summation of Intensities for each filtered smFISH spot (Figure 3.5).

First, we employed D.o.G filtering which uses a high-pass filter and a low-pass filter. The high-pass filter enhances the boundaries and local intensities of spots in the image, while the low-pass filter smooths out the image and reduces noise. Subtracting the blurred image from the sharpened image enhances the spot objects from the surrounding noise. As expected, applying the D.o.G filter led to a significant reduction in background and sharper boundaries of smFISH spots (Figure 3.5) Second, we identified the local maxima or the intensity peaks of spot objects in 2D for every z-slice, and collected the XY coordinates of the peaks collected. We set our maximum number of peaks for each z-slice to be 600, an estimate for the maximum number of transcriptionally active spots in a single slice based on the number of nuclei and the composition of cell types that can be captured in a single z-slice. Third, we filtered out the coordinates of peaks that do not appear in three consecutive z-slices. This is because the

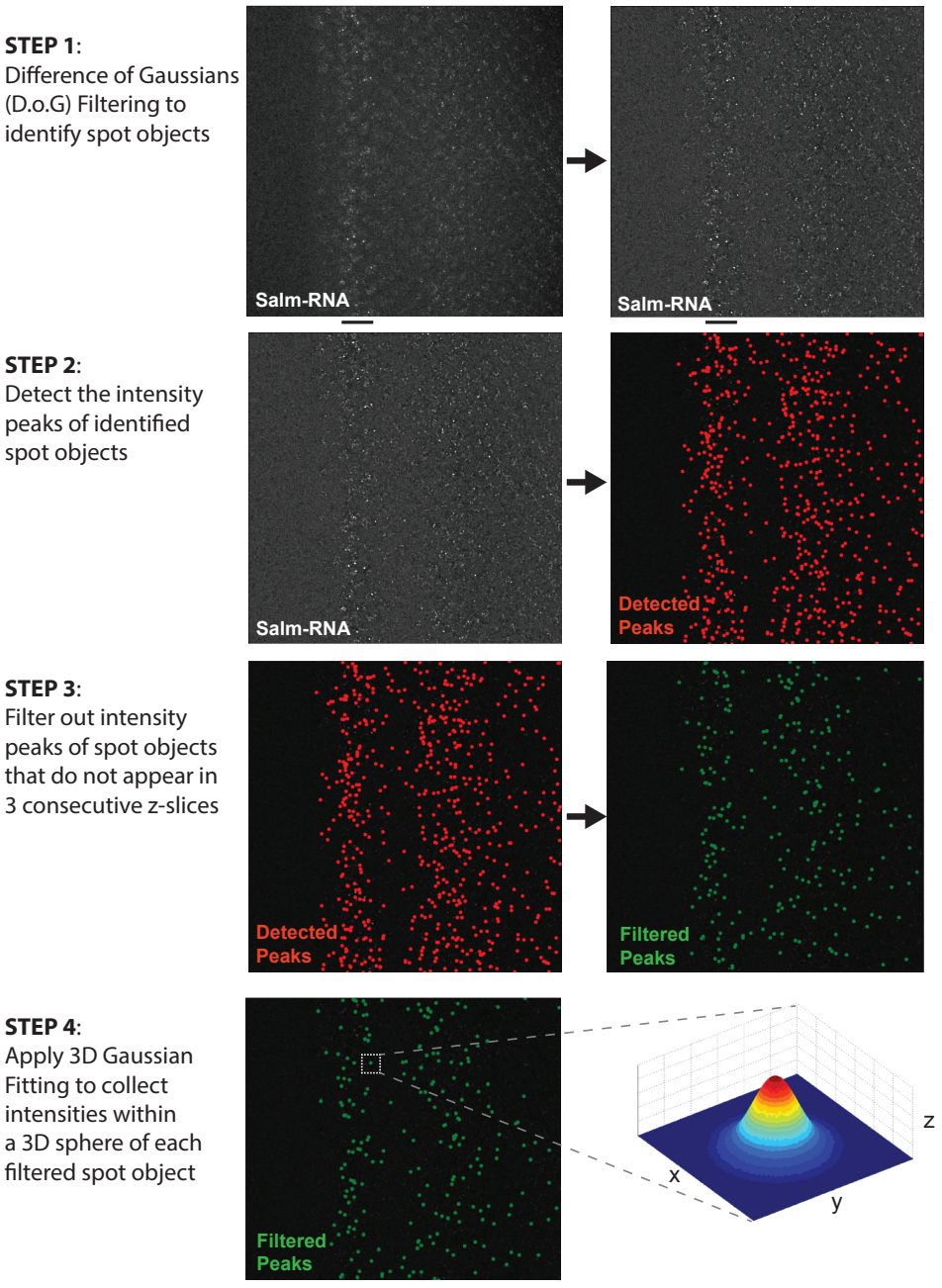

Figure 3.6: Overview of smFISH spot detection and quantification pipeline smFISH channel (against Salm mRNA) of a single z-slice from the mid-basal region of a wild type eye disc is shown to demonstrate the analysis pipeline. Black bars below the top panel of images show the first wave of Pnt expression that coincides with the transition of a subset of progenitor cells into R3/R4 photoreceptor cells that begin to express *salm*.

Pointed Spread Function (PSF), or irradiance distribution, of a single nascent transcription spot, is calculated to span at the depth of least three z-slices. Finally, we use 3D Gaussian Fitting to draw a hypothetical 3D sphere using each intensity peak as a centroid, and sum up all pixel intensities within that region to measure the total fluorescence signal of each spot presumed to be nascent transcripts from actively transcribing DNA loci.

3.4.3 Quantifying the effect of nuclear Pnt and Yan protein levels on photoreceptor fate-determining gene transcription

I focused my analysis to Pros expression at the P vs. R7 decision window. In this temporal window, "young" or transitioning R7 cells are identified by the previously characterized hallmarks of cells undergoing P to R fate transition. First, their nuclei are located more apically compared to their unspecified progenitor cell neighbors along the apicobasal plane, right below the specified R1 and R6 cells (TOMLINSON, 1988). These young R7 cells do not yet express neuronal markers such as Elav, and are therefore distinct from mature R7 photoreceptor neurons. This early population marks the beginning of the P to R7 cell trajectory and show elevated Pnt/Yan ratio (Peláez et al., 2015; Bernasek et al., 2023).

First, I asked whether young R7 cells can be distinguished from their neighboring progenitor cells that remain in the unspecified state based on Salm and/or Pros expression. The difference may be based on the active vs. inactive promoter state, or on expression levels. To this end, I plotted the expression level of either Salm or Pros nascent RNA in the young R7 vs. the P cell population. This revealed a significant difference in both Salm and Pros nascent RNA expression between the two populations (Figure 3.7). For Salm expression, this difference was binary - all R7 cells had at least one actively transcribing Salm locus in their nuclei while none of the progenitor neighbors engaged in active Salm transcription (Figure 3.7 A). The result was more nuanced for Pros expression. While a significant difference in Pros expression was observed with the same trend, few R7 cells and progenitors displayed

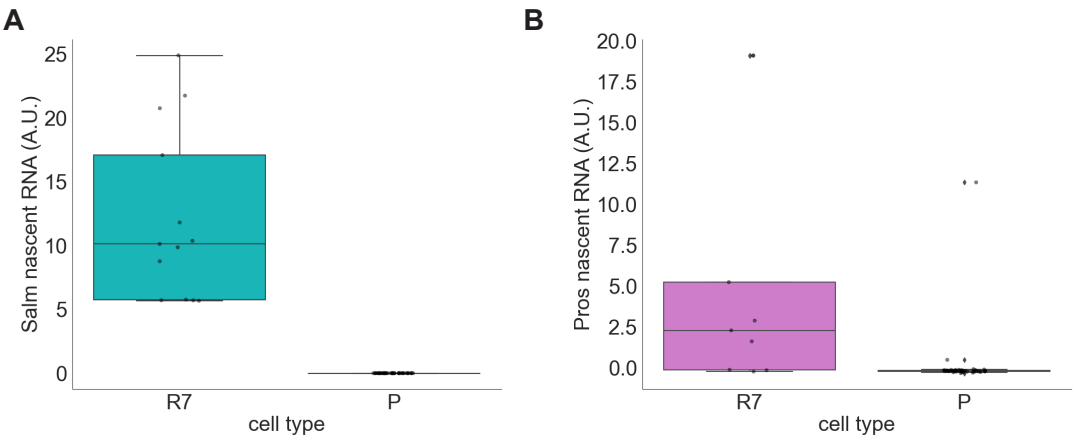

Figure 3.7: Photoreceptor gene transcription during P vs. R7 cell fate decision Box plot showing distributions for single nuclei measurements of (A) Salm nascent RNA expression and (B) Pros nascent RNA expression in young R7 vs. neighboring P cell population. In both cases, the expression distributions between the two populations were significantly different (Student's t-test $p<0.01$).

inactive and active Pros transcriptional state respectively (Figure 3.7 B).

Next, I examined how input Pnt and Yan abundance affect output transcription at target loci. I first plotted Salm nascent RNA level as a function of each one of the three measured quantities (1) Pnt level, (2) Yan level, and (3) the Pnt/Yan ratio (Figure 3.8). Salm transcription showed weak positive correlation to both Pnt level and the Pnt/Yan ratio. Meanwhile, the relationship between Salm transcription and Yan level showed a more threshold-like trend. All cells with at least one actively transcribing Salm locus (sum intensitiy >0 A.U.) came from the lowest Yan expressing population (the first quartile for mean Yan expression). When Yan expression in individual cells exceeded a certain threshold, this was accompanied by the transcriptional "Off" state of the Salm promoter.

I next examined the same input-output relationships for Pros transcription (Figure 3.9). Plotting Pros nascent RNA levels as a function of varying Pnt levels showed no discernible trend, indicating that Pnt concentration alone does not predict Pros transcriptional output (Figure 3.9A). This is consistent with existing literature that another transcriptional activator, Lozenge, is used to regulate Pros expression (Xu et al., 2000). Meanwhile, Pnt/Yan

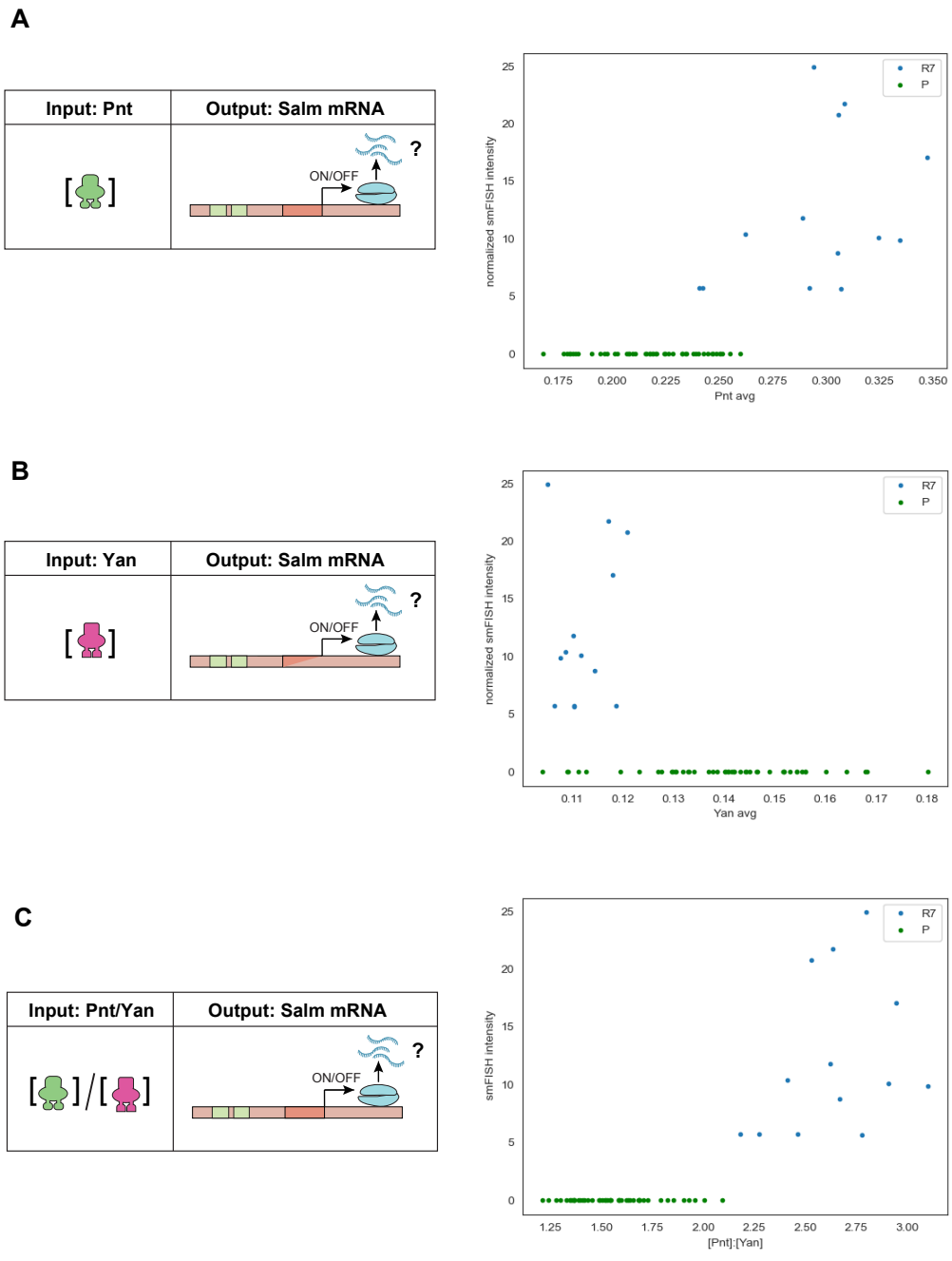

Figure 3.8: Salm transcription as a function of Pnt levels, Yan levels, and Pnt/Yan ratio

A scatterplot linking input (A) average Pnt-GFP expression, (B) average Yan expression, and (C) Pnt/Yan expression ratio on the x-axis to output nascent mRNA expression at the Salm locus on the y-axis. Each spot represents measurements from an individual nucleus from a single eye imaginal disc.

ratio showed a weak positive correlation with Pros RNA levels (Figure 3.9B). This suggested that Yan protein levels play a role in regulating Pros expression. Strikingly, when Pros RNA levels were plotted as a function of Yan protein levels, it produced the same switch-like trend observed for Salm transcription. All nuclei with at least one transcribing allele of *pros* belonged to the population with the lowest Yan levels.

3.5 Discussion

Here, I describe the first study that combines smFISH with immunostaining to link input transcription factor expression to output nascent RNA level at single nuclei resolution in a developing Drosophila eye disc.

I first show that at the R7 vs. P decision point, cells that are adopting the R7 fate turn on Salm and Pros transcription while most progenitors do not. In Chapter 2, I show that transitioning R7 cells start to elevate their Pnt/Yan ratio, and that this elevated ratio is necessary for stable transition of young R7 cells to mature R7 cells. My observation that young R7 cells have a much higher probability for initiating Salm and Pros transcription compared to their neighboring progenitor cells supports our previous finding that ratiometric sensing of Pnt and Yan is key to this fate transition. It also adds a previously missing intermediate layer in the flow of information: changes in transcription factor levels *are integrated at individual gene loci leading to distinct transcriptional responses*, which ultimately lead to cell fate transitions.

Second, I show that while transcription initiation of photoreceptor genes does not depend on the complete removal of Yan from the nucleus, Yan appears to control the On/Off state of its target gene promoters by a threshold-based mechanism. My results suggest that moderate to high Yan presence in the nucleus blocks target genes from entering a transcriptionally active state, and that this level must be reduced below some threshold for target genes to enter a permissive state for transcription.

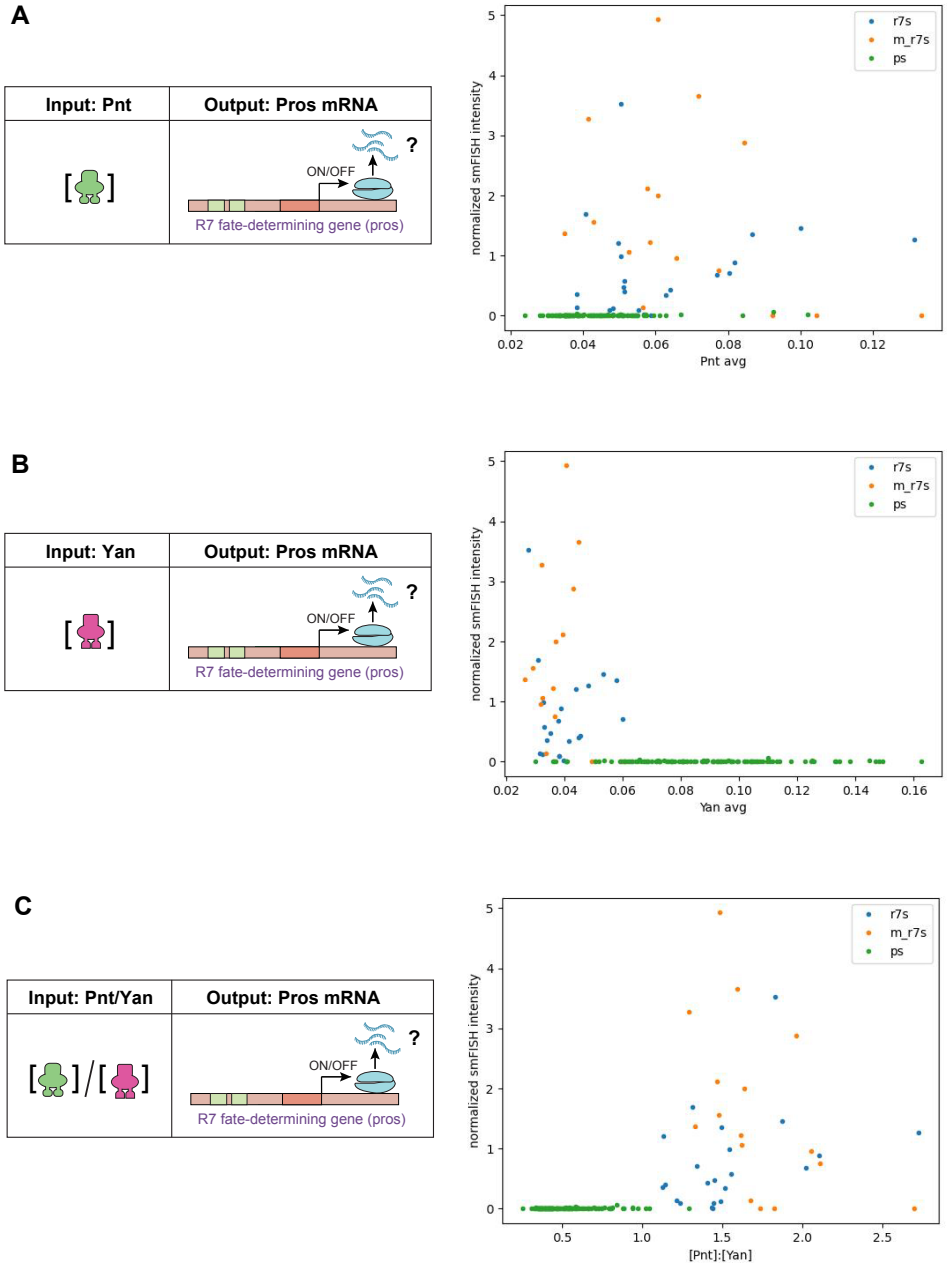

Figure 3.9: Prospero transcription as a function of Pnt levels, Yan levels, and Pnt/Yan ratio

A scatterplot linking input (A) average Pnt-GFP expression, (B) average Yan expression, and (C) Pnt/Yan expression ratio on the x-axis to output nascent mRNA expression at the Pros locus on the y-axis. Each data point represents measurements from an individual nucleus. Data points are pooled from n=4 eye imaginal discs after normalizing with a smFISH spike-in control (probe against Senseless mRNA).

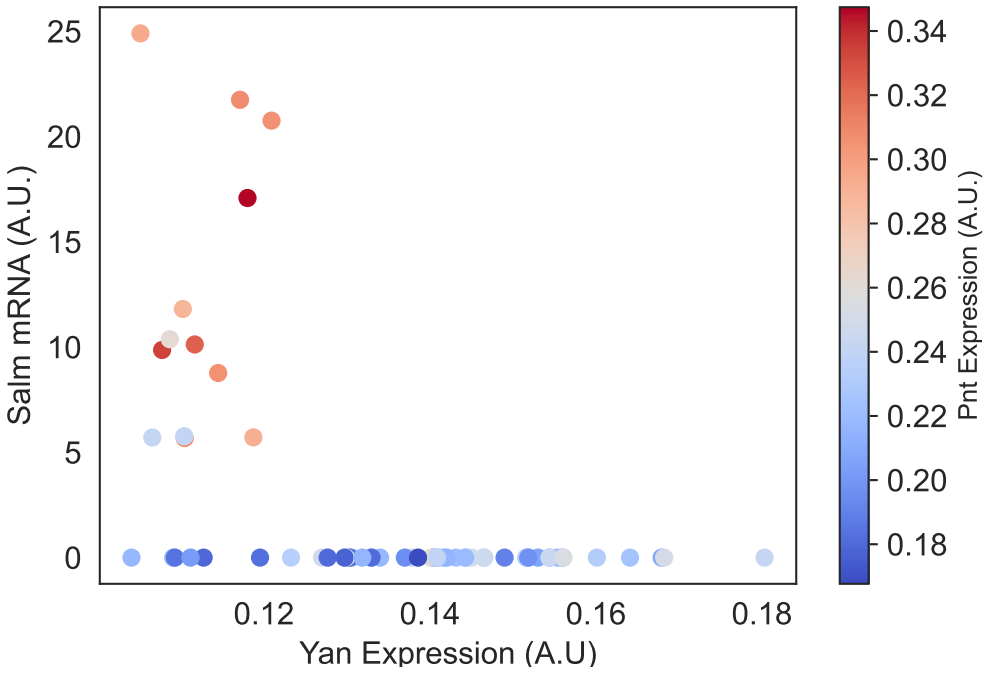

Figure 3.10: Combined effect of Pnt and Yan levels on Salm nascent RNA expression

A scatterplot linking input average Yan expression (A.U.) on the x-axis to output nascent mRNA expression at the Salm locus (A.U.) on the y-axis as shown in Figure 3.8, with an additional heatmap component indicating Pnt-GFP expression (A.U.) in each data point. The color bar on the right indicates a spectrum of colors from red to blue corresponding to the spectrum of Pnt-GFP expression.

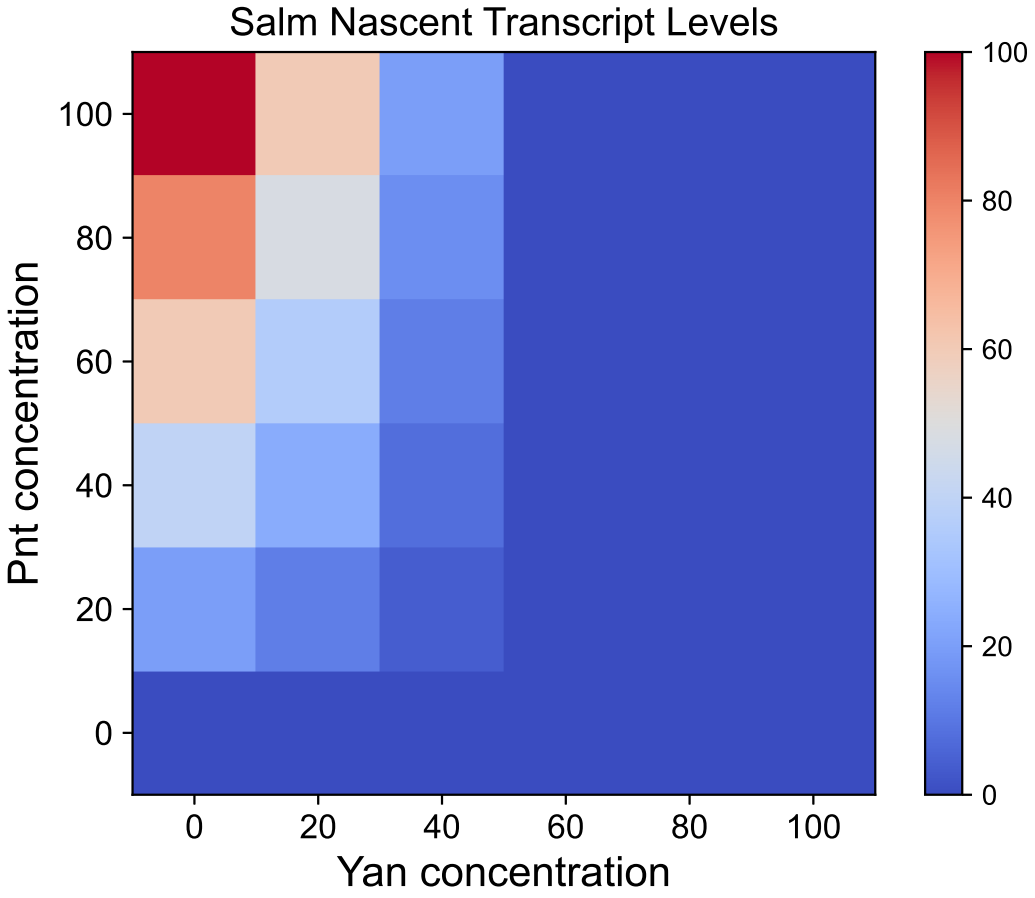

Figure 3.11: Modeling potential switch and rheostat mechanism of Pnt and Yan regulation
A toy model simulating the output of Salm nascent RNA levels as a response to Pnt and Yan interactions at the Salm locus.

Third, I show that Pnt expression is weakly correlated to its target gene expression. Nevertheless, the level of the entire Pnt protein population in the nucleus was not a strong predictor of transcriptional activities at target genes, especially for the Pros locus. I speculate that the positive correlation would be stronger if the abundance of only the active forms of Pnt (i.e., phosphorylated PntP2 and P3 isoforms and the constitutively active P1 isoform) were measured and plotted against the nascent RNA levels of Salm and Pros. This is discussed further in Chapter 4.

One limitation of this study is that both nascent RNA and protein levels are not measured in terms of their absolute concentrations, but as fluorescent intensity levels. This approach enabled us to investigate how relative changes in the abundance of Pnt and Yan elicit changes to the transcriptional states and levels of their target genes. However, the exact mathematical relationships between input protein concentrations and output transcript levels cannot be deduced from this approach since our measurements are unitless – i.e., we do not know how intensity values translate to the number of protein and mRNA molecules in the nucleus. In the future, calibrating intensity measurements into absolute units will allow analogous fitting of our data and parameterizations to simulate the two-state model of Salm and Pros transcriptional bursting, and the extraction of further quantitative insights. This idea is further explored in Chapter 4.

How might multiple opposing inputs be integrated at individual genes to produce a coordinated response? One intriguing mechanism is the "switch and rheostat" mechanism demonstrated in the yeast galactose-responsive (GAL) pathway (Ricci-Tam et al., 2021). Here, expression of the *gal1* gene is under direct regulation by a single Gal4 activator. Nevertheless, Gal4 activity at the *gal1* locus is under bipartite regulation. The concentration of the Gal4 activator is transcriptionally regulated by the Mig1 repressor. As the Mig1 repressor level decreases, the transcription and protein levels of Gal4 increase. Yet, Gal4 activity is further tuned by the level of Gal80 that sequesters its activity. These interactions

result in a rheostat-like control of Gal4 expression and a switch-like control of Gal4 activity (Ricci-Tam et al., 2021). As such, the downstream transcription of the gal1 gene is under both binary and analog (or continuous) regulation, allowing for fine-tuning of its expression. Additionally, this bipartite mode of regulation was shown to produce two independently controlled features: (1) the fraction of cells that initiate transcription of the *gal1* gene and (2) the level of *gal1* expression.

It is tempting to speculate that regulation of *pros* and *salm* genes are under similar bipartite control by Pnt and Yan. Indeed, examining the relationships between Pnt levels, Yan levels, and Salm nascent RNA at the same time (Figure 3.10) produces results that is predicted by modeling the switch and rheostat mechanism by Yan and Pnt (Figure 3.11). This suggests that Pnt controls the expression level while Yan controls the on/off switch of these photoreceptor-fate determining genes. This would allow Yan to ensure only transitioning photoreceptor cells turn on the transcription of photoreceptor genes (since Yan levels are significantly reduced in these cells due to sustained RTK/MAPK signaling), while activated Pnt levels would ensure that the expression levels of these genes are sufficiently high for a stable fate transition to take place.

Future investigations should focus on exploring genetic perturbations that alter Pnt and Yan concentrations to test the robustness of photoreceptor gene expression to varying levels of Pnt and Yan. For example, varying the allelic copy number of each Pnt and Yan from one to four copies result in robust progenitor to photoreceptor fate transitions (Chapter 2). It remains to be seen whether these perturbations also produce wildtype expression of photoreceptor fate-determining genes, or whether fate transitions are robust to moderate changes in the expression state and/or level of these genes.

Meanwhile, genetic perturbations that alter the Pnt/Yan ratio produce defects such as too many or too few progenitor cells that transition to photoreceptor cells. For instance, the Sev>RasV12 mutant, which drives hyperactivation of MAPK signaling during the P vs. R7

decision window, leads to ectopic transition of progenitor cells to R7 cells. Hyperactivation of MAPK results in both increase in Pnt and decrease in Yan expression. Teasing apart how each of these perturbations affect the transcriptional state and level at Pros and Salm loci by plotting individual relationships (i.e., Pnt vs. nascent RNA, Yan vs. nascent mRNA, and ratio vs. nascent mRNA) at single nuclei resolution would provide valuable insights into how altered transcriptional responses downstream of altered transcription factor expression levels contribute to cell fate defects.

Another intriguing avenue for research would involve revisiting the Sev>YanAct mutation, which results in the overexpression of a constitutively active form of Yan that remains in the nucleus, independent of MAPK signaling input. By decoupling high MAPK signaling and resulting increase in Pnt from Yan levels, it becomes possible to observe cells with both high Yan and Pnt presence. In these cells, investigating whether the high Yan presence leads to an inactive transcriptional state at target genes, even in the presence of high Pnt, or if elevated Pnt levels can outcompete high Yan levels and activate transcription in some cells, could provide valuable insights into the underlying regulatory mechanisms.

3.6 Materials and Methods

3.6.1 Fly Genetics

All fly lines used in this Chapter are a subset of those used in Chapter 2. Fly stocks from the Bloomington Stock Center were as follows: pnt-GFP BAC transgene, BL42680; pntΔ88 (O'Neill et al., 1994); sev>RasV12 (Fortini et al., 1992); sev>YanACT (Rebay and Rubin, 1995). The genotype of the wildtype control is the following: w1118; pnt-gfp/+; pntΔ88/+. Experiments with RasV12 or YanACT, the following genotypes were used: w1118; pnt-gfp/sev>RasV12; pntΔ88/+ relative to w1118; pnt-gfp/+; pntΔ88/+; and w1118; pnt-gfp/sev>YanACT; pntΔ88/+ relative to w1118; pnt-gfp/+; pntΔ88/+.

3.6.2 Generation of smFISH probes

Each smFISH probe set contains 48 non-overlapping, 20 nucleotides-long cDNA oligonucleotides that are complementary to a region in the primary mRNA of interest (Sens, Salm, Pros). Probe sequences were designed and obtained using the Stellaris smFISH Designer tool provided by Biosearch Technologies (https://www.biosearchtech.com/). These sequences were obtained with 5' amine modification from IDT. These oligonucleotides were crosslinked to the NHS-ester ATTO 633 dye for 2 hours at 37 °C at pH 8.5 using freshly prepared NaHCO3 solution. Crosslinked probes were ethanol precipitated, resuspended in 30μM of water, and purified from unbound ATTO-633 using G-25 spin columns (GE Illustra) according to manufacturer's instructions. Probes were stored at -20°C and protected from light until use.

3.6.3 Combined Immunostaining and smFISH

All eye-antennal discs were dissected from white prepupae. Samples were fixed in 4% paraformaldehyde (w/v)/PBS-Triton X-100 0.3% (v/v) for 25 min in a glass bottom plate on a rocking platform at 90rpm at room temperature. After fixation, eye discs were washed in 1x PBS-Tween 20 0.3% (v/v) three times, and in 1x PBS for the final wash in a glass bottom plate. Discs were transferred to a 1.5ml eppendorf tube, and incubated in 1ml wash buffer (4X SSC buffer + 0.1% Tween 20 + 1% v/v vanadryl ribonucleoside) for 1h at 37C. Then, discs were incubated with hybridization buffer containing sense smFISH probe (conjugated to ATTO633, 1:15), *salm* or *pros* smFISH probe (conjugated to ATTO633, 1:15), primary monoclonal Yan antibody (DHSB, 1:200) and DAPI for 12-16 hr at 37C. Discs were washed with 1ml wash buffer for 1hr, followed by incubation with 1ml wash buffer containing secondary antibody (goat anti-mouse Cy3; 1:1000) and DAPI for 1 hr at 37C. Discs were then washed with 0.1% Tween 20 for 1 hr rocking at room temperature. Samples were mounted in VectaShield Antifade mounting medium (Vector Laboratories). To avoid rapid changes

in cell volume, which would affect relative fluorescence measurements, the substitution of PBS-Triton X-100 with VectaShield was carried out gradually using increments of ∼33% in VectaShield concentration.

3.6.4 Confocal Imaging

Image stacks were collected using the Zeiss 880 confocal microscope, using a 40X oil immersion objective (NA = 1.4) with a digital zoom of 1.2-1.5, and a frame size of 2400 x 2400. The pixel size for each image stack was 70nm (X) by 70nm (Y). This ensures at least 2 pixels cover the Point-Spread-Function (PSF) radius of a nascent transcription spot. Approximately 60-80 optical slices were collected for each disc, with a z-interval of 0.20μM. This ensures that the transcription spot appears in a minimum of three consecutive Z-slices, which is used to distinguish true signal from imaging noise. DAPI, GFP, Cy3, and ATTO 633 were excited by 405, 488, 555, and 630 nm lasers respectively. The pinhole size for each laser was set at 1 Airy unit.

3.6.5 smFISH Channel Processing

Processing of the smFISH channel was performed with custom python scripts. First, Difference of Gaussian (D.o.G) filtering was applied using minimum and maximum sigmas of 2 and 6 pixels respectively to find spot objects that indicate nascent transcription spots in each z-slice. This was followed by peak detection, using the "peak_local_max" function from sci-kit image library (https://scikit-image.org). This function returns the pixel coordinates of local intensity peaks (maxima) in each z-slice. These peaks were then filtered based on the presence of shadows - the appearance in three consecutive z-slices. Finally, filtered peaks are used as centroids to draw circles of an expected radius of a transcription spot. Pixel intensities within each drawn circle is summed to calculate the signal output from individual filtered smFISH spots.

3.6.6 Linking Expression Levels to Individual Segmented Nuclei

Individual nuclei were segmented, their cell types annotated, and the normalized average Pnt-GFP and Yan expression levels from each nucleus were quantified using the annotation software FlyEye Silhouette, and the analysis packages FlyEye and the FlyQMA described in Chapter 2. The FlyEye package was modified to add a third expression measurement for each nucleus, the expression level of filtered smFISH spots.

3.7 References

Acampora, D., Di Giovannantonio, L. G., Garofalo, A., Nigro, V., Omodei, D., Lombardi, A., Zhang, J., Chambers, I., and Simeone, A. (2017). Functional antagonism between otx2 and nanog specifies a spectrum of heterogeneous identities in embryonic stem cells. *Stem cell reports*, 9(5):1642–1659.

Bakker, R., Mani, M., and Carthew, R. W. (2020). The wg and dpp morphogens regulate gene expression by modulating the frequency of transcriptional bursts. *Elife*, 9:e56076.

Bernasek, S. M., Hur, S. S., Peláez-Restrepo, N., Lachance, J.-F. B., Bakker, R., Navarro, H. T., Sanchez-Luege, N., Amaral, L. A., Bagheri, N., Rebay, I., et al. (2023). Ratiometric sensing of pnt and yan transcription factor levels confers ultrasensitivity to photoreceptor fate transitions in drosophila. *Development*, pages dev–201467.

Cook, T., Pichaud, F., Sonneville, R., Papatsenko, D., and Desplan, C. (2003). Distinction between color photoreceptor cell fates is controlled by Prospero in Drosophila. *Developmental Cell*, 4(6):853–864.

Doe, C. Q., Chu-LaGraff, Q., Wright, D. M., and Scott, M. P. (1991). The prospero gene specifies cell fates in the drosophila central nervous system. *Cell*, 65(3):451–464.

Domingos, P. M., Brown, S., Barrio, R., Ratnakumar, K., Frankfort, B. J., Mardon, G., Steller, H., and Mollereau, B. (2004a). Regulation of R7 and R8 differentiation by the spalt genes. *Developmental Biology*, 273(1):121–133.

Domingos, P. M., Mlodzik, M., Mendes, C. S., Brown, S., Steller, H., and Mollereau, B. (2004b). Spalt transcription factors are required for r3/r4 specification and establishment of planar cell polarity in the drosophila eye.

Gregor, T., Tank, D. W., Wieschaus, E. F., and Bialek, W. (2007). Probing the Limits to Positional Information. *Cell*, 130(1):153–164.

Hayashi, T., Xu, C., and Carthew, R. W. (2008). Cell-type-specific transcription of prospero is controlled by combinatorial signaling in the drosophila eye.

Li, Y., Pang, Z., Huang, H., Wang, C., Cai, T., and Xi, R. (2017). Transcription factor antagonism controls enteroendocrine cell specification from intestinal stem cells. *Scientific reports*, 7(1):988.

Little, S. C. and Gregor, T. (2018). Single mRNA Molecule Detection in Drosophila. *Methods in Molecular Biology (Clifton, N.J.)*, 1649:127–142.

Little, S. C., Tikhonov, M., and Gregor, T. (2013). Precise developmental gene expression arises from globally stochastic transcriptional activity. *Cell*, 154(4):789–800.

Peláez, N., Gavalda-Miralles, A., Wang, B., Navarro, H. T., Gudjonson, H., Rebay, I., Dinner, A. R., Katsaggelos, A. K., Amaral, L. A., and Carthew, R. W. (2015). Dynamics and heterogeneity of a fate determinant during transition towards cell differentiation. *eLife*, 4:e08924. Publisher: eLife Sciences Publications, Ltd.

Raj, A., Van Den Bogaard, P., Rifkin, S. A., Van Oudenaarden, A., and Tyagi, S. (2008). Imaging individual mrna molecules using multiple singly labeled probes. *Nature methods*, 5(10):877–879.

Ricci-Tam, C., Ben-Zion, I., Wang, J., Palme, J., Li, A., Savir, Y., and Springer, M. (2021). Decoupling transcription factor expression and activity enables dimmer switch gene regulation. *Science*, 372(6539):292–295.

Schaffer, A. E., Freude, K. K., Nelson, S. B., and Sander, M. (2010). Nkx6 transcription factors and ptf1a function as antagonistic lineage determinants in multipotent pancreatic progenitors. *Developmental cell*, 18(6):1022–1029.

TOMLINSON, A. (1988). Cellular interactions in the developing drosophila eye. *Development*, 104(2):183–193.

Walsh, J. C., DeKoter, R. P., Lee, H.-J., Smith, E. D., Lancki, D. W., Gurish, M. F., Friend, D. S., Stevens, R. L., Anastasi, J., and Singh, H. (2002). Cooperative and antagonistic interplay between pu. 1 and gata-2 in the specification of myeloid cell fates. *Immunity*, 17(5):665–676.

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(1):87–97.

Zoller, B., Little, S. C., and Gregor, T. (2018). Diverse Spatial Expression Patterns Emerge from Unified Kinetics of Transcriptional Bursting. *Cell*, 175(3):835–847.e25.

CHAPTER 4

CONCLUSIONS AND DISCUSSION

4.1 Summary and Implications

In this thesis, I present a quantitative investigation into the highly conserved RTK/MAPK signaling pathway, that ensures the accuracy and robustness of cell state transitions in numerous contexts of organismal development. Using the transition to progenitor cells to differentiated photoreceptor neurons in the developing Drosophila eye as a model system, I investigate how the RTK/MAPK transcription factors Pnt and Yan dynamically influence each other's expression levels and DNA binding to inform accurate and robust fate decisions, providing new quantitative insights and developing opportunities for further exploration.

In Chapter 2, we discover that cell state decisions are regulated by the relative levels of Pnt and Yan, rather than by the absolute level of either protein. We uncover a positive regulatory relationship between Pnt and Yan to ensure that the activator-to-repressor ratio remains robust to fluctuating levels of either protein in progenitor cells, possibly to prevent aberrant state transitions. This novel finding emphasizes the importance of considering the combinatorial effects of co-expressed transcription factors in accurately predicting and understanding whether a cell transitions to a differentiated state.

In Chapter 3, I delved deeper into the regulatory mechanism of Pnt and Yan at individual target loci, producing insights into how the co-expressed activator and repressor pair ensures that fate-determining genes are expressed at the right time during photoreceptor cell state transitions. To this end, we developed a quantitative microscopy approach that enables simultaneous measurements of mRNA and protein levels at single nuclei resolution in the developing Drosophila eye disc. This approach represents a significant advancement in experimental and computational analysis of the Drosophila eye disc, providing a powerful tool to unravel the intricate dynamics of gene expression regulation during cell state transitions.

Overall, work presented in this thesis advances our knowledge of the Pnt and Yan regulatory network downstream of RTK/MAPK signaling by providing further quantitative insights into their regulatory mechanisms. The developed methodologies and insights provide a strong foundation for further exploration, including the development of theoretical models to simulate cell state transitions based on our experimental data. Finally, this work is not without its limitations. In subsequent sections of this chapter, I discuss these limitations and propose future experiments.

4.2 Future Directions

4.2.1 Pnt isoform/ phosphorylation state-dependent regulation

In my thesis, I studied how input Pnt protein concentrations affect output mRNA expression and cell fate decisions. For this I used a recombineered Pnt genomic transgene (Boisclair Lachance et al., 2014) in which GFP was inserted in-frame at the carboxy-terminus, tagging all three Pnt isoforms. A major limitation of this approach is our inability to assess isoforom-specific as well as phosphorylation-dependent control of transcriptional activation.

To date, three distinct Pnt protein isoforms encoded by the pnt locus have been characterized: PntP1, PntP2, and PntP3. Structurally, all Pnt isoforms share a common C-terminal region that includes three exons, one containing the ETS DNA-binding domain (Scholz et al., 1993; Klämbt, 1993; Wu et al., 2020). The isoform-specific differences are found at the N-terminal region: PntP1 is a smaller protein with a single exon in its N-terminal region, while PntP2 and PntP3 have five and four exons respectively (Scholz et al., 1993; Klämbt, 1993; Wu et al., 2020). Functionally, PntP1 is constituitivey active while PntP2 and PntP3 have been shown to require activation by MAPK phosphorylation (Brunner et al., 1994; Shwartz et al., 2013; Wu et al., 2020).

Transcriptionally, PntP1's expression is regulated by PntP2 and PntP3 (Shwartz et al.,

2013; Wu et al., 2020). The removal of PntP1 alone leads to the loss of photoreceptors R1-R7, and the removal of PntP2 affects the second wave of transitioning R fates (R1/R6 and R7). Meanwhile, removal of the PntP3 isoform alone does not produce abnormal photoreceptor cell counts suggesting that PntP3 is not necessary for photoreceptor differentiation (Wu et al., 2020). Nevertheless, PntP3 is suggested to function redundantly with PntP2 to provide robustness to fluctuating levels of Pnt to ensure normal photoreceptor cell transitions (Wu et al., 2020).

A potential avenue for further exploration is the contribution of different Pnt isoforms and their phosphorylation states on the transcriptional output at individual target loci. Transgenes containing isoform-specific tagging of the Pnt locus may allow input-output measurement for each Pnt isoform and their target genes. However, it would be difficult to remove the endogenous Pnt species in a isoform-specific manner. Meanwhile, developing a monoclonal antibody against the phosphorylated form of Pnt would allow one to distinguish between the activating input of phosphorylated Pnt species from the total protein population. This might be an involved process however, since high quality anti-phospho primary antibodies suited for quantitative analysis are difficult to generate.

4.2.2 Yan oligomerization and Cell State Transitions

Under given transcription factor concentrations in the nucleus, gene expression outputs are further regulated by protein-DNA and protein-protein interactions. In chapter 3, I explored two different genes to explore potential enhancer-specific DNA-protein interactions between these genes and the transcription factors. Another critical aspect of gene regulation by Yan is its protein-protein interaction. Yan monomers are able to form homotypic oligomers through its SAM domain (Qiao et al., 2004; Hope et al., 2018). Mutations that block this interaction thus limiting Yan in its monomeric state perturbs Yan's repressive ability (Qiao et al., 2004; Zhang et al., 2010). Interestingly, mutations that strongly increase the SAM-SAM affinity

in Yan also results in a loss of function phenotype (Hope et al., 2018). Results from this work suggest that dialling up Yan's oligomerization affinity by four orders of magnitude can drive Yan polymerization off DNA, resulting reduced DNA binding by Yan (Hope et al., 2018). Yan nuclear and cytoplasmic aggregation is also positively correlated with increasing SAM affinity above the wildtype range (Hope et al., 2018). One interesting finding is that monomerizing vs. polymerizing mutants of Yan manifest in nuanced differences in their loss of function phenotypes. The monomerizing Yan mutant allele Yan^{V105R} predominantly causes excessive outer photoreceptor cells R3/R4 (70% of ommatidia) and less frequent extra inner photoreceptor cell R7 (28%) (Hope et al., 2018). The reverse is observed for the polymerizing Yan mutant alleles SuperYan(SY) 3, SY4 and SY5, in which extra outer photoreceptor cells are more common (Hope et al., 2018).

The loss of function phenotypes caused by monomerizing and polymerizing series of mutants suggest a possible differential requirement for Yan protein-protein interactions for genes that determine outer vs. inner photoreceptor fates. Alternatively, it is possible that Yan target genes require a common oligomerization minimum of Yan, for example its binding as dimers. Under this scenario, the difference in susceptibility of outer vs. inner photoreceptor fates to Yan oligomerization mutants may be a reflection of their RTK/MAPK signaling and concentration differences in these cells. R7 cells experience higher RTK/MAPK signaling due to the presence of two types of RTKs (EGFR and Sevenless) to receive the signal, while the outer photoreceptor cells lack the Sevenless receptor. Higher RTK/MAPK signaling could result in reduced concentration of Yan monomers in these cells, such that aggregates or excessive polymers of Yan are less likely to form in the SY mutants. To distinguish between these two scenarios (i.e. differential Yan oligomer requirement by specific genes vs. differential concentration of Yan in specific cell types), one could perform the combined smFISH and immunostaining experiment to link the input-output relationship between Yan nuclear levels and nascent RNA levels of Salm (expressed in R3/R4 as well as R7 cells) and

Pros (expressed in R7 cells).

4.2.3 Dissecting Transcriptional Kinetics Downstream of Pnt and Yan

Advances in live imaging of transcription (Elowitz et al., 2002), single molecule mRNA counting (Zenklusen et al., 2008), and computational modeling tools provided compelling evidence that transcription of most genes from yeast to mammalian cells occurs in discontinuous bursts. The most commonly used model to understand transcriptional dynamics is the two-state telegraph model. This model posits that a given promoter switches between active and inactive states with rates *Kon* and *Koff*. In the active state, nascent RNA is transcribed with a constant rate, *r*. This model has been solved analytically (Peccoud and Ycart, 1995), and predicts bursty gene expression (Raj et al., 2006). Another feature of bursty gene expression is the non-Poissonian nature of mRNA distributions from single cells at a fixed time point. This contrasts with rare instances of constitutively active gene transcription, such as housekeeping genes in yeast cells, in which the resulting mRNA distributions follow a Poissonian distribution (Zenklusen et al., 2008). For a Poisson distribution, the Fano factor (the ratio of variance to the mean) is 1, meaning that expression variability is equal to mean expression. For non-Poisson mRNA distributions resulting from bursty transcription, the Fano factor is higher than 1.

The Fano factor from mRNA distributions can be used to deduce which parameter of the two state model, *Kon*, *Koff*, or *r*, is being regulated for a given gene transcription. This is because plotting the Fano factor against mean mRNA number gives rise to unique Fano factor "signatures" or shapes in the curve depending on which one of these parameters is under regulatory control (H.G. Garcia; unpublished book chapter). This approach was used in a paper studying bacterial transcription to understand how transcription factors regulate bursty transcription dynamics (So et al., 2011). A distinct Fano factor signature was also seen when this method was applied to data from mammalian cells (Chen et al., 2015).

The Fano factor approach can be used to determine which parameter of the two-state model is regulated by Pnt and Yan during Salm and Pros transcription. The idea would be to bin progenitor and photoreceptor cell populations based on specific ranges of cellular mRNA counts, and plot the Fano factor against the mean expression level for each bin's mRNA distribution. I could not perform this analysis due to the compromised resolution of my smFISH data - a trade off to perform 4-channel imaging without photobleaching the tissues. While the resolution was high enough to detect bright nascent transcription spots, it was too low to detect single molecule mRNAs. As such, I was unable to convert the arbitrary unit (A.U.) of spot intensities into number of mRNAs. In the future, resolution should be further optimized by trialling different microscopes and detectors. For instance, the Gregor lab had great success using the hybrid detector which combines photomultiplier tubes (PMT) and avalanche photodiodes (APD) detectors to obtain a wide range of signal intensities with high resolution (Zoller et al., 2018).

4.3 References

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. *Developmental Biology*, 385(2):263–278.

Brunner, D., Dücker, K., Oellers, N., Hafen, E., Scholzi, H., and Klambt, C. (1994). The ets domain protein pointed-p2 is a target of map kinase in the sevenless signal transduction pathway. *Nature*, 370(6488):386–389.

Chen, K. H., Boettiger, A. N., Moffitt, J. R., Wang, S., and Zhuang, X. (2015). Spatially resolved, highly multiplexed rna profiling in single cells. *Science*, 348(6233):aaa6090.

Elowitz, M. B., Levine, A. J., Siggia, E. D., and Swain, P. S. (2002). Stochastic gene expression in a single cell. *Science*, 297(5584):1183–1186.

Hope, C. M., Webber, J. L., Tokamov, S. A., and Rebay, I. (2018). Tuned polymerization of the transcription factor Yan limits off-DNA sequestration to confer context-specific repression. *eLife*, 7:e37545.

Klämbt, C. (1993). The drosophila gene pointed encodes two ets-like proteins which are involved in the development of the midline glial cells. *Development*, 117(1):163–176.

Peccoud, J. and Ycart, B. (1995). Markovian modeling of gene-product synthesis. *Theoretical population biology*, 48(2):222–234.

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: structural insights into the regulation of yan by mae. *Cell*, 118(2):163–173.

Raj, A., Peskin, C. S., Tranchina, D., Vargas, D. Y., and Tyagi, S. (2006). Stochastic mrna synthesis in mammalian cells. *PLoS biology*, 4(10):e309.

Scholz, H., Deatrick, J., Klaes, A., and Klämbt, C. (1993). Genetic dissection of pointed, a drosophila gene encoding two ets-related proteins. *Genetics*, 135(2):455–468.

Shwartz, A., Yogev, S., Schejter, E. D., and Shilo, B.-Z. (2013). Sequential activation of ets proteins provides a sustained transcriptional response to egfr signaling. *Development*, 140(13):2746–2754.

So, L.-H., Ghosh, A., Zong, C., Sepúlveda, L. A., Segev, R., and Golding, I. (2011). General properties of transcriptional time series in escherichia coli. *Nature genetics*, 43(6):554–560.

Wu, C., Boisclair Lachance, J.-F., Ludwig, M. Z., and Rebay, I. (2020). A context-dependent

bifurcation in the pointed transcriptional effector network contributes specificity and robustness to retinal cell fate acquisition. *PLoS genetics*, 16(11):e1009216.

Zenklusen, D., Larson, D. R., and Singer, R. H. (2008). Single-rna counting reveals alternative modes of gene expression in yeast. *Nature structural & molecular biology*, 15(12):1263–1271.

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. *Molecular and Cellular Biology*, 30(5):1158–1170.

Zoller, B., Little, S. C., and Gregor, T. (2018). Diverse Spatial Expression Patterns Emerge from Unified Kinetics of Transcriptional Bursting. *Cell*, 175(3):835–847.e25.

APPENDIX A

APPENDIX:THE ROLE OF PNT AND YAN IN OMMATIDIAL PATTERNING: VIEW FROM E-CADHERIN SUBCELLULAR LOCALIZATION

A.1 Introduction

During animal development, gene expression changes that drive cell fate decisions often coincide with cell shape and arrangement changes that drive tissue morphogenesis. Feedback interactions between mechanical and chemical signals are thought to drive tissue self-organization (Chan et al., 2017; Kim et al., 2018).

The Drosophila retina is a highly tractable tissue whose development relies on complex epithelial remodeling and simultaneous specification of multiple cell types. Ommatidial patterning is initiated at the apical plane driven by progressive remodeling of non-muscle myosin II (myo-II) and E-cadherin (E-cad) at adherens junctions (AJs) (Robertson et al., 2012; Brown et al., 2006). Concurrently, EGF receptor (EGFR) - ligand interaction between neighbouring cells triggers the sequential specification of eight photoreceptor neurons (described in detail in Chapter 1).

Epithelial remodeling at the apical plane is driven by rearrangement of non-muscle myosin II (myo II) and E-cadherin (E-cad) at the adherens junctions (AJs)(Robertson et al., 2012; Brown et al., 2006; Escudero et al., 2007). Distinct stages of epithelial remodeling manifest in regularly spaced multicellular structures with unique packing geometry (Figure A.1). The first structure to emerge is the "line" at the posterior boundary of the furrow. This early structure contains five to seven cells connected by myo-II (total as well as phosphorylated (Robertson et al., 2012; Escudero et al., 2007)) enrichment at the posterior AJs. The line transitions into an "arc" as myo-II exerts a contractile force and bends the structure (Robertson et al., 2012; Escudero et al., 2007). E-cad starts to accumulate at mediolateral

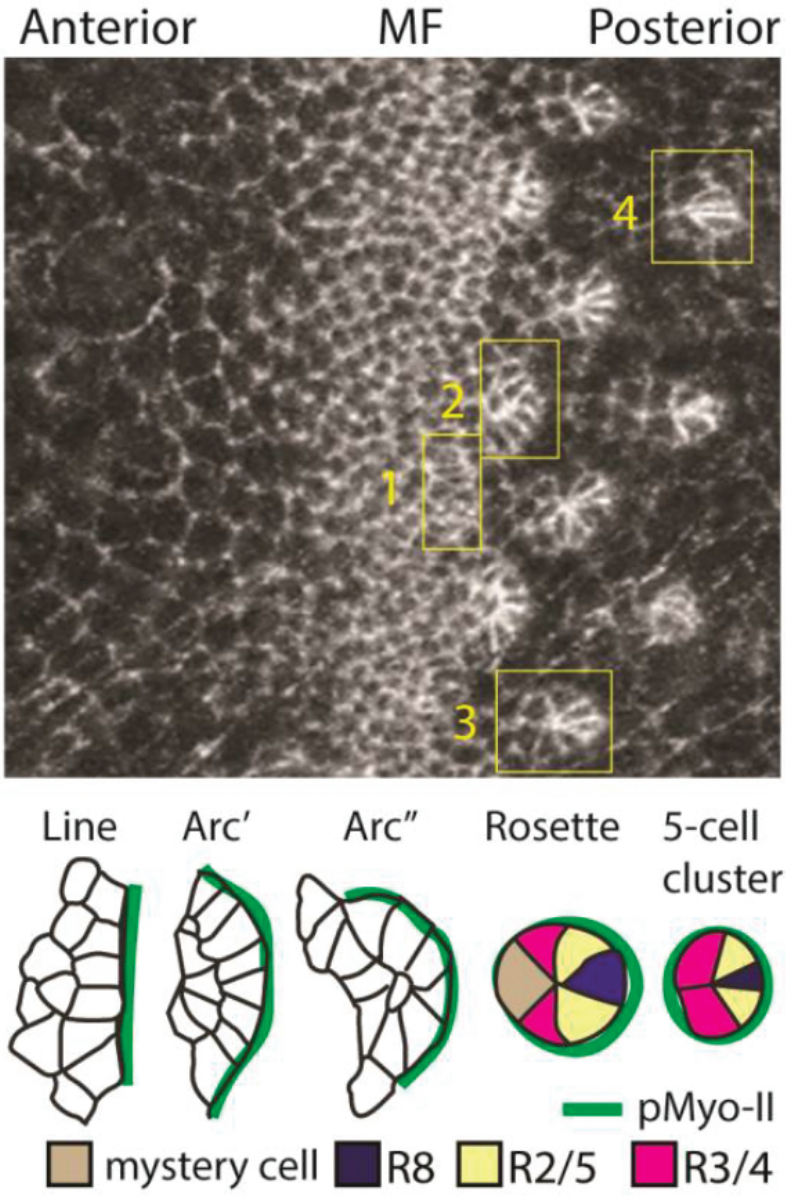

Figure A.1: Ommatidial patterning involves line-arc-five cell cluster transition. E-cad immunostaining of the 3rd instar larval imaginal eye disc with marked structures: 1. Line 2. Arc. 3. Rosette 4. Founding five-cell cluster. Graphic representation of progressive epithelial remodeling leading up to the five-cell cluster formation

AJs between these cells. The "arc" then zippers shut along the midline extruding multiple core cells in the process. This produces a six to seven cell "rosette" joined at the single vertex composed of five future photoreceptors and one or two extra cells, the mystery cells (Robertson et al., 2012; Escudero et al., 2007). Selective down-regulation of E-cad at AJs extrudes the mystery cells from the cluster (Robertson et al., 2012; Escudero et al., 2007). This forms the founding 5-cell cluster that is marked by supracellular myo-II ring along the posterior AJs and high E-cad at mediolateral AJs (Brown et al., 2006). In this cluster are the first five photoreceptors to differentiate: R8, R2/R5 and R3/R4 (Ready et al., 1976; Wolff, 1993; Brown et al., 2006). Photoreceptors R1/R6 and R7 are recruited after a final round of cell division, the second mitotic wave (SMW) (Ready et al., 1976; Wolff, 1993; Brown et al., 2006).

Atonal and EGFR signaling were shown to regulate apical epithelial remodeling by exerting regulatory roles on E-cad and myo-II. To summarize current knowledge: 1) Atonal promotes E-cad transcription and myo-II enrichment at the posterior AJs during the initial "line" formation (Brown et al., 2006; Robertson et al., 2012); 2) EGFR through the transcription factor Pnt remodels of E-cad and myo-II for the formation of "arcs" (Robertson et al., 2012); 3) EGFR hyperactivation can upregulate E-cad4 and myo-II at AJs to produce ommatidial "mega-clusters" containing ectopic photoreceptors (Escudero et al., 2007); 4) In myo-II dominant negative clones, neural marker (Elav) positive cells are present but typical ommatidial patterns are lost (Escudero et al., 2007).

A.2 Rationale

Here, I sought to understand how Pnt and Yan, the transcriptional regulators of RTK/EGFR signaling, mediate AJ remodeling posterior to the MF and thereby contribute to the morphogenesis of the developing eye in addition to their role in regulating the expression of photoreceptor fate-determining genes. Results from this investigation may inform how cell

fate specification and tissue morphogenesis might be coupled in space and time.

A.3 Results and Discussion

I used RNA interference (RNAi) against Yan or Pnt to reduce their expression levels at the furrow. The UAS-RNAi and UAS-Pnt RNAi transgenes were each overexpressed using the Eyeless(Ey)-Gal4 driver that is broadly expressed prior to the MF. Both RNAi against Yan and Pnt led to a striking defect in ommatidal patterning. Antibody staining against E-cad showed that in Yan RNAi tissues, there were less selective upregulation of E-cad. Instead, E-cad expression seemed uniform across cells at the MF. In wild type tissues, apical constriction of cells typically mark the MF (Figure A.2.A) however this was not seen in Yan RNAi tissues (Figure A.2.B).

In Pnt knockdown tissues, cells immediately posterior to the MF showed increased selective upregulation of E-cad (Figure A.2.C). This resulted in irregular spacing of the "arcs" due to an increased number of cells that participate in the photoreceptor cluster formation.

It would be interesting to see if Pnt and Yan regulate photoreceptor fate transitions and ommatidial patterning independently from each other through directly regulating the expression of photoreceptor fate-determining genes and regulating cytoskeletal and junction remodeling genes. Alternatively, it is possible that differentiation of progenitor cells into photoreceptor neurons leads to AJ remodeling in these cells, or vice versa.

A.4 References

Brown, K. E., Baonza, A., and Freeman, M. (2006). Epithelial cell adhesion in the developing drosophila retina is regulated by atonal and the egf receptor pathway. *Developmental biology*, 300(2):710–721.

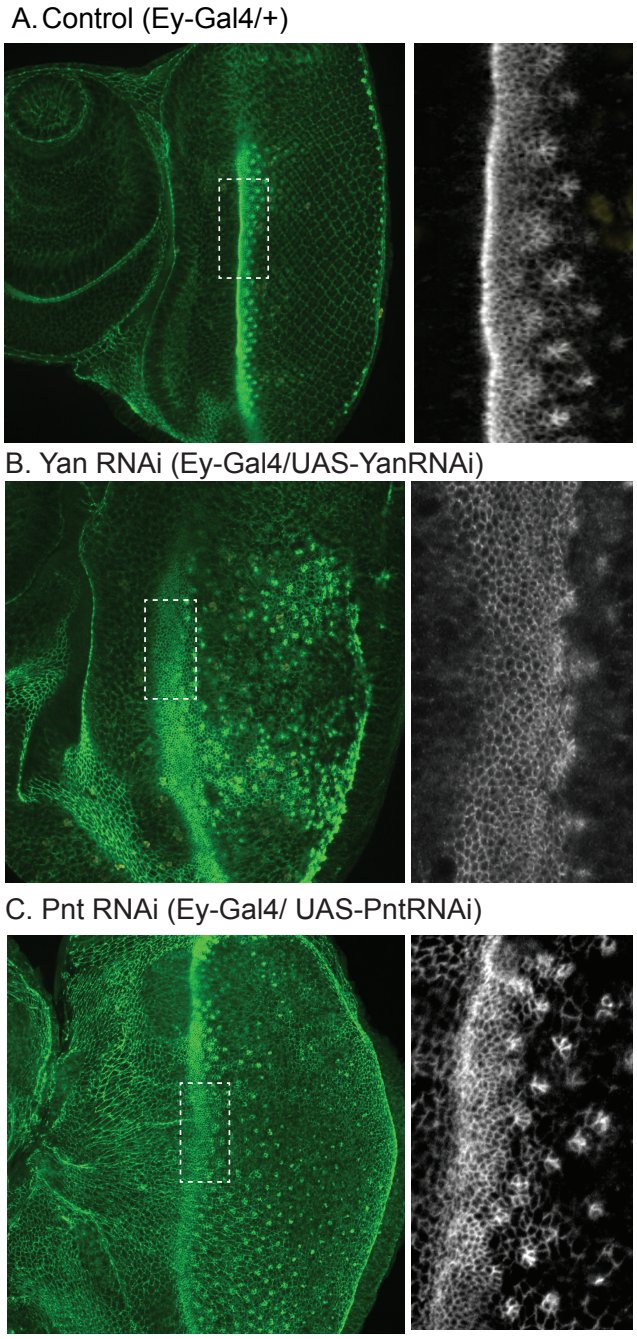

Figure A.2: Effects of perturbing Pnt and Yan on E-cad Remodelling During Ommatidia Morphogenesis

3rd instar eye imaginal discs of (A) Control (B) Yan RNAi and (C) Pnt RNAi driven by Ey-Gal4 with immunostaning against E-cad antibody. Images on the left show the entire disc proper (green: anti E-cad), and images on the right show a magnified view of the insets (dashed box) which mark a region spanning from immediately anterior to the MF, the MF, and immediately posterior (grey: anti E-cad).

Chan, C. J., Heisenberg, C.-P., and Hiiragi, T. (2017). Coordination of morphogenesis and cell-fate specification in development. *Current Biology*, 27(18):R1024–R1035.

Escudero, L. M., Bischoff, M., and Freeman, M. (2007). Myosin ii regulates complex cellular arrangement and epithelial architecture in drosophila. *Developmental cell*, 13(5):717–729.

Kim, E. J. Y., Korotkevich, E., and Hiiragi, T. (2018). Coordination of cell polarity, mechanics and fate in tissue self-organization. *Trends in cell biology*, 28(7):541–550.

Ready, D. F., Hanson, T. E., and Benzer, S. (1976). Development of the drosophila retina, a neurocrystalline lattice. *Developmental biology*, 53(2):217–240.

Robertson, F., Pinal, N., Fichelson, P., and Pichaud, F. (2012). Atonal and egfr signalling orchestrate rok-and drak-dependent adherens junction remodelling during ommatidia morphogenesis. *Development*, 139(18):3432–3441.

Wolff, T. (1993). Pattern formation in the drosophila retina. *The development of Drosophila melanogaster*, pages 1277–1325.

APPENDIX B

APPENDIX: THE EFFECT OF INCREASED YAN SELF-ASSOCIATION AFFINITY ON EMBRYONIC MESODERM SPECIFICATION

B.1 Introduction

Cardiac mesoderm specification is another developmental context in Drosophila that requires gene expression regulation by Yan. Specifically, downstream of RTK/MAPK signaling, Yan regulates the transcriptional activity at the *eve* locus. In the stage 10 embryo, a field of dorsal mesoderm cells of about twelve cells initially express the *eve* gene and are equally competent to form cardiac progenitor cells (Carmena et al., 1998). Eve expression becomes progressively refined to two to four Eve-positive cells (an average of three cells) per embryonic hemisegment by late stage 11 (Carmena et al., 1998). Loss of Yan results in excessive Eve-positive cells (Halfon et al., 2000; Carmena et al., 2002). Yan's indispensable repressive function in the larval eye disc and in the embryonic mesoderm provides an opportunity to compare and contrast the effect of tuning Yan's SAM-SAM affinity that drives its homotypic multimerization.

A previous theoretical study by the Rebay lab suggested that the interplay between the concentration of Yan monomers and it's SAM affinity for self-association fine-tunes the repressive activities of Yan (Hope et al., 2017, 2018). For example. if the concentration of Yan monomers is too low, regardless of how high its self-association affinity is, Yan monomers are less likely to encounter each other in its nuclear milieu. Meanwhile, at moderate to high concentration, increasing the SAM affinity would result in excessive self-association of Yan monomers. Indeed, Yan aggregates were observed in cells of the developing eye disc carrying endogenous Yan mutant alleles (Super Yan, SY) with increased SAM affinity (Hope et al., 2018). This was accompanied by Yan hypomorphic phenotype - excessive

progenitor cells transitioning to photoreceptor fates (Hope et al., 2018). How increased Yan SAM affinity influences its repressive function in the developing mesoderm during cardiac progenitor specification has not been investigated.

B.2 Rationale

It is unknown whether tuning the self-association capacity of Yan monomers produces tissue-specific outcomes on Yan's repressive activity and the cell fate decisions it regulate. Comparing the consequences of endogenous Yan mutant alleles that increase Yan's SAM affinity between the developing eye and the mesoderm is likely to reveal novel insights on how the intricate interplay between the subcelluar context and the protein-protein interaction of a transcription factor produce distinct cell fate outcomes.

B.3 Results and Discussion

All of the Super Yan alleles (SY3, SY4, or SY5) produced wildtype number of Eve+ cells in Stage 11 embryos. This is in contrast to the developing eye disc, in which all SY alleles produced a hypomorphic phenotype. This indicates that cells in different developmental contexts display differential susceptibility to increased Yan self-association.

Differential requirement of Yan self-association strength is also seen between different cell fates in the developing eye disc: the hypomorphic phenotype i.e. excessive photoreceptor cells affect outer photoreceptor cells (R3/R4) more frequently than the inner photoreceptor cell R7. One likely possibility for this difference is due to the different amount of RTK/MAPK signaling received by transitioning R3/R4 cells vs. R7 cells. Transitioning R7 cells receive higher MAPK signaling compared to R3/R4 cells due to the presence of a second source of RTK signaling from the Sevenless (Sev) receptor (in addition to the EGF receptor). RTK/MAPK signaling facilitates the breaking down of Yan polymers into Yan monomers,

and its degradation. It is therefore likely that higher MAPK signaling leads to reduced number of Yan multimers.

It remains to be seen if a similar mechanism is responsible for the observed wildtype phenotype in the embryonic mesoderm. In support of this, I did not observe the presence of Yan aggregates in the mesoderm cells. It would be interesting to quantify the absolute concentration of Yan in different contexts, and see if the concentration and Yan SAM affinity together can predict the repressive activity of Yan at target gene loci, as well as the downstream cell fate outcome.

B.4 References

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. *Developmental Biology*, 244(2):226–242.

Carmena, A., Gisselbrecht, S., Harrison, J., Jimenez, F., and Michelson, A. M. (1998). Combinatorial signaling codes for the progressive determination of cell fates in the Drosophila embryonic mesoderm. *Genes & Development*, 12(24):3910–3922.

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(1):63–74.

Hope, C. M., Rebay, I., and Reinitz, J. (2017). DNA Occupancy of Polymerizing Transcription Factors: A Chemical Model of the ETS Family Factor Yan. *Biophysical Journal*, 112(1):180–192.

Hope, C. M., Webber, J. L., Tokamov, S. A., and Rebay, I. (2018). Tuned polymerization

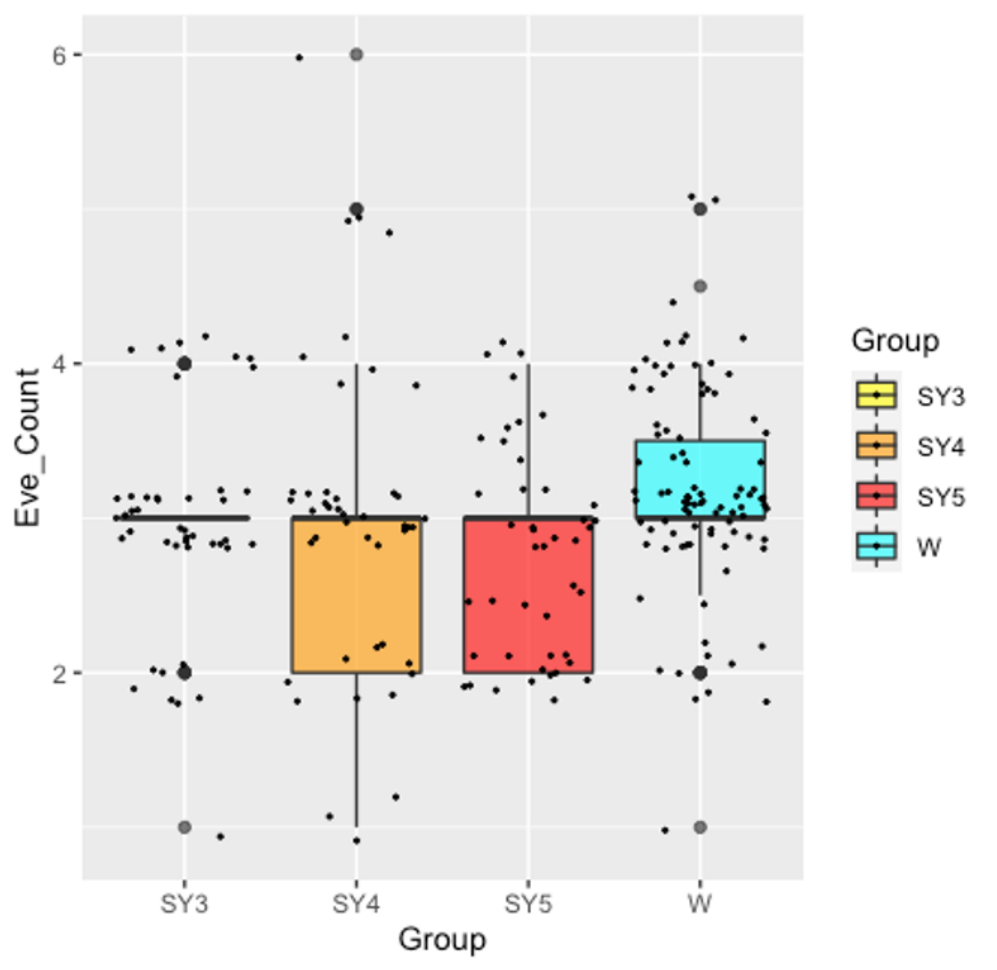

Figure B.1: Increasing Yan SAM affinity produces no significant effect on the number of cardiac muscle progenitor cells in stage 11 embryonic mesoderm
Box plot showing the distributions of Eve+ cell count in Stage 11 embryonic hemisegments for each indicated genotype. None of the SY allele produced a significant difference compared to the wildtype (Student's T-test; p=0.89, p=0.62, p=0.62 for SY3, SY4, SY5 respectively compared to the wild type)

of the transcription factor Yan limits off-DNA sequestration to confer context-specific repression. *eLife*, 7:e37545.

APPENDIX C

APPENDIX: SEQUENCES FOR SMFISH PROBES

The following tables contain smFISH oligonucleotide sequences against Pros, Salm, and Sens transcripts. See Chapter 3 for details on its design, crosslinking with ATTO633 dye, and subsequent purification.

Name	Sequence	Name	Sequence
Pros1	tgctgctgttggtttttag	Pros25	ttgatgtccgggaagtacat
Pros2	agttcgagttcgagttcttg	Pros26	caattgtgctgtgttgttct
Pros3	aggatctttgacgcgaacga	Pros27	ttctcggaagttcgagaacc
Pros4	tgaccgattgtggttggaac	Pros28	cgtgcatatttctccatttg
Pros5	aaactggtgagttgctcgtg	Pros29	atcaacagatcatcgggtgt
Pros6	acaaggacgatgccgaactg	Pros30	agcacgcgatacaattcact
Pros7	accaaacaggttgccaaagg	Pros31	gttattgcgattgtagtgca
Pros8	caaacagctcgttcagcatc	Pros32	aattctggggtacctcaatg
Pros9	gatgcgagtgactggaactg	Pros33	gagtcgattcgaccacgaag
Pros10	tgttgcaactgagactggtg	Pros34	gtattgcacggaagaactcc
Pros11	atattgcgcagcatgtgatg	Pros35	atatggacttcttccacgac
Pros12	tccgaatcgtcgatcagatt	Pros36	attgggcgacttgaagtact
Pros13	ctcttgatcttgatgtcagc	Pros37	ctcgattttatttgctgggt
Pros14	tcttacaaccatttacctgc	Pros38	gtgtgtggatactgtttgtc
Pros15	atcatgctctgcagattgag	Pros39	aaacgcgttttcgactcgtg
Pros16	acacgcttctgttggatttg	Pros40	ctggatttgtatatccgtgg
Pros17	acatatttctgctgcatctc	Pros41	cttgctgctctatgttaagt
Pros18	ttgctcatcatctgggatag	Pros42	gtaaattacgctcggtgtgt
Pros19	aagaactgctggtacatggc	Pros43	acaaaccaacggcgagtaca
Pros20	agcatttggaacttttccga	Pros44	gtgtgtcaactttcggtaga
Pros21	cgcatcattgagctattgtt	Pros45	tgtgtaggtatttccttgtg
Pros22	aaaagaccttggactcgtgc	Pros46	aactgcttaccggaaagtcg
Pros23	cggattgaagaacggcgagt	Pros47	cctgcttttgtaaagcagtg
Pros24	tgcatcggtgtcaatgtaga	Pros48	agattaattcccgatttccg

Figure C.1: Pros smFISH probe set

Name	Sequence	Name	Sequence
Salm1	tgcggcactgagagaaatcg	Salm25	cggtggtaaagttaaggggg
Salm2	gcgcgtgtttagctaattta	Salm26	tcgtgcttgatggggttaaa
Salm3	cagtgggaaacgcgtagtac	Salm27	aggtgttcgaaacctcgatg
Salm4	tcagcttgatcattgcacta	Salm28	atcacagaccacacactgat
Salm5	acacagaacgttggacaggt	Salm29	cactggtggaagtttctcat
Salm6	tatccttgaagtcactacgc	Salm30	gcattcgagtacttcttgtg
Salm7	ttttattgatggtctcctgg	Salm31	ataaagtgaccgggcatcat
Salm8	ccttcagctgtttgacaatg	Salm32	acgacacatcatcgtcgtag
Salm9	aaccattcgcaatggtcttg	Salm33	ctcagcttttgctcgaacaa
Salm10	tcataggtatctgtttcctg	Salm34	cgaagagctgcgcttgataa
Salm11	tttattagatgctgacgcgc	Salm35	ctgagaagcctctgaacgag
Salm12	cttcagagattcagctggag	Salm36	tcacagatgctgcacttgaa
Salm13	tgtgttatccaaactcggtc	Salm37	tacttattcgcgatcttcgt
Salm14	tggttggtgataatgctgga	Salm38	tcttgatagcggattaagcc
Salm15	tttctcaccgaaggcaaagt	Salm39	gcgagaatcgattgcatctt
Salm16	tgaagaacggctcattgcga	Salm40	ccccggctagatcataaaaa
Salm17	atcttatgtggatctggagc	Salm41	gcatctttggatacgtttca
Salm18	aagtgaaccttaaggttgcc	Salm42	agattccgttctcagttttg
Salm19	tgaaacttgtccatgtgctc	Salm43	tcggcgatcgcatctaaatt
Salm20	gggaatgattgggagagcta	Salm44	gcaaagcgtttggatttcga
Salm21	tagagattctgaaggccagg	Salm45	gtggttcttcatctgtgata
Salm22	ttaaggatttccataggcgg	Salm46	ttaatccttccttccaatga
Salm23	tctgatgaaggagagggtgt	Salm47	gtacgatcgctcatagactt
Salm24	gaagaaatcttcgcggggtg	Salm48	acattggcatccgtgtataa

Figure C.2: Salm smFISH probe set

Name	Sequence	Name	Sequence
Sens1	agacgaacacgtccgtatcg	Sens25	ttgctgtggtgtactcgaac
Sens2	cggttttatcttgactcact	Sens26	ggtgaccagatgatgttacc
Sens3	cttcgatgggtcttattcag	Sens27	cgatctctcgcacatcgagg
Sens4	gtgatcttggattcgaacgg	Sens28	acaataggatccacctggtg
Sens5	caaatctcactttcttggcg	Sens29	gcgctcatacttgaactttc
Sens6	ttttctcaatactgttcact	Sens30	gcgaggatatggatgcagtt
Sens7	ttgctttaacagtctttgga	Sens31	agcgaggagatcggcgattg
Sens8	tcactggatctgatttacgt	Sens32	tcgaactccagatcctggac
Sens9	ctggatctttacttttgga	Sens33	atgggcatacaactgctgtt
Sens10	cggcgataggtgattcattt	Sens34	cgtcaatccggccatgaatg
Sens11	tatcacagtgttgaaggcgg	Sens35	tgagcagctccagattgttg
Sens12	gcgggttgtacaaaaagggc	Sens36	ctcttcagcttcaagtgctg
Sens13	agtagggcactcgagtagag	Sens37	ttgatgcgatgctgttgctg
Sens14	cgtggccgaggacaacaaaa	Sens38	tgccaccaaattcatgagtg
Sens15	ctgacccaaaaccacagagg	Sens39	gtgttgattacgcatgtagc
Sens16	tcaaggcaaagtcacgatcc	Sens40	tgatgcaactgttgttgctg
Sens17	ttcatttggagttcatgctc	Sens41	tgatgctgatgttgctgcgg
Sens18	cttgctgttctcattgttat	Sens42	cgtcgagtcaggatgttgct
Sens19	gtgcttagatttagtggcat	Sens43	gacgaggaggacgaacgacg
Sens20	gaatcatccgatgtgatgcg	Sens44	ttcgttttctccttggtaag
Sens21	ctgctgtgatactgatctcg	Sens45	gtggatgagcagatctggaa
Sens22	cacttcactgctggaactgc	Sens46	ccacaatattggcagggata
Sens23	taacgttcagcgaggtcatc	Sens47	atgtccgacttttggtggaa
Sens24	tttagattcacagcgctcag	Sens48	gtatgtatgtacgtgtgctt

Figure C.3: Senseless smFISH probe set