UNIVERSITY OF CALIFORNIA, SAN DIEGO

How Would I Know That This Could Be My Fate: Transcriptional Regulation of Specification and Differentiation in the *Drosophila* Mechanosensory Organ Lineage

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy

in

Biology

by

Steven Walter Miller

Committee in charge:

Professor William J. McGinnis, Chair
Professor Lawrence S. B. Goldstein
Professor Cornelius Murre
Professor James W. Posakony
Professor John B. Thomas

2009

The Dissertation of Steven Walter Miller is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____

_____

_____

Chair

University of California, San Diego

2009

DEDICATION

For Allison, My Love and Inspiration

EPIGRAPH

Bill Watterson, 30 November 1985.
Universal Press Syndicate.

TABLE OF CONTENTS

Signature Page……..... iii

Dedication…….. iv

Epigraph……. v

Table of Contents……. vi

List of Figures…… viii

Preface……... xiv

Acknowledgements…….. xv

Vita…….. xix

Abstract of the Dissertation…..…….. xxi

Chapter One: Introduction: Transcriptional regulation of cell fate specification and differentiation……..... 1

Chapter Two: Cell non-autonomous auto-regulation of *neuralized* during *Drosophila* mechanosensory organ precursor lateral inhibition ……..…… 46

Chapter Three: Complex interplay of three transcription factors in controlling the tormogen differentiation program of *Drosophila* mechanoreceptors..…… 173

Chapter Four: Auto-regulation, proneural synergy, and mutual inhibitory interactions shape expression of *shaven* and influence the tormogen and tricogen differentiation programs..... 189

Chapter Five: pMemor: A highly versatile system for rapid enhancer transgenesis and beyond..... 223

LIST OF FIGURES AND TABLES

Chapter One

Figure 1.1: Cell type diversity is dependent upon differential fate specification…… 6

Figure 1.2: Simplified gene regulation…….. 10

Figure 1.3: Combinatorial input of cell type specific transcription factors and extra-cellular signaling pathways regulating gene expression…….. 14

Figure 1.4: The Notch signaling pathway promotes binary cell fate specification signaling between adjacent cells……. 19

Figure 1.5: The mechanosensory organs of *Drosophila melanogaster* require the combinatorial function of the Notch signaling pathway mediated lateral inhibition and the proneural proteins, Achaete and Scute……..... 23

Figure 1.6: Iterative Notch signaling and differential transcription factor expression within the mechanosensory organ lineage…… 28

Table 1.1: Summary of transcription factors expressed in the adult mechanosensory organ lineage…….. 29

Figure 1.7: The specification-differentiation interface…….. 33

Chapter Two

Figure 2.1: The *neur* 4D enhancer module directs SOP expression downstream of proneural activity…… 69

Figure 2.2: Functional requirements of bHLH-R and proneural binding sites in the *neur* 4D SOP enhancer…….. 73

Figure 2.3: Activation of the *neur* 4D enhancer includes input from the proneural and at least two other inputs…… 77

Figure 2.4: Embryonic expression of the *neur* 4DWT enhancer and mutants…… 81

Figure 2.5: Chn is capable of binding SMCα sites and can activate the 4D enhancer……. 84

Figure 2.6: *Drosophila* NF-κB proteins are not nuclear localized in the SOP…….. 87

Figure 2.7: *neur* nascent transcript as well as protein is detectable in multiple cells prior to SOP specification……. 91

Figure 2.8: Nascent *Dl* expression in SOPs……... 95

Figure 2.9: Persistent non-SOP *neur* expression causes lateral inhibition defects…… 99

Figure 2.10: Inhibition of Neur function in the SOP causes ectopic *neur* mRNA accumulation and neur4DWT activity……. 102

Figure 2.11: Phenotypes of a bHLH-R site mutant *neur* rescue construct in pCaSPer…… 105

Figure 2.12: A second enhancer in the *neur* locus has SOP specificity…… 109

Figure 2.13: The NRS1 region, although exhibiting inversions throughout evolution, contains a conserved bHLH-R consensus site and several conserved PN consensus sequences……... 113

Figure 2.14: Consequences of mutation of proneural and bHLH-R inputs upon the 1B subfragment……. 116

Table 2.1a: neurRC bristle counts, total head and thorax……. 119

Table 2.1b: neur4D,1BRm, dorsocentral and scutellar bristles… 119

Figure 2.15: Mutation of the bHLH-R binding sites in the *neur* locus causes ectopic expression of *neur*…….. 123

Figure 2.16: Functional redundancy in the two *neur* SOP enhancers during lateral inhibition…….. 127

Supplemental Figure 2.17: Additional examples of *neur* expressing cells in the presumptive SOP domain……. 155

Supplemental Figure 2.18: *neur* and *Dl* expression in the embryo…… 157

Supplemental Figure 2.19: *Dl* expression in heterochronic SOP pairs……... 158

Supplemental Figure 2.20: The 1B enhancer directs reporter expression that accumulates in the lineage independently of the intact proneural binding site……. 159

Supplemental Figure 2.21: Low level, ubiquitous expression of neurRC-WT-GFP…….. 160

Chapter Three

Figure 1: *Sox15* is expressed specifically in socket cells of external sensory organs……... 175

Figure 2: Identification of the *Sox15* socket enhancer…….. 177

Figure 3: Analysis of conserved sequence motifs in the *Sox15* socket enhancer reveals Su(H) and Vvl as potential regulatory inputs…….. 178

Figure 4: Loss of Su(H) and Vvl binding in cis and in trans affects *Sox15* socket enhancer activity…….. 180

Figure 5: *Sox15* loss-of-function phenotypes……. 181

Figure 6: Socket cell phenotype of the *Su(H)* ASE-*Sox15* double mutant……... 183

Figure 7: Summary model of the collaborative roles of *Su(H)* and *Sox15* as N pathway targets required for socket cell differentiation……. 184

Chapter Four

Figure 4.1: Identification of and conservation within a sheath/shaft enhancer upstream of *sv*…….. 203

Figure 4.2: *sv* expression in the lineage is controlled by auto regulatory activity…….. 208

Figure 4.3: Mutual inhibition of the activities of Su(H), Sox15, and Sv……. 212

Chapter Five

Figure 5.1: The Memor system for enhancer transgenesis…….. 242

Table 5.1: Memor Docking Site Starter Line Mapping Information……. 245

Figure 5.2: Recombination mediated cassette exchanges (RMCEs) at Memor docking sites……... 254

Figure 5.3: Enhancer activity in Memor docking site #39…….. 257

Table 5.2: ΦC31 RMCE into Memor docking sites..... 260

Table 5.3: Memor Enhancer Integrant Orientation and Expression…….. 262

Table 5.4: Cre RMCE of Memor Docking Site Architecture……. 265

Supplemental Figure 5.4: The crossing scheme for collecting RMCE integrants in the Memor system…….. 269

PREFACE

I first would like to issue a disclaimer that I am horribly naïve, to a fault. I hold on to the belief that at least one person out there who is not at all an expert any the field of biology will some day flip through the pages of this dissertation. Perhaps they are a relative, or happened upon it accidentally while looking for another Steve Miller, or came looking for material to write my future biography ("Idiocy Disguised: the Life and Failures of Steve Miller—No, Not That One"). Whatever the purpose, it is for this reason that my introductory chapter contains a large portion of basic explanation leading into the meat of the background information for this manuscript. I think very highly of those scientists who take the time and can effectively communicate complex topics to the novice public, and you will see from these pages that I am not one of those scientists. I say this not only to warn you, that one person whom I'm talking about, but also to warn those expert biologists who look at those initial pages and think to themselves, "What is this worthless drivel?" To the latter (especially my committee members) I apologize.

ACKNOWLEDGEMENTS

I wish to above all acknowledge all those who have made contributions to the breadth of both our knowledge of Biology and the tools we use for perpetual discovery. Everyday is a journey that begins upon the footsteps of others.

Along this path I have benefited immensely from the influence and support of others along the way. For all my middle and high school science teachers, I wish to thank you for the tireless work you have done; you have often been in my thoughts, especially during the construction of this document. To my parents, sources of great personal support and direction, I wonder where I might be had you not seen the biologist in me at a time when I had yet to find it. I also wish to acknowledge the enormous impact upon both my personal and scientific career by the late Danny Brower, the sudden loss of whom has left me with great sadness, but the memory of whom has instilled me with great strength for what the future may hold. To the members of my thesis committee, I can think of no others with whom I would wish to walk the path would that I repeat it again; I thank you for your time, your thoughts, your critiques, your encouragement, and your

patience. I cannot forget the joy I have had in my interactions with members of the Posakony lab throughout this work, especially Joe Fontana, Nick Reeves, MarkRebeiz, Mariano Loza Coll, Feng Liu, Sui Zhang, Scott Barolo, and Par Towb; science is less fulfilling without great colleagues. Lastly, I wish to thank James Posakony, for your support not often acknowledged, for your direction not always followed, for your thoughts not often sought (but always appreciated), and for your influence not soon to be forgotten.

I would be ungrateful to end without extending my appreciation for my wife, Allison. This journey has been less burdensome with you as a traveling companion. I shall truly relish each future step as we continue along together, wherever the path may lead.

Chapter Two, in full, is a manuscript in preparation for publication: Cell non-autonomous auto-regulation of *neuralized* during *Drosophila* mechanosensory organ precursor lateral inhibition. Miller, S. W. and Posakony, J. W. The dissertation author was the primary researcher and author. James Posakony provided critical comments and oversight through the course of this investigation. Nick Reeves

cloned the recombinant Chn construct, and Joseph Fontana generated the *E(spl)m4, E(spl)mα* double deletion lines.

Chapter Three, in full, is a reprint of the material as it appears in Developmental Biology 2009. Miller, S.W., Avidor-Reiss, T., Polyanovsky, A., and Posakony, J. W., Complex interplay of three transcription factors in controlling the tormogen differentiation program of Drosophila mechanoreceptors. *Dev Biol*. 2009 May 15;**329**(2):386-399. The dissertation author was the primary researcher and author of this paper. James W. Posakony provided critical oversight and discussion throughout this work. Tomer Avidor-Reiss performed the electrophysiological studies, and Andrey Polyanovsky performed the transmission electron microscopy analysis.

Chapter Four, in full, is a manuscript in preparation for publication: Auto-regulation, proneural synergy, and mutual inhibitory interactions shape expression of *shaven* and influence the tormogen and tricogen differentiation programs. Miller, S. W. and Posakony, J. W. The dissertation author was the primary researcher and author. James Posakony provided critical comments and oversight through the course of this investigation. Scott Barolo originally located the

enhancer, Jamy Peng generated the UAS-Sox15 construct, and Tammie Stone performed the random binding site selection that was used to deduce a Sv binding consensus.

Chapter Five, in full, is a manuscript in preparation for publication: pMemor: A highly versatile system for rapid enhancer transgenesis and beyond. Miller, S. W. and Posakony, J. W. The dissertation author was the primary researcher and author. James Posakony provided key inputs into the initial designs of the system and provided crucial oversight throughout the study. Scott Barolo also contributed to construct design.

VITA

2000	Bachelor of Science, University of Arizona
2000-2001	Research Technician, University of Arizona
2001-2009	Research and Teaching Assistant, University of California, San Diego
2009	Doctor of Philosophy, University of California, San Diego

PUBLICATIONS

Miller SW, Avidor-Reiss T, Polyanovsky A, Posakony JW. Complex interplay of three transcription factors in controlling the tormogen differentiation program of Drosophila mechanoreceptors. *Dev Biol.* 2009 May 15;**329**(2):386-99.

Bunch TA, Miller SW, Brower DL. Analysis of the Drosophila betaPS subunit indicates that regulation of integrin activity is a primal function of the C8-C9 loop. *Exp Cell Res.* 2004 Mar 10;**294**(1):118-29.

Miller SW, Hayward DC, Bunch TA, Miller DJ, Ball EE, Bardwell VJ, Zarkower D, Brower DL. A DM domain protein from a coral, Acropora millepora, homologous to proteins important for sex determination. *Evol Dev.* 2003 May-Jun;**5**(3):251-8.

Baker SE, Lorenzen JA, Miller SW, Bunch TA, Jannuzi AL, Ginsberg MH, Perkins LA, Brower DL. Genetic interaction between integrins and moleskin, a gene encoding a Drosophila homolog of importin-7. *Genetics.* 2002 Sep;**162**(1):285-96.

Jannuzi AL, Bunch TA, Brabant MC, Miller SW, Mukai L, Zavortink M, Brower DL. Disruption of C-terminal cytoplasmic domain of betaPS

integrin subunit has dominant negative properties in developing Drosophila. *Mol Biol Cell.* 2002 Apr;**13**(4):1352-65.

FIELDS OF STUDY

Major Field: Cellular and Developmental Biology and Genetics

Studies in Developmental Genetics in the *Drosophila* Mechanosensory Lineage
Professor James W. Posakony

Major Field: Molecular and Cellular Biology and Genetics

Studies in Molecular Genetics of *Drosophila* Integrin Biology
Professor Danny L. Brower

Studies in Molecular and Developmental Biology of the coral *Acropora millepora*
Professors Danny L. Brower and Eldon E. Ball

ABSTRACT OF THE DISSERTATION

How Would I Know That This Could Be My Fate:
Transcriptional Regulation of Specification and Differentiation in the *Drosophila* Mechanosensory Organ Lineage

by

Steven Walter Miller

Doctor of Philosophy in Biology

University of California, San Diego, 2009

Professor William J. McGinnis, Chair

Molecular and genetic studies over the past two decades have established that animals ranging from the fruit fly to the human utilize a common “toolkit” of genes to create the myriad cell types during

development. For even some of the most well studied developmental paradigms, however, little is known of the transcriptional regulatory networks activated by these core components. Moreover, how these networks manifest the initial cell fate decision remains a rigorous topic of investigation.

The *Drosophila melanogaster* mechanosensory organ lineage has long been a model for binary cell fate specification mechanisms. While it is well-established that the Notch signaling pathway is the dominant cell fate specification mechanism in this lineage, little is known of the battery of genes downstream of the canonical pathway components that enforce the cell fate decision.

In this body of work I analyze the Notch signaling events at opposite ends of the lineage with the specification of the mechanosensory organ precursor cell (SOP) at the beginning and the post-mitotic socket and shaft cells at the penultimate of the lineage. These studies focus on the transcriptional regulation of three genes, *neuralized*, *Sox15* and *sv*. I utilize a number of loss of function mutants, misexpression studies, and transgenic reporter constructs to identify enhancer modules and tease out their regulatory factors.

Chapter 2 comprises an analysis of the transcriptional regulation of *neuralized* during SOP specification where I present evidence that *neuralized* is both a participant in and a target of Notch signaling, requiring two partially redundant. Chapter 3 focuses on the transcriptional regulation and function of *Sox15*, a transcription factor specifically expressed in the post-mitotic socket cell. I show that *Sox15* is a target of Notch signaling in the socket cell and is required for the proper differentiation of the socket cell as it relates to mechanosensory function. Chapter 4 dovetails off *Sox15* and examines regulation of *sv* in the shaft cell, where I find the dominant theme is auto-regulation. Lastly, Chapter 5 describes the development of a new transgenic reporter system I developed for the rapid screening of genomic fragments for enhancer activity.

CHAPTER ONE:

Introduction:

Transcriptional regulation of cell fate specification and differentiation

Listening to the genome

Whether much of the world's population realizes, we are living in the "Genomics Age" (Daly et al., 2001; Walsh, 2001; Sharp et al., 2004; Smith, 2004; Cleary et al., 2005; Xu, 2006; Gelvin, 2009). The flood of complete sequenced genomes in the past decade has not diminished the imagination for applications of genomics information, especially from the human genome. Since its completion, the fantasy persists for the potential to use a single person's DNA sequence to predict their risk(s) to specific diseases then apply this knowledge to tailor specific treatments (Subramanian et al., 2001). Furthermore, this information might even be used to "fix" a person's DNA sequence, effectively replacing the suspect nucleotides. From a broad clinical standpoint, such practices for the moment are still largely in the realm of science fiction, both in terms of the costs involved with sequencing a person's entire genome and in the completeness in our understanding of what changes in the DNA sequence translate to in the clinic. This may be changing in the coming years, however.

Part of the problem is that a significant portion of the genome remains a mystery. Less than two percent of the human genome

contains sequences that code for proteins, the main structural and functional cellular components (Venter et al., 2001). If these regions are mutated, it is fairly straightforward to determine the consequences, and these protein-coding sequences have long been targeted for detailed functional analyses. So much focus has been placed on the importance of the protein coding sequences that the remaining ninety-eight percent of the genome has been referred to as "junk DNA" (Smith, 1972). Not surprisingly, it turns out that much of this non-coding sequence is not "junk" after all, but contributes significantly to the function of the genome in the diverse cell types that make up our body, and may also be part of what makes us human (Bird et al., 2006; Ponting and Lunter, 2006; Venkatesh et al., 2006; Hahn, 2007; Cobb et al., 2008). Indeed, significant efforts are continuing to sieve the information out of the "junk" (2004; Celniker et al., 2009).

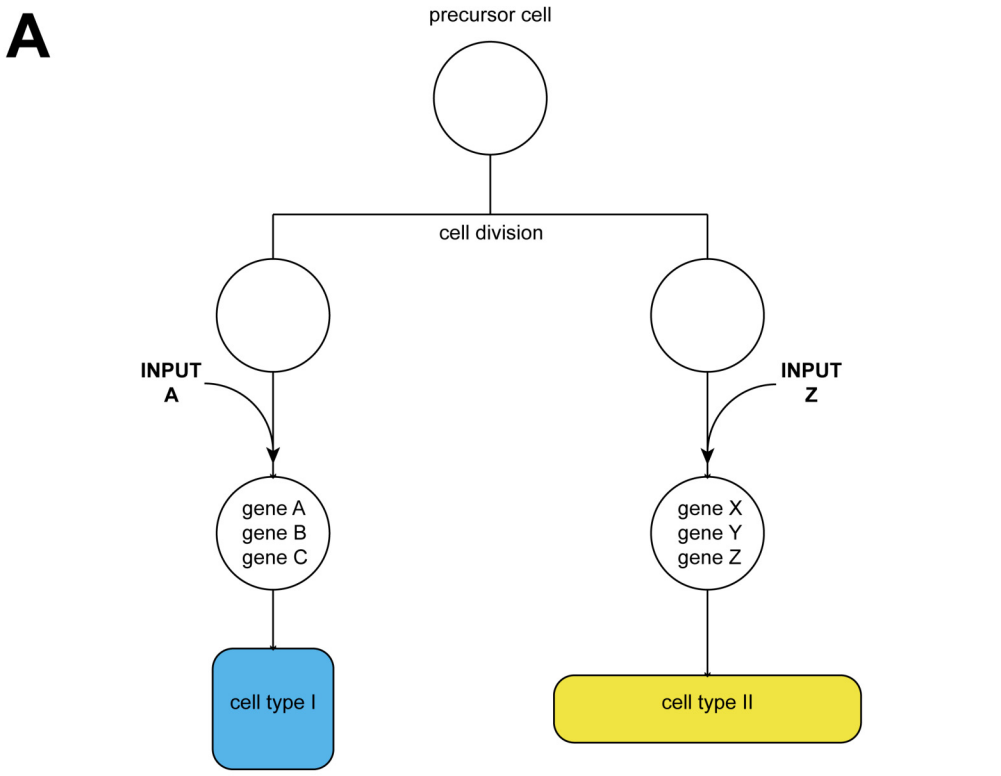
Needing "junk" to make the trunk

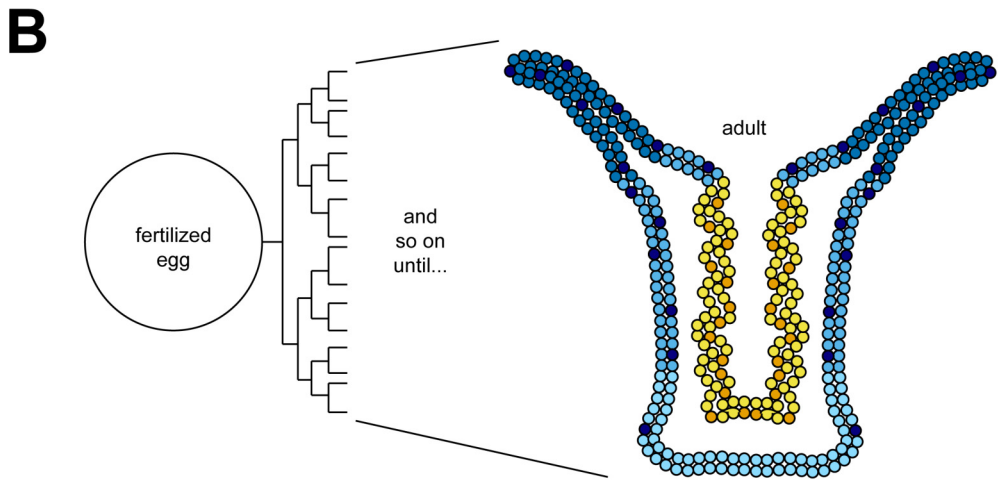
A major focus of developmental biology is understanding how the entire complement of cell types comes to be during the growth of an organism from a fertilized egg to an adult. The basic picture that

has emerged is illustrated in Figure 1.1. The development of an organism is essentially a large number of divisions of precursor cells. As each precursor cell divides, its daughter cells are exposed to cues that direct them to adopt unique fates. These cues manifest themselves as differential gene expression between the two different daughter cells. It is this change in gene expression throughout development that drives the diversity of cell types in a mature organism. Thus, identifying both the signals that an individual cell receives at a specific developmental time point and how those signals manifest as differential gene expression is fundamental to understanding the biology of an organism—both in terms of how it works and how to fix problems—be it person or pachyderm.

In general, gene expression is controlled by regulatory proteins known as transcription factors (since they regulate the making of an RNA molecule from a DNA template, a process called transcription), many of which can recognize specific nucleotide sequences in the DNA. Much of the functional binding takes place in the "junk DNA," outside of the coding sequence, in regions of the genome referred to as promoters and enhancers. A large number of proteins bind to the

Figure 1.1: Cell type diversity is dependent upon differential fate specification.
(A) Cell type diversity is achieved through cells receiving and responding to different signals. In this basic example, a single precursor cell divides to produce two daughter cells. Each daughter receives a different input, which encourages each daughter cell to express different genes. The action of these genes in response to the differential inputs results in the two daughter cells manifesting two distinct types, or fates (type I or type II). (B) Throughout the development of an entire multicellular organism, these types of responses to different inputs as cells divide and differentiate lead to the compilation of a multitude of cell fates in the adult (represented by different colors).


A
precursor cell
cell division
INPUT
A
INPUT
Z
gene A
gene B
gene C
gene X
gene Y
gene Z
cell type I
cell type II


B
fertilized
egg
and
so on
until...
adult

DNA sequences at the promoter (Juven-Gershon et al., 2006; Juven-Gershon et al., 2008), assembling a multi-protein complex whose purpose is to initiate transcription of the gene under the right conditions (Thomas and Chiang, 2006). RNA polymerase, which generates the RNA molecule encoded by each gene, is a key member of the complex bound at the promoter sequences. For many so-called "housekeeping" genes—those who perform basic metabolic functions necessary in every cell regardless of type—the promoter may be sufficient to activate gene expression (Farre et al., 2007). Often, however, this is not the case, and the level of expression due to proteins binding to promoter sequences at certain genes needs to be "enhanced;" this is where enhancers come into play.

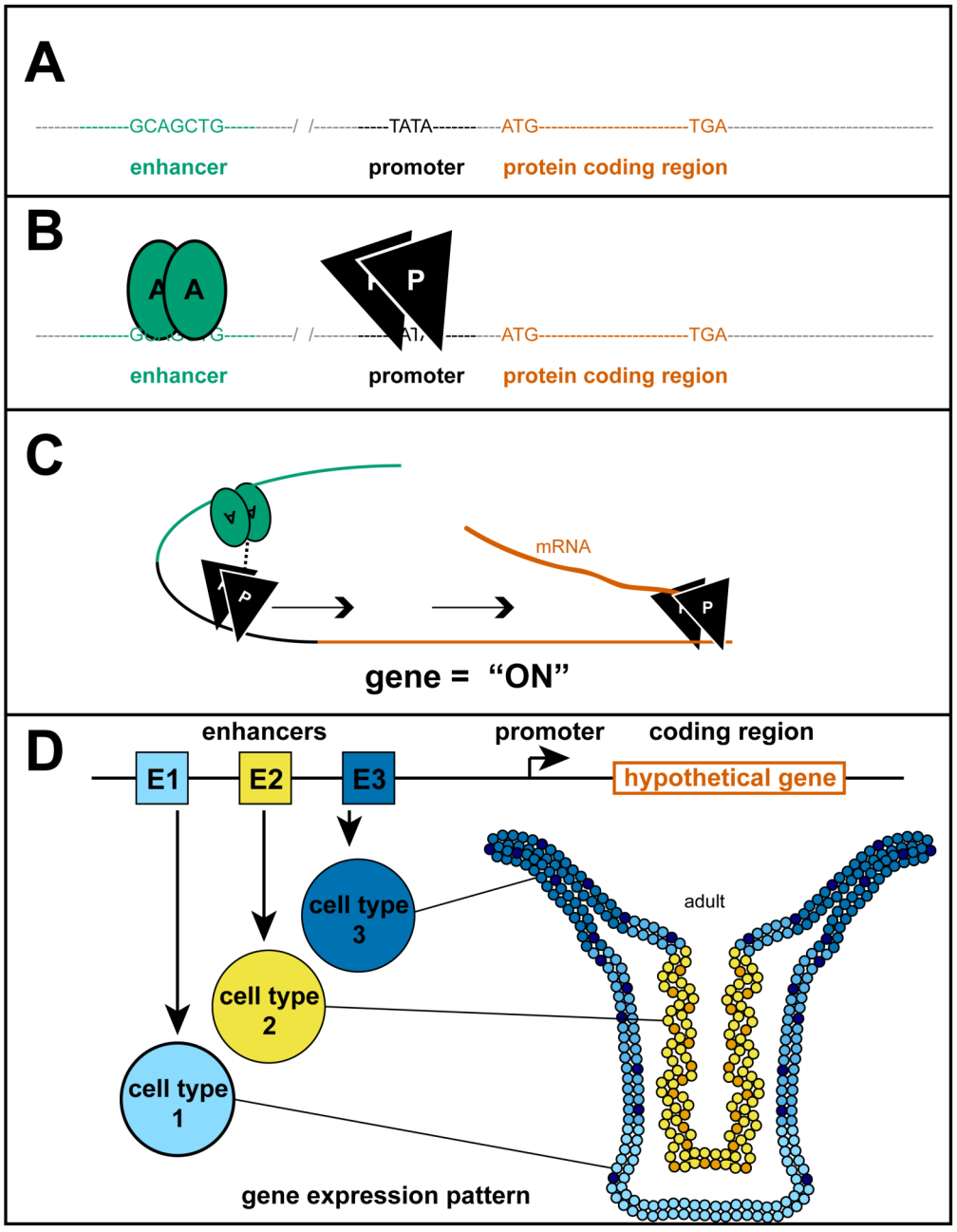
Enhancers are sequences that can sometimes be a long distance from the gene which they regulate, and are defined by regions of DNA which, when placed in combination with a promoter, can provide a significant increase in transcriptional output than that which occurs with the promoter alone (Crenshaw et al., 1989). In this sense, enhancers are like molecular switches; allowing genes to be turned on at particular times during the development of an organism. Moreover,

this enhancement may be required not only to a specific time, but to only a small subset of cells in the organism. To continue with the switch analogy, imagine a house with many rooms. In these times of energy consciousness, it would be wasteful to keep on all the lights in the house at once day and night; rather it makes more sense to have all the lights with their own switch to be able to control which rooms are lit at any given time.

Similar to the promoter, enhancers are regulated by the binding of transcription factors. These transcription factors can be expressed either in every cell or only in very specific cell types; the combinatorial activity of different transcription factors on the same enhancer promotes activation of the enhancer only when the precise complement is present, thereby restricting activation of the target gene (Maniatis et al., 1987). While the exact molecular nature of enhancer operation is still unknown, a common model involves a change in DNA architecture (de Laat and Grosveld, 2003; de Laat et al., 2008) that allows proteins bound to the enhancer to make contact with those bound to the promoter, resulting in transcription of the gene (Figure 1.2). What is known is that individual genes can in fact have multiple

Figure 1.2: Simplified gene regulation.

(A) Many genes contain 3 major components: coding sequence, promoter(s), and enhancer(s), often with identifiable DNA sequence components. (B) Different proteins recognize specific DNA sequences at the promoter and enhancer. (C) The interaction between promoter and enhancer proteins leads to the transcription of the coding sequence into mRNA, which will be subsequently translated into protein. (D) Often a single gene can have several enhancers, each of which can act as a switch, to promote transcription of that gene in a different cell type in the body, as well as at different developmental times


A
GCAGCTG
TATA
ATG
TGA
enhancer
promoter
protein coding region
B
A
A
P
ATG
TGA
enhancer
promoter
protein coding region
C
mRNA
P
gene = "ON"
D
enhancers
promoter
coding region
E1
E2
E3
hypothetical gene
cell type 3
cell type 2
cell type 1
adult
gene expression pattern

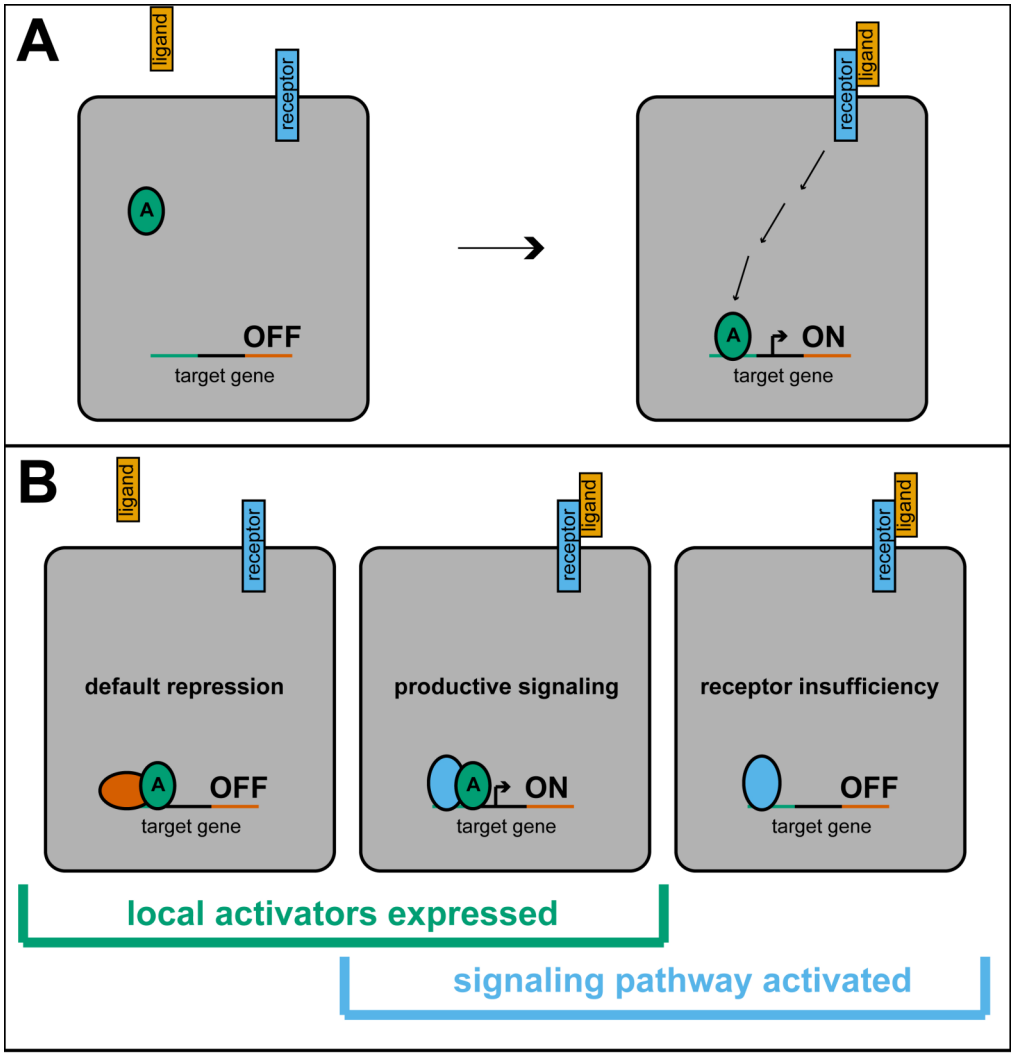
enhancers, each with a different spatial and temporal specificity. It is clear that these so-called “junk” components in our genome are the major players that determine how each cell in our body activates only the genes it needs to do its own specific job so that our bodies not only work properly, but also are put together properly.

Talk to me

As important as enhancers are to gene expression, it is equally important that they are properly regulated. During development, the cells of an organism are in constant communication with one another. Each cell has a battery of proteins, called receptors, which extend through the cell membrane and recognize specific molecules (known as ligands—they can include proteins or other molecules present in the extracellular environment). Certain classes of receptors, when bound to ligand outside the cell, initiate a chain of events inside the cell that leads to transcription factors binding to enhancers and turning on certain genes. This chain of events, from receptor activation to gene expression, is known as a signaling pathway (Figure 1.3 A).

The number of signaling pathways that control the development of complex organisms with anywhere from thousands to trillions of cells turns out to be astonishingly small. The theme that has emerged is that a single signaling pathway is used repeatedly in different contexts to lead to a huge diversity of outcomes specific to each situation (Barolo and Posakony, 2002; Affolter et al., 2008). What this means is that a single signaling pathway can activate a different set of genes in different situations. The different outcomes are often the result of the influence of different transcription factors (known as local activators) present in the different contexts that cooperate with the signaling pathway to active only the genes specific for that cell type (Figure 1.3 B). The integration of the two inputs can occur at the level of the enhancer: these types of target genes would have enhancers containing binding sites for both the specific local activator and the signaling pathway transcription factor. In this situation the signaling pathway on its own would be unable to activate these target gene enhancers in the absence of local activators. Similarly, the local activators would be unable to turn on the enhancers without the activation of the signaling pathway due to “default repression” of

Figure 1.3: Combinatorial input of cell type specific transcription factors and extra-cellular signaling pathways regulating gene expression.
(A) Often, gene expression is regulated through extra-cellular signaling pathways. In this example, the binding an extra-cellular ligand molecule to a transmembrane receptor protein causes a series of events inside the cell (arrows) that leads to a specific activating transcription factor (green oval "A") binding to DNA sequences in the enhancer or a target gene, activating transcription of that gene (turning it "ON"). (B) Strict specificity of activation can be achieved through the cooperative action of cell type specific (a.k.a. "local activators") and signaling pathway-regulated transcription factors on target gene enhancers. In the absence of ligand, the transcription factors downstream of the signaling pathway maintain an inactive state (a.k.a. "default repression"), even in the presence of local activators. Conversely, without the presence of local activators, the activated signaling pathway is incapable of inducing target gene expression on its own (a.k.a. "receptor insufficiency"). Thus the specific target genes are only activated in the subset of cells for which both the signaling pathway has been activated by ligand and the cell type specific transcription factors are expressed (adapted from Barolo *et al.*, 2002).


A
ligand
receptor
A
OFF
target gene
receptor
ligand
A
ON
target gene
B
ligand
receptor
default repression
A
OFF
target gene
receptor
ligand
productive signaling
A
ON
target gene
receptor
ligand
receptor insufficiency
OFF
target gene
local activators expressed
signaling pathway activated

signaling pathway targets. For many signaling pathways, the terminal transcription factor is able to not only activate genes in response to ligand binding, but also keep them off in the absence of ligand (i.e. the default state—no ligand—is to be repressed). This dual ability often involves groups of proteins binding to the transcription factor, and the activation of the signaling pathway leads to a change in the type of protein complex bound from one of repression to one of activation. Thus, target genes can only be activated when two conditions are met: the presence of the correct local activator and the signal-induced activation complex bound to the terminal transcription factor of that signaling pathway. These types of context-specific, signal-dependent changes in gene expression, occurring at multiple times and in many places, help to ensure the correct types of cells are born in the correct place in the body with the correct genes activated for the job that cell has to do in the organism.

A Notch Above the Rest

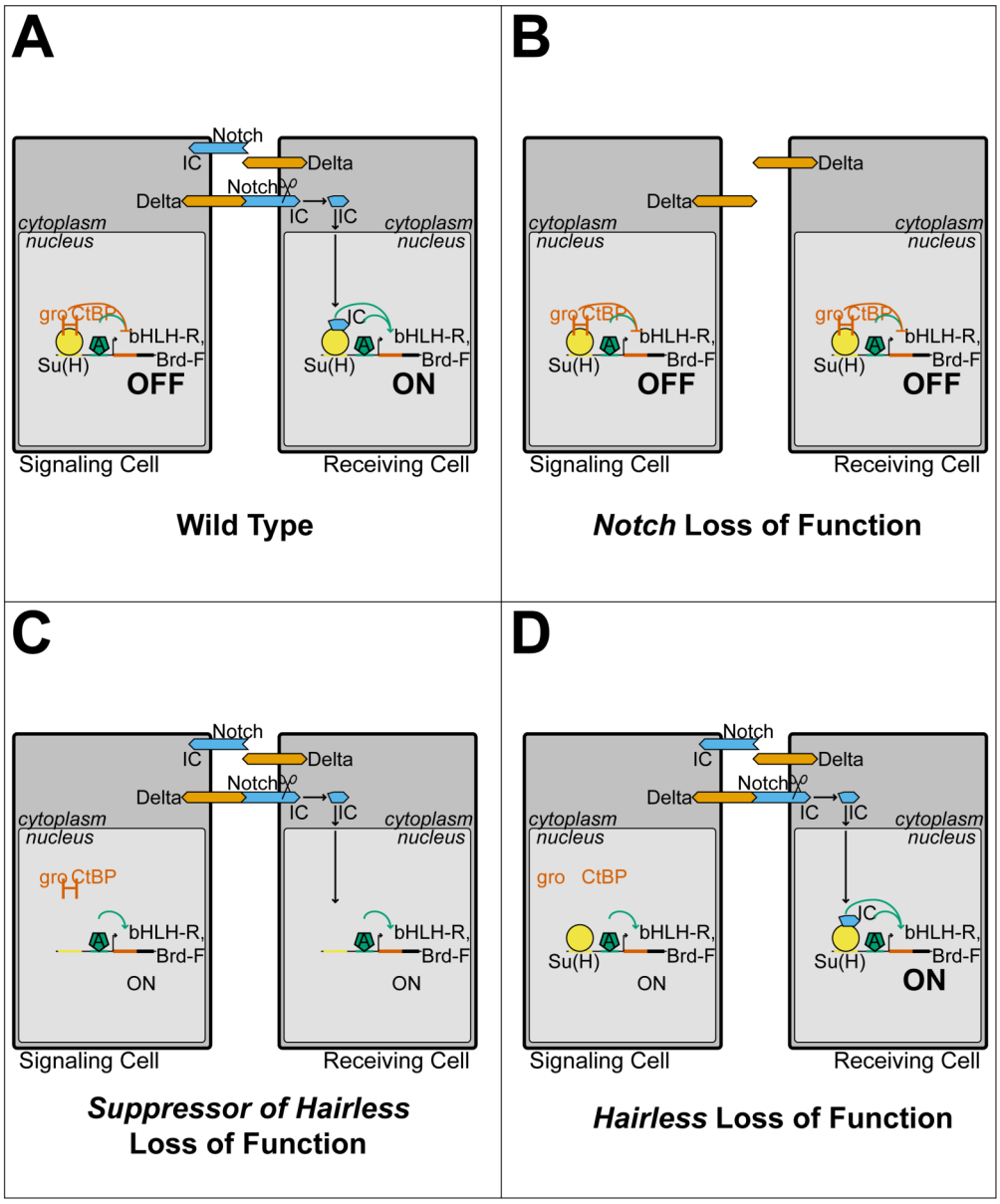
The Notch signaling pathway is just one of the handful of developmental signaling pathways, and is used over and again during

the development of multicellular organisms (Callahan and Egan, 2004; Weng and Aster, 2004; Cornell and Eisen, 2005; Louvi and Artavanis-Tsakonas, 2006; Hurlbut et al., 2007). Notch signaling facilitates communication between adjacent cells, both of which express membrane-spanning proteins extending to both the extracellular and intracellular space (Bray, 2006; Ehebauer et al., 2006; Fiuza and Arias, 2007; Kovall, 2008; Poellinger and Lendahl, 2008; Fortini, 2009; Kopan and Ilagan, 2009). The signaling facilitates a binary cell fate decision (one fate or another fate); in most cases, each cell has the potential to adopt either of two fates, and the Notch signaling event directs one cell to adopt a particular fate (called the "Notch-dependent" fate), while the other cell adopts the other, "Notch-independent" fate. Every multicellular organism utilizes Notch signaling at some stage during development, but the pathway has been most well studied in the fruit fly, *Drosophila melanogaster* (Figure 1.5 A).

In *Drosophila*, the receptor is encoded by the *Notch* gene (*N*); the ligands are encoded by the genes *Delta* (*Dl*) and *Serrate* (*Ser*); and gene *Suppressor of Hairless* [*Su(H)*] encodes the terminal transcription factor (Figure 1.4). In the absence of a signaling event, Su(H) protein

is contained within a repression complex containing Hairless (H) adaptor protein and the co-repressor proteins Groucho (Gro) and C-terminal Binding Protein (CtBP) (Barolo et al., 2002; Castro et al., 2005). This complex prevents the expression of N target genes, including those found in the regions of the genome known as the *Enhancer of split Complex* [*E(spl)-C*] and the *Bearded Complex* (*Brd-C*). These two regions contain two types of proteins, basic-helix-loop-helix (bHLH) transcription factors [found in the *E(spl)-C*] and a group of small proteins referred to as Bearded proteins or Bearded Family Members [BFMs, found both in the *E(spl)-C* and *Brd-C*] (Klambt et al., 1989; Delidakis and Artavanis-Tsakonas, 1992; Knust et al., 1992; Schrons et al., 1992; Jennings et al., 1994; Bailey and Posakony, 1995; Lai et al., 2000). Like Hairless, the bHLH transcription factors also contain specific amino acids that bind to the Groucho-repressor (Paroush et al., 1994), allowing them to specifically turn off their target genes (unlike H, they contain a domain of positively charged amino acids that allow them to bind to a specific DNA sequence—H does not bind DNA); thus, they are often referred to as bHLH-repressors (bHLH-Rs).

Figure 1.4: The Notch signaling pathway promotes binary cell fate specification signaling between adjacent cells.
(A) The Notch signaling pathway operates between two adjacent cells, with one designated as the Notch signal sending cell and the other the Notch signal receiving cell. In the absence of a signal, the transcription factor Suppressor of Hairless [Su(H)] is part of a repressive complex containing the co-repressor proteins Hairless (H), Groucho (gro) and C-terminal Binding Protein (CtBP). In such a state, target genes [e.g. the basic-helix-loop-helix transcription factors and Bearded family proteins found in the Enhancer of Split and Bearded complexes (bHLH-R and Brd-F, respectively)] are actively prevented from being transcribed. Upon binding of the Notch receptor with one if its ligand molecules on the sending cell (in this case Delta) the receptor is cleaved at the cell membrane, releasing the intracellular portion of the protein (IC) to be transported to the nucleus where it facilitates the exchange of the repressive complex for an activation complex, initiating strong target gene expression. (B) In the absence of the Notch receptor, the repressive complex is maintained and target genes are prevented from being transcribed. (C) In a *Su(H)* mutant, with neither the repressive or activating complexes forming, low level, de-repressed target gene expression is detectable. (D) In a *H* mutant, target genes are de-repressed only in the signaling cell but activated normally in the signal receiving cell.


A
Notch
IC
Delta
Notch
Delta
IC
IC
cytoplasm
nucleus
cytoplasm
nucleus
gro CtBP
bHLH-R,
Brd-F
Su(H)
OFF
IC
bHLH-R,
Brd-F
Su(H)
ON
Signaling Cell
Receiving Cell
Wild Type
B
Delta
Delta
cytoplasm
nucleus
cytoplasm
nucleus
gro CtBP
bHLH-R,
Brd-F
Su(H)
OFF
gro CtBP
bHLH-R,
Brd-F
Su(H)
OFF
Signaling Cell
Receiving Cell
Notch Loss of Function
C
Notch
IC
Delta
Notch
Delta
IC
IC
cytoplasm
nucleus
cytoplasm
nucleus
gro CtBP
bHLH-R,
Brd-F
ON
bHLH-R,
Brd-F
ON
Signaling Cell
Receiving Cell
Suppressor of Hairless
Loss of Function
D
Notch
IC
Delta
Delta
Notch
IC
IC
cytoplasm
nucleus
cytoplasm
nucleus
gro CtBP
bHLH-R,
Brd-F
Su(H)
ON
IC
bHLH-R,
Brd-F
Su(H)
ON
Signaling Cell
Receiving Cell
Hairless Loss of Function

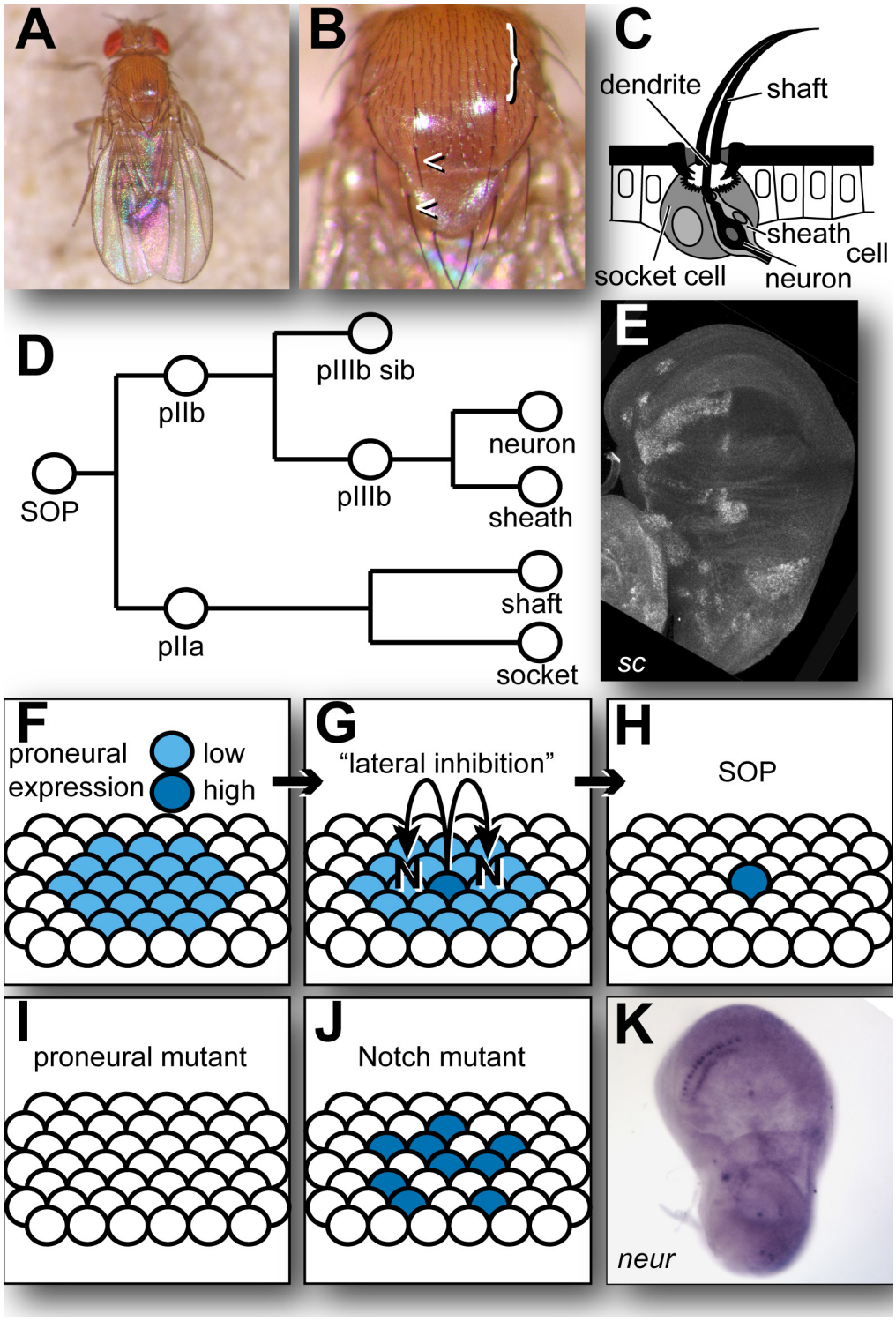
Binding of ligand by the N receptor initiates a series of protein cleavage events that results in the intracellular portion of the protein (NICD) being released from the membrane and transported to the nucleus. Once there NICD combines with the protein Mastermind (Mam) to form an activation complex bound to Su(H) (Helms et al., 1999; Schuldt and Brand, 1999; Petcherski and Kimble, 2000; Wu et al., 2000), displacing the Hairless repression complex (Morel et al., 2001; Castro et al., 2005) and—with the aid of local activators—initiating the transcription of bHLH-R and BFM genes, as well as additional context-specific targets.

Standard Operating Procedure for a Specific Context

Many of the details of the Notch signaling pathway, as well as general themes of signal-dependent changes in gene expression described on these pages to this point have been illuminated in the context of *Drosophila* mechanosensory bristle development. Flies are covered by a large number of bristles over different parts of the body that are used to gather a variety of sensory information, including taste (feet, proboscis, labial palps), smell (antennae), and touch (legs, head,

thorax). The bristles on the dorsal head and thorax, called the notum (Figure 1.5 B), are comprised of the larger macrochaetes and the smaller microchaetes, both of which display a similar composition and developmental program. Each bristle structure contains four functional cellular components (Figure 1.5 C): the shaft structure is connected to a bipolar neuron, aided by a supporting sheath cell. The internal cells are surrounded by a large socket cell that secretes an external cuticular structure as well as creates a large internal extracellular cavity with a particular ion content necessary for the mechanotransduction of the neuron (Grunert and Gnatzy, 1987; Keil, 1997; Jarman, 2002). Each of these four distinct cell types descends from a single precursor cell, the sensory organ precursor cell (SOP, Figure 1.5D) (Hartenstein and Posakony, 1989; Gho et al., 1999; Reddy and Rodrigues, 1999b). The bristles in the adult are arranged in a highly reproducible, stereotypic pattern; similarly, the SOPs arise in predictable locations in the developing thoracic epithelium (Figure 1.5 K) (Bryant, 1975; Hartenstein and Posakony, 1989; Huang et al., 1991). Cells in the epithelium are endowed with the capacity to become an SOP by the expression of a set of local activators, Achaete

Figure 1.5: The mechanosensory organs of *Drosophila melanogaster* require the combinatorial function of the Notch signaling pathway mediated lateral inhibition and the proneural proteins, Achaete and Scute.
(A) The commonly-used model organism, the fruit fly *Drosophila melanogaster*. (B) The fly is covered by a large number of mechanosensory bristle structures. The most well-studied are those bristles on the dorsal thorax, or notum, which consist of the large macrochaetes (arrowheads) and the smaller microchaetes (bracket). (C) Each bristle in the adult is composed of four different cell types, the socket, shaft, sheath and neuron, the latter of which has a ciliary dendrite extending to and maintaining contact with the shaft structure. (D) These four cells all descend from a single sensory organ precursor, or SOP, through a series of asymmetric cell divisions. The potential to adopt an SOP fate is conferred by the activity of the so-called "proneural" transcription factors, Achaete and Scute. (E,F) The proneurals, with *scute* mRNA in the wing imaginal disc (E) shown as an example, are expressed in discrete patches of cells known as proneural clusters (PNCs,F). (G) SOPs arise from these equivalence groups and signal to their neighbors via the Notch signaling pathway, a process referred to as lateral inhibition. (H) Lateral inhibition causes down-regulation of proneural activity in non-SOP PNC cells leaving the SOP with strong proneural expression. (I) In the absence of proneural activity (e.g. $sc^{10\text{-}1}$), SOPs fail to become specified; no bristles develop. (J) In a Notch mutant, PNC cells fail to be inhibited and supernumerary SOPs develop. (K) mRNA from the *neuralized* gene exemplifies an SOP expression in the larval wing imaginal disc.


A
B
C
dendrite
shaft
sheath cell
socket cell
neuron
D
pIIIb sib
pIIb
neuron
SOP
pIIIb
sheath
shaft
pIIa
socket
E
sc
F
proneural expression
low
high
G
"lateral inhibition"
N
N
H
SOP
I
proneural mutant
J
Notch mutant
K
neur

and Scute, known as proneural transcription factors for their ability to confer a neural fate. Ac and Sc are expressed in discrete groups of cells in the epithelium called proneural clusters (PNCs, Figure 1.5 E,F), from which the SOPs arise and can be detected by a further increase in the level of proneural protein expression. In the absence of proneural function (e.g. a sc^{10-1} mutant, Figure 1.5 I), no SOPs are specified and subsequently no bristles develop (Ruiz-Gomez and Modolell, 1987; Romani et al., 1989; Jimenez and Campos-Ortega, 1990; Rodriguez et al., 1990; Simpson, 1990). Conversely, ectopic expression of the proneurals outside of the normal PNC domain results in production of supernumerary bristles. While the proneurals establish a field of potential SOPs, N signaling is used to inhibit the majority of PNC cells from adopting an SOP fate, a process known as lateral inhibition (Figure 1.5 G, H) (Cabrera, 1990; Simpson, 1990). Without N pathway function, extra SOPs develop from within the PNC field at the expense of epidermal cell fates, while activation of the N pathway within the SOP can extinguish its fate.

This overview of SOP specification and the role of two major components, the N pathway and the proneural proteins is quite elegant

for both its simplicity and for the ease with which it fits into the present discussion of signal-regulated gene expression. Indeed, gene regulatory analyses have led to a straightforward model integrating these core components (Bailey and Posakony, 1995; Heitzler et al., 1996). However, for reasons to be elaborated later, this model is dissatisfying if not lacking, and is the main focus of Chapter 2.

NOTch AGAIN?!

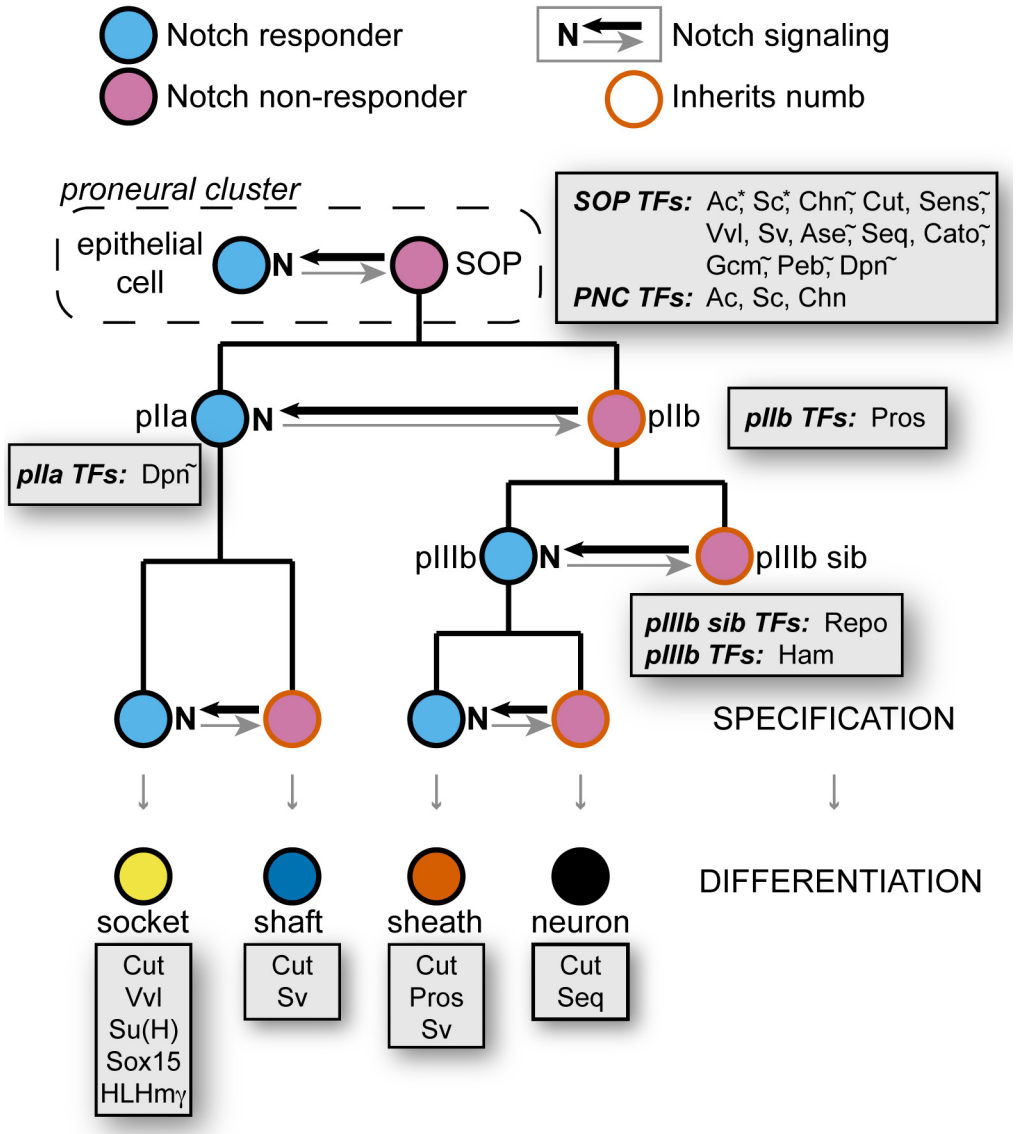
As fruitful as the studies of N signaling in SOP specification have been, the biology of events subsequent to the SOP have also contributed to the current understanding. After the SOP is specified, it gives way to a series of four binary cell fate specification events that ultimately leads to the creation of the four cell types mentioned above. Each of these four cell fate decision events involves input from the N pathway (Figure 1.6), and interruption of N pathway function can affect the outcome of each developmental decision in the lineage (Parks et al., 1997).

One of the outstanding features of N signaling in the context of the lineage is that it represents four cases for which a signaling event

occurs, yet each event results in the specification of an entirely different fate. Based upon what has been presented thus far, the prediction would be that different N signaling outcomes, even in the restricted groups of cells descended from the SOP, would be the result of the function of different local activators present in the different cells. Do date there has been little experimental effort that has been directed toward this specific hypothesis, however.

Support for such a local activator-mediated distinction in the mechanosensory organ lineage comes from the current available knowledge of the expression patterns of different transcription factors in the lineage that have been identified, as collected in Figure 1.6 and Table 1.1. Each of the four post-mitotic cell types has a unique transcription factor profile upon differentiation. While this makes it convenient to distinguish between the different cellular identities using antibodies directed against these different factors, it also suggests the importance of specific profiles to activate complements of target genes unique to each cell type. Intriguingly, several of the factors contributing to the individual cellular profiles begin their expression prior to the birth of that cell, sometimes even as early as the SOP. Moreover, even

Figure 1.6: Iterative Notch signaling and differential transcription factor expression within the mechanosensory organ lineage. Beginning in the proneural cluster, the Notch pathway mediates a series of binary cell fate decisions as the cells of the mechanosensory organ are specified. After the SOP divides, each daughter cell pair is contains both a Notch responsive sister and a Notch non-responsive sister, the latter of which is due to the asymmetric inheritance of Notch-antagonistic Numb protein. Furthermore, the identities of the cells throughout the elaboration of the lineage can be distinguished by the combinatorial expression of a number of transcription factors. Thematic in the divisions is the inheritance of the transcription factor complement present in the preceding precursor cell, often combined with the novel expression of an additional transcription factor in one of the sister cells but not the other. Unless otherwise indicated, if a cell is noted to express a particular transcription factor then it is also inherited by its immediate daughters. This continues through the specification of the post-mitotic cells. However, as the cells differentiate, some factors are down-regulated in specific daughters, ultimately leading to the transcription factor profiles indicated for each cell type.



*Expression of Ac and Sc is downregulated in the SOP before it divides (Skeath and Carroll, 1991).
~Adult mechanosensory organ lineage expression of these factors have not been thoroughly analyzed.

Table 1.1: Summary of transcription factors expressed in the adult mechanosensory organ lineage. Accompanies Figure 1.6.

Transcription Factor	Description	Reference
Achaete (Ac)	Expressed in proneural clusters. Required for SOP fate. Misexpression can induce ectopic SOPs.	Romani et al., 1989.
Asense (Ase)	Expressed in SOP and lineage. Loss of function causes SOP loss and defects in post-mitotic cells. Misexpression can induce ectopic SOPs.	Dominguez and Campuzano, 1993.
Cousin of Atonal (Cato)	Expressed in SOP and lineage. Required for neuronal differentiation. Misexpression can induce ectopic SOPs.	Goulding et al., 2000.
Charlatan (Chn)	Expressed in proneural clusters and SOP lineage. Required for Ac, Sc SOP up-regulation. Misexpression can induce ectopic SOPs and Ac, Sc up-regulation.	Escudero et al., 2005.
Cut (Ct)	Expressed in SOP and lineage. Required for external versus internal (chordotonal) sense organ identity. Misexpression can convert chordotonal organs to external sensory organs.	Bodmer et al., 1987.
Deadpan (Dpn)	Expressed in SOP and pIIa; excluded from pIIb.	Andrews et al., 2009; Bier et al., 1992.
Glial Cells Missing (Gcm)	Expressed in SOP and some lineage. Required for pIIIb sib differentiation.	Fichelson and Gho, 2003.
Hamlet (Ham)	Expressed in pIIIb and inherited by sheath and neuron. Loss of function causes switch from single to multiple dendritic neuron	Moore et al., 2002.
HLHmγ (mγ)	Expressed in a subset of proneural clusters; expressed in socket cells.	Nellesen et al., 1999.
Pebbled (Peb)	Expressed in SOPs; a common marker, though not functionally studied in the lineage. In trachea, involved in cell morphogenesis and differentiation.	Wilk et al., 2001.
Prospero (Pros)	Expressed in pIIb and descendents. Required for internal cell (sheath, neuron, pIIIb sib) identity. When misexpressed can cause external to internal fate conversion.	Reddy and Rodrigues, 1999.
Reversed Polarity (Repo)	Expressed in pIIIb sib. Required for differentiation markers in embryo.	Halter et al., 1995.
Scute (Sc)	Expressed in proneural clusters. Required for SOP fate. Misexpression can induce ectopic SOPs.	Romani et al., 1989
Senseless (Sens)	Expressed in SOP and lineage. Required for Ac, Sc SOP up-regulation, and for external cell (socket, shaft) fate. Misexpression can induce ectopic SOPs.	Nolo et al., 2000; Jafar-Nejad et al., 2006.
Sequoia (Seq)	Expressed in SOP and lineage. Down-regulated from socket, shaft, and sheath over time to remain in neuron. Required for neuron differentiation.	Andrews et al., 2009; Brenman et al., 2001.
Shaven (Sv)	Expressed in SOP and lineage; refined to only shaft and sheath. Required for shaft differentiation; misexpression in socket can induce partial shaft differentiation implementation.	Kavaler et al., 1999. This study.
Sox15	Expressed in socket cells. See Chapter 3 and 4 for more exciting details!	Cremazy et al., 2001.
Suppressor of Hairless [Su(H)]	Expressed ubiquitously at low levels; high levels specific to socket cells. Required for socket mechanosensory function.	Barolo et al., 2000.
Ventral Veins Lacking (Vvl)	Expressed in SOP and lineage; refined to only socket cell. Required for mitotic pattern and post-mitotic differentiation.	Inbal et al., 2003.

though a number of transcription factors become newly transcribed in the SOP following lateral inhibition, loss of any these factors does not affect the SOP versus epithelial fate decision, suggesting that these factors are expressed in the SOP but functionally required in its daughters. Indeed, loss of lineage transcription factors instead can affect both the fates as well as the proper differentiation of the cells descended from the SOP (Bodmer et al., 1987; Kavaler et al., 1999; Reddy and Rodrigues, 1999a; Moore et al., 2002; Inbal et al., 2003; Jafar-Nejad et al., 2006).

You Know Who You Are, Now Do Your Job Like You're Supposed To!

Every terminally differentiated cell type in the body has a specific function, and executes a genetic program to perform those functions. It is believed that this differentiation program is the result of a number of cell type specific gene activation events, all dependent upon the initial cell fate determination. Very few studies, however have elucidated the exact steps linking the cell fate transcriptional changes to those controlling the differentiation of that cell; it is even unclear how often they can be functionally separated. Little is known about how

deep the connection between the differentiation program and the cell fate machinery is, but one study regarding the socket cell (Barolo et al., 2000) differentiation program suggests that some aspects of the program are rather closely linked.

As discussed above, the terminal transcription factor of the N pathway is Su(H). Su(H) protein is expressed at low basal levels in a fairly ubiquitous cellular pattern in the fly, with the exception of a single cell type—the socket cell. In this cell, Su(H) is expressed at very high levels immediately after the socket cell is specified. This high-level socket expression is controlled by an enhancer found in the region downstream of the *Su(H)* coding sequence, and contains a number of binding sites for Su(H) protein itself. Without this region, referred to as the auto-regulatory socket enhancer (ASE), the specification of the socket cell is unaffected and even the cuticular socket structure appear completely normal. The function of this activity of Su(H) is revealed in electrophysiological assays; both the transepithelial potential (TEP; a measurement of the ionic capacity of the receptor lymph cavity created and maintained by the socket cell) and the mechanoreceptor current (MRC; a measure of the current initiated by the neuron in response to

moving the bristle shaft structure, assaying the ability of the organ to respond to a mechanical stimulus) are abolished in these flies. These data implicate the same transcription factor in both the cell fate specification mechanism as well as aspects of the differentiation program and yet these two activities are entirely separable (Figure 1.7). Furthermore, even though this activity of Su(H) is initiated from the birth of the socket cell, there are many aspects of socket differentiation that are not controlled by ASE function and thus must be regulated by other factors in the socket cell. A putative such factor is the Sox family transcription factor Sox15, and its role in the socket cell is the subject of Chapter 3.

This picture of socket differentiation can be contrasted with what is known about the differentiation of the socket cell sister, the shaft. This cell contains high levels of the Pax family transcription factor, Shaven (Sv). Sv is initially expressed in the SOP and the rest of the lineage, persisting even into the post-mitotic cells. As these cells are differentiating, however, Sv becomes highly expressed in both the shaft and sheath cells, at the same time being downregulated in the socket and neuron (Kavaler et al., 1999). This type of expression

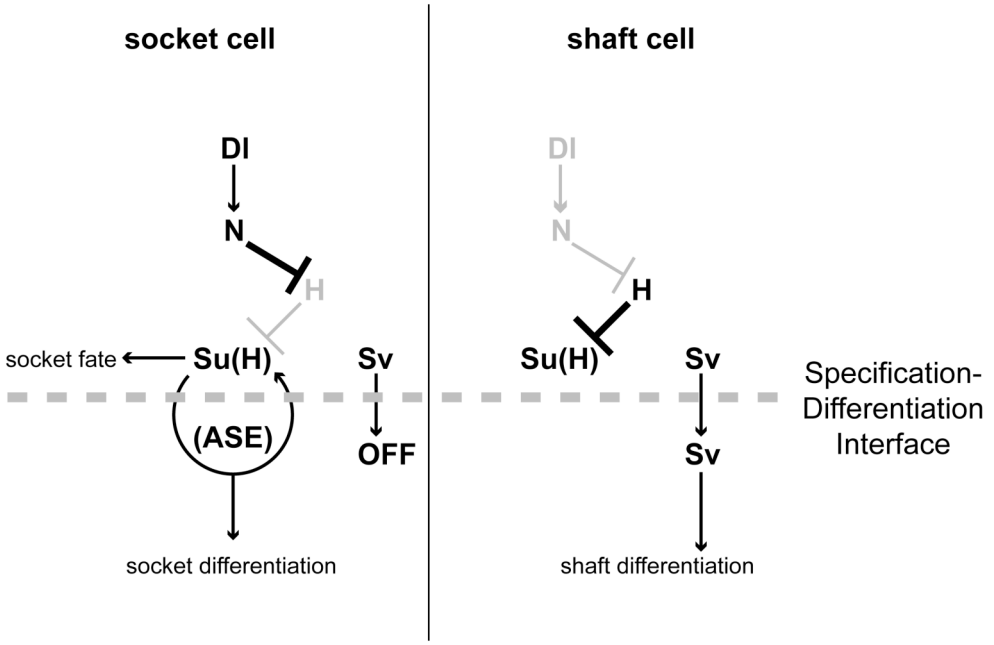


Figure 1.7: The specification-differentiation interface. Differentiation of cell types can be functionally separable from their specification, as illustrated by the example of the socket and shaft sister cells. At birth, both cells contain basal levels of Su(H) and inherit low level Shaven expression. As the Notch independent cell fate, the shaft cell fails to alleviate the Hairless co-repressor complex, leading to the default shaft cell fate. Shaven expression is maintained and activates the shaft-specific differentiation program. In the other cell, Notch is activated, alleviating Hairless-mediated default repression and activating the socket cell fate. Su(H) activates a Notch-independent auto-regulatory loop through its Auto-regulatory Socket Enhancer (ASE). ASE function is required for socket differentiation, but not for fate specification. Also as a consequence of socket fate specification, Shaven fails to be maintained and cannot activate the shaft differentiation program in the socket cell. The separable fate and differentiation functions of Su(H) exemplify a gene operating at the specification-differentiation interface.

pattern is particularly intriguing since it does not track with the N signaling activities of these two cells, and is in part addressed in Chapter 4. In the shaft cell, loss of Sv activity greatly diminishes the shaft structure, suggesting Sv is a high-level regulator of the shaft cell differentiation program. Consistent with this model, ectopic expression of Sv in the socket cell is sufficient to induce the socket cell to project shaft structures (S. Barolo and J. Posakony, unpublished data). This also indicates that the socket cell is in part a permissive environment to the activities of Sv function. The socket cell does not completely convert to a shaft upon ectopic Sv expression, however, and can even wrap around the endogenous shaft like normal, with the induced shaft structures projecting on either side. This would suggest aspects of socket differentiation can still proceed normally in spite of this perturbation; it remains to be seen if the programs downstream of Sv are completely independent of socket differentiation or if ectopic Sv expression can negatively affect transcription of socket-specific programs.

Bring on the Enhancers

The experimental analyses of individual enhancers such as is presented in this body of work are key to understanding how each cell in the body utilizes the information contained in the genome to correctly and reproducibly construct a functioning intact organism. Such studies are inherently biased toward gleaning as much information as possible about a single node in an enormously complex regulatory network. Compiling these types of studies to generate a network map of a complete organism would require a multitude of labs several decades at the current rate. Therefore, over the last few years more effort has been put forth to adopting more genome-wide regulatory analyses that take a more encompassing yet cursory approach to deducing regulatory information (2004; Celniker et al., 2009). These types of efforts will generate predictions of thousands of putative enhancers throughout the genomes of nematodes, flies, mice, and humans. Determining how successful these prediction approaches are requires a way of testing putative enhancers in as rapid a manner as possible. To this end, Chapter 5 discusses a potential approach to rapid enhancer validation that could apply to organisms as diverse as flies and humans. Certainly, the next few

years hold exciting potential for us to more completely sift through all the “junk” and be truly able to listen to all that the genome is telling us.

References

Affolter, M. *et al.* (2008). Signal-induced repression: the exception or the rule in developmental signaling? Dev Cell *15*, 11-22.

Andrews, H. K. *et al.* (2009). Sequoia regulates cell fate decisions in the external sensory organs of adult Drosophila. EMBO Rep *10*, 636-641.

Bailey, A. M., and Posakony, J. W. (1995). Suppressor of hairless directly activates transcription of enhancer of split complex genes in response to Notch receptor activity. Genes Dev *9*, 2609-2622.

Barolo, S., and Posakony, J. W. (2002). Three habits of highly effective signaling pathways: principles of transcriptional control by developmental cell signaling. Genes Dev *16*, 1167-1181.

Barolo, S. *et al.* (2002). Default repression and Notch signaling: Hairless acts as an adaptor to recruit the corepressors Groucho and dCtBP to Suppressor of Hairless. Genes Dev *16*, 1964-1976.

Barolo, S. *et al.* (2000). A notch-independent activity of suppressor of hairless is required for normal mechanoreceptor physiology. Cell *103*, 957-969.

Bier, E. *et al.* (1992). deadpan, an essential pan-neural gene in Drosophila, encodes a helix-loop-helix protein similar to the hairy gene product. Genes Dev *6*, 2137-2151.

Bird, C. P. *et al.* (2006). Functional variation and evolution of non-coding DNA. Curr Opin Genet Dev *16*, 559-564.

Bodmer, R. *et al.* (1987). Transformation of sensory organs by mutations of the cut locus of D. melanogaster. Cell *51*, 293-307.

Bray, S. J. (2006). Notch signalling: a simple pathway becomes complex. Nat Rev Mol Cell Biol *7*, 678-689.

Brenman, J. E. *et al.* (2001). Sequoia, a tramtrack-related zinc finger protein, functions as a pan-neural regulator for dendrite and axon morphogenesis in Drosophila. Dev Cell *1*, 667-677.

Bryant, P. J. (1975). Pattern formation in the imaginal wing disc of Drosophila melanogaster: fate map, regeneration and duplication. J Exp Zool *193*, 49-77.

Cabrera, C. V. (1990). Lateral inhibition and cell fate during neurogenesis in Drosophila: the interactions between scute, Notch and Delta. Development *110*, 733-742.

Callahan, R., and Egan, S. E. (2004). Notch signaling in mammary development and oncogenesis. J Mammary Gland Biol Neoplasia *9*, 145-163.

Castro, B. *et al.* (2005). Lateral inhibition in proneural clusters: cis-regulatory logic and default repression by Suppressor of Hairless. Development *132*, 3333-3344.

Celniker, S. E. *et al.* (2009). Unlocking the secrets of the genome. Nature *459*, 927-930.

Cleary, J. D. *et al.* (2005). Antimycotic drug discovery in the age of genomics. Am J Pharmacogenomics *5*, 365-386.

Cobb, J. *et al.* (2008). Searching for functional genetic variants in non-coding DNA. Clin Exp Pharmacol Physiol *35*, 372-375.

Cornell, R. A., and Eisen, J. S. (2005). Notch in the pathway: the roles of Notch signaling in neural crest development. Semin Cell Dev Biol *16*, 663-672.

Cremazy, F. *et al.* (2001). Genome-wide analysis of Sox genes in Drosophila melanogaster. Mech Dev *109*, 371-375.

Crenshaw, E. B. r. *et al.* (1989). Cell-specific expression of the prolactin gene in transgenic mice is controlled by synergistic

interactions between promoter and enhancer elements. Genes Dev *3*, 959-972.

Daly, D. C. *et al.* (2001). Plant systematics in the age of genomics. Plant Physiol *127*, 1328-1333.

de Laat, W., and Grosveld, F. (2003). Spatial organization of gene expression: the active chromatin hub. Chromosome Res *11*, 447-459.

de Laat, W. *et al.* (2008). Three-dimensional organization of gene expression in erythroid cells. Curr Top Dev Biol *82*, 117-139.

Delidakis, C., and Artavanis-Tsakonas, S. (1992). The Enhancer of split [E(spl)] locus of Drosophila encodes seven independent helix-loop-helix proteins. Proc Natl Acad Sci U S A *89*, 8731-8735.

Dominguez, M., and Campuzano, S. (1993). asense, a member of the Drosophila achaete-scute complex, is a proneural and neural differentiation gene. EMBO J *12*, 2049-2060.

Ehebauer, M. *et al.* (2006). Notch signaling pathway. Sci STKE *2006*, cm7.

Escudero, L. M. *et al.* (2005). Charlatan, a Zn-finger transcription factor, establishes a novel level of regulation of the proneural achaete/scute genes of Drosophila. Development *132*, 1211-1222.

Farre, D. *et al.* (2007). Housekeeping genes tend to show reduced upstream sequence conservation. Genome Biol *8*, R140.

Fichelson, P., and Gho, M. (2003). The glial cell undergoes apoptosis in the microchaete lineage of Drosophila. Development *130*, 123-133.

Fiuza, U. M., and Arias, A. M. (2007). Cell and molecular biology of Notch. J Endocrinol *194*, 459-474.

Fortini, M. E. (2009). Notch signaling: the core pathway and its posttranslational regulation. Dev Cell *16*, 633-647.

Gelvin, S. B. (2009). Agrobacterium in the genomics age. Plant Physiol

Gho, M. *et al.* (1999). Revisiting the Drosophila microchaete lineage: a novel intrinsically asymmetric cell division generates a glial cell. Development *126*, 3573-3584.

Goulding, S. E. *et al.* (2000). cato encodes a basic helix-loop-helix transcription factor implicated in the correct differentiation of Drosophila sense organs. Dev Biol *221*, 120-131.

Grunert, U., and Gnatzy, W. (1987). K+ and Ca++ in the receptor lymph of arthropod cuticular mechanoreceptors. J Comp Physiol A *161*, 329-333.

Hahn, M. W. (2007). Detecting natural selection on cis-regulatory DNA. Genetica *129*, 7-18.

Halter, D. A. *et al.* (1995). The homeobox gene repo is required for the differentiation and maintenance of glia function in the embryonic nervous system of Drosophila melanogaster. Development *121*, 317-332.

Hartenstein, V., and Posakony, J. W. (1989). Development of adult sensilla on the wing and notum of Drosophila melanogaster. Development *107*, 389-405.

Heitzler, P. *et al.* (1996). Genes of the Enhancer of split and achaete-scute complexes are required for a regulatory loop between Notch and Delta during lateral signalling in Drosophila. Development *122*, 161-171.

Helms, W. *et al.* (1999). Engineered truncations in the Drosophila mastermind protein disrupt Notch pathway function. Dev Biol *215*, 358-374.

Huang, F. *et al.* (1991). The emergence of sense organs in the wing disc of Drosophila. Development *111*, 1087-1095.

Hurlbut, G. D. *et al.* (2007). Crossing paths with Notch in the hyper-network. Curr Opin Cell Biol *19*, 166-175.

Inbal, A. *et al.* (2003). Multiple roles for u-turn/ventral veinless in the development of Drosophila PNS. Development *130*, 2467-2478.

Jafar-Nejad, H. *et al.* (2006). Senseless and Daughterless confer neuronal identity to epithelial cells in the Drosophila wing margin. Development *133*, 1683-1692.

Jarman, A. P. (2002). Studies of mechanosensation using the fly. Hum Mol Genet *11*, 1215-1218.

Jennings, B. *et al.* (1994). The Notch signalling pathway is required for Enhancer of split bHLH protein expression during neurogenesis in the Drosophila embryo. Development *120*, 3537-3548.

Jimenez, F., and Campos-Ortega, J. A. (1990). Defective neuroblast commitment in mutants of the achaete-scute complex and adjacent genes of D. melanogaster. Neuron *5*, 81-89.

Juven-Gershon, T. *et al.* (2006). Perspectives on the RNA polymerase II core promoter. Biochem Soc Trans *34*, 1047-1050.

Juven-Gershon, T. *et al.* (2008). The RNA polymerase II core promoter - the gateway to transcription. Curr Opin Cell Biol *20*, 253-259.

Kavaler, J. *et al.* (1999). An essential role for the Drosophila Pax2 homolog in the differentiation of adult sensory organs. Development *126*, 2261-2272.

Keil, T. A. (1997). Functional morphology of insect mechanoreceptors. Microsc Res Tech *39*, 506-531.

Klambt, C. *et al.* (1989). Closely related transcripts encoded by the neurogenic gene complex enhancer of split of Drosophila melanogaster. EMBO J *8*, 203-210.

Knust, E. *et al.* (1992). Seven genes of the Enhancer of split complex of Drosophila melanogaster encode helix-loop-helix proteins. Genetics *132*, 505-518.

Kopan, R., and Ilagan, M. X. (2009). The canonical Notch signaling pathway: unfolding the activation mechanism. Cell *137*, 216-233.

Kovall, R. A. (2008). More complicated than it looks: assembly of Notch pathway transcription complexes. Oncogene *27*, 5099-5109.

Lai, E. C. *et al.* (2000). Antagonism of notch signaling activity by members of a novel protein family encoded by the bearded and enhancer of split gene complexes. Development *127*, 291-306.

Louvi, A., and Artavanis-Tsakonas, S. (2006). Notch signalling in vertebrate neural development. Nat Rev Neurosci *7*, 93-102.

Maniatis, T. *et al.* (1987). Regulation of inducible and tissue-specific gene expression. Science *236*, 1237-1245.

Moore, A. W. *et al.* (2002). hamlet, a binary genetic switch between single- and multiple- dendrite neuron morphology. Science *297*, 1355-1358.

Morel, V. *et al.* (2001). Transcriptional repression by suppressor of hairless involves the binding of a hairless-dCtBP complex in Drosophila. Curr Biol *11*, 789-792.

Nellesen, D. T. *et al.* (1999). Discrete enhancer elements mediate selective responsiveness of enhancer of split complex genes to common transcriptional activators. Dev Biol *213*, 33-53.

Nolo, R. *et al.* (2000). Senseless, a Zn finger transcription factor, is necessary and sufficient for sensory organ development in Drosophila. Cell *102*, 349-362.

Parks, A. L. *et al.* (1997). The dynamics of neurogenic signalling underlying bristle development in Drosophila melanogaster. Mech Dev *63*, 61-74.

Paroush, Z. *et al.* (1994). Groucho is required for Drosophila neurogenesis, segmentation, and sex determination and interacts directly with hairy-related bHLH proteins. Cell *79*, 805-815.

Petcherski, A. G., and Kimble, J. (2000). Mastermind is a putative activator for Notch. Curr Biol *10*, R471-3.

Poellinger, L., and Lendahl, U. (2008). Modulating Notch signaling by pathway-intrinsic and pathway-extrinsic mechanisms. Curr Opin Genet Dev *18*, 449-454.

Ponting, C. P., and Lunter, G. (2006). Signatures of adaptive evolution within human non-coding sequence. Hum Mol Genet *15 Spec No 2*, R170-5.

Reddy, G. V., and Rodrigues, V. (1999a). Sibling cell fate in the Drosophila adult external sense organ lineage is specified by prospero function, which is regulated by Numb and Notch. Development *126*, 2083-2092.

——— (1999b). A glial cell arises from an additional division within the mechanosensory lineage during development of the microchaete on the Drosophila notum. Development *126*, 4617-4622.

Rodriguez, I. *et al.* (1990). Competence to develop sensory organs is temporally and spatially regulated in Drosophila epidermal primordia. EMBO J *9*, 3583-3592.

Romani, S. *et al.* (1989). Expression of achaete and scute genes in Drosophila imaginal discs and their function in sensory organ development. Genes Dev *3*, 997-1007.

Ruiz-Gomez, M., and Modolell, J. (1987). Deletion analysis of the achaete-scute locus of Drosophila melanogaster. Genes Dev *1*, 1238-1246.

Ryder, E., and Russell, S. (2003). Transposable elements as tools for genomics and genetics in Drosophila. Brief Funct Genomic Proteomic *2*, 57-71.

Schrons, H. *et al.* (1992). The Enhancer of split complex and adjacent genes in the 96F region of Drosophila melanogaster are required for segregation of neural and epidermal progenitor cells. Genetics *132*, 481-503.

Schuldt, A. J., and Brand, A. H. (1999). Mastermind acts downstream of notch to specify neuronal cell fates in the Drosophila central nervous system. Dev Biol *205*, 287-295.

Sharp, R. R. *et al.* (2004). Shaping science policy in the age of genomics. Nat Rev Genet *5*, 311-316.

Simpson, P. (1990). Lateral inhibition and the development of the sensory bristles of the adult peripheral nervous system of Drosophila. Development *109*, 509-519.

Smith, G. (2004). The Genomics Age: How DNA Technology is Transforming the Way We Live and Who We Are AMACOM).

Smith, H. (1972). Evolution of Genetic Systems: Brookhaven Symposium in Biology, Number 23 Routledge).

Subramanian, G. *et al.* (2001). Implications of the human genome for understanding human biology and medicine. JAMA *286*, 2296-2307.

(2004). The ENCODE (ENCyclopedia Of DNA Elements) Project. Science *306*, 636-640.

Thomas, M. C., and Chiang, C. M. (2006). The general transcription machinery and general cofactors. Crit Rev Biochem Mol Biol *41*, 105-178.

Venkatesh, B. *et al.* (2006). Ancient noncoding elements conserved in the human genome. Science *314*, 1892.

Venter, J. C. *et al.* (2001). The sequence of the human genome. Science *291*, 1304-1351.

Walsh, B. (2001). Quantitative genetics in the age of genomics. Theor Popul Biol *59*, 175-184.

Weng, A. P., and Aster, J. C. (2004). Multiple niches for Notch in cancer: context is everything. Curr Opin Genet Dev *14*, 48-54.

Wilk, R. *et al.* (2000). The hindsight gene is required for epithelial maintenance and differentiation of the tracheal system in Drosophila. Dev Biol *219*, 183-196.

Wu, L. *et al.* (2000). MAML1, a human homologue of Drosophila mastermind, is a transcriptional co-activator for NOTCH receptors. Nat Genet *26*, 484-489.

Xu, J. (2006). Microbial ecology in the age of genomics and metagenomics: concepts, tools, and recent advances. Mol Ecol *15*, 1713-1731.

CHAPTER TWO:

Cell Non-Autonomous Auto-Regulation of *neuralized* During *Drosophila* Mechanosensory Organ Precursor Lateral Inhibition

Abstract

A key developmental process is the progressive restriction of cell fate potential by which extrinsic signals influence positional information to activate specific gene regulatory programs in subsets of cells with equivalent potential. In understanding this process, an important focus is the ensuing changes in gene expression and how those changes can be directly linked with the two initial inputs. One of the most characterized models for such a developmental process is the specification of mechanoreceptor sensory organ precursor cells (SOPs) from proneural clusters (PNCs) in the *Drosophila melanogaster* peripheral nervous system. Expression of the proneural transcription factors Achaete and Scute confer SOP potential upon groups of adjacent cells. Only a subset of these cells achieves this potential due to the activity of the Notch signaling pathway, through a process known as lateral inhibition. Here, we present evidence of transcriptional regulation of the N pathway component *neuralized* (*neur*) as a key integration point for positional cues endowed by the proneural proteins and fate potential restriction mediated by N signaling. As such, we propose a model where the proneural proteins activate *neur*

expression, activating the Delta ligand in the SOP. The resulting N signaling event cell non-autonomously activates expression of the protein products of the genes from the *E(spl) Complex* and *Bearded Complex* in non-SOP cells. These proteins then provide a two-pronged mechanism to both inhibit *neur* expression and function in the non-SOP cells, enabling *neur* expression amplification in the SOP that permits the rapid creation of a sharp *neur* expression demarcation and SOP establishment.

Introduction

Subdivision as a developmental strategy

As most complex, multicellular organisms navigate the process of development, rarely do highly specialized cell types randomly emerge from an otherwise carte blanche field of undifferentiated tissue. Rather, it is more often the case that beginning at early stages in development, an organism is progressively divided into domains of distinct identity (Desai and McConnell, 2000; Pinaudeau et al., 2000; Guo et al., 2003). Each of these domains then becomes further subdivided into groups of cells that influence their local environment to coordinate the development of both cells in their immediate area, as well as the organism as a whole.

As subdivision occurs, it is accompanied by changes in gene expression. Most organisms use a common toolkit of regulatory genes that confer spatial identity upon the different organismal subdomains (Meyerowitz, 1999). Cells within a domain of common gene expression often respond to spatial cues in the local environment to deploy additional regulatory genes that promote the activation of discrete developmental programs that alter gene expression in further

restricted patterns. In the organisms for which we have the most experience, both the genes that define organismal subdomains are known, as well as the signals that are used to further specialize these regions. The current challenge is to describe how the integration of these two inputs happens mechanistically to establish gene expression differences that lead to progressively specialized cell types in the adult.

One of the types of genes that are part of the common toolkit is the proneural genes, which help subdivide the developing epithelium into domains of neural competence (or pro-neural). These genes are utilized in organisms as diverse as anemone, anchovies, and anteaters; any animal that has a nervous system (Ruiz-Gomez and Modolell, 1987; Romani et al., 1989; Rodriguez et al., 1990; Grens et al., 1995). The proneurals encode for transcriptional factors with an architecture of a basic-helix-loop-helix (bHLH) type that allows them to heterodimerize and bind to DNA, directly activating expression of target genes (Murre et al., 1989; Cabrera and Alonso, 1991; Van Doren et al., 1991; Jarman et al., 1994).

The context in which the proneurals have been most extensively investigated has been in the developing peripheral nervous system

(PNS) of the fruit fly, *Drosophila melanogaster*. The fly PNS deploys the activity of primarily two proneural proteins encoded by the genes *achaete* and *scute*, which not only reside adjacent to one another in the genome but also share a number of cis-regulatory elements (Leyns et al., 1989). Ac and Sc proteins are expressed in patches of cells, termed proneural clusters (PNCs), which are endowed with the competence to adopt a neural fate (Romani et al., 1989; Rodriguez et al., 1990; Skeath and Carroll, 1991). Competence does not equal commitment, however, and only a small subset of cells realizes their neural potential.

Notch-mediated lateral inhibition subdivides the field of neural competence

Much of the structures in the adult fruit fly come from groups of cells, called imaginal discs, which are set aside during the larval phases and which undergo rapid phases of growth and differentiation during the pupal metamorphosis. The adult fly is covered by an array of sensory bristles structures that are wired to the nervous system to alert the fly to a number of external cues. The dorsal thorax of the fly,

or notum, contains an ordered pattern of mechanosensory bristles, whose development originates with PNCs in the larval wing imaginal disc. The cells that compose each bristle structure all descend from a single sensory organ precursor cell (SOP) born from a PNC.

The number of SOPs that are specified from a single PNC in the wing imaginal disc is regulated through the activity of the Notch (N) signaling pathway (Cabrera, 1990). The ligand Delta on the surface of the cell that will become the SOP contacts Notch receptor molecules on surrounding cells, leading to a series of proteolytic cleavage events of the receptor that results in the release of the Notch intracellular domain (NICD) from the membrane. NICD proceeds to enter the nucleus where it forms a complex with Mastermind (Mam) and Suppressor of Hairless [Su(H)] to activate target gene expression (Bray, 2006); in the proneural clusters these include genes encoded within the *Enhancer of split Complex* [*E(spl)-C*] and *Bearded Complex* (*Brd-C*) (Klambt et al., 1989; Delidakis and Artavanis-Tsakonas, 1992; Knust et al., 1992; Leviten et al., 1997; Lai et al., 2000). Within these two complexes reside genes encoding bHLH-repressor transcription factors (bHLH-R) and small proteins known as Bearded Family

Members (BFMs). Activation of bHLH-R in the non-SOP cells functions to repress expression of target genes, among which are believed to be the proneurals themselves (Jennings et al., 1994; Singson et al., 1994; Bailey and Posakony, 1995; Heitzler et al., 1996). In this manner, Notch signaling restricts the neural competence of most of the cells of the cluster, allowing only a small subset to become an SOP, a process known as lateral inhibition (Cabrera, 1990; Simpson, 1990).

Feedback integration of proneural and Notch inputs

As the SOP is specified during lateral inhibition, it is believed to develop two key features: both strong signaling and weak responding capacities. It is currently unclear how distinct these two capabilities are, although some studies suggest that they may both be strongly linked to the ligand Delta. An activity known as cis-inhibition has been proposed to contribute to the latter, by which Delta molecules make contact with Notch molecules on the same cell and thus prevent activation by Delta on neighboring cells (de Celis and Bray, 1997; Klein et al., 1997; Micchelli et al., 1997; Sakamoto et al., 2002; Li and Baker,

2004). The strong signaling capability of the SOP has been also been proposed to rely on Delta expression, by way of a proneural feedback loop. Kunisch *et al* identified a number of proneural binding sites upstream of the *Delta* promoter that were required to activate gene expression in embryonic neuroblasts (Kunisch et al., 1994). This result helped to promote a model by which Delta in the SOP signals through Notch to activate the bHLH-Rs in non-SOP cells in which they inhibit proneural activity. Without such repression in the SOP, the proneurals undergo self-stimulation and accumulate at higher levels. As these SOP proneural levels increase, they bind to these sequences upstream of *Dl* and promote elevated Dl protein levels in the SOP, leading to both increased N activating capacity and resistance to N signal receipt (Bailey and Posakony, 1995; Heitzler et al., 1996).

Problems with a Delta*-centric model*

Over the years, however, a number of key experiments have cast doubt on the existence of exactly such a mechanism. While subtle differences *Dl* transcript accumulation in the embryonic neurogenic field can reveal the locations of neuroblasts, no such

expression differences have been seen for the adult proneural clusters when assaying either *Dl* mRNA or protein levels . Moreover, even if bulk mRNA and protein levels remain constant, the model would predict that if the SOP has greatly elevated levels of proneural proteins which are activating *Dl* expression, differences between the SOP and non-SOPs should be observed when examining nascent transcript; this has turned out not to be the case (Parks et al., 1997). Finally, in another crucial test of this model, when *Dl* is detached from differences in proneural activity by independently driving ubiquitous *Dl* expression, both neuroblast (Seugnet et al., 1997) and SOP (Pitsouli and Delidakis, 2005) specification appear to largely operate without consequence. In total, these data would support the production of an alternative model for proneural feedback upon N signaling during lateral inhibition.

The attractiveness of cis-regulatory control of neuralized

An alternative focus for proneural regulation of N signaling in the SOP may be the gene *neuralized* (*neur*), which encodes and E3 ubiquitin ligase (Deblandre et al., 2001; Lai et al., 2001; Yeh et al., 2001). In the N pathway, Neur functions in the sending cell where it

mono-ubiquitinates Dl, promoting ligand endocytosis, an event required for proper N signaling. Loss of function mutants in *neur* result in a strong defect in lateral inhibition, leading to the production of supernumerary SOPs (Lehmann et al., 1981; Lehmann et al., 1983; Yeh et al., 2000). *neur* expression is remarkably specific: it accumulates at high levels in the SOP and is virtually undetectable in non-SOPs (Boulianne et al., 1991; Boulianne et al., 1993). This expression requires proneural function, as transcript is abolished in a sc^{10-1} background (Reeves and Posakony, 2005). Thus, understanding how *neur* achieves such a strict specificity of expression despite its apparent involvement in early events during lateral inhibition could be a potentially crucial step toward constructing a more accurate model for integration of proneural activity and N signaling.

With such a goal in mind, in this study we have begun to investigate the cis-regulatory mechanism for achievement of *neur* SOP specificity. We report the identification of conserved proneural and bHLH-R binding sites in the *neur* locus and reveal activities for both proteins. Using high resolution *in situ* detection, we also demonstrate expression of *neur* in the proneural cluster prior to SOP specification,

further implicating Neur function during early aspects of the lateral inhibition. Unlike has been shown for Dl, we demonstrate that persistent expression of Neur in non-SOP cells has consequences for lateral inhibition. As shown, our results support an integration model with *neur* as a central regulatory focus during SOP specification, rather than was previously proposed for *Dl*.

Materials and Methods

Fly strains and mosaic analysis

FLPase stock for generating neur MARCM clones was generously provided by Christos Delidakis (Pitsouli and Delidakis, 2005). UAS-NF-κB fluorescent conjugates were provided by Steven Wasserman (Bettencourt et al., 2004). mα>GAL4 was described previously in Castro *et al*., 2005 (Castro et al., 2005). UAS-neur, UAS-Tom was constructed by Eric Lai, and UAS-FLAGm4 by Joseph Fontana (Fontana and Posakony, 2009). All *neur* rescue construct variants were injected and transformants collected by Genetic Services, Inc. The *mα, m4* deletion line was constructed by Joseph Fontana through two independent homologous recombination events as described in Rong and Golic (Rong and Golic, 2001). Mosaic analyses using the FLP/FRT and MARCM variant have been described (Golic and Lindquist, 1989; Golic, 1991; Xu and Rubin, 1993; Lee et al., 2000; Pitsouli and Delidakis, 2005).

Reporter constructs

Reporter constructs for neur4DWT and mutants as well as NRS1 and NRS2 were cloned into pH-Stinger (Barolo et al., 2000) or pH-RedStinger (Barolo et al., 2004). Mutations were generated by overlap extension PCR (Ho et al., 1989). At least three independent transformant lines were analyzed before a representative line was selected for all further analysis. Constructs were injected using standard transformation techniques (Rubin and Spradling, 1982), with w^{1118} as the recipient strain. NRS1 subfragments were initially cloned and analyzed using the pMemor system (see Chapter 5); wild type and mutant versions of neur1B were generated also and cloned into pH-Stinger-attB and injected using the ΦC31 integrase system (Bischof et al., 2007) into the docking site VK00037 (Venken et al., 2006).

Rescue constructs

neurRC-WT-pCaSPer was constructed by cloning a 20 kb Eag I fragment cut from the BAC clone BACR09F04 with a Eag I, Bsu 36I double digest into the pCaSPer variant pNotCaSPer cut with Not I. 4DRm and 4Ddel versions were constructed by generating mutations in and subcloning a 4 kb Not I , Sac II into neurRC-WT-pCaSPer cut

with Not I and Sac II. Rescue constructs were injected using standard methods (Rubin and Spradling, 1982) by Genetic Services, Inc. using w^{1118} as the recipient strain.

neurRC-WT-P[acman] constructs were generated by BACR09F04-mediated gap repair of attB-P[acman]-AmpR via recombineering, as described (Venken et al., 2006). These constructs were subsequently injected into the docking sites attP40 and attP16 (Markstein et al., 2008) by Genetic Service, Inc., using the ΦC31 integrase system (Bischof et al., 2007). Mutant and tagged variants of this starting construct were generated by recombineering using galK-mediated selection (Warming et al., 2005), and injected into the attP40 docking site. neurRC-WT-GFP was integrated into attP40, attP2, attP16, and VK00037 (Groth et al., 2004; Venken et al., 2006; Markstein et al., 2008); neurRC-WT-V5 and neurRC-4D,1BRm-GFP were integrated into the VK00037 docking site.

Gene diagrams and alignment cartoons

Gene diagrams and alignment cartoons were constructed using the latest version of GenePalette, http://www.genepalette.org, (Rebeiz and Posakony, 2004), and were edited in Adobe Illustrator.

In situ hybridization

Single probe in situ hybridizations were performed as previously described (O'Neill and Bier, 1994; Lai et al., 2000; Reeves and Posakony, 2005; Miller et al., 2009). Quantification of in situ signal area for the neurRC-4D,1BRm experiment was performed using the ImageJ software, taking the average of 9 discs for the WT and 21 discs for the Rm constructs. Statistical significance was assayed by an ANOVA test. Multiplex fluorescent in situ hybridizations in third instar wing imaginal discs were performed basically as described (Kosman et al., 2004); anti-hapten antibodies were used at a 1:5000 dilution in 1X PBS + 0.1% Triton X-100 (PBT) to reduce background, without using a block solution. Probes were constructed by cloning an intronic DNA fragment into pGEM-T, linearizing, and transcribing RNA using the T7 RNA polymerase following the Kosman protocol. The following probes were used: DNP-*sca*, DIG-*neur*, BIO-*neur*, DIG-*Dl*, BIO-*CG32150*,

DNP-*hb*, BIO-*hb*. Images were captured as described below, adjusting the gain to maximally reveal any coincidence between *neur* probes.

Immunohistochemistry

With the exception of GFP antibody staining, immunohistochemistry was performed essentially as described (Miller et al., 2009). Discs from neurRC-WT-GFP discs also included a blocking step after fixation in 0.3% milk in PBT. Blocking was done overnight at 4°C, with primary antibodies added the next morning, also in the milk blocking solution. Secondaries for this stain were added in PBT only. The following antibodies were used: guinea pig anti-Sens (generously provided by Hugo Bellen), 1:2000; mouse anti-Cut [Developmental Studies Hybridoma Bank, University of Iowa (DSHB)], 1:100; mouse anti-Hnt (a.k.a. pebbled), DSHB, 1:100; rabbit anti-GFP (Invitrogen), 1:500. All secondaries used were of the AlexaFluor varieties from Invitrogen and included anti-guinea pig-Alexa555 conjugate, anti-mouse-Alexa555 conjugate, and anti-mouse-Alexa647 conjugate. Secondaries were always used in staining at a 1:1000 dilution in PBT.

Electro Mobility Shift Assays (EMSAs)

chn coding sequence was cloned by Nick Reeves into the pGEX-5X-2 vector. Optimum conditions for purification were determined as described (Mercado-Pimentel et al., 2002), in the presence of 100 μM ZnSO4. 30 bp complimentary oligos were γ-^{32}P end-labeled and annealed, then used in gel shift assays as described (Escudero et al., 2005).

Preparation of adult cuticle (light microscopy, SEM)

Preserved adult notum cuticle for light microscopy and whole adult flies for scanning electron microscopy were prepared as described (Bang et al., 1991; Miller et al., 2009).

Bristle/SOP phenotypic analysis

For analysis of ectopic GFP reporter expression due to cis-mutations, 30 third instar larvae were dissected and all discs with discernible DC and SC positions were analyzed, noting the presence of any cell expressing nuclear GFP but not Sens at only these positions.

Scoring included aSC, pSC, aDC, and pDC positions that either had Sens expression as well as unspecified aSC and aDC positions that did not have Sens expression.

For scoring of extra or missing bristles, all macrochaete positions on the dorsal head and thorax were analyzed, and each position documented as either missing or extra for each fly, over a total of 25 males and 25 females unless otherwise noted. Statistical significance was determined by a pairwise ANOVA.

Confocal microscopy

Description of confocal microscopy has been described (Miller et al., 2009). Images of fluorescent in situ hybridizations were collected as series of 1 micron sections; antibody stains were collected at low magnification as 2 micron sections with high magnification images as 1 micron sections. For the collection of z-sections to generated the cross sectional view in Figure 2.7E, we shorted the distance to 0.75 micron sections. Images were collected using the confocal software and cropped with Adobe Photoshop and combined into figures using Adobe Illustrator.

Primers

Oligonucleotide sequences for all constructs are available upon request.

Results

Identification of a neuralized *SOP Enhancer*

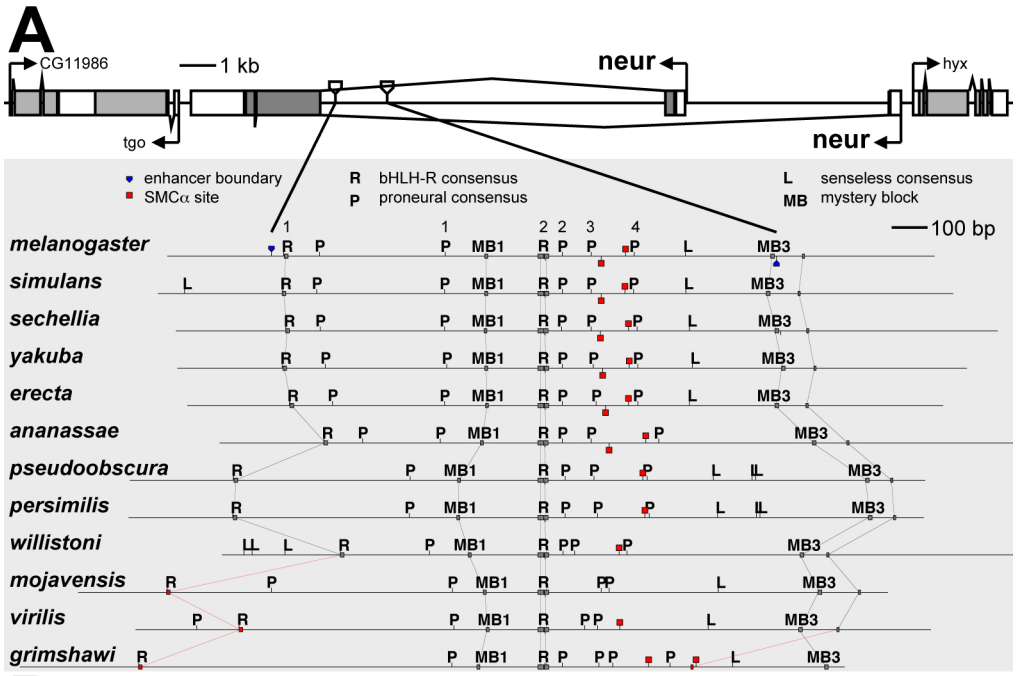
Initial studies to locate the *neuralized* SOP enhancer restricted focus to a 5.4 kb region in the 3' end of the 2nd intron (Barolo, S and J. W. Posakony, unpublished data). This expression was revealed by a 1.4 kb enhancer construct in the most 3' end of this sequence. Under control of this enhancer, GFP is expressed in third instar wing imaginal discs in the majority of positions co-staining for the SOP marker Senseless, with the exception of incomplete wing margin SOP expression. GFP is also express in the embryo, as well as the pupal notum in a manner recapitulating aspects of the endogenous gene expression (Figure 2.2A, F).

The microchaete SOPs in the notum are confined to a series of rows of proneural gene expression. SOPs within a single row are specified coordinately, and individual rows develop in a particular order (Usui and Kimura, 1993). In the notum, the 1.4 kb fragment (henceforth referred to as 4DWT) is first detected in the developing microchaete rows around 12 hrs APF, mirroring Sens expression. Intriguingly, GFP expression can repeatedly be observed at a position

in a row which does not express Sens, but which appears properly spaced from another Sens-positive SOP (Figure 2.2F). Furthermore, these Sens-negative, GFP-positive positions exhibit often two adjacent cells both of whose nuclear diameter is less than the Sens-positive SOPs. Huang *et al*. had documented similar instances in the analysis of the order of macrochaete SOP emergence in the wing imaginal disc using A101-LacZ (Huang et al., 1991). This observation had been proposed to potentially represent cases of A101 expression in a subset of PNC cells prior to specification of a single SOP. Consistent with this hypothesis, we observe Sens-negative, GFP-positive cells neither adjacent to Sens-positive pIIa:pIIb pairs, nor in such SOP-like positions within microchaete rows after the majority of SOPs have proceeded to division.

As a further crude test of the ability of this regulatory sequence to recapitulate endogenous *neur* expression, tested enhancer dependency on the function of the proneural proteins. In a sc^{10-1} hemizygous background, GFP fails to express, other than in the Atonal-dependent cluster in the anterior ventral hinge and in the ectopically expressing trachea (Figure 2.1 C,D). This pattern is

Figure 2.1: The *neur* 4D enhancer module directs SOP expression downstream of proneural activity.
(A) Diagram of the *neur* locus indicating neighboring genes, both *neur* transcription start sites, and the location of the 4D enhancer module. Below the locus diagram is an expansion of the 4D region, and an alignment of the orthologous regions from the eleven other sequenced Drosophila genomes. Indicated on the alignment are sequences fitting the consensuses for binding sites for the *E(spl)-C* bHLH-R proteins [R], the proneural proteins [P], and Senseless [L]; also indicated are sequences of unknown identity "mystery blocks" [MB], and SMCα sites. Sequences identical over at least 8 bp between all the orthologues are connected by a line in each sequence. bHLH-R and proneural sites displayed in (B) are numbered as such above the Drosophila melanogaster sequence. Local inversions are represented in red. (B) Aligned sequences of the conserved "words" in the 4D enhancer, presented in the same linear order as they occur in (A). Sequences contained in binding site consensuses have grey background box, and local inversions are represented in red. (C) neur4DWT>GFP is expressed in an SOP pattern in the third instar wing imaginal disc. (D) Expression is dependent upon *ac* and *sc* function, as no GFP expression is observed in SOPs in a $sc^{10\text{-}1}$ mutant, with the anterior ventral atonal-dependent cluster the exception.



B

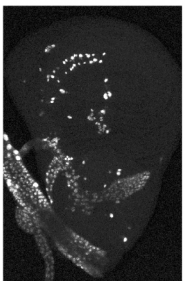
```
                bHLH-R site 1          PN site 1                MB1                      MB2                      bHLH-R site 2
mel ACAGGGCACGTGACTGTATT TATCAGCAGCTGCTGACA ATGGCGACAGGTTTCTGC CTGGTGATTACCATGTTGATATGCGGCGCGTGGCACTGAA
sim ACAGGGCACGTGACTGGATT TATCAGCAGCTGCTGACA ATGGCGACAGGTTTCTGC CTGGTGATTACCATGTTGATATGCGGCGCGTGGCACTGAA
sec AAAGGGCACGTGACTGGATT AATTAGCAGCTGCTGACA ATGTCGACAGGTTTCTGC ATGGTGATTACCATGTTGATATGCGGCGCGTGGCACTGAA
yak ACAGGGCACGTGACTGGATG TATCAGCAGCTGCTGACA ATGGCGACAGGTTTCTGC ATGGTGATTACCATGTTGATATGCGGCGCGTGGCACTGAA
ere ACAGGGCACGTGACTGGATT CATCAGCAGCTGCTGACA ATGGCGACAGGTTTCTGC CTGGTGATTACCATGTTGATATGCGGCGCGTGGCACTGGA
ana TCCAGGCACGTGACTAGATC CCTCAGCAGCTGCTGACA ATGGCGACAGGTTTCTCC AGGTTGATTACCATGTTGATATGCGGCGCGTGGCACTTGG
pse CAGCAGCACGTGACTTCGAT TTGCAACAGCTGCTGGCG CAGAGGACAGGTTTTTCC AAGATGATTACCATGTTGATATGCGGCGCGTGGCAGCAGC
per CAGCAGCACGTGACTTCGAT TTGCAACAGCTGCTGGCG CAGAGGACAGGTTTTTCC AAGATGATTACCATGTTGATATGCGGCGCGTGGCAGCAGC
wil AAATAGCACGTGACTTTGAT TGGCCGCACCTGCTGTGA ACTGCGACAGGTTTTTTG AAGTTGATTACCATGTTGATATGCGGCGCGTGGCACAGTG
moj GCACAGCACGTGACTTTGAT CAACAGCAGCTGCAAGAC TGGGCGACAGGTTCGAGG AGACTGATTACCATGTTGATACGCGGCGCGTGGCAGCGTC
vir GCGCAGCACGTGACTTTGAT TAACAGCAGCTGCCTGAC TTGGCGACAGGTTTCGGG GCGCCGATTACCATGTTGATATGCGGCGCGTGGCAGCGTC
gri TAAGAGCACGTGACTTTGAT GGGCAGCAGCTGTCTGAC CAGACGACAGGTTTGTGG GTCTTGATTACCATGTTGATATGCGGCGCGTGGCAGTGTC

    PN site 2                PN site 3          SMC site 2          PN site 4
mel TTGGAGCGGCATGTGGCAGCTGTCCTTT AGAAAGCAGCTGTACGGT CGCTTAGGGGTTGCTGCA AGTCTGCAGCTGTTCAAATTTCGAAAA
sim TTGGAGCGGCATGTGGCAGCTGTCCTTT AGAAAGCAGCTGTACGGC CGCTTAGGGGTTGCTGCA AGTCTGCAGCTGTTCAAATTTCGAAAA
sec TTGGAGCGGCATGTGGCAGCTGTCCTTT AGAAAGCAGCTGTACGGC CGCTTAGGGGTTGCTGCA AGTCTGCAGCTGTTCAAATTTCGAAAA
yak TTGGAGCGGCATGTGGCAGCTGTCCTTT AGAAAGCAGCTGTACGAC CGCTTAGGGGTTGCTGCA GGTCTGCAGCTGTTCAAATTTCGAAAA
ere TTGGAGCGGCATGTGGCAGCTGTCCTTT AGAAAGCAGCTGTACGGC CGCTTAGGGGTTGCTGCA AGTCTGCAGCTGTTCAAATTTCGAAAA
ana TTAAAGCGGCATGTGGCAGCTGCCTTTT AGAAAGCAGCTGTACACT GACTGAGGGGTTGTTGCT ACTCGGCAGCTGTTCAAATTTCGACAA
pse CAGCGGCGGCATGTGGCAGCTGCCACAC CAGAAGCAGCTGTACATT GTCCTAGGGGTTGGCGGC TGGCGGCAGCTGTTCAAATTTCATGAA
per CAGCGGCGGCATGTGGCAGCTGCCACAC CAGAAGCAGCTGTACATT GTCCTAGGGGTTGGCGGC TGGCGGCAGCTGTTCAAATTTCATGAA
wil TCTATGCGGCATGTGGCAGCTGCCTTCT AGAAAGCAGCTGTACGCT TCACTAGGGGTTGCTAGT CAGCGGCAGCTGTTCAAATTTTCAGAA
moj ATCAGGCGGCATGTGGCAGCTGCTCGTG GCGAGGCAGCTGCACCGA TCAACAAGGGTTGCTAGA AAACGGCAGATGTTCAAATTACGAAAA
vir ATCAGGCGGCATGTGGCAGCTGCCCGTG GCGAGGCAGCTGCAGCGA TCAACAGGGGTTGCTGGC AAACGACAGATGTTCAAATTACGAAAA
gri AGGCTGCGGCATGTGGCAGCTGCCCGTG AAAAAGCAGCTGCAGCAA TCAACAGGGGTTGCTTGC AATCGGCAGCTGTTCAAATTTCAAAAA

    sens site          MB3
mel ACACGAAATCTGTACTGT AAGTAGGAAATCACGTACTAC
sim ACACGAAATCTGTACTGT AAGTAGGAAATCACGTACTAC
sec ACACGAAATCTGTACTGT AAGTAGGAAATCACGTACTAC
yak ACACGAAATCAGTACTGT AAGTAGGAAATCACGTACTAC
ere ACACGAAATCAGTACTGT AAGTAGGAAATCACGTACTAC
ana ACACAAAATCTCGTCTCT AAGTAGGAAATCACGTCTTGG
pse AGCAGAAATCTGTACAGA AAGTAGGAAATCACGTATATT
per AGCAGAAATCTGTACAGA AAGTAGGAAATCACGTATATT
wil TTGAAAAATCTTTCGCCT AAGTAGGAAATCACGTTTTCT
moj CAGAAAAATCTGTGACTG AAGTCGGAAATCACGTTTTGC
vir CAGAAAAATCTGTGTGGT AAGTCGGAAATCACGTTTGGG
gri CAAGAAAATCTGTGGCTG AAATCGGAAATCACGTTTTTT
```

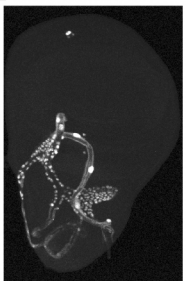
C

w^{1118}



D

sc^{10-1}



identical to *neur* transcript (Reeves and Posakony, 2005), and indicates the 4DWT enhancer requires proneural activity for activation by either a direct or indirect mechanism.

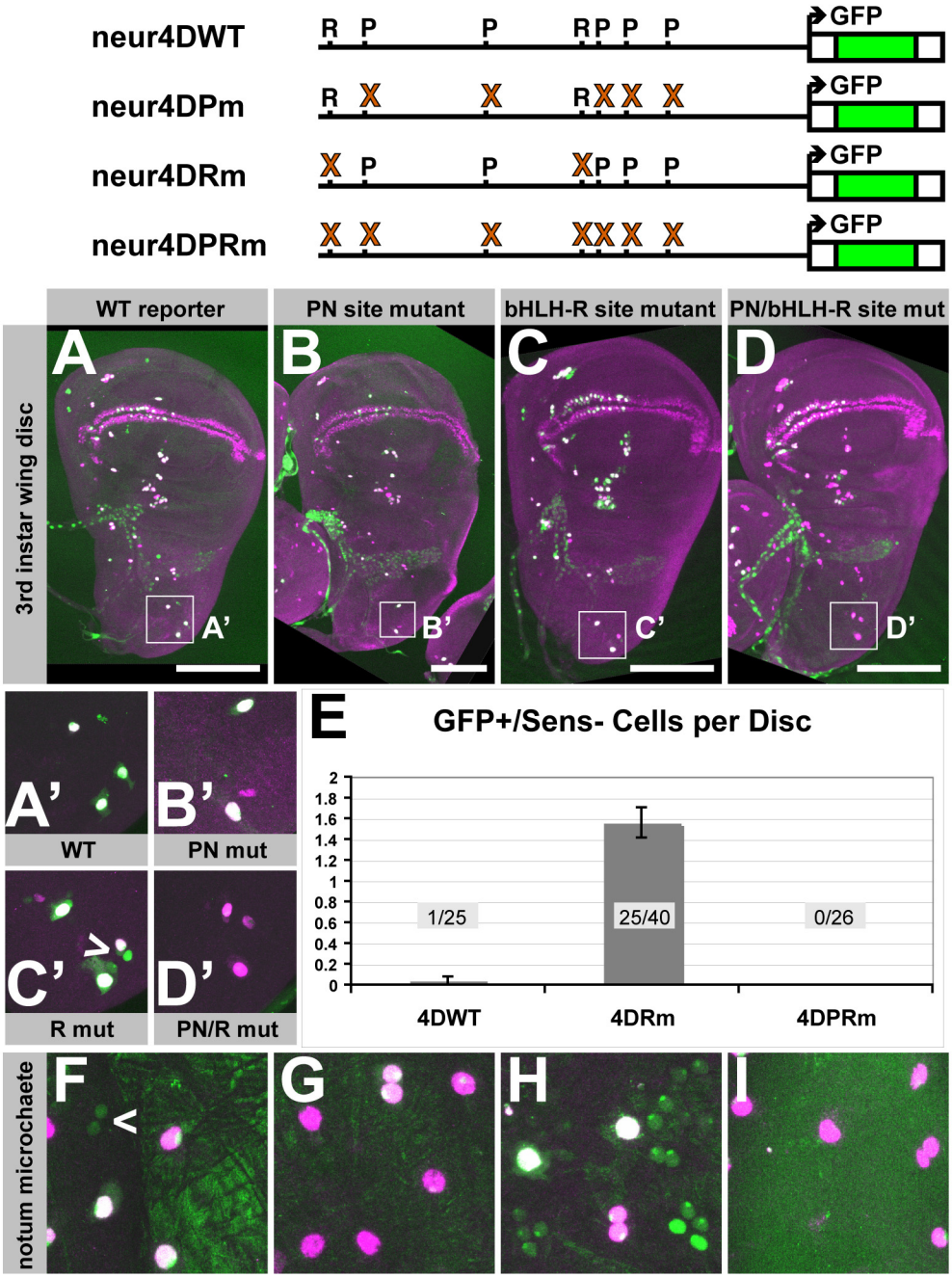
Conservation and Function of Proneural and bHLH-Repressor Binding Sites

We turned to evolutionary conservation as an approach to reveal regulatory inputs required for the 4DWT enhancer. Making use of the twelve sequenced *Drosophila* genomes, we focused on short sequences that were preserved throughout all species, spanning a distance of approximately 70 million years. Within 4DWT are five binding sites for the proneural proteins, suggesting the proneural input is direct (Figure 2.1 A, B). Four of these proneural sites are conserved throughout all twelve genomes, while the fifth appears to have only been conserved through the last common ancestor between *Drosophila melanogaster* and *Drosophila ananassae*. Also preserved throughout all twelve genomes are two binding sites for the bHLH-repressor proteins of the *E(spl)-C*, although one of these sites is inverted in *Drosophila mojavensis*, *Drosophila virilis*, and *Drosophila*

grimshawi relative to the others. To assess the requirements of these conserved binding sites, we mutated them in the context of the reporter construct.

Previous analysis of the enhancers of the *E(spl)-C* genes, which are expressed in the non-SOP cells of the PNCs, revealed the presence of a "P + S" code, with strong requirements for binding sites for both the proneural proteins and Su(H) (Singson et al., 1994; Bailey and Posakony, 1995; Nellesen et al., 1999; Castro et al., 2005). Mutation of the proneural sites abolishes enhancer activity, while mutating the Su(H) sites causes loss of activation in the non-SOPs and derepression in the SOPs in a proneural-dependent manner. In contrast, the mutation of the proneural binding sites in the 4D enhancer appears to have at best a partial requirement for these sites; 4DPm causes stochastic loss of GFP in some SOPs in the third instar larval wing imaginal disc as well as in the microchaetes of the pupal notum (Figure 2.2B, G; Figure 2.3B22). The most significant difference is seen by 24hrs after puparium formation (APF), where there is significant loss of GFP expression in the post-mitotic cells. Since we cannot be certain the lineage expression of 4DWT is due to enhancer

Figure 2.2: Functional requirements of bHLH-R and proneural binding sites in the *neur* 4D SOP enhancer.
At the top are diagrams of the reporter constructs used in this figure. Site mutation are represented by a red X. (A – D) Reporter variant expression in the third instar wing imaginal disc and (F – I) in the 12 hr APF notum. (A, F) The wild type reporter construct expresses in SOPs of the macrochaetes and microchaetes. In the notum, sometimes expression in two small nuclei in expected SOP positions that do not yet express Sens (magenta), the SOP marker (arrowhead in F). (B, G) The proneural binding site mutations cause a small reduction in GFP expression. (C, H) Mutating the bHLH-R binding sites causes ectopic GFP expression in both the wing disc (C) and notum (H). (C') Higher magnification view of the DC and SC regions, noting the extra GFP-expressing cells (arrowhead). (D, I) The bHLH-R site-dependent ectopic expression requires intact proneural binding sites. (E) Quantification of ectopic GFP expression (GFP+/Sens-) from third instar imaginal discs. Only DC and SC positions were scored.


neur4DWT
R P P R P P P GFP
neur4DPm
R X X R X X X GFP
neur4DRm
X P P X P P P GFP
neur4DPRm
X X X X X X X GFP
WT reporter
PN site mutant
bHLH-R site mutant
PN/bHLH-R site mut
3rd instar wing disc
A
B
C
D
A'
B'
C'
D'
WT
PN mut
R mut
PN/R mut
E
GFP+/Sens- Cells per Disc
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
1/25
25/40
0/26
4DWT
4DRm
4DPRm
notum microchaete
F
G
H
I

activation in the lineage or due to perdurance of the GFP from the initial SOP expression, we can similarly not distinguish the weakened GFP expression in the post-mitotic cells of the 4DPm mutant from reduced activity in the SOP or in the lineage.

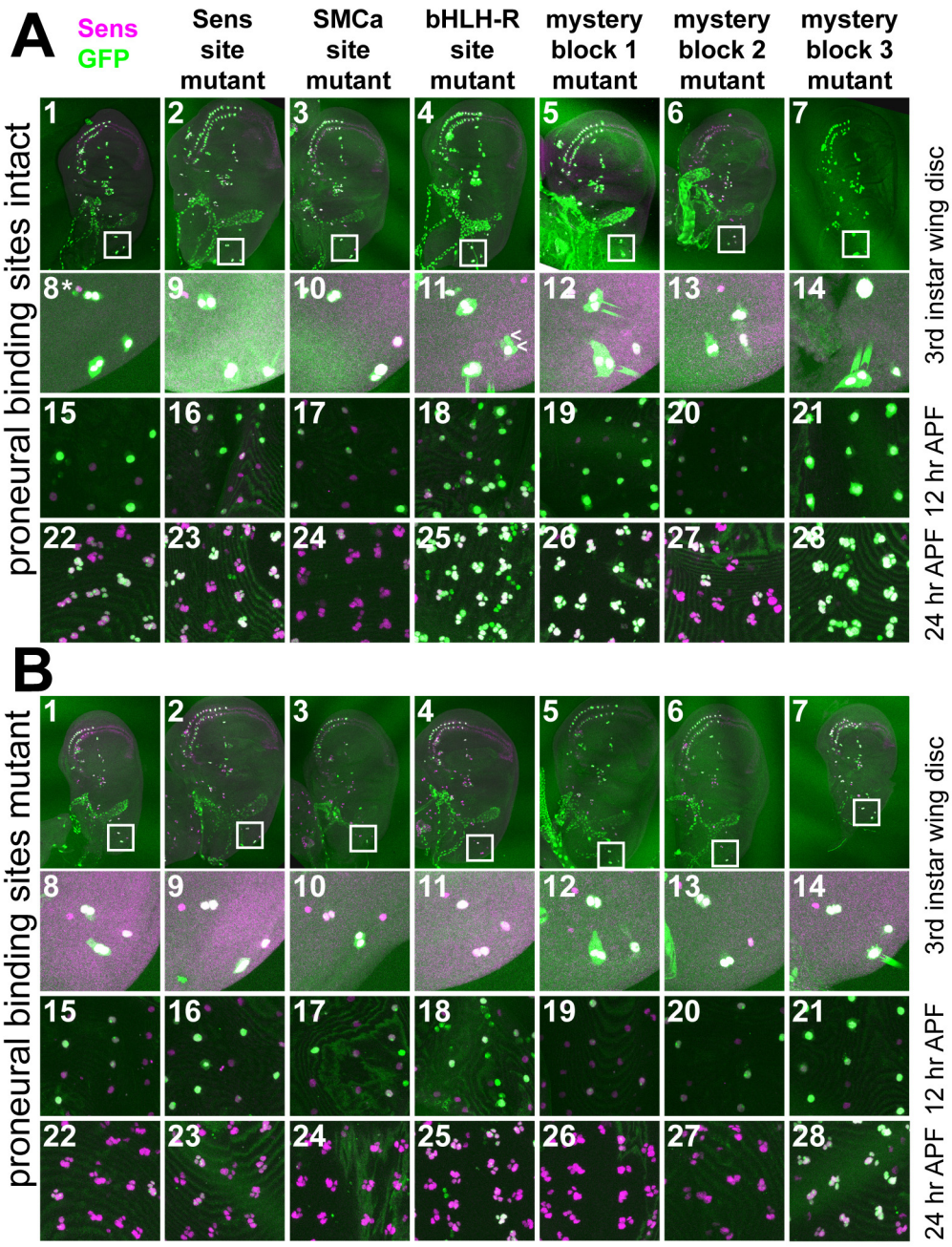
The non-SOP expression of the bHLH-repressor (henceforth bHLH-R) proteins results in a repression of target genes downstream of N signaling. Although few such targets have been identified, they would be predicted to be de-repressed in non-SOP cells in response to bHLH-R binding site mutagenesis; therefore, we mutated these sites in the context of the 4D enhancer. Upon making these mutations we see that 4DRm causes ectopic expression in non-SOP cells (Figure 2.2C, H; Figure 2.3A4,11,18,25), consistent with *neur* being a bHLH-R target. However, expression does not expand through the entirety of the PNCs, and instead is restricted to only a few cells in addition to the SOP. We chose the macrochaete SOPs for the dorsocentral (DC) and scutellar (SC) positions to quantify this effect and found that there were 0.04 ±0.04 SEM extra GFP-positive cells per disc (1/25 discs) for 4DWT, 1.56 ±0.14 SEM (25/40 discs) for 4DRm, and 0 extra GFP-positive cells per disc (0/26) for 4DPRm (Figure 2.2E). These data

indicate not only a statistically significant ($p<<0.01$) requirement for the bHLH-R sites, but also suggest the ectopic expression caused by the R site mutagenesis is dependent upon direct proneural activation.

Requirements of Additional Inputs for Enhancer Activation

The differences in the proneural-dependent transcription factor expression profiles between SOP and non-SOP cells (see Chapter 1) could permit targets in the SOP access to several different factors for activation in addition to the proneurals. In this case, targets would require a combinatorial input involving multiple factors. To test evidence for this model, we examined the 4D enhancer for additional conserved sequences that might provide such additional input (Figure 2.1A, B). Among these conserved sequences are binding sites for the SOP-specific transcription factor Senseless (orthologous in 10/12 sequences), SMCα sites (Culi and Modolell, 1998) (two sites in *Drosophila melanogaster*, one orthologous in 11/12 sequences, the other conserved only through *Drosophila ananassae*), and three blocks of sequence of unknown function—"mystery blocks" (conserved throughout all twelve species). Following the model that one or more

Figure 2.3: Activation of the *neur* 4D enhancer includes input from the proneural and at least two other inputs.
(A) Individual site mutations in the 4D enhancer compared with Sens expression, magenta, in (A1 – A14) third instar wing imaginal discs, (A15 – A21) 12 hr APF notum and (A22 – A28) 24 hr APF notum. A1, A8, A15, A22 represent wild type enhancer, the rest are indicated above. A8 – A14 represent higher magnification views of the boxed regions in A1 – A7. Asterisk in 8, arrowheads in A11 indicate a GFP+/Sens- cell adjacent to the (A8) aDC and (A11) aSC. (B) Samples correspond to the same mutations in A, with the proneural bindings sites also mutated. Note reduced expression of GFP, especially evident in B22 – B28. In some panels, over saturation of the green channel due to extremely high GFP signal resulted in one or more streaks projecting or adjacent from one or more SOP positions (e.g. A12). All images corresponding to the same developmental time point and magnification were collected at the same gains for both channels.


A
Sens
GFP
Sens site mutant
SMCa site mutant
bHLH-R site mutant
mystery block 1 mutant
mystery block 2 mutant
mystery block 3 mutant
proneural binding sites intact
3rd instar wing disc
12 hr APF
24 hr APF
1
2
3
4
5
6
7
8*
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
B
proneural binding sites mutant
3rd instar wing disc
12 hr APF
24 hr APF
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28

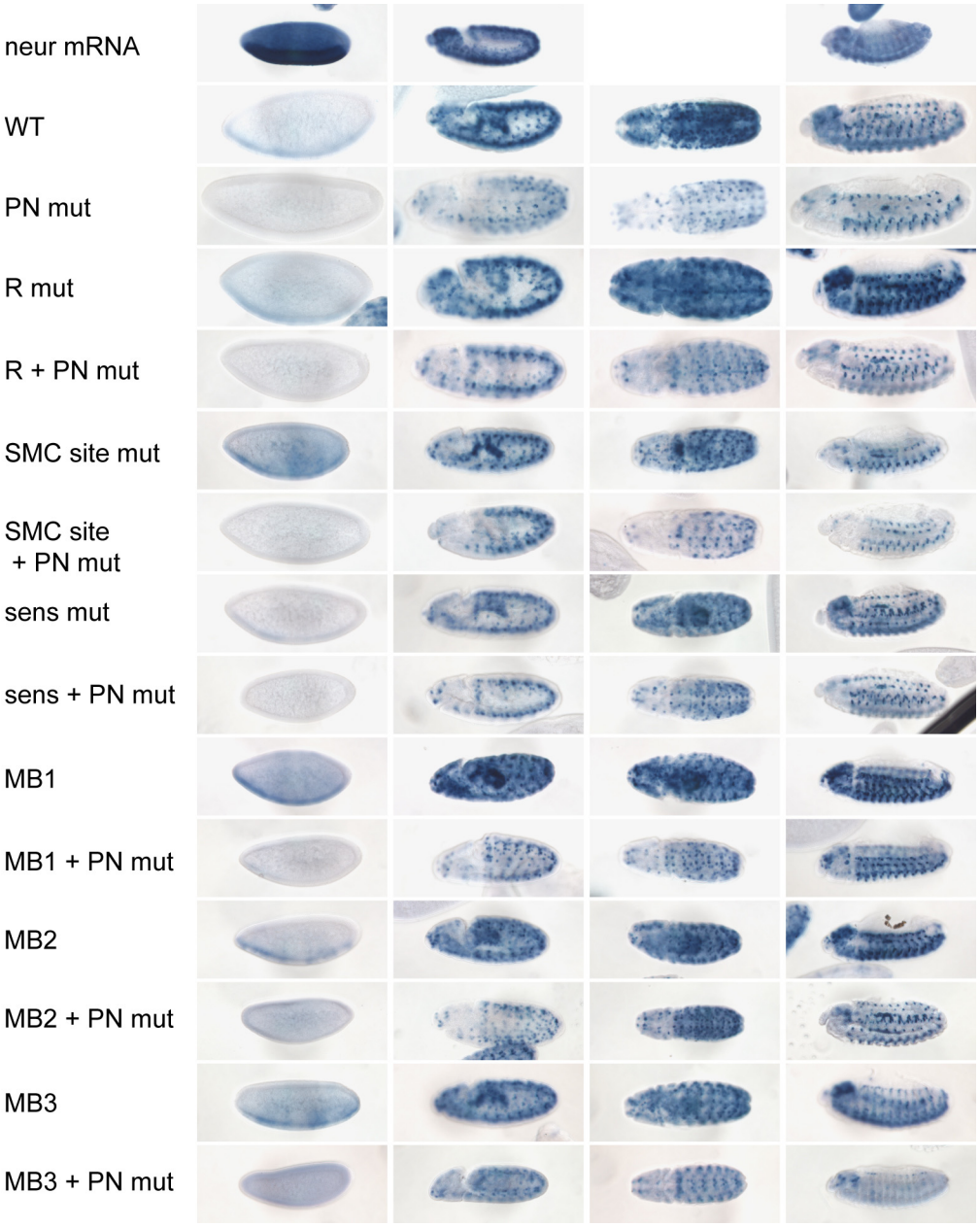
of these sequences would operate in combinatorial fashion with the proneural binding sites, we created reporter constructs with each of these sites mutated individually (Figure 2.3 A) or in combination with complete mutation of all proneural binding sites (Figure 2.3 B).

Mystery blocks 1 and 3 appeared to give increased GFP expression in SOPs compared with 4DWT (Figure 2.3). Non-SOP expression was not observed, suggesting these sites are required for repression within the SOP, but we cannot rule out that factors binding these sequences operate in domains beyond the SOP. The increased expression was no longer observed upon mutation of the proneural binding sites, further supporting the activating role of the proneurals in the SOP, although the MB3 + Pm enhancer still drove GFP expression at levels higher than the Pm mutant alone.

Mutation of the MB2, Sens, and SMCα sites produced very similar phenotypes. Neither mutation significantly reduced GFP expression and all three mutation types in combination with proneural binding site mutation appeared not significantly weaker than the proneural binding site mutation alone. Senseless is a well known contributor to events downstream of SOP specification, though not

required for SOP formation (Jafar-Nejad et al., 2006). We cannot speculate upon what factor binds MB2; however, this sequence contains an octomer shown to be bound by the POU domain protein Brn-1, promoting activation synergy with this factor and mash1 to activate Delta1 in mouse neurogenesis (Castro et al., 2006). The combined mutation of the SMCα and proneural sites appeared to give the strongest phenotype, although still not eliminating reporter expression. The contribution of the SMCα sites, although minor, is significant based upon the previous characterization of these sequences in the regulation of the *sc* SMC enhancer (Culi and Modolell, 1998). The SMC enhancer is a 356 bp sequence that when added to a minimal *sc* rescue construct is sufficient to both induce elevated Sc protein level in single cells (the presumed SOP) and rescue macrochaete formation. Moreover, a synthetic enhancer containing only proneural sites and SMCα sites is sufficient to direct SOP expression. The presence and action of these sites in the regulation of *neuralized* raises the possibility that *neuralized* may be activated prior to the specification of the singular SOP, in parallel to the proneurals.

Figure 2.4: Embryonic expression of the *neur* 4DWT enhancer and mutants.
GFP *in situ* hybridization indicating the expression pattern of the various enhancer mutants as labeled at left; *neur* mRNA is shown for comparison at top. Each row represents different stage embryos stained with the same probe. Of note is the pattern of reduced expression by comparison for all PN mutant enhancers, as well as the expression in early mesectoderm (left most panels). All reporter *in situs* were developed on the same time scale in parallel.


neur mRNA
WT
PN mut
R mut
R + PN mut
SMC site mut
SMC site
+ PN mut
sens mut
sens + PN mut
MB1
MB1 + PN mut
MB2
MB2 + PN mut
MB3
MB3 + PN mut

Charlatan is Capable of SMCα Site Binding and Can Activate neuralized *Expression*

Despite the identification of SMCα sites more than ten years ago (Culi and Modolell, 1998), no protein has been demonstrated to bind to these sequences either in vivo or in vitro. Culi described their similarity to half-sites for NF-κB type proteins, but yet no binding data has been generate nor had there been any documentation of a bristle phenotype for any of the Drosophila NF-κB genes. For this reason, we chose to investigate the possibility that these sequences might be bound by a different class of proteins. Querying the JASPAR database (Sandelin et al., 2004) for proteins that bind a similar sequence produced a strong hit for C2H2 type zinc finger proteins, of which several are known to be expressed in the bristle lineage. Of these, the most attractive candidate was encoded by the *charlatan* gene. *charlatan* mRNA is expressed in proneural clusters and is a direct target of the proneural proteins (Reeves and Posakony, 2005). Ectopic Charlatan expression is capable of inducing the formation of extra SOPs as well as ectopic *ac* and *sc* expression. Moreover, clones of

chn loss of function prevent SOP formation, redirecting it to zones of heterozygosity (Escudero et al., 2005). For these reasons, we asked whether Chn was capable of binding SMCα sites, as well as promoting 4DWT activation.

Chn was shown to bind to sequences within the DC proneural enhancer, but no strong consensus sequence has been determined (Escudero et al., 2005; Tsuda et al., 2006). We therefore tested the ability of Chn to bind shorter 30 bp sequences containing all of the described SMCα sequences as well as the two sites from the *neur* 4D enhancer (Figure 2.5A). Chn was capable of weakly binding to each of these probes but not mutant versions that follow the theme AGGGGTTG → CGTTGTTG. The lone exception was the SMCα1 probe, for which the wild type and mutant probes display the same shift pattern. This pattern is particularly noteworthy as the SMCα1 sequence is the most disparate from the rest of the characterized SMCα sequences (Culi and Modolell, 1998). These in vitro binding data raises the possibility that Chn is the SMCα site binding protein, which would be highly supported by its genetic data.

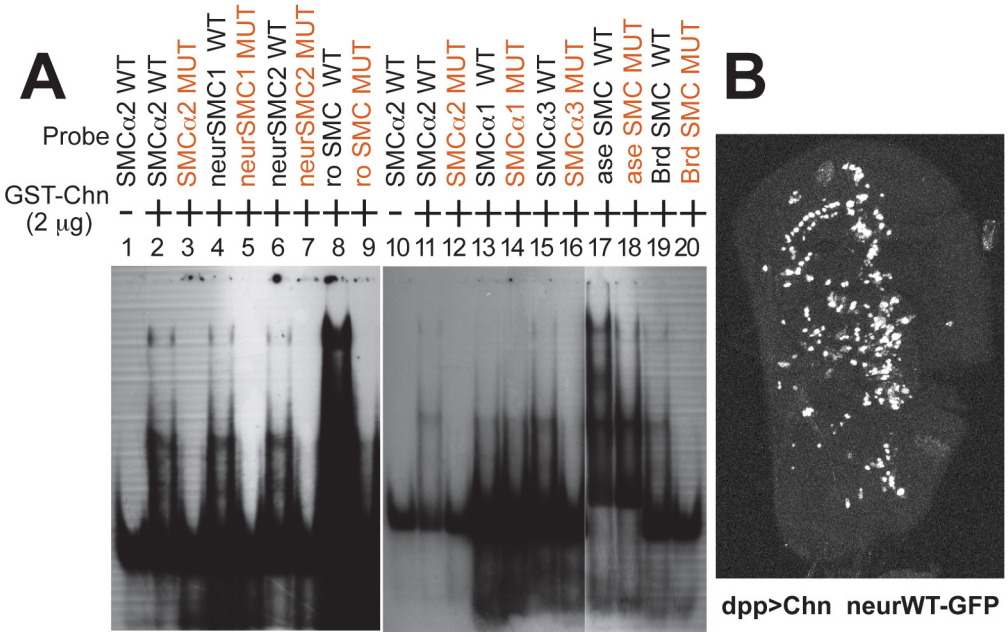


Figure 2.5: Chn is capable of binding SMCα sites and can activate the 4D enhancer.
(A) EMSA assays for the ability of recombinant GST-Chn to bind different probes containing SMCα sites or corresponding mutant probes indicated in red. With the exception of SMCα1 (lanes 13 & 14), all probes show weak binding by GST-Chn that is dependent upon the intact SMCα sequence. With the exception of the neur SMCa sequences, all other SMCα sites were from Culi et al., 1998. (B) Wing imaginal disc from a neur4DWT>GFP/+; dpp>GAL4/+; UAS-chn/+ animal. Note the central longitudinal GFP expression along the anterior-posterior boundary; compare with Figures 2.1C, 2.2A, 2.3A.

present evidence for a model of NF-κB activity downstream of Toll-8 acting in non-SOP cells degrading the mRNA of SOP genes, and synergizing with the proneurals in their auto-regulatory upregulation in the SOP. A major prediction of this model is that if NF-κB factors are involved in transcriptional activation in the SOP, they should be activated and found within the nucleus. To test this model, therefore, we examined the localization of fluorescent fusion proteins representing each of the NF-κB factors from *Drosophila* when expressed along the anterior-posterior boundary with dpp>GAL4. While we were able to observe instances where these proteins exhibited a nuclear localization, never did we observe nuclear localization coincident with SOPs as marked with 4DWT-RFP (Figure 2.6). As such, these data do not favor a model for NF-κB transcriptional activation in the SOP.

neuralized *is Expressed in the Proneural Cluster Prior to SOP Specification*

Ultimately, the true identity of the SMCα site binding protein is of small significance for the overall conclusions presented in this study.

While Chn is capable of inducing ectopic proneural expression, we sought to determine a link between Chn function and *neur* expression. Therefore we examined 4DWT activity in regions where Chn has been ectopically express with dpp>GAL4 (Figure 2.5B). Indeed, in these wing imaginal discs we find high level reporter expression along the anterior-posterior boundary, consistent with activation of the 4D enhancer downstream of Chn expression. However, we recognize that the ectopic SOPs produced by ectopic Chn may also be inducing 4D expression merely by encouraging the SOP transcriptional program and not through direct transcriptional activation of 4D. Nevertheless, Chn remains the only protein demonstrated as binding SMCα sites in vitro and whose gain and loss of function phenotypes are consistent with a role in SOP specification in cooperation with the proneural proteins.

NF-κB Protein Localization in Proneural Clusters

The role of NF-κB proteins in SOP specification has been the study of a recent investigation (Ayyar et al., 2007), creating a controversy in light of the Chn binding data presented above. Ayyar

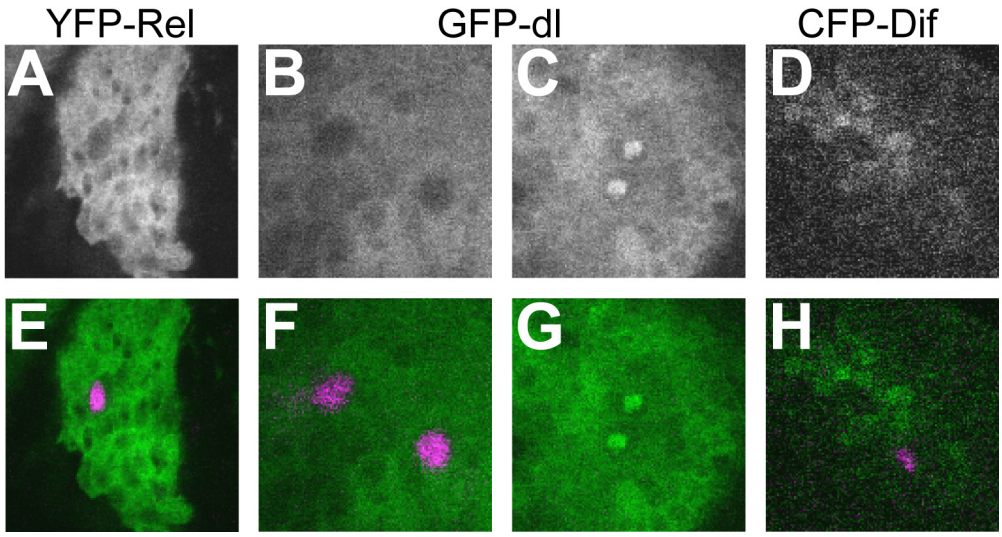


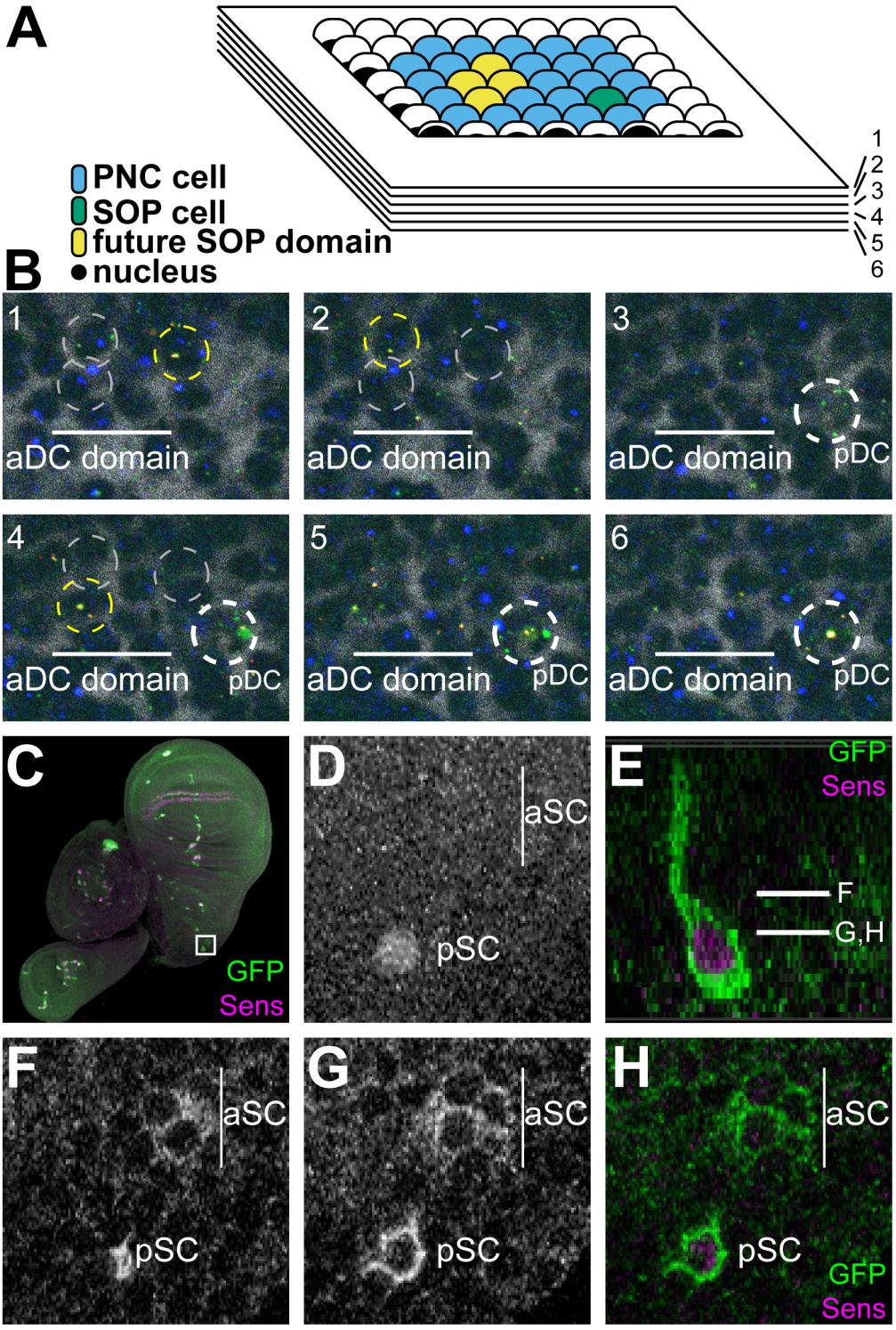
Figure 2.6: *Drosophila* NF-κB proteins are not nuclear localized in the SOP.
dpp>GAL4 driving fluorescent protein fusions to different NF-κB proteins as indicated at figure top. (A – D) GFP expression. (E – F) Merge showing both GFP expression (green) and SOP positions marked by neur4DWT-RFP (magenta). While C and G indicate the detection of nuclear NF-κB, similar localization is not observed overlapping with the SOP marker.

The capability of the 4DWT enhancer to direct reporter expression in microchaete positions prior to Sens expression, the presence and activity of the SMCα sequences (regardless of which protein binds them), as well as the reported “pre-SOP” expression of A101 (Huang et al., 1991) suggests that *neur* may be activated during, rather than as a result of, SOP formation. To test this hypothesis, we looked at *neur* nascent transcript expression in proneural clusters with “heterochronic” SOPs (Huang et al., 1991). In these positions, such as the DC and SC clusters, one SOP is specified a significant amount of time before a second SOP develops from the same cluster within a predictable location relative to the first. If *neur* is expressed only after SOP specification, we should never observe more than two *neur* expressing cells in these clusters. Quite the contrary, we frequently are able to observe two or three nuclei with *neur* nascent transcript “dots” in the region where the second SOP will develop, in addition to the already specified first SOP (Figure 2.7B, Supplemental Figure 2.17). Making use of the multiplex fluorescent in situ technique developed by Kosman *et al*. (Kosman et al., 2004) using intron probes to detect nascent transcript, we were confident in detection of *neur* by the coincidence of

differentially labeled *neur* RNA probes while simultaneously identifying membership in the cluster by *sca* expression and SOP specification by co-expression of both *neur* and *CG32150* (Reeves and Posakony, 2005). After examining numerous such positions we noted the observations of four types of clusters. In the first class, three or more cells in the presumptive second SOP region are positive for *neur* and *sca* expression. The second class has two adjacent cells positive for *neur* and *sca*; in these cells the *neur* transcript dots appear brighter than those of the first class. The third class of cluster has only a single *neur* and *sca* expressing cell, although this is not positive for *CG32150*. The final class encompasses those clusters that clearly have two identifiable SOPs based on detectable expression of *neur*, *sca*, and *CG32150*. In classes three and four, the *neur* transcript dots appear brighter than those in either class one or two. These four classes are very attractive to use to describe a temporal progression of *neur* expression in the developing SOP region, but we wish to be clear that we have no direct evidence of such a progression.

As further support for these *in situ* data, we also made use of a 20 kb genomic rescue construct into which we recombineered a GFP

Figure 2.7: *neur* nascent transcript as well as protein is detectable in multiple cells prior to SOP specification.
(A) Schematic of a proneural cluster with a heterochronic SOP pair. Within the cluster (blue), one SOP develops first (green), and the area where the next SOP will develop (yellow) is predictable. Such a cluster can be subject to confocal sectioning and examined for nascent transcript. (B) Six adjacent 1 micron confocal sections through a DC cluster, which contains a heterochronic SOP pair; the pDC SOP is specified first, followed by the aDC. Multiplex fluorescent *in situ* hybridization using intron probes for *sca* (blue, marks PNC nuclei), *CG32150* (green, marks the specified SOP), and *neur* (two probes with different haptens; coincident probe hybridization cause green and red overlap, displaying yellow). DNA is stained by Hoechst, indicated by inverted grey (i.e. black = DNA). The pDC is indicated by the white circle in panels 3 – 6, and shows *sca*, *CG32150*, and *neur* expression. In the area where the aDC will develop, 3 cells are detected with *neur* probe coincidence (grey circles, yellow in the sections where *neur* probe is visible). (C – H) Visualization of endogenous Neur protein through the use of a GFP-tagged rescue construct supports nascent transcript data. Tissue was stained with anti-GFP and anti-Sens antibodies. (C) Wing, haltere, and leg discs heterozygous for an insertion of the neurRC-WT-GFP construct, displaying the SOP pattern observed with both Sens and GFP expression. Boxed is the scutellar region magnified in D – H. Image is the result of a collapsed z-series. (D) Sens expression is only evident in the pSC position as seen in this collapsed z-series. (E) A single longitudinal section through the pSC noting the extent of GFP signal and Sens expression. While lines indicate planes of the single sections shown in F through H. (F) Single section through the scutellar cluster reveals two cells expressing elevated GFP levels in the domain where the aSC will develop. (G) A deeper section shows elevated GFP expression in three cells. (H) GFP (green) and Sens (magenta) expression from the same section.


A
PNC cell
SOP cell
future SOP domain
nucleus
1
2
3
4
5
6
B
1
aDC domain
2
aDC domain
3
aDC domain
pDC
4
aDC domain
pDC
5
aDC domain
pDC
6
aDC domain
pDC
C
GFP
Sens
D
aSC
pSC
E
GFP
Sens
F
G,H
F
aSC
pSC
G
aSC
pSC
H
aSC
pSC
GFP
Sens

tag upstream of the *neur* stop codon. Flies carrying this construct express a Neur::GFP fusion construct that is readily detectable with an anti-GFP antibody (Figure 2.7C). As expected, we detect GFP expression in an SOP pattern, with the signal primarily localized to the membrane as has been reported previously for such fusions under GAL4/UAS control (Yeh *et al*., 2000). Examining many such discs co-stained for the SOP marker Sens, we find strong coincidence with Sens and GFP signals. As we had done with nascent transcript, we closely examined the GFP signal in heterochronic SOP positions, looking for evidence of GFP signal in areas that lack Sens expression. Indeed, as shown in Figure 2.7F-H we can show such an example for an scutellar position where the anterior scutellar, lacking Sens (Figure 2.7D), shows GFP positivity for multiple cells. While a single plane reveals two adjacent cells expressing GFP, a lower confocal section identifies three additional adjacent cells also positive, indicating at least five cells positive for GFP. Together, these data from both protein and nascent transcript expression, prior to emergence of a single SOP, clearly establish *neur* expression as a participant in singling out an SOP, rather than as a consequence of SOP specification.

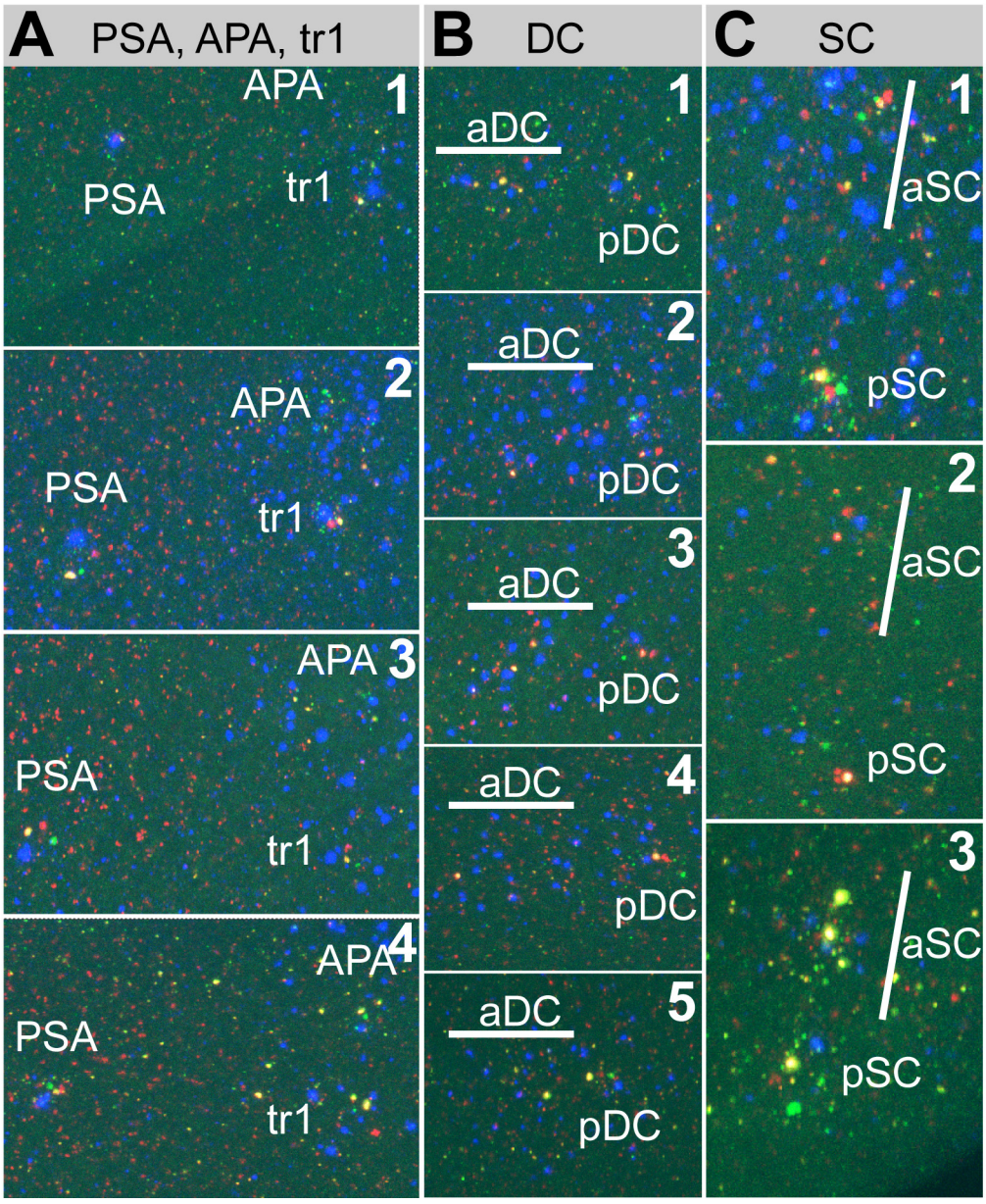
Delta *Upregulation in SOPs*

With the degree of detail we could attain using this multiplex fluorescent *in situ* technique, we thought we may be suited to address the possibility of *Delta* upregulation in the SOP. To this end, we examined heterochronic SOP pairs for their simultaneous expression of *sca*, *neur*, *CG32150*, and *Dl*. *Dl* nascent transcript probes expressed ubiquitously throughout the disc in a pattern reminiscent of Dl protein accumulation. We used a 100 fold diluted probe to potentially reveal differential expression patterns between neighboring cells. With this technique we were able to detect apparent differences in *Dl* nascent transcript (Figure 2.8, Supplemental Figures 2.18, 2.19). However, these differences were only most readily detected in specified SOPs, also expressing high *neur* and *CG32150*, suggesting that if *Dl* is expressed differentially between SOPs and non-SOPs, it is a consequence of SOP specification.

Persistent Non-SOP Neuralized Expression Causes Lateral Inhibition Defects Moderated by Bearded Family Function

Figure 2.8: Nascent *Dl* expression in SOPs.

Collapsed z series from several wing imaginal discs examining different SOP domains. (A) PSA, APA, tr1 regions from four different discs. (B) DC clusters from five different discs. (C) SC clusters from three different discs. For all panels, *sca* = blue; *CG32150* = green; *Dl* = red; *neur* = red + green overlap = yellow.


A PSA, APA, tr1
APA
1
PSA
tr1
2
APA
PSA
tr1
APA
3
PSA
tr1
4
APA
PSA
tr1
B DC
1
aDC
pDC
2
aDC
pDC
3
aDC
pDC
4
aDC
pDC
5
aDC
pDC
C SC
1
aSC
pSC
2
aSC
pSC
3
aSC
pSC

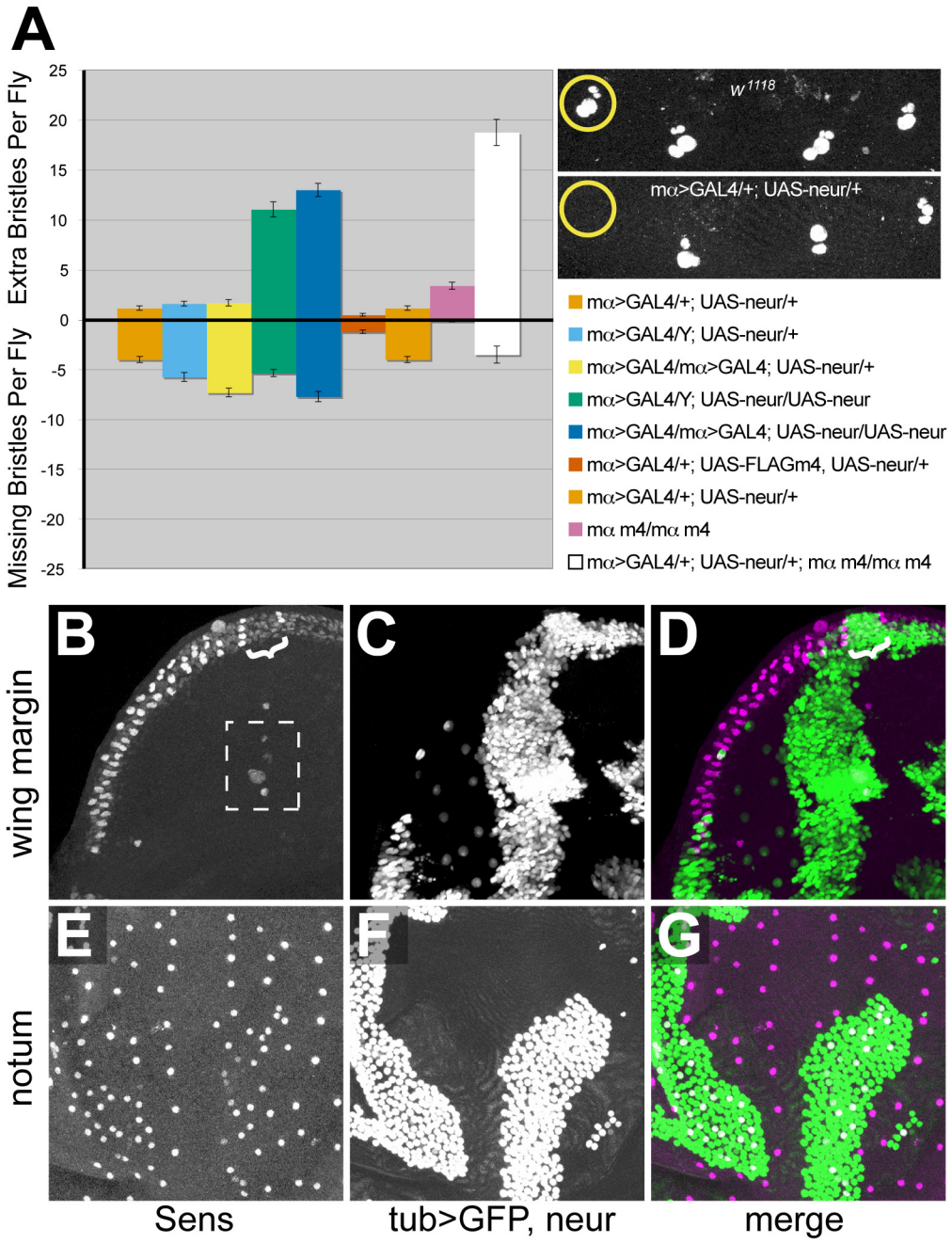
The data presented to this point suggest transcriptional regulation of *neur* is a central focus in SOP specification. This hypothesis would make a number of predictions, one of which is that persistent non-SOP expression of *neur* would cause lateral inhibition defects. We may expect two different phenotypic classes, both of which are consistent with a defect in lateral inhibition. Bristle loss would be expected if ectopic signal coming from non-SOP cells is sufficient to activate N signaling in the SOP, thereby extinguishing the SOP without any further SOPs being able to be specified from the non-SOP pool. Bristle gain could occur if Neur expression is able to make one or more non-SOP cells resistant to the N signal coming from the SOP, preventing them from being inhibited from adopting an SOP fate. It is unclear how these two properties of being an SOP—resistant to N signaling and acting as a strong N signaling cell—are at all separable experimentally, and therefore we could anticipate either or both outcomes.

We chose to test this hypothesis by inducing ectopic, non-SOP expression of *neur* of through two different flavors of the GAL4/UAS system (Fischer et al., 1988): non-SOP expression in all clusters with

the *E(spl)mα>GAL4* driver line (Castro et al., 2005) and clonal ubiquitous expression using the MARCM system with *tub>GAL4* (Lee et al., 2000). Using *mα>GAL4*, we observed a mild lateral inhibition phenotype that became increasingly stronger with increasing copies of both driver and responder (Figure 2.9A). Based upon the hypothesis, we might expect more drastic consequences of non-SOP expression in this manner but there are at least two reasons why it may not be the case.

The first reason involves the non-SOP expression of the Bearded family member (BFM) proteins downstream of N signaling. BFMs have been shown to bind to Neur and inhibit its ability to bind to Dl (Bardin and Schweisguth, 2006; De Renzis et al., 2006; Fontana and Posakony, 2009). Therefore, any non-SOP expression of Neur would be subject to inhibition by BFMs in these cells, neutralizing the effect. To test if this interaction is lessening the potential effect of non-SOP *neur* expression, we repeated this cross in a background homozygous for two deletions eliminating transcription of the BFMs *E(spl)m4* and *E(spl)mα*. In this background we observe a strong increase in especially the extra bristle phenotype with even a single

Figure 2.9: Persistent non-SOP *neur* expression causes lateral inhibition defects.
(A) ma>GAL4, UAS-neur causes both bristle loss and bristle gain, with strength of phenotype dependent on the copy number of driver, responder, and BFMs. Bar graph labels at right correspond by both color and order to graph X axis. Images at right are 24 hr APF nota stained for Cut protein, and indicate bristle loss corresponds to SOP loss (e.g. of aSC in ma>GAL4/+; UAS-neur/+). (B – G) Clones of cells ubiquitously expressing *neur* by tub>GAL4 using the MARCM system, labeled by co-expression of GFP (C, F, green in D, G); SOP positions are indicated by Sens immunoreactivity (B, E, magenta in D, G). (B – D) A clone overlapping the wing margin in a third instar wing imaginal disc results in SOP loss, indicated by bracket. SOPs in boxed region (B) are in a different plane from the GFP expressing cells. (E – G) A clone overlapping the microchaete SOPs in a 12 hr APF notum results in increased microchaete density.


A
Extra Bristles Per Fly
Missing Bristles Per Fly
25
20
15
10
5
0
-5
-10
-15
-20
-25
w1118
mα>GAL4/+; UAS-neur/+
mα>GAL4/+; UAS-neur/+
mα>GAL4/Y; UAS-neur/+
mα>GAL4/mα>GAL4; UAS-neur/+
mα>GAL4/Y; UAS-neur/UAS-neur
mα>GAL4/mα>GAL4; UAS-neur/UAS-neur
mα>GAL4/+; UAS-FLAGm4, UAS-neur/+
mα>GAL4/+; UAS-neur/+
mα m4/mα m4
mα>GAL4/+; UAS-neur/+; mα m4/mα m4
wing margin
notum
B
C
D
E
F
G
Sens
tub>GFP, neur
merge

copy of each of the GAL4 driver and UAS responder, suggesting BFMs are indeed operating to antagonize the function of any ectopic Neur expression in non-SOPs.

An additional reason for the apparently weak phenotype of the *E(spl)mα>GAL4* driven *neur* expression may be the nature of this particular driver itself. The mα enhancer directing GAL4 expression is itself N responsive and any reduction in N signaling in the non-SOP cells would reduce GAL4 expression (Castro et al., 2005). Furthermore, if the driver is activated in response to a N signal emanating from the SOP, the established SOP may already be resistant to any N insults. Thus, for these reasons we sought additional methods to induce persistent non-SOP expression of *neur*.

By inducing large patches of tissue ubiquitously expressing *neur* through the MARCM system we were able to observe similar types of defects as seen with *E(spl)mα>GAL4*, namely both SOP loss and gain. Examining clones in the third instar wing imaginal disc, the dominant phenotype is loss of Sens expression within in the domains of ubiquitous *neur* expression that overlap SOP positions (Figure 2.9B-D); this SOP loss phenotype was also seen using dpp>GAL4 to drive *neur*

expression (data not shown). However, looking in the pupal notum we observe that in the microchaete field domains of ubiquitous *neur* have an increased density of SOPs, even though they are still separated from one another by at least a single cell nuclei (Figure 2.9E-G). Taken together, these data using both types of GAL4 drivers provide clear evidence for the requirement for restricting *neur* to the SOP, with the consequences of failure ranging from failure to specify the SOP to inability to establish proper spacing between SOPs.

Neuralized Function in the SOP is Required to Cell Non-Autonomously Prevent neuralized *Non-SOP Expression*

If proper restriction of *neur* expression does play such a critical role during lateral inhibition, inhibition of Neur function in N signaling within the SOP would be predicted to induce ectopic *neur* expression in non-SOP cells. To test this prediction, we inhibited Neur function specifically in the SOP through the expression of BFMs in the SOP using *neur>GAL4*. Indeed, consistent with the hypothesis, we observe expanded expression of both endogenous *neur* transcript and neur4DWT in response to ectopically driving BFMs in the SOP (Figure

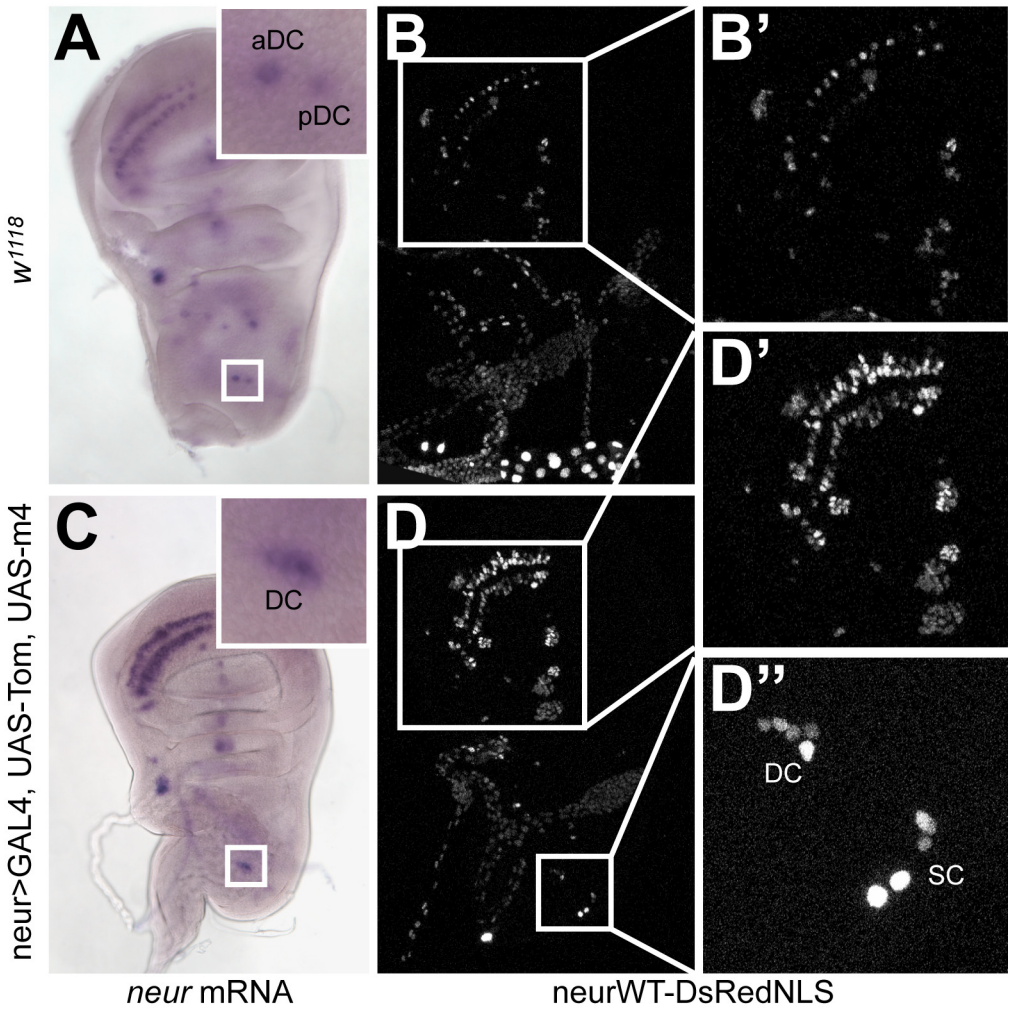


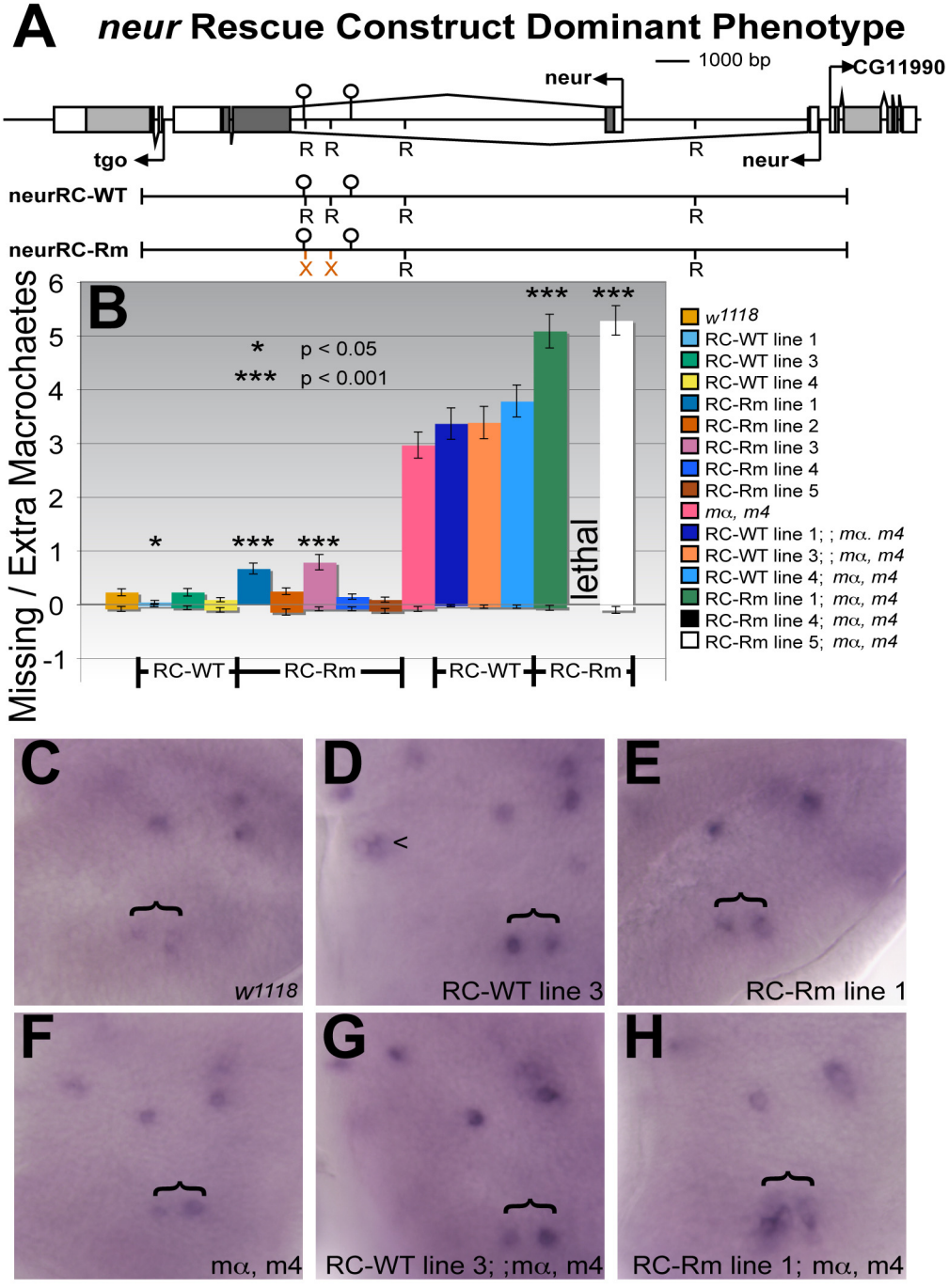
Figure 2.10: Inhibition of Neur function in the SOP causes ectopic *neur* mRNA accumulation and neur4DWT activity.
(A – B') w^{1118} third instar wing imaginal discs. (C – D'') UAS-Tom, UAS-m4/+; neur>GAL4/+ third instar wing imaginal discs. (A, C) *neur* mRNA expression is expanded beyond a single SOP in discs misexpressing m4 and Tom. Insets in A and C at same magnification. (B, D) neur4DWT>DsRed expression mimics that seen for *neur* mRNA, with expansion clearly seen at the wing margin (D') and in the DC cluster (D'').

2.10C, D). These data are consistent with a regulatory loop that links Neur function in the SOP with cell non-autonomous inhibition of *neur* expression in non-SOPs.

Consequences of bHLH-Repressor Binding Site Mutagenesis Upon neuralized *Expression*

To further validate the model of cell non-autonomous self inhibition of Neur, we chose to test the specific requirements for the bHLH-R binding sites in the context of the endogenous *neur* locus. To this end we cloned at 20 kb genomic rescue fragment in to pCaSpeR (Thummel et al., 1988) and either left this fragment intact or subcloned in fragments that either lacked the 4D enhancer (see the subsequent section) or contained mutations in the bHLH-R binding sites within this enhancer. We then examined the phenotype of flies containing either the WT or 4DRm rescue construct (Figure 2.11). If the bHLH-R binding sites are required to prevent *neur* expression in non-SOPs, and that this ectopic expression has consequences for lateral inhibition, then we should be able to observe these defects in a dominant fashion. While we fail to detect any lateral inhibition phenotype or ectopic *neur*

Figure 2.11: Phenotypes of a bHLH-R site mutant *neur* rescue construct in pCaSPer.
(A) Diagram of the *neur* locus, indicating the boundaries of the WT and Rm mutant rescue constructs, and noting the mutation of the bHLH-R sites in the 4D enhancer. (B) Bar graph of the bristle gain and loss phenotypes of multiple genomic insertions of rescue construct variants. Labels indicated at right are the same color and order of the bars, from left to right. Note the genetic interaction between the *mα, m4* genotype and each of the RC-Rm lines, last three samples. (C – H) *neur* mRNA expression in the notum region of third instar wing imaginal discs of (C) w^{1118}, (D) RC-WT line 3, (E) RC-Rm line 1, (F) *mα, m4*, (G) RC-WT line 3; *mα, m4*, (H) RC-Rm line 1; *mα, m4*. DC cluster indicated by bracket, for reference.


A
neur Rescue Construct Dominant Phenotype
1000 bp
CG11990
neur
tgo
neur
R
neurRC-WT
neurRC-Rm
X
B
Missing / Extra Macrochaetes
p < 0.05
p < 0.001
lethal
RC-WT
RC-Rm
w1118
RC-WT line 1
RC-WT line 3
RC-WT line 4
RC-Rm line 1
RC-Rm line 2
RC-Rm line 3
RC-Rm line 4
RC-Rm line 5
mα, m4
RC-WT line 1; ; mα. m4
RC-WT line 3; ; mα, m4
RC-WT line 4; mα, m4
RC-Rm line 1; mα, m4
RC-Rm line 4; mα, m4
RC-Rm line 5; mα, m4
C
w1118
D
RC-WT line 3
E
RC-Rm line 1
F
mα, m4
G
RC-WT line 3; ;mα, m4
H
RC-Rm line 1; mα, m4

expression with independent transformant lines for either the WT or 4DRm constructs, when these lines are place in the *m4, ma* mutant background, we can detect both mild ectopic expression and an increase in lateral inhibition defects only with the 4DRm lines, consistent with a requirement for these sites to prevent non-SOP expression of *neur*.

The neuralized *4D SOP Enhancer is Dispensable for Development, Including Lateral Inhibition*

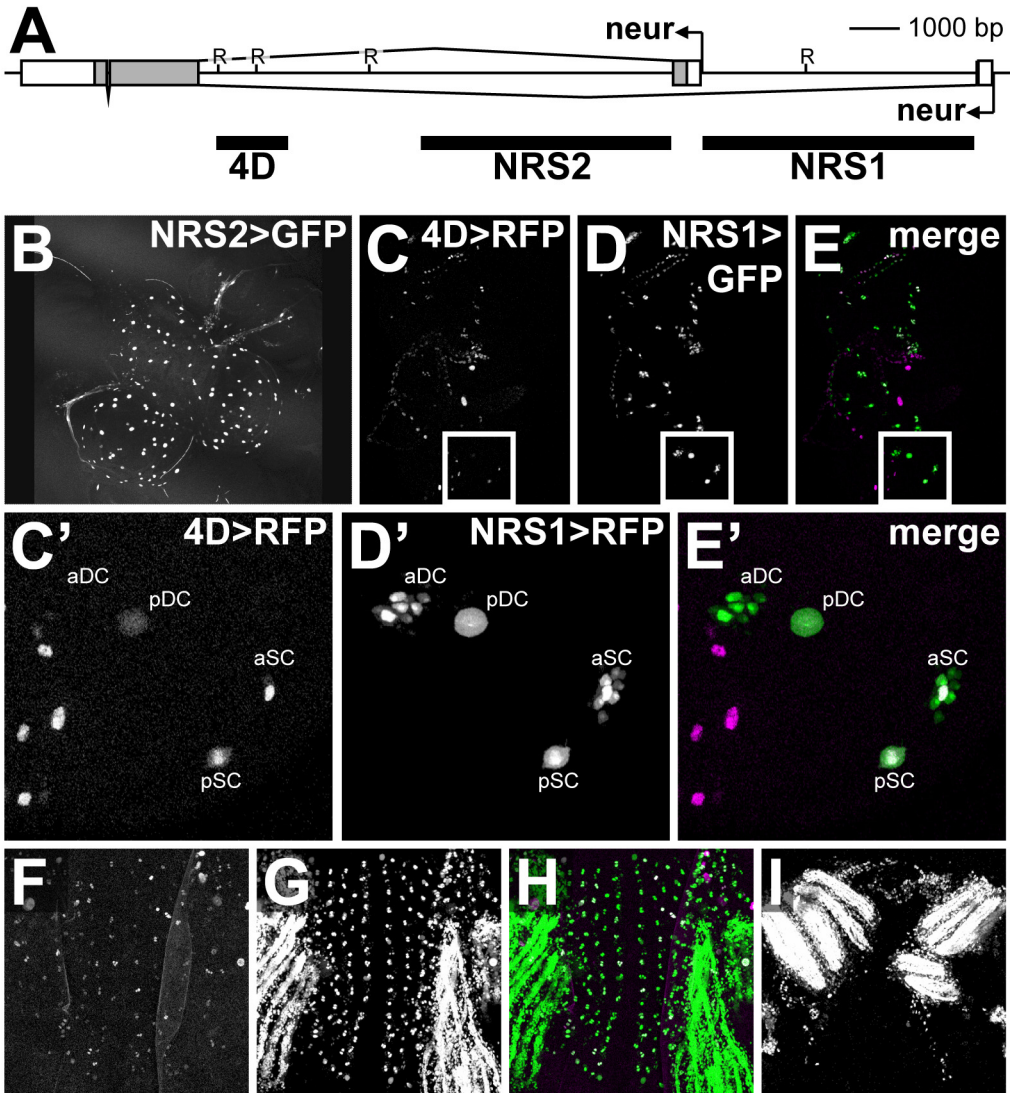
In addition to testing the requirement for the bHLH-R binding sites, we also sought to confirm that the 4D enhancer was necessary for lateral inhibition. To our surprise, *neurRC-4Ddel; neur*^{lF65} homozygous embryos displayed no defects in their neural pattern as stained for with Elav, and when allowed to reach adulthood they exhibit no aberrant bristle pattern (data not shown). These data suggest that this enhancer is dispensable for lateral inhibition, perhaps due to the presence of another enhancer with partially overlapping specificity. Moreover, *neurRC-4DWT; neur*^{lF65} flies actually exhibited a neurogenic phenotype which we subsequently mapped to a single nucleotide

deletion at Valine 245, resulting in a frameshift and rendering the rescue construct ineffective. These two unfortunate results forced us to construct a new series of rescue constructs and to search for any additional enhancer(s) within the locus that direct expression in the PNS.

neuralized *Contains Two SOP Enhancers with Overlapping Specificity*

We constructed two new reporter fragments downstream of each of the two alternative promoters for *neur* (Figure 2.12A). A 4 kb fragment downstream of the proximal promoter (NRS2) directed reporter expression in a regularly spaced pattern in the brain, but failed to drive expression in SOPs in the third instar wing imaginal disc (Figure 2.12B). A fragment downstream of the distal promoter (NRS1), however, directed strong expression in SOPs, very similar to 4DWT (Figure 2.12C-E). Moreover, this expression was abolished in a sc^{10-1} mutant background, also like both 4DWT and endogenous *neur* (Figure 2.12I; Figure 2.13D). Unlike 4DWT, however, we could also detect NRS1 expression in cells surrounding the SOPs in the DC and SC cluster (Figure 2.12D, E). The differences between NRS1 and 4DWT

Figure 2.12: A second enhancer in the *neur* locus has SOP specificity.
(A) Diagram of *neur* locus, indicating positions of the 4D enhancer, and the additional reporters used to survey the region (NRS = Neuralized Reporter Survey). (B) The NRS2 fragment directs reporter expression in a subset of cells of the brain. (C) Expression DsRed (RFP) via control of the neur4DWT enhancer in the third instar wing imaginal disc. Boxed region is magnified in C', highlighting expression in the dorsocentral (DC) and scutellar regions (SC). (D) The NRS1 fragment directs GFP reporter expression in a primarily SOP pattern strikingly similar to neur4D. Boxed region is magnified in D'; not the expanded expression at SOP positions relative to C'. (E, E') Merged images of C, D, C', and D'; GFP signal is in green, DsRed is in magenta. (F – I) The NRS1 region also encodes for muscle specificity independent of proneural activity. (F, magenta in H) neur4DWT>RFP expression in a 24 hr APF notum. (G, green in H) The same notum expressing NRS1>GFP; note the expression in muscle not seen with neur4DWT. (I) NRS1>GFP expression in a $sc^{10\text{-}1}$ mutant; while GFP is not detectable in a the microchaete field, the muscle expression remains.


A
neur
1000 bp
R
R
R
R
neur
4D
NRS2
NRS1
B
NRS2>GFP
C
4D>RFP
D
NRS1>
GFP
E
merge
C'
4D>RFP
aDC
pDC
aSC
pSC
D'
NRS1>RFP
aDC
pDC
aSC
pSC
E'
merge
aDC
pDC
aSC
pSC
F
G
H
I

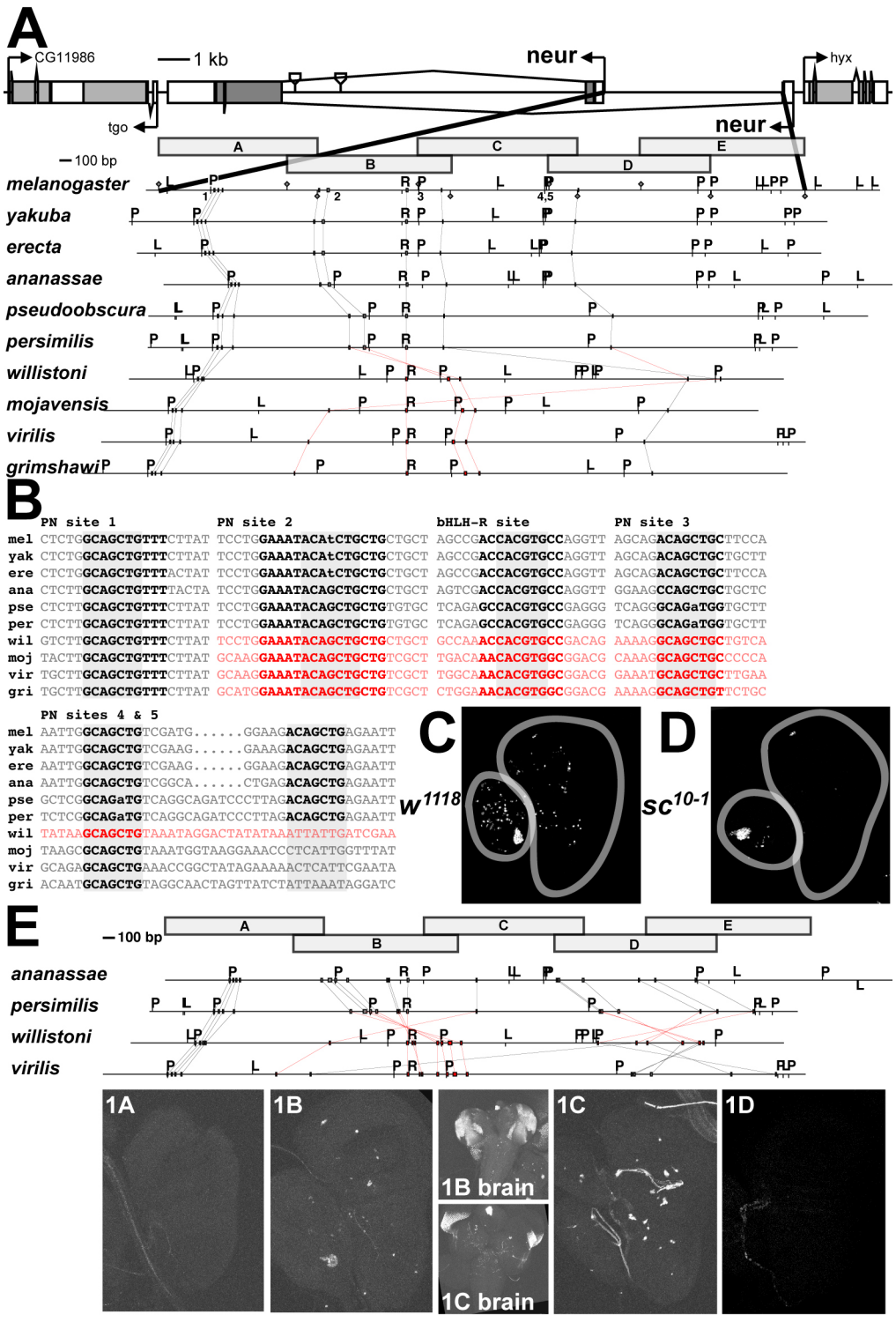
can be extended further into the pupal notum. At 24 hrs APF we find that NRS1 no only expresses in the bristle field, but also gives strong expression in a muscle pattern (Figure 2,12G, H). This muscle pattern does not require proneural protein function, as reporter expression is maintained in a sc^{10-1} background (Figure 2.12I). We sought to examine the pattern of sequence conservation in this upstream region to potentially direct our focus to a smaller fragment with SOP specificity.

The Upstream SOP Regulatory Region Has Been Subject to Recent Inversions

To narrow the upstream region, we divided the sequence in to five overlapping fragments of approximately 1.3 kb each (Figure 2.13A). Two adjacent fragments, 1B and 1C, each gave expression in a subset of SOPs, while only 1B gave appropriate expression in the brain (Figure 2.13E). The larger region contains a number of proneural binding sites, Sens binding sites, and a single bHLH-R site. Looking at the pattern of conservation throughout the 10 *Drosophila* genomes for this region (*Drosophila simulans* and *Drosophila sechelia* were

excluded due to a stretch of Ns in the assembled sequence), we find evidence of mainly four of the proneural binding sites and the single bHLH-R site having been preserved since the last common ancestor between *Drosophila melanogaster* and *Drosophila grimshawi* (Figure 2.13B). Furthermore, this upstream region appears to not have been under as tight of evolutionary constraint as the 4D enhancer regions, since we observe two regions that have undergone inversions in at least one of the branches of the evolutionary tree. The first region encompasses much of the 1B region and has been inverted relative to the *Drosophila melanogaster* sequence in the clade containing *Drosophila willistoni* through *Drosophila grimshawi*. The second region was discovered by performing an alignment between only *Drosophila ananassae*, *Drosophila persimilis*, *Drosophila willistoni*, and *Drosophila virilis*. With this reduced set of species, we can observe another large inverted block mainly contained by the 1D fragment that is unique to *Drosophila willistoni* (Figure 2.13E). Since these inversions suggest that sequences that are inverted together are required functionally, we would interpret the inversions to direct our focus solely to the 1B fragment for subsequent analysis, as sequences within 1C that overlap

Figure 2.13: The NRS1 region, although exhibiting inversions throughout evolution, contains a conserved bHLH-R consensus site and several conserved PN consensus sequences.
(A) Diagram of the *neur* locus, indicating the position of the 4D enhancer (compare with Figure 2.1) and the NRS1 region. Below the gene diagram is an expanded view of the NRS1 region from Drosophila melanogaster aligned with the orthologous regions from the eleven other sequences Drosophila genomes. Indicated are the locations of Senseless binding sites consensuses, L; bHLH-R consensus, R; and proneural consensuses, P. Blocks of 8 bp or more identical across all genomes are connected by lines and noted along the sequence, while the different proneural sites are numbered below the melanogaster line to be referred to in B. Inversion are represented by red. Above the melanogaster sequence are five overlapping subfragments tested to refine enhancer dimensions. (B) Sequence alignments of each of the orthologous proneural and bHLH-R sites. Text in red notes an inversion relative to melanogaster. (C) Expression of the NRS1 enhancer in a wild type third instar wing imaginal disc. (D) Expression of the NRS1 enhancer in a sc10-1 mutant third instar wing imaginal disc. Leg and wing imaginal discs are outlined for both C and D images. (E) Alignment of the sequences more closely related than melanogaster to willistoni. Subfragment boundaries are as in A. Note that two regions of sequences linked by inversion map well to subfragments B and D. Below the alignment are larval third instar tissue from representative reporter constructs using fragments A through D. Note the overlapping fragments 1B and 1C both drive SOP expression but that 1B has brain expression more representative of endogenous *neur*.


A
CG11986
1 kb
neur
hyx
tgo
neur
A
B
C
D
E
100 bp
melanogaster
yakuba
erecta
ananassae
pseudoobscura
persimilis
willistoni
mojavensis
virilis
grimshawi
B
PN site 1
PN site 2
bHLH-R site
PN site 3
mel CTCTGGCAGCTGTTTCTTAT TCCTGGAAATACAtCTGCTGCTGCT AGCCGACCACGTGCCAGGTT AGCAGACAGCTGCTTCCA
yak CTCTGGCAGCTGTTTCTTAT TCCTGGAAATACAtCTGCTGCTGCT AGCCGACCACGTGCCAGGTT AGCAGACAGCTGCTGCTT
ere CTCTGGCAGCTGTTTACTAT TCCTGGAAATACAtCTGCTGCTGCT AGCCGACCACGTGCCAGGTT AGCAGACAGCTGCTTCCA
ana CTCTTGCAGCTGTTTTACTA TCCTGGAAATACAGCTGCTGCTGCT AGTCGACCACGTGCCAGGTT GGAAGCCAGCTGCTGCTC
pse CTCTTGCAGCTGTTTCTTAT TCCTGGAAATACAGCTGCTGTGTGC TCAGAGCCACGTGCCGAGGG TCAGGGCAGaTGGTGCTT
per CTCTTGCAGCTGTTTCTTAT TCCTGGAAATACAGCTGCTGTGTGC TCAGAGCCACGTGCCGAGGG TCAGGGCAGaTGGTGCTT
wil GTCTTGCAGCTGTTTCTTAT TCCTGGAAATACAGCTGCTGCTGCT GCCAAACCACGTGCCGACAG AAAAGGCAGCTGCTGTCA
moj TACTTGCAGCTGTTTCTTAT GCAAGGAAATACAGCTGCTGTCGCT TGACAAACACGTGGCGGACG CAAAGGCAGCTGCCCCA
vir TGCTTGCAGCTGTTTCTTAT GCAAGGAAATACAGCTGCTGTCGCT TGGCAAACACGTGGCGGACG GAAATGCAGCTGCTTGAA
gri TGCTTGCAGCTGTTTCTTAT GCATGGAAATACAGCTGCTGTCGCT CTGGAAACACGTGGCGGACG AAAAGGCAGCTGTTCTGC
PN sites 4 & 5
mel AATTGGCAGCTGTCGATG.....GGAAGACAGCTGAGAATT
yak AATTGGCAGCTGTCGAAG.....GAAAGACAGCTGAGAATT
ere AATTGGCAGCTGTCGAAG.....GGAAGACAGCTGAGAATT
ana AATTGGCAGCTGTCGGCA.....CTGAGACAGCTGAGAATT
pse GCTCGGCAGaTGTCAGGCAGATCCCTTAGACAGCTGAGAATT
per TCTCGGCAGaTGTCAGGCAGATCCCTTAGACAGCTGAGAATT
wil TATAAGCAGCTGTAAATAGGACTATATAAATTATTGATCGAA
moj TAAGCGCAGCTGTAAATGGTAAGGAAACCCTCATTGGTTTAT
vir GCAGAGCAGCTGAAACCGGCTATAGAAAAACTCATTCGAATA
gri ACAATGCAGCTGTAGGCAACTAGTTATCTATTAAATAGGATC
C
w1118
D
sc10-1
E
100 bp
A
B
C
D
E
ananassae
persimilis
willistoni
virilis
1A
1B
1B brain
1C
1C brain
1D

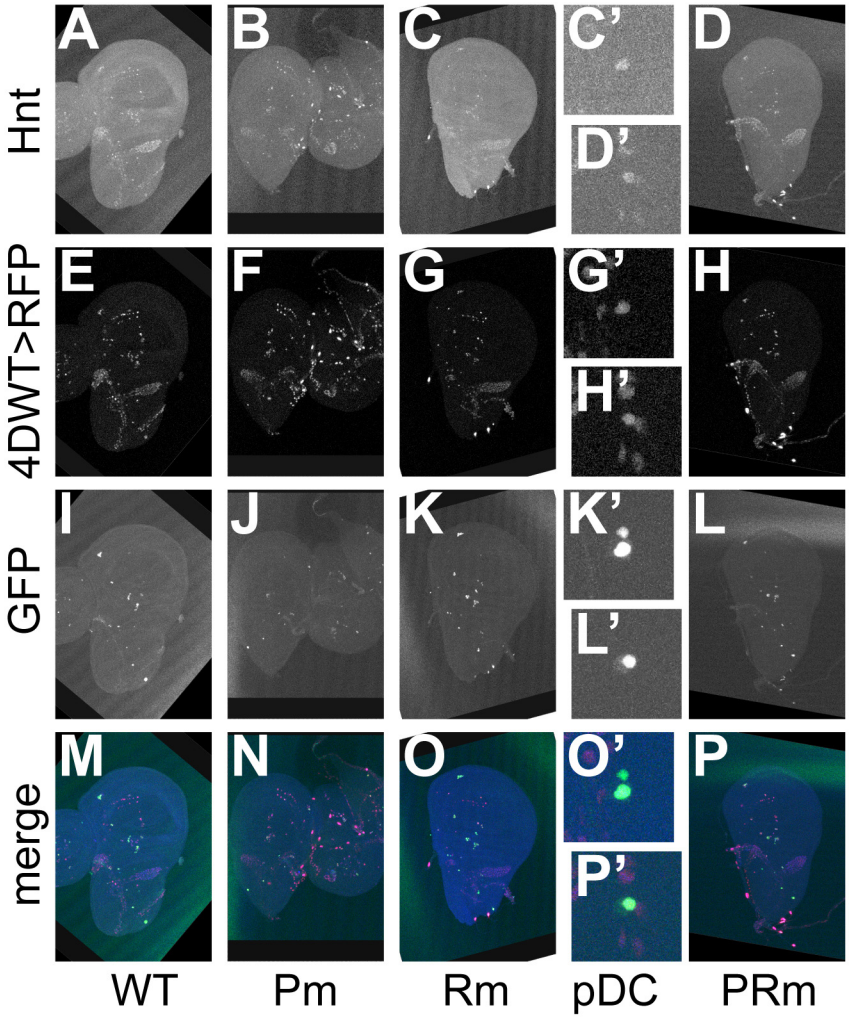
1B and 1D appear to not be required to be maintained as a functionally discrete unit.

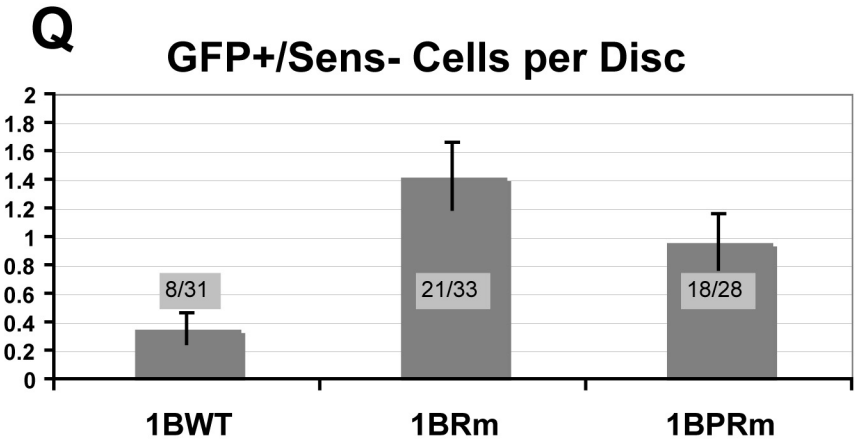
Conservation of Proneural and bHLH-Repressor Inputs on a Minimal Upstream Enhancer Fragment

The pattern of conservation within the 1B fragment suggests a requirement for the bHLH-R site, but puts less importance upon the proneural binding sites (Figure 2.13B). One of the proneural sites in the 1B fragment has a G > T mutation that results in a deviation from the RCAGSTG proneural consensus in *Drosophila melanogaster*, *Drosophila yakuba*, and *Drosophila erecta*, and the second proneural site, while present in *Drosophila melanogaster*, contains a C > A mutation in *Drosophila pseudoobscura* and *Drosophila persimilis* that also prevents it from fitting the consensus. Nevertheless, to test the requirements for these sequences on the 1B enhancer, we examined the consequences of mutating these sequences (Figure 2.14).

First, we noticed that like the full length fragment, we often observe non-SOP reporter expression driven by the 1B fragment (0.35 ±0.11 extra GFP-positive, Sens-negative cells per disc; 8/31 discs

Figure 2.14: Consequences of mutation of proneural and bHLH-R inputs upon the 1B subfragment.
(A – P) Third instar wing imaginal discs expressing either (A, E, I, M) 1BWT>GFP; (B, F, J, N) 1BPm>GFP; (C, G, K, O) 1BRm>GFP; (D, H, L, P) 1BPRm>GFP. (A – D) Hnt expression, indicating SOP positions. (E – H) 4DWT>RFP in the same discs illustrating differences between the strength of 4D and 1B. (I – L) GFP expression directed from the different 1B variants. (M – P) Merge of all three channels. (C' – P') High magnification view of the pDC cell in the (C', G', K', O') 1BRm disc or the (D', H', L', P') 1BPRm disc. (Q) Quantification of extra GFP cells directed by 1B mutant fragments. Ratios above each bar indicate number of discs exhibiting extra GFP cells over the total number of discs scored. Reduction of extra GFP cells in 1BPRm discs is not statistically different from 1BRm.


A
B
C
C'
D
D'
Hnt
E
F
G
G'
H
H'
4DWT>RFP
I
J
K
K'
L
L'
GFP
M
N
O
O'
P
P'
merge
WT
Pm
Rm
pDC
PRm


Q
GFP+/Sens- Cells per Disc
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
8/31
21/33
18/28
1BWT
1BRm
1BPRm

display extra GFP cells). Mutating the bHLH-R binding site increases both the average number of extra GFP-positive cells per disc (1.42 ±0.23 extra GFP-positive cells, $p>0.01$) and the frequency of discs with extra GFP-positive cells (21/33). Furthermore, consistent with the lack of strong conservation of proneural binding sites in this region, we fail to detect a strong decrease in GFP expression upon mutation of the lone proneural binding site in Drosophila melanogaster and we also fail to detect proneural dependency of the non-SOP expression of the bHLH-R mediated ectopic GFP-positive cells (1BPRm: 0.96 ±0.20 extra GFP-positive cells; 18/28 discs with extra GFP-positive cells).

Loss of bHLH-R Input on neuralized *Expression*

Our analyses with both the 4D and 1B enhancers suggest a functional requirement for the bHLH-R sites in each of these enhancers. We chose to assess this question in the context of a different series of rescue constructs incorporated into the attB-P[acman] backbone (Venken et al., 2006), so that all the mutants could be examined at the same docking site to minimize the position effect. neurRC-WT, neurRC-4DRm, neurRC-1BRm, and neurRC-4D,1BRm

were generated by recombineering using a galK substitution and replacement strategy (Warming et al., 2005). These constructs were all inserted into a single docking site, attP40 (Markstein et al., 2008), using the ΦC31 integrase system (Groth et al., 2004). Analyzing bristle counts on each of these mutants revealed no statistically significant difference between the wild type and Rm variants (Table 2.1a). Interestingly, with these constructs we were not able to assay the effect of removing copies of the *E(spl)mα* and *E(spl)m4* genes due to lethality of the neurRC-WT construct in this background. While noting the absence of this key positive control, we did observe a strong tendency for extra bristles to appear in the aDC and aSC positions, only in flies carrying the Rmut rescue construct and lacking copies of *E(spl)mα* and *E(spl)m4*. This result was statistically significant from an *E(spl)mα* and *E(spl)m4* mutant alone ($p>0.01$, see Table 2.1b); but we could only take this as a hint of a phenotype.

In spite of the lack of a strong lateral inhibition phenotype with an intact *E(spl)-C*, we reasoned that the level of ectopic *neur* required to produce such a phenotype might be relatively high given our observations of dosage dependent with UAS-neur and mα>GAL4.

Table 2.1a: neurRC bristle counts, total head and thorax.

Genotype (n = 50 unless noted)	Missing Bristles	Extra Bristles
w^{1118}	0.08 ±0.05	0.22 ±0.07
neurWT.V5.VK37~	0.18 ±0.07	0.70 ±0.12
neur4D,1B-RM.GFP.VK37	0.40 ±0.11	0.28 ±0.08
neurWT-attP40(#1)	0.00 ±0.00	0.62 ±0.12
neur1B-RM-attP40(#1)	0.04 ±0.04	0.64 ±0.12
neur1B-RM-attP40(#2)	0.04 ±0.03	0.88 ±0.14
neur4D-RM-attP40	0.00 ±0.00	0.30 ±0.08
neur4D,1B-RM-attP40(#1)	0.10 ±0.04	0.92 ±0.15
neur4D,1B-RM-attP40(#2)	0.12 ±0.05	0.60 ±0.11
m4mα	0.06 ±0.03	4.84 ±0.27
neurWT.V5.VK37;m4mα	0.04 ±0.03	3.24 ±0.30
neur4D,1B-RM.GFP.VK37;m4mα	0.12 ±0.06	2.36 ±0.23
neurWT-attP40(#1);m4mα (n = 9)	0.11 ±0.11	3.22 ±0.70
neur1B-RM-attP40(#1);m4mα (n = 28)	0.14 ±0.07	6.82 ±0.45
neur4D-RM-attP40;m4mα	0.46 ±0.09	1.54 ±0.21
neur4D,1B-RM-attP40(#1);m4mα	0.04 ±0.04	2.50 ±0.26
neurWT-attP40 (line 1)/+;$neur^{IF65}$/Df(3R)ED5330	0.02 ±0.02	0.04 ±0.03
neur1B-RM-attP40(#1)/+;$neur^{IF65}$/Df(3R)ED5330	0.02 ±0.02	0.06 ±0.03
neur1B-RM-attP40(#2)/+;$neur^{IF65}$/Df(3R)ED5330	0.12 ±0.06	0.18 ±0.06
neur4D-RM-attP40/+;$neur^{IF65}$/Df(3R)ED5330	0.03 ±0.03	0.18 ±0.07
neur4D,1B-RM-attP40(#2)/+;$neur^{IF65}$/Df(3R)ED5330 (n = 33)	0.06 ±0.04	0.16 ±0.05

~VK37 and attP40 indicate ΦC31 docking sites

Table 2.1b: neur4D,1BRm, dorsocentral and scutellar bristles.

Genotype	Missing Bristles	Extra Bristles
w^{1118}	0.02 ±0.02	0.02 ±0.02
neur4D,1B-RM-attP40(#1)	0.00 ±0.00	0.16 ±0.05
m4mα	0.00 ±0.00	0.44 ±0.09
neur4D,1B-RM-attP40(#1);m4mα	0.00 ±0.00	1.08 ±0.17*

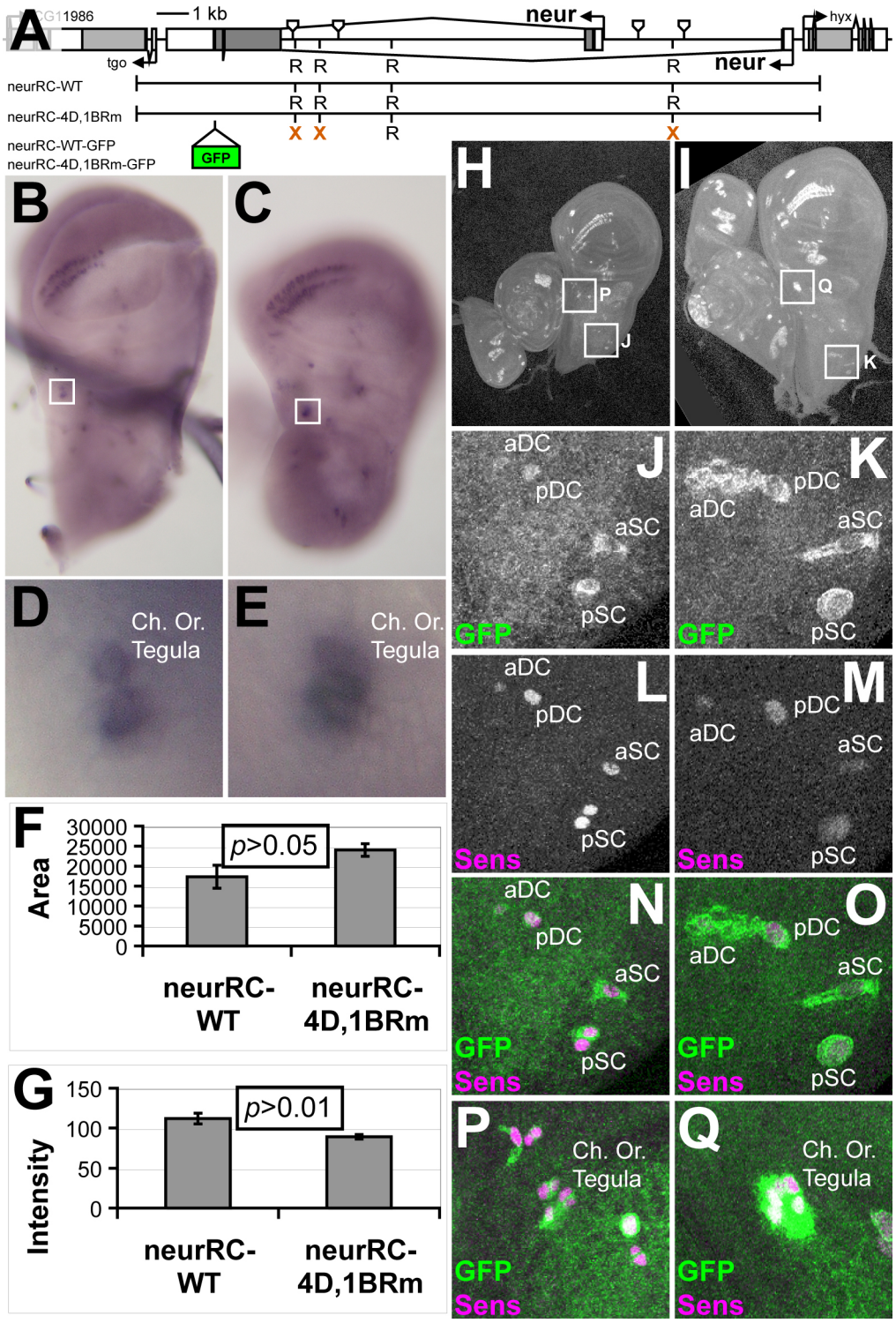
*$p>0.01$ between *m4mα* and *neur4D,1B-Rm;m4mα*

Therefore, we sought to determine whether any ectopic *neur* transcript could be detected in flies carrying homozygous rescue construct insertions. While most positions do not appear to reveal any striking ectopic expression, we often observed a much darker *in situ* signal in the cluster corresponding to the chordotonal organ of the tegula, a balancing organ on the proximal wing, in the neurRC-4D,1BRm discs (Figure 2.15). This suspicion was confirmed by measuring both the average area as well as the average intensity of *in situ* signal for this position. The neurRC-4D,1BRm tegula chordotonal clusters had an average area of 24040 ±1575 SEM pixels (n = 21), compared to 17340 ±2888 SEM pixels (n = 9) for the neurRC-WT discs ($p>0.05$ by ANOVA). Analyzing the brightness intensity for these areas serves a dual purpose. First, since we determined the areas of *in situ* probe detection by hand, we could introduce bias by selecting a greater area for the Rm discs than for the WT discs, a bias that would be reflected by increasing the average brightness in the Rm discs due to the incorporation of extra area not darkened by the probe signal. Second, we can also assay a potential increase in not only the area of *in situ* signal but also the density to further support ectopic *neur* expression.

When analyzing the intensity, we find the neurRC-WT tegula chordotonal clusters have an average intensity of 112 ±6.68 SEM, and the neurRC-4D,1BRm discs have an average intensity of 89.2 ±2.8 SEM, a very significant difference (p>0.001 by ANOVA). Had our area determination been biased, we would expect to find the Rm tegula clusters to have a higher average intensity, but this was not the case. Moreover, the Rm clusters had a statistically significant lower intensity, indicating that the *in situ* signal was darker within the selected area than in the WT clusters. These data support that at least for the tegula chordotonal clusters for which we performed this analysis, mutation of the bHLH-R sites in the 4D and 1B enhancers results in both an expansion of, as well as a greater average level of *neur* transcript expression in the cluster.

While the above results indicate that at least one cluster has a detectable difference in *neur* mRNA expression, we chose to also examine Neur protein accumulation consequential to mutation of the bHLH-R sites through the GFP-tagged rescue constructs. Because Neur expression is very age dependent, we chose to examine discs from larvae dissected at the white prepupae stage. This not only

Figure 2.15: Mutation of the bHLH-R binding sites in the *neur* locus causes ectopic expression of *neur*.
(A) Diagram of the corresponding constructs used in this figure: B – G utilize data from untagged rescue constructs (either wild-type or 4D,1BRm), while H – Q are C-terminal GFP tagged rescue constructs. (B - E) In situ hybridization detection of *neur* transcript in third instar wing imaginal discs from animals homozygous for a *neur* rescue construct variant integrated into docking site attP40. (B, D) neurRC-WT disc; chordotonal organ of the tegula position is boxed in B and a high magnification view is shown in D. (C, E) neurRC-4D,1BRm disc; boxed region is same as for B, and a high magnification view is shown in E. (F) Quantification of the area of *neur* transcript from the chordotonal organ of the tegula cluster from either neurRC-WT or neurRC-4D,1BRm flies. (G) Quantification of the brightness of the regions in (F). A lower value indicates a darker average in situ signal within the area. Images in B – E represent samples near the mean area for each genotype. (H – I) Immunohistochemical detection of the GFP tag on either a (H, J, L, N, P) wild-type or (I, K, M, O, Q) 4D,1BRm mutant *neur* rescue construct variant integrated into docking site VK00037. Note the expansion of GFP expression in the RC-RmGFP tissues compared to their RC-WTGFP counterparts. Number of GFP positive cells in aDC region in K, O is not representative of average phenotype but shown for example; most discs have 1 or 2 extra GFP positive cells at either the aDC or aSC position. P and Q are representative.

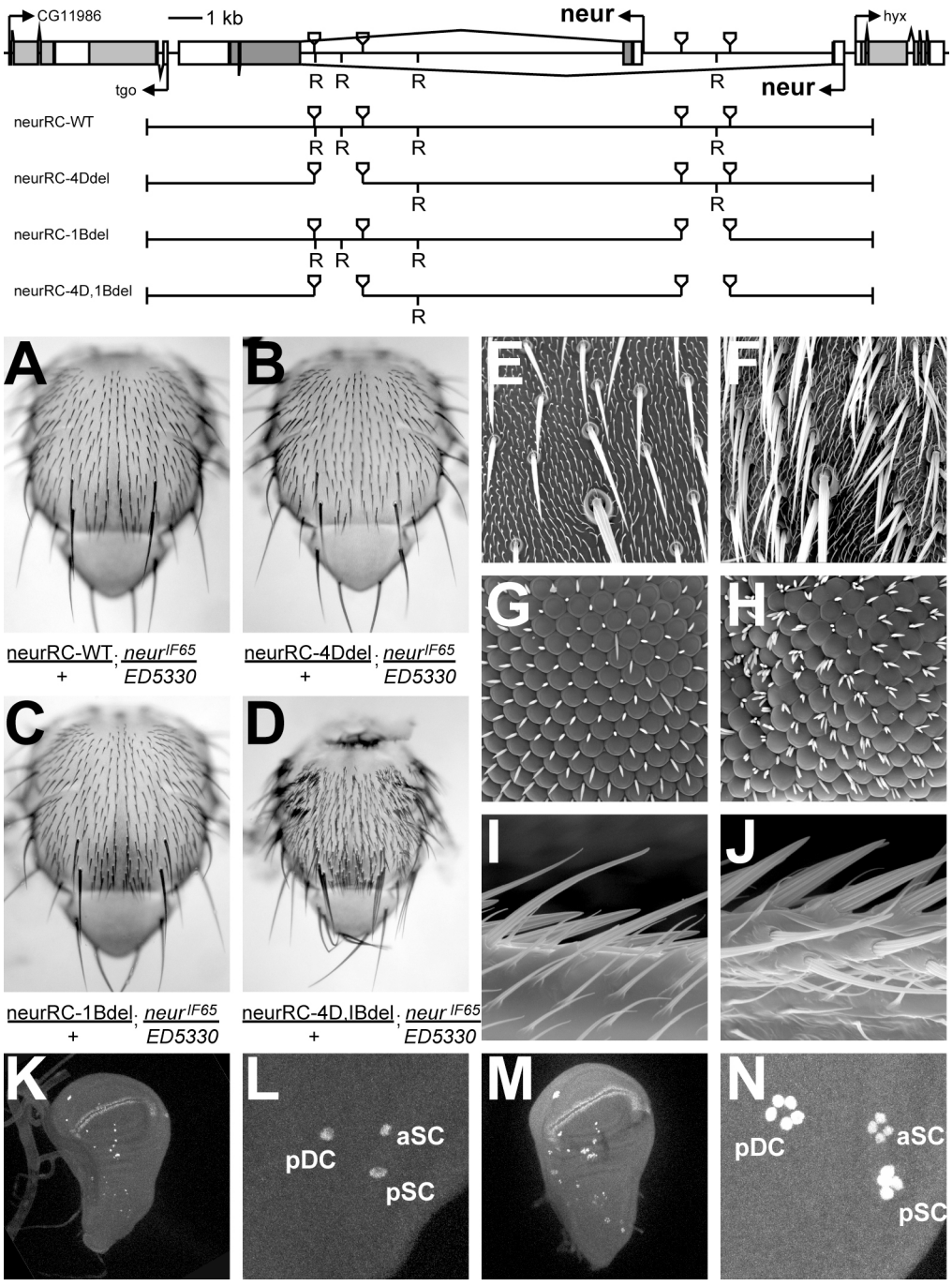

A
CG11986
1 kb
neur
hyx
tgo
R R R R
neur
neurRC-WT
neurRC-4D,1BRm
X X
neurRC-WT-GFP
neurRC-4D,1BRm-GFP
GFP
B
C
D
Ch. Or.
Tegula
E
Ch. Or.
Tegula
F
Area
30000
25000
20000
15000
10000
5000
0
p>0.05
neurRC-WT
neurRC-4D,1BRm
G
Intensity
150
100
50
0
p>0.01
neurRC-WT
neurRC-4D,1BRm
H
P
J
I
Q
K
J
aDC
pDC
aSC
pSC
GFP
K
aDC
pDC
aSC
pSC
GFP
L
aDC
pDC
aSC
pSC
Sens
M
aDC
pDC
aSC
pSC
Sens
N
aDC
pDC
aSC
pSC
GFP
Sens
O
aDC
pDC
aSC
pSC
GFP
Sens
P
Ch. Or.
Tegula
GFP
Sens
Q
Ch. Or.
Tegula
GFP
Sens

serves as a control for age differences between the different rescue construct variants, but also is an age at which the two main heterochronic SOP positions we have followed for much of this study would be fully specified. At this stage of development, we find a high proportion of discs (6/10) expressing the neurRC-4D,1BRm-GFP variant which express GFP in cells neighboring either the aDC or the aSC positions (Figure 2.15K); we could find no such cases with the WT-GFP variant (Figure 2.15J). We also noticed that similar to the detectable mRNA with the untagged constructs, we find an increase in GFP expression in the chordotonal organ of the tegula in the 4D,1BRm-GFP constructs but not with the WT-GFP variant (Figure 2.15P, Q); this pattern also holds for the dorsal radius cluster (data not shown). Indeed, in these very rigorous tests for bHLH-R site dependency in the context of the *neur* locus, these sites do appear to be required for the proper expression pattern of *neur* in the wing imaginal disc.

The Combined Activities of Both SOP Enhancers Are Required for Lateral Inhibition

Finally, given our observation that a rescue construct lacking the 4D enhancer module still was capable of rescuing the *neur* amorphic phenotype, we asked how much additional activity was being supplied by the 1B enhancer module. To this end we generated additional attB-P[acman] rescue constructs deficient for either the 4D module (neurRC-4Ddel), the 1B module (neurRC-1Bdel), or both (neurRC-4D,1Bdel), and examined their ability to rescue a *neur*^{lF65}*/Df(3R)ED5330* genotype (Figure 2.16). Similar to what we had observed with the 4Ddel construct (Figure 2.16B), a 1Bdel rescue construct successfully rescues the *neur* mutant phenotype (Figure 2.16C), suggesting that either enhancer module is sufficient to substitute for the other. When both enhancers have been deleted, however, we observe robust bristle tufting of both the macrochaete and microchaete bristles along the head and thorax (Figure 2.16D, F, H, J). The extra bristles can be traced back to an overproduction of SOPs, as observed by Sens staining in the third instar larval wing disc (Figure 2.16M, N), indicating that the activity of both of these enhancers is indeed required to establish the appropriate expression profile of *neur* to participate in lateral inhibition.

Figure 2.16: Functional redundancy in the two *neur* SOP enhancers during lateral inhibition.
Top is schematic of *neur* locus and relevant rescue constructs. (A - D) Preparations of adult nota from *neur*^{IF65}*/Df(3R)ED5330* with one copy of a *neur* rescue construct variant. (A) neurRC-WT. (B) neurRC-4Ddel. (C) neurRC-1Bdel. (D) neurRC-4D,1Bdel. (E – J) Scanning electron micrographs from *neur*^{IF65}*/Df(3R)ED5330* adults with either (E, G, I) neurRC-WT or (F, H, J) neurRC-4D,1Bdel. (E, F) Notum, including the aDC macrochaete and surrounding microchaetes. (G, H) Compound eye; note irregularities in ommatidia pattern and interommatidial bristles. (I, J) Anterior wing margin. (K – N) Extra bristles in neurRC-4D,1Bdel rescued animals result from overproduction of SOPs as marked by Sens immunoreactivity. (K, L) Third instar wing imaginal disc from neurRC-WT rescued animal. (M, N) Third instar wing imaginal disc from neurRC-4D,1Bdel rescued animal.


CG11986
1 kb
neur
hyx
tgo
neur
R
neurRC-WT
neurRC-4Ddel
neurRC-1Bdel
neurRC-4D,1Bdel
A
B
E
F
G
H
neurRC-WT / + ; neurIF65 / ED5330
neurRC-4Ddel / + ; neurIF65 / ED5330
C
D
I
J
neurRC-1Bdel / + ; neurIF65 / ED5330
neurRC-4D,lBdel / + ; neurIF65 / ED5330
K
L
M
N
pDC
aSC
pSC

Discussion

neuralized *needs to be restricted to the SOP*

Much of the previous work on the role of *neur* in lateral inhibition to date has focused on defining its molecular function and contribution to N signaling. The exquisite SOP specificity of *neur* has largely only been appreciated for its utility as a cell fate marker, especially for the enhancer trap A101. In this study, however, we have established that the beauty of *neur* transcript specificity is quite crucial to its role in lateral inhibition. We have demonstrated through multiple means that failure to restrict *neur* expression to the SOP can result in either of two lateral inhibition phenotypes: extinguishment of the SOP, or the production of supernumerary SOPs. We find intriguing the observation that different SOP classes behave differently under conditions of uniform *neur* expression, with macrochaete SOPs becoming extinguished, while microchaete SOPs persist and permit more of their non-SOP neighbors to adopt the SOP fate. This difference in phenotype may be simply due to the difference in time between specification of the SOP and its first division between these two types of bristles, or may reflect differences in additional regulatory molecules

that sway the outcome into a particular direction. Nevertheless, it is clear that PNC cells need to not only activate *neur* to generate the correct bristle pattern, but also need to restrict *neur* expression from cells that are not destined to become SOPs.

Endogenous neuralized *expression defines a proneural cluster subdomain relevant for SOP specification*

Multiple analyses presented in this study suggest that over the course of development, *neur* transcript is not absolutely restricted to only those cells that will ultimately become SOPs. We can observe neur4DWT reporter expression in positions in the microchaete rows, which fail to express SOP markers (Sens-), that show spacing from Sens+ SOPs similar to intra-row SOP spacing. Such cells often occur in pairs and have nuclei smaller than SOP nuclei, consistent with an epithelial cell fate. Moreover, these cells are observed in tissue collected around the time of early microchaete SOP specification, and are not found at later time points. We also note the occasional observation of 4DWT>GFP expressing cells next to a GFP-/Sens+ cell in third instar wing imaginal discs, especially at the aDC position.

Similar observations have been reported during a careful analysis of A101-LacZ expression, an enhancer trap in the *neur* locus, in the developing macrochaete SOP positions in the third instar wing imaginal disc (Huang et al., 1991). Thus, two independent reporters of *neur* cis-regulatory elements suggest the capability of enhancers to direct expression in cells of a seemingly "pre-SOP" nature.

Examining nascent transcript of *neur* expression in heterochronic PNCs provided further evidence for the expression of *neur* pre-mRNA in more cells than would be predicted based upon the known SOP number for that cluster. We believe the utilization of *neur* transcript detection by the coincident signal of two differently labeled fluorescent probes suggest the multiple cells exhibiting *neur* signal is not due to non-specific probe binding. Moreover, the cluster cells that have coincident signal are often adjacent, mimicking the reporter assays described above. We were relieved to see that the expression of the GFP-tagged wild type rescue construct indeed confirmed the results with *neur* nascent transcript and reporter expression. In total, these data present a strong case for *neur* expression in proneural clusters in "pre-SOPs."

A number of additional data further support the existence of a subpopulation of PNC cells from which the SOP develops. The expression of *emc*, an HLH-only protein believed to be an inhibitor of proneural function by promoting the formation of non-productive heterodimers, in the wing imaginal disc is nearly complementary to *ac* and *sc* expression. However expression of *emc* does infiltrate the PNCs, and the SOP appears to develop in the subset of the cluster that has the lowest levels (Skeath and Carroll, 1991; Van Doren et al., 1992; Cabrera et al., 1994). Mutational analysis of a proneural responsive enhancer of *chn* further supports non-equivalence of PNC cells (Reeves and Posakony, 2005). This enhancer mimics *chn* transcript accumulation in the majority of PNC cells, and contains a number of proneural binding sites. When all the proneural binding sites are mutated, reporter expression is not abolished, but is maintained in a small subset of PNC cells surrounding the SOP. This result is in stark contrast to a number of other reporters for which mutation of proneural sites completely abolishes activity: edl PE, Traf1 PE (Reeves and Posakony, 2005), E(spl)mα PE (Castro et al., 2005). Additionally the rho PE, while largely restricted to SOPs, does express

in a few cells surrounding the SOP (Reeves and Posakony, 2005). We suggest that these "SOP+" expressing reporters may to be reflecting a non-uniform trans-regulatory environment within the proneural cluster that are established due to the activity of Emc and other factors.

neuralized *is both a participant in and a target of N signaling during lateral inhibition*

The "pre-SOP" expression of *neur* resolves an apparent paradoxical relationship between expression and function of Neur and BFMs. Previous studies have established a physical interaction between these two proteins (Bardin and Schweisguth, 2006; De Renzis et al., 2006; Fontana and Posakony, 2009), resulting in an inhibition of the ability of Neur to bind Dl. However, until now these factors were observed to be expressed mainly in the two complementary patterns in the PNC; Neur is restricted to the SOP, from which the BFMs (and bHLH-Rs) are excluded (Jennings et al., 1994; Kramatschek and Campos-Ortega, 1994; Jennings et al., 1995; Lai et al., 2000; Nolo et al., 2000; Zaffran and Frasch, 2000; Castro et al., 2005). As shown in this study, Neur and BFM function in non-SOP cells support this

antagonistic relationship. Ectopic expression of *neur* using the regulatory region from BFM *E(spl)mα* can cause lateral inhibition defects that can be enhanced by the loss of two BFMs, *E(spl)m4* and *E(spl)mα*, and mitigated by the coexpression of *E(spl)m4*. Moreover, expression of BFMs in the SOP results in ectopic *neur* expression and SOP formation, further demonstrating the potent ability of BFMs to mitigate cell autonomously the inhibitory capacity of cells within the PNC. These data support a mechanism by which Neur function in the pre-SOPs would activate BFM expression in receiving cells, allowing Neur to inhibit its own function in a cell non-autonomous manner, preventing these cells from signaling back.

This study has also revealed an important overlaying regulatory relationship controlling *neur* expression in the proneural cluster through the bHLH-R genes of the *E(spl)-C*. Both the 4D and 1B enhancers contain well-conserved bHLH-R binding site consensuses that, when mutated, cause ectopic enhancer expression in non-SOP cells. In both cases, this ectopic expression does not expand throughout the entire PNC, but is limited to a few cells neighboring the SOP, again in favor of an underlying non-uniform trans-regulatory network throughout the

cluster. Moreover, applying the strict test of site functionality in the context of a genomic rescue construct validates the requirements of these sequences in preventing ectopic PNC expression of *neur* transcript, at the very least in the chordotonal organ of the tegula cluster. Lastly, we also observe a stark contrast in the expression of *neur* in response to ectopic activity of the two major classes of N pathway targets. While ectopic SOP expression of either BFMs or bHLH-Rs are able to cause a strong balding phenotype, neur>BFM causes ectopic *neur* expression, while neur>m7 causes loss of *neur* expression (data not shown). In total, our data are consistent with a model for *neur* regulation within the PNC in which Neur upregulation in the pre-SOP domain parallels proneural upregulation, and inhibits cells outside of this domain from becoming SOPs. Furthermore, to singularize the SOP, a dominant cell is able to initiate a productive signaling event that activates both BFMs and bHLH-Rs in the other pre-SOP cells. These two families provide a two-pronged approach for preventing an SOP fate in these N receiving cells, by quickly inactivating already expressed Neur and preventing the continued expression of *neur* transcript. We believe such an approach would

result in a rapid creation of asymmetry between the SOP and surrounding cells, and reliably generate a proper pattern of bristle number and spacing.

A P + R code for identifying SOP enhancers?

It is perhaps not surprising that both of the *neur* SOP enhancers possess the dominating feature of the presence of conserved binding sites for both the proneural proteins and the bHLH-Rs, given activities of these proteins in the PNC. The proneural proteins are well-established high-level activators within the proneural cluster likely activating a significant network of targets within the cluster, both on their own ("P only") (Reeves and Posakony, 2005) and in combination with other trans-machinery ("P + X"). One example of the P + X code has been established for genes in the *E(spl)-C* (Castro et al., 2005), with the additional input being provided by the N pathway through Su(H) (P + S). This allows N signaling to facilitate a sharp asymmetry of activation of *E(spl)-C* members between the SOP and non-SOP cells, and promoting the strong repression of N pathway targets specifically in non-SOP cells. The most simple model then would be in

strong favor of a similar P + R code, where the absence of bHLH-R repression (by nature of their P + S regulation) would permit the proneurals to be dominant activators in the SOP.

However, the data described in this study do not favor such an elegant P + R code of regulation, in contrast to the P + S code. In that latter case, the proneural binding sites are absolutely required for activity; and eliminating the dual activation-in-non-SOPs and inhibition-in-SOPs activity of Su(H) results in a complete reversal of enhancer activity: causing ectopic SOP and loss of non-SOP expression (Castro et al., 2005). In the case of the *neur* enhancers, while bHLH-R site mutation does cause ectopic reporter expression, it is not persistent throughout all non-SOP cells as a simple P + R model might predict. Similarly, the simple model would also predict that elimination of proneural binding sites would completely abrogate the enhancer, as occurs in the examples of P + S enhancers; indeed this is not what we observe.

Examining other SOP-specific genes adds an additional puzzling layer: the apparently predictive value of the P + R code. Using a two-factor SCORE approach (Rebeiz et al., 2002), we were

able to identify statistically significant clusters of proneural and bHLH-R binding sites in regions nearby several known SOP-specific loci, including *phyl*, *nvy*, *wor*, and *sens* (M. Rebeiz and J. Posakony, unpublished data). In each of these cases, the significant cluster of sites has revealed SOP expression. Of these additional P + R enhancers, the *phyl* SOP enhancer (phylSE) and the nvySE display a significant expansion of expression upon bHLH-R binding site mutation; the sensSE Rm mutant appears similar in its activity to the *neur* Rm enhancers. Moreover, mutation of the proneural sites in these enhancers also mimics the results we observe with the *neur* Pm enhancers: the lack of an absolute proneural site requirement for enhancer activity in the SOP (M. Rebeiz and J. Posakony, unpublished data). These data are also not in support of a simple P + R model and would even appear to present a paradox: why does a P + R code both reliably predict SOP enhancer activity and exhibit strong evolutionary conservation but strong requirements for each of these sites have not been observed in enhancer analyses?

The additional combinatorial enhancer mutagenesis we have performed for the neur4D enhancer may provide a small degree of

illumination. While the Rm enhancer only results in a small subset of non-SOP cells becoming derepressed, the ectopic expression absolutely requires intact proneural binding sites. Furthermore, mutation of several of the other conserved sequences appears to further weaken a Pm enhancer, as is observed for the SMCα sites. We propose that these additional requirements reflect a trans-regulatory network within the SOP operating on P + R enhancers that is at least one level down from the network regulating P + S enhancers. Enhancers of the P + S type that are directly downstream of the fate specification signal would see a much more uniform trans-regulatory environment across the cluster; to establish sharp expression differences between cell fates would favor a strong all-or-none translation of the specification signal and the local activators in the cluster. By the very nature of the directionality of the N signal, the non-responding cell activates the first-level of SOP gene expression in a permissively N-independent manner. The prediction would be that the enhancers for these genes would be activated in response to high level proneural activity and may indeed be strongly dependent upon intact proneural binding sites for their activation; these enhancers may not

even be strongly SOP-specific yet (the chnPE may represent such an example). Just one transcriptional activator regulated in such a manner would create a non-uniform trans-regulatory environment not present just a single network level prior—the “pre-SOP” subgroup. Moreover, the presence of such an additional activator in these cells would put less pressure on SOP-specific enhancers to rely only on proneural input and may in fact favor combinatorial input from the proneurals and one or more pre-SOP activator. The P + X mutations in neur4D may be revealing both types of pre-SOP activation: the proneural site-dependent expression in the Rm enhancer illuminates cells for which the enhancer primarily responds to the highest levels of proneural proteins while the SOP-specific expression requires the combination of high proneural levels (a requirement that may support their persistent conservation) and input from at least one additional factor. Such a combinatorial environment within the pre-SOP cells would only require bHLH-R activity to distinguish the SOP out of this cluster subset, and mutation of these sites would not result in de-repression throughout the entire PNC; nevertheless, this limited requirement may be enough to favor conservation of these sites.

The subdivision of the cluster and combinatorial activation to achieve SOP specificity would favor eventual loss of either the proneural sites or the bHLH-R sites, or both, in some enhancers, however. It may be the case the that the 1B enhancer is demonstrating this transition. This enhancer still does demonstrate requirement of bHLH-R input, as illustrated by the observation of derepression upon mutating the single site present, similar to the 4D enhancer. Unlike the 4D enhancer, however, the derepression is not dependent upon intact proneural sites, of which there is only one in *Drosophila melanogaster*. This would support a more complete transition of the enhancer away from strict proneural requirements. Indeed, the loss of a second proneural consensus sequence since the last common ancestor between *Drosophila melanogaster* and *Drosophila ananassae* may be indicative of such a shift occurring in the relatively recent evolutionary past.

NF-κB regulation of SOP specification: done in by a charlatan*?*

Even in this rather extended study (thanks for reading this far, by the way), we have still not resolved the identity of the factor(s) that

bind to the SMCα sequences. Here we have presented further evidence for the importance of these sequences in establishing an SOP specificity, but have raised two-fold doubt over the proposed NF-κB + P synergy within the SOP (or pre-SOP). The first was demonstrated by the apparent lack of nuclear localization in the SOP of any of the fluorescently-tagged *Drosophila* NF-κB paralogues, even though such localization is detectable in the third instar wing imaginal disc with these reagents. We acknowledge, however, that only a small subset of the protein may need to be nuclear in order to achieve the necessary synergy on SOP-specific targets, and such an amount may be below the level of detection. Moreover, lack of SOP nuclear localization does not preclude NF-κB proteins from participating in the non-SOP mRNA degradatory role as has been proposed; indeed, such an activity may also not even require nuclear localization, and our results therefore cannot rule out such a function.

The additional level of doubt cast upon the NF-κB + P SOP synergy comes from the first demonstration of any protein interacting with the SMCα sites *in vitro*, in this case Chn. In our hands, the interaction is very weak, although for such a synergistic interaction to

occur as has been proposed, a strong protein-to-DNA interaction may not be absolutely required and may in fact require combinatorial protein-DNA and protein-protein contacts with additional factors to facilitate optimum binding—such a case has been seen for other factors (Muller et al., 2003; Barrick and Kopan, 2006; Nam et al., 2006). Furthermore, *chn* loss of function and misexpression phenotypes, as well as the expression and regulatory information seem too optimal for what would be envisioned for the SMCα site binding factor: expressed in the proneural cluster downstream of the proneurals; receiving direct input from the proneurals for full activity (Reeves and Posakony, 2005); the ability to promote proneural transcript accumulation upon misexpression (as well as neur4DWT expression—this study); the requirement for full proneural activation; and the ability to affect SOP specification (Escudero et al., 2005). Nevertheless, we have not presented here the evidence sufficient to either fully crown Chn as the SMCα site binding factor, or to fully eliminate a NF-κB mediated contribution to SOP specification; indeed the truth may even reveal a combination of the two models.

Refining the Dl model

We introduced this study in part with a dissatisfaction of the current *Dl*-focused model for the combinatorial activities of the proneural proteins and the N signaling pathway to establish the two ultimate fates that arise from the PNC. We believe our work provides further support against such a model and rather favors one that has *neur*, rather than *Dl*, as a major focus. Moreover, a *neur*-centric model fits one of the major pieces of evidence against the *Dl*-centric model; that uniform Dl expression appears not to have consequences for lateral inhibition. If instead the focus is on establishing strict *neur* specificity, Dl could certainly be expressed uniformly as long as it was being activated by Neur-mediated endocytosis in a highly restricted manner. But now the question may be shifted back to *Dl*: if uniform Dl expression is non-essential for lateral inhibition and SOP-specific upregulation of *Dl* has not been observed, then why have regulatory DNA sequences been identified that provide SOP and/or neuroblast-specific expression? Perhaps our second-take at nascent *Dl* expression may provide at least some superficial explanation. Unlike *neur*, which appears to accumulate prior to SOP singularization, locally

elevated *Dl* expression in the SOP seems to coincide with the expression of markers for SOP specification, raising the possibility that *Dl* transcript may indeed be transiently upregulated in the SOP, despite what has been described previously. We acknowledge this contrasts the situation observed in neuroblasts, which downregulate *Dl* upon specification (Haenlin et al., 1990), and also we have not determined if this is merely due to a cellular response to the massive degradation of Dl protein taking place after endocytosis in the SOP or is in preparation for binary signaling in the lineage or is just misinterpretation of an *in situ* artifact.

Shadowing neuralized *enhancer evolution*

Perhaps the most surprising--and from our perspective annoying--result that has come from our analysis of *neur* regulation is the discovery of not just one, but two enhancers with strong overlapping specificity regulating *neur* expression in the SOP. The notion of a single gene receiving input from multiple, independent cis-regulatory modules has been well-established; a more recent observation has been the identification of a number of enhancers of

segmentation pathway genes which contain both similar functional trans inputs as demonstrated by ChIP-chip but also similar specificity when placed in front of a reporter gene (Hong et al., 2008). These "shadow enhancers," have been found both nearby and at great distance from the genes they are thought to control; none, however, have been tested for their ability to demonstrate functional requirements in regulation of their cognate gene in a representative genomic context. Thus, this study represents the first such case in *Drosophila* for which duplicate enhancers with overlapping specificity have not only been identified, but also shown to be functionally redundant. A similar demonstration has been described only for the regulation of the TCRγ locus (Xiong et al., 2002), but we imagine that similar analyses will emerge in the coming years; duplicate, partially-redundant regulatory modules may not prove to be exceedingly rare.

The current challenge is to satisfactorily explain why redundant enhancers exist in the genome in the first place. The model that has been suggested is that duplicate cis-regulatory modules (dCRMs—we do not favor the term "shadow enhancer" because of the implied bias toward one enhancer over the other) are a stepping-stone toward

regulatory novelty without compromising necessary functions (Hong et al., 2008). With this model, however, one of the two dCRMs should be able to rapidly evolve new specificity or even lose functionality entirely; yet the described shadow enhancers all appear to exhibit strong conservation across the 12 genomes. The neur1B module described here may actually be illustrating the selective pressure dCRMs should be subject to. Not only do we observe complete inversion of the DNA sequence of this model within the last common ancestor of all twelve species, but also recent loss of a proneural site in *Drosophila melanogaster* and even the apparent loss of proneural site-dependent derepression. Moreover, we also observe both the presence of Twist binding consensuses (data not shown) and presumed muscle expression of the 1B enhancer in the pupal notum. Perhaps the 1B enhancer in *Drosophila melanogaster* may be in the process of transitioning to increasing emphasis on muscle specificity, or vice versa. Muscle founder cells in the embryo are known undergo a process similar to lateral inhibition (Park et al., 1998); it may not been too surprising if over the course of evolution muscle and PNS dCRMs are able to swap specificities.

dCRMs as a theme in peripheral nervous system development?

In the course of this study, combined with the investigation of the P + R code, we may have stumbled upon a theme of dCRMs in SOP genes. In the cases of *phyl* and *sens*, both the enhancers we have found by the P + R dual-factor SCORE approach are actually dCRMs distinct from the published enhancers for these genes (M. Rebeiz and J. Posakony, unpublished data) (Jafar-Nejad et al., 2003; Pi et al., 2004). As with the two *neur* SOP enhancers, these dCRMs appear to have similar overall specificity, but with some differences in relative expression level. With all three of these factors participating in high level regulation within the SOP, it is not surprising that they may evolve mechanisms of maintaining multiple layers of regulation to achieve sharp, reliable SOP activation. However, at this stage we do not yet understand if these represent true functional redundancies, transitioning enhancers, or modules representing discrete yet necessary subtle temporal specificities. The neur1B enhancer could be an example of this last point. Since we observe a much greater degree of non-SOP expression in both the wild type versions of the

NRS1 and 1B regulatory fragments than we ever observe with 4DWT, the 1B region may actually be contributing the majority of the "pre-SOP" expression that we observe with *neur* transcript. Under such a model, 1B would modulate the low level PNC *neur* expression followed by a transition to use of 4D to facilitate strong SOP-specific upregulation.

Studying enhancers in their genomic context

As a final parting note, our study may provide a cautionary tale toward regulatory analyses in the future. To date the majority of functional enhancer studies have been performed with reporter construct subfragments taken out of the surrounding genomic context. Such work has proved enormously valuable to our construction of gene regulatory networks as we continue to decode the mystery of the genome, as well as our understanding of enhancer function on the most basic terms. It is not gone without note, then, the differences between our results using such reporter constructs, compared with a translation to a larger genomic context. Our first observation was that mutation of the repressor sites within a single enhancer, while clearly

resulting in mis-regulation of the reporter construct, failed to generate the corresponding phenotype using the rescue construct when assayed by *neur* mRNA expression. Even what should be the most basic test of bHLH-R site functionality in which all the bHLH-R consensus sequences in each of the two enhancers are mutated did not produce as strong a phenotype as would have been predicted based upon the reporter assays. One possible explanation for this apparent discrepancy could simply be differences in detection. Part of the reason for the popularity of reporter gene tools is their strong expression, often much more easily detectable than a large number of endogenous gene transcripts. Indeed, this does appear to be a contributing factor since we were able to more readily detect ectopic Neur expression using the GFP moiety. But one still might wonder why the lateral inhibition phenotype upon bHLH-R site mutation is not more robust.

Part of the reason why a mutant phenotype is more readily seen could be due to the very nature of the SOP program downstream of the proneurals, and the influence of the bHLH-R proteins on the rest of the program. Independent of *neur* expression the proneurals are

accumulating at higher levels in a subset of the cluster and trying to activate their downstream SOP program. Once one of these downstream factors is activated, our hypothesis would suggest it would further contribute to *neur* activation, which may be just enough of a boost to create a difference in expression levels that between the SOP and the cells ectopically expressing *neur*. This would create just enough difference between the two that the ectopically expressing cell may activate a greater amount of BFM expression that quells its potential to fully become an SOP, even though it briefly maintains *neur* expression. The fact that the same bristle positions in which we observe ectopic Neur-GFP expression are more susceptible to lateral inhibition defects upon having lost both bHLH-R regulation of *neur* and depletion of two BFMs may be indicative of such a model. We note that recently all BFMs have been systematically deleted from the genome with no apparent consequence (Chanet et al., 2009); perhaps this genetic background may reveal even more striking phenotypes upon loss of bHLH-R regulation of *neur*.

A final contributing factor may be one of the least well understood phenomena in gene regulation: that of the local

architecture within the nucleus. We note that while we mutated all the bHLH-R sites within the enhancers that provide SOP specificity, there still exists a strongly conserved site upstream of neur4D that had remained intact in the rescue context. If multiple regulatory regions are able to loop toward the promoter, it is unclear if enhancers of different specificity are able to influence one another. It has been documented that multiple tandem modules that direct *brk* expression appear to be able to overcome derepression within one of the modules (Yao et al., 2008), although all these modules have the same specificity. Nevertheless, this would certainly explain a weak phenotype for Rm versions of individual enhancers in the rescue context, and may be contributing to weakening the 4D,1BRm phenotype as well.

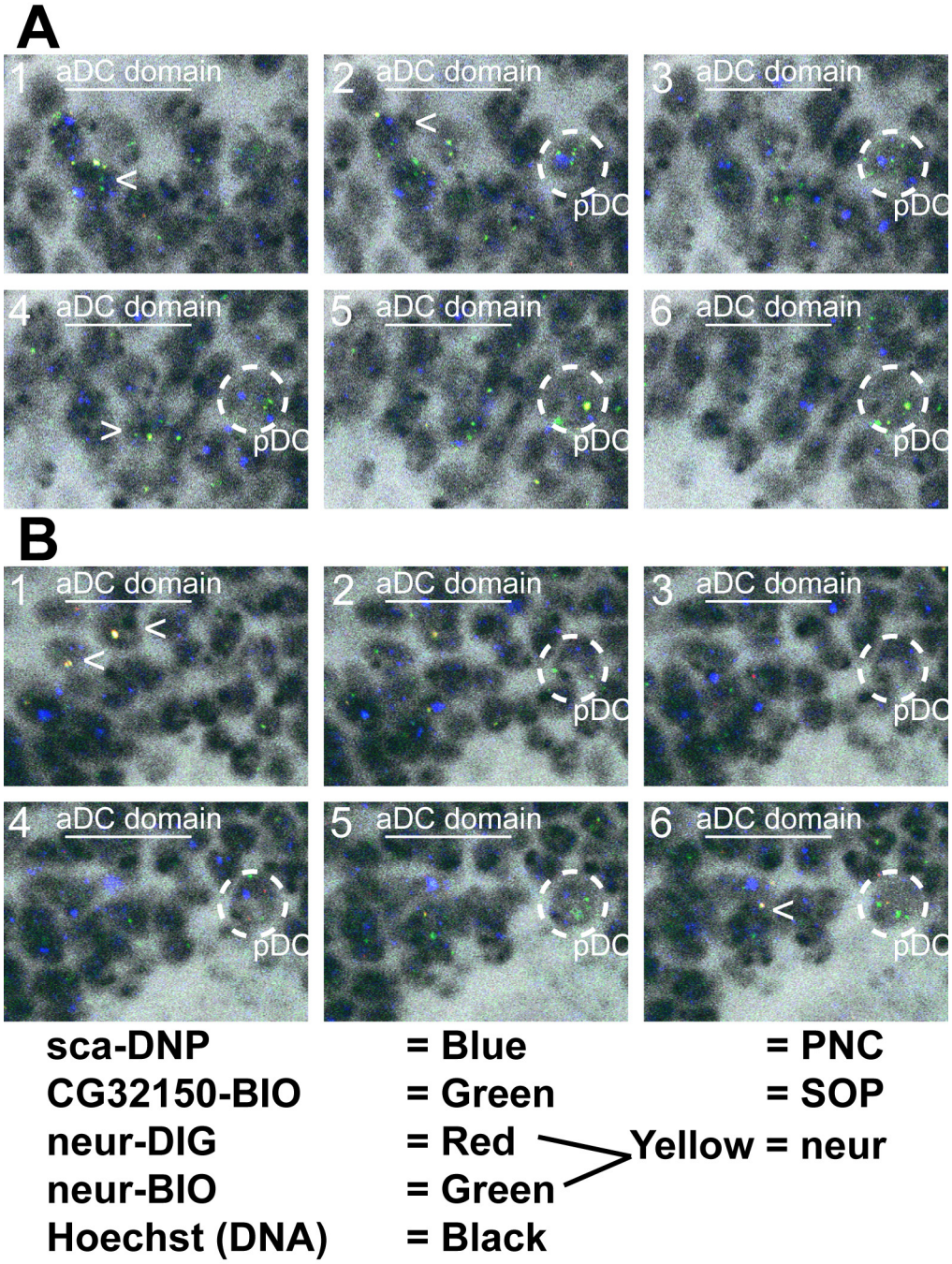
The other surprising result to come from the rescue construct experiments is the phenotypic disparity between complete deletion of both the 4D and 1B enhancers with that of a *neur* loss of function phenotype in the SOP. In the latter case, loss of Neur function causes a failure of lateral inhibition and the subsequent production of supernumerary SOPs. Each of these cells divides, but lack of Neur function in the pIIb cell fails to specify a pIIa cell, causing all progeny to

adopt an internal cell fate at the expense of external cells. However, the 4D,1Bdel construct only appears to affect the lateral inhibition function of *neur*. Apparently all the SOPs that develop proceed with lineage specification completely normally. This phenotype may be the result of incomplete deletion of either enhancer, or from a low basal level of expression of *neur* that is insufficient for lateral inhibition. The differences in reporter expression strength between NRS1 and 1B may suggest the former as the more plausible explanation, and that seems the more attractive model to confirm. However, we do note that in our *neur* fluorescent *in situ* experiments we did observe the apparent coincidence of weak probe signal in broad regions outside the proneural clusters. As with the signal within clusters, this low level expression appears to have been confirmed using the GFP-tagged rescue construct (Supplemental Figure 2.21).

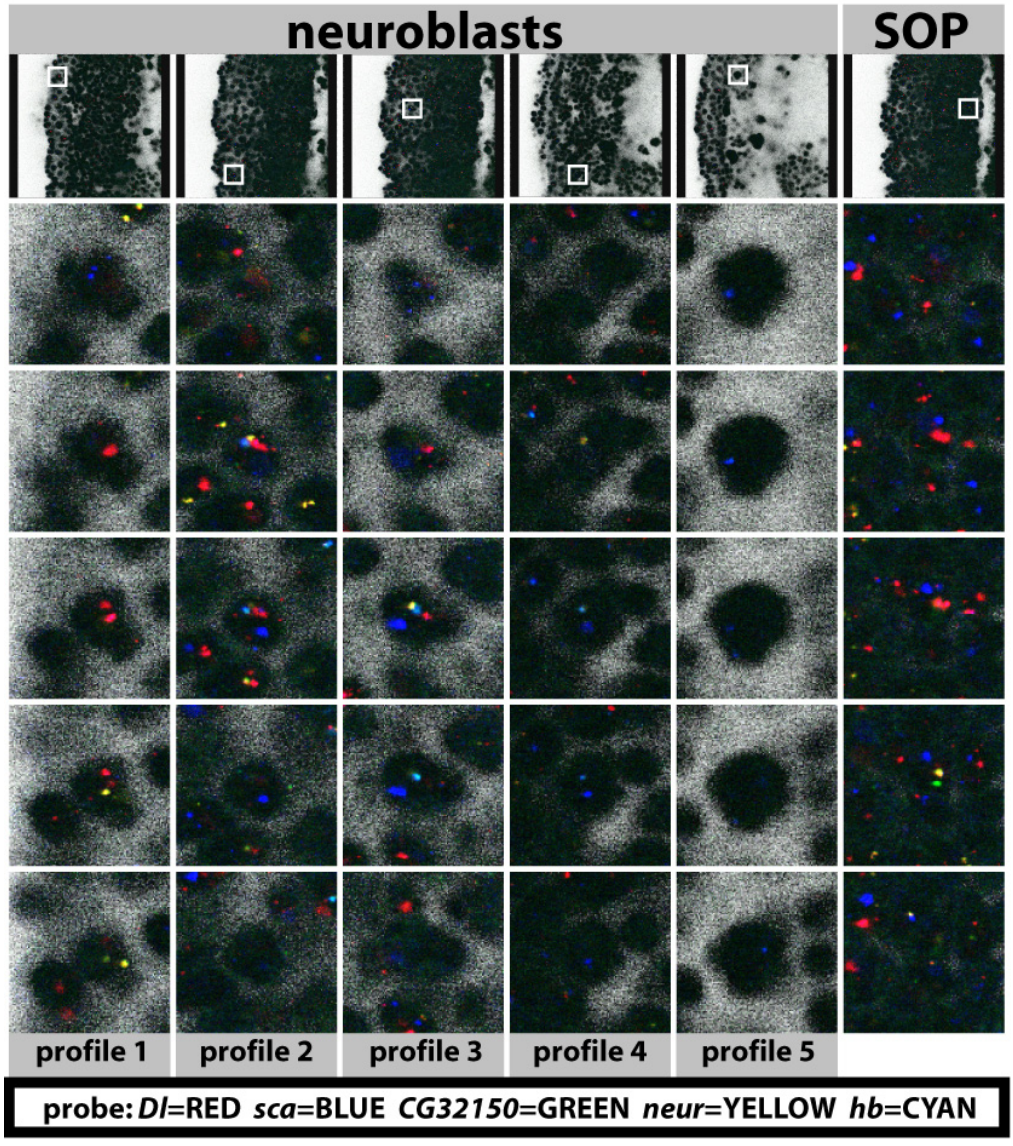
Lastly, the lack of a lineage phenotype may have directed our attention to the existence of an additional module that directs *neur* expression in a specific pattern in the cells downstream of the SOP division. As the lineage precursors divide, Neur protein is asymmetrically segregated to one of the two daughters (Le Borgne and

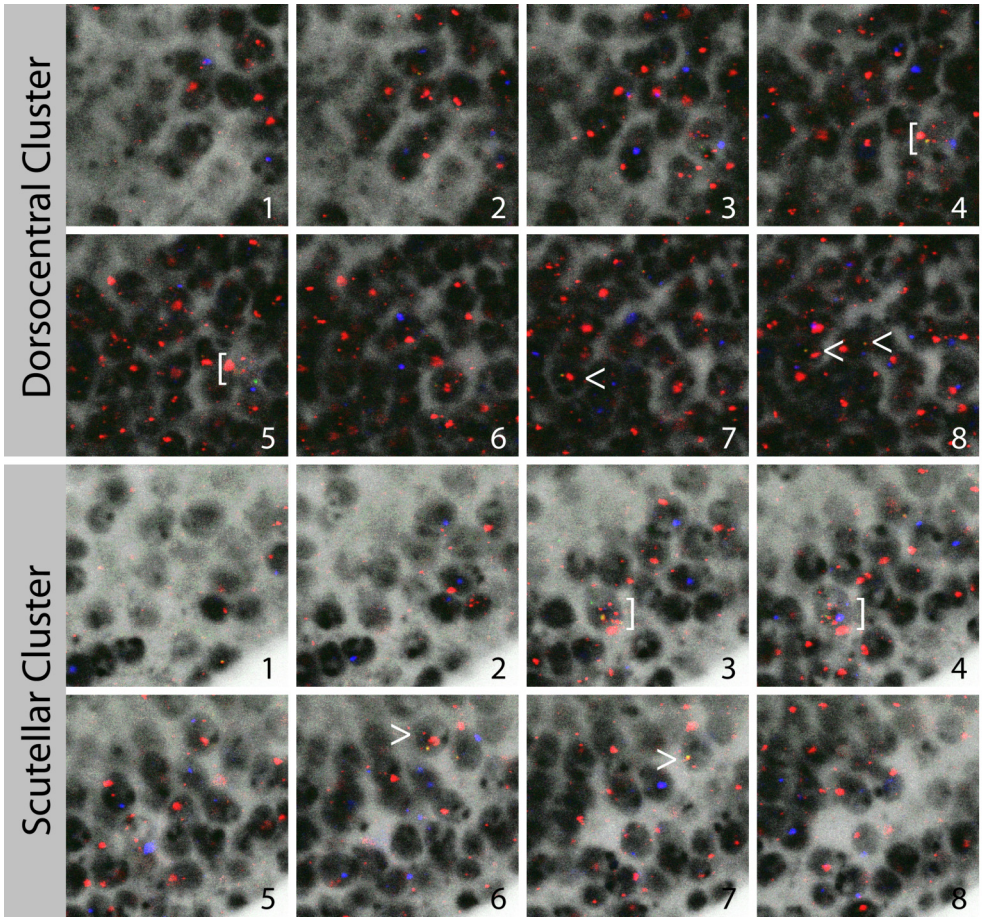
Schweisguth, 2003), the N-independent cell fate, similar to what has been observed for the protein Numb. Like Numb, the N-dependent daughter precursor still would need to make new Neur so it can be segregated to one of its daughters. This activity has been discovered to be directed by a N-responsive enhancer in the *numb* locus (M. Rebeiz, S. Miller, and J. Posakony, unpublished data); since Neur encounters a similar problem, it would not be surprising if it has developed a similar such CRM. Even if such an enhancer exist for *neur*, the prediction would still be that the first lineage signaling event--between pIIa and pIIb—would be affected by the 4D,1Bdel mutant if it fails to activate *neur* expression in the SOP; thus the more plausible explanations remain the two described above.

Supplemental Figure 2.17: Additional examples of *neur* expressing cells in the presumptive SOP domain.
Both A and B are additional discs as described in Figure 2.7B. Panels are six adjacent 1 micron section through DC cluster. Probes are labeled below the images with the indicated genes, haptens, and resulting color in the images. The specified pDC SOP is circled, and when cells in the aDC domain have a *neur* nascent transcript dot they are pointed to with an arrowhead.

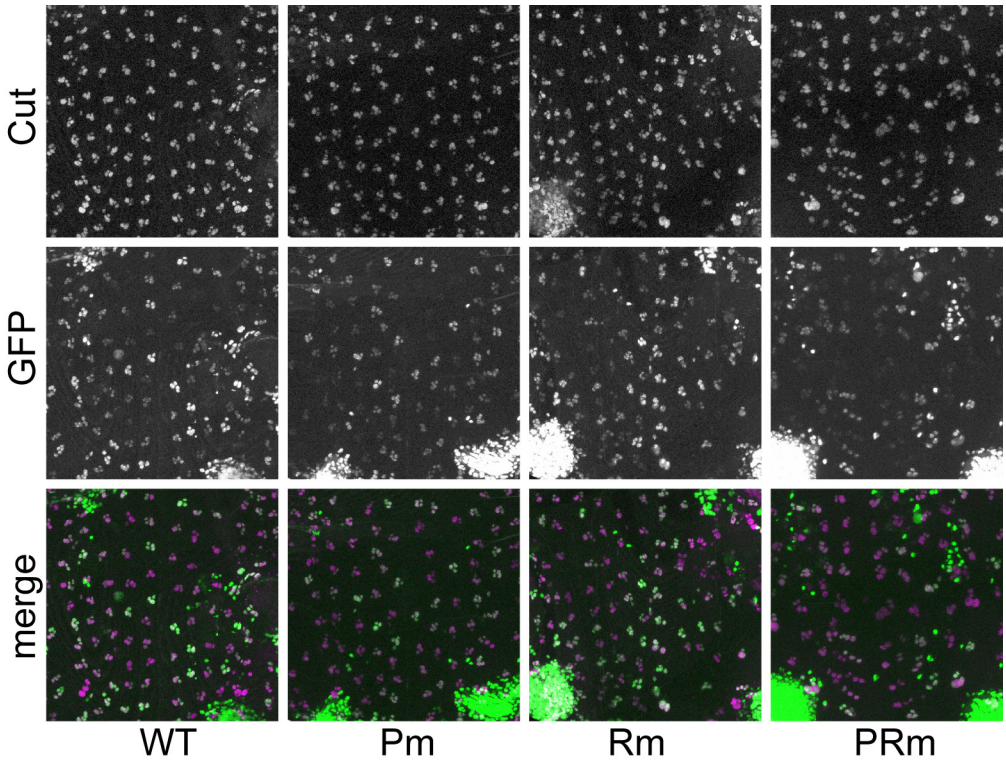

A
1 aDC domain
2 aDC domain
3 aDC domain
4 aDC domain
5 aDC domain
6 aDC domain
pDC
B
1 aDC domain
2 aDC domain
3 aDC domain
4 aDC domain
5 aDC domain
6 aDC domain
pDC
sca-DNP = Blue = PNC
CG32150-BIO = Green = SOP
neur-DIG = Red
neur-BIO = Green
Yellow = neur
Hoechst (DNA) = Black

Supplemental Figure 2.18: *neur* and *Dl* expression in the embryo. Presented are five different profiles of *neur* and *Dl* expression observed in embryonic neuroblasts in a single embryo. For each profile, the first row indicates the location of the cell, of which five z sections are shown to illustrate all expressing genes. The last column at the right is taken from the lateral peripheral nervous system. The neuroblast profiles are as follows: (1) Strong *sca*, *neur*, *Dl*, but no detectable *hb* expression. (2) Similar to 1, with stronger *Dl* signal and strong *hb* expression. (3) Similar to 1 in all respects but strong *hb* expression. (4) Weaker *neur*, detectable *sca* and *hb*, but no *Dl* signal. (5) Only detectable *hb*. The SOP pattern in the CNS displays strong signal for *sca*, *CG32150*, *neur*, and *Dl* appear elevated relative to adjacent cells, although not quantitative.

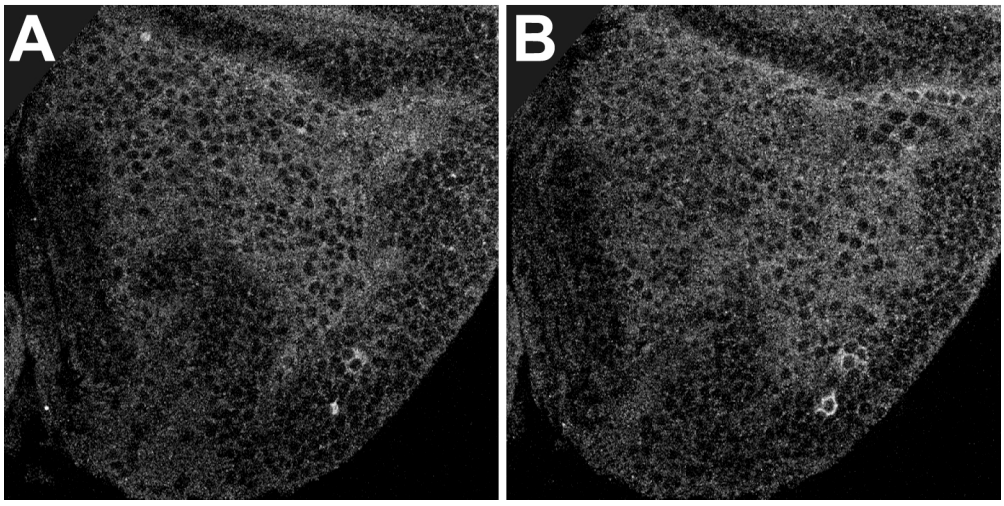

neuroblasts
SOP
profile 1
profile 2
profile 3
profile 4
profile 5
probe: Dl=RED sca=BLUE CG32150=GREEN neur=YELLOW hb=CYAN



Supplemental Figure 2.19: *Dl* expression in heterochronic SOP pairs.
Shown are eight adjacent confocal sections through either the DC cluster or the SC cluster. Probes are the same as in Figure 2.8, although in 2.8 the *Dl* probe was 100 fold more dilute to better highlight apparent differences in expression level between adjacent cells. For each cluster, the specified SOP (pDC or pSC) is indicated by a bracket. Cells in the presumptive domain for the second SOP which express *neur* are pointed to with an arrowhead.



Supplemental Figure 2.20: The 1B enhancer directs reporter expression that accumulates in the lineage independently of the intact proneural binding site.
Expression of 1B reporter variants in 24 hr APF nota, assaying for both expression of Cut (top row) to mark location of microchaete positions, and GFP (bottom row). Reporter variant is labeled at the bottom.



Supplemental Figure 2.21: Low level, ubiquitous expression of neurRC-WT-GFP.
(A) Wide view of GFP signal from the same single confocal section as is seen in Figure 2.7, panel F. (B) As with A, wider view of the same confocal section as in Figure 2.7, panels G and H. In both panels, note the expression throughout the notum region in view, with the same subcellular localization as is seen with the high expressing SOP regions.

References

Ayyar, S. *et al.* (2007). NF-kappaB/Rel-mediated regulation of the neural fate in Drosophila. PLoS One *2*, e1178.

Bailey, A. M., and Posakony, J. W. (1995). Suppressor of hairless directly activates transcription of enhancer of split complex genes in response to Notch receptor activity. Genes Dev *9*, 2609-2622.

Bang, A. G. *et al.* (1991). Hairless is required for the development of adult sensory organ precursor cells in Drosophila. Development *111*, 89-104.

Bardin, A. J., and Schweisguth, F. (2006). Bearded family members inhibit Neuralized-mediated endocytosis and signaling activity of Delta in Drosophila. Dev Cell *10*, 245-255.

Barolo, S. *et al.* (2000). GFP and beta-galactosidase transformation vectors for promoter/enhancer analysis in Drosophila. Biotechniques *29*, 726, 728, 730, 732.

Barolo, S. *et al.* (2004). New Drosophila transgenic reporters: insulated P-element vectors expressing fast-maturing RFP. Biotechniques *36*, 436-40, 442.

Barrick, D., and Kopan, R. (2006). The Notch transcription activation complex makes its move. Cell *124*, 883-885.

Bettencourt, R. *et al.* (2004). Hemolymph-dependent and -independent responses in Drosophila immune tissue. J Cell Biochem *92*, 849-863.

Bischof, J. *et al.* (2007). An optimized transgenesis system for Drosophila using germ-line-specific phiC31 integrases. Proc Natl Acad Sci U S A *104*, 3312-3317.

Boulianne, G. L. *et al.* (1991). The Drosophila neurogenic gene neuralized encodes a novel protein and is expressed in precursors of larval and adult neurons. EMBO J *10*, 2975-2983.

——— (1993). The Drosophila neurogenic gene neuralized encodes a novel protein and is expressed in precursors of larval and adult neurons. EMBO J *12*, 2586.

Bray, S. J. (2006). Notch signalling: a simple pathway becomes complex. Nat Rev Mol Cell Biol *7*, 678-689.

Cabrera, C. V. (1990). Lateral inhibition and cell fate during neurogenesis in Drosophila: the interactions between scute, Notch and Delta. Development *110*, 733-742.

Cabrera, C. V., and Alonso, M. C. (1991). Transcriptional activation by heterodimers of the achaete-scute and daughterless gene products of Drosophila. EMBO J *10*, 2965-2973.

Cabrera, C. V. *et al.* (1994). Regulation of scute function by extramacrochaete in vitro and in vivo. Development *120*, 3595-3603.

Castro, B. *et al.* (2005). Lateral inhibition in proneural clusters: cis-regulatory logic and default repression by Suppressor of Hairless. Development *132*, 3333-3344.

Castro, D. S. *et al.* (2006). Proneural bHLH and Brn proteins coregulate a neurogenic program through cooperative binding to a conserved DNA motif. Dev Cell *11*, 831-844.

Chanet, S. *et al.* (2009). Genome Engineering-based Analysis of Bearded Family Genes Reveals Both Functional Redundancy and a Non-essential Function in Lateral Inhibition in Drosophila. Genetics

Culi, J., and Modolell, J. (1998). Proneural gene self-stimulation in neural precursors: an essential mechanism for sense organ

development that is regulated by Notch signaling. Genes Dev *12*, 2036-2047.

de Celis, J. F., and Bray, S. (1997). Feed-back mechanisms affecting Notch activation at the dorsoventral boundary in the Drosophila wing. Development *124*, 3241-3251.

De Renzis, S. *et al.* (2006). Dorsal-ventral pattern of Delta trafficking is established by a Snail-Tom-Neuralized pathway. Dev Cell *10*, 257-264.

Deblandre, G. A. *et al.* (2001). Xenopus neuralized is a ubiquitin ligase that interacts with XDelta1 and regulates Notch signaling. Dev Cell *1*, 795-806.

Delidakis, C., and Artavanis-Tsakonas, S. (1992). The Enhancer of split [E(spl)] locus of Drosophila encodes seven independent helix-loop-helix proteins. Proc Natl Acad Sci U S A *89*, 8731-8735.

Desai, A. R., and McConnell, S. K. (2000). Progressive restriction in fate potential by neural progenitors during cerebral cortical development. Development *127*, 2863-2872.

Escudero, L. M. *et al.* (2005). Charlatan, a Zn-finger transcription factor, establishes a novel level of regulation of the proneural achaete/scute genes of Drosophila. Development *132*, 1211-1222.

Fischer, J. A. *et al.* (1988). GAL4 activates transcription in Drosophila. Nature *332*, 853-856.

Fontana, J. R., and Posakony, J. W. (2009). Both inhibition and activation of Notch signaling rely on a conserved Neuralized-binding motif in Bearded proteins and the Notch ligand Delta. Dev Biol

Golic, K. G. (1991). Site-specific recombination between homologous chromosomes in Drosophila. Science *252*, 958-961.

Golic, K. G., and Lindquist, S. (1989). The FLP recombinase of yeast catalyzes site-specific recombination in the Drosophila genome. Cell *59*, 499-509.

Grens, A. *et al.* (1995). Evolutionary conservation of a cell fate specification gene: the Hydra achaete-scute homolog has proneural activity in Drosophila. Development *121*, 4027-4035.

Groth, A. C. *et al.* (2004). Construction of transgenic Drosophila by using the site-specific integrase from phage phiC31. Genetics *166*, 1775-1782.

Guo, Q. *et al.* (2003). Fate map of mouse ventral limb ectoderm and the apical ectodermal ridge. Dev Biol *264*, 166-178.

Haenlin, M. *et al.* (1990). The pattern of transcription of the neurogenic gene Delta of Drosophila melanogaster. Development *110*, 905-914.

Heitzler, P. *et al.* (1996). Genes of the Enhancer of split and achaete-scute complexes are required for a regulatory loop between Notch and Delta during lateral signalling in Drosophila. Development *122*, 161-171.

Ho, S. N. *et al.* (1989). Site-directed mutagenesis by overlap extension using the polymerase chain reaction. Gene *77*, 51-59.

Hong, J. W. *et al.* (2008). Shadow enhancers as a source of evolutionary novelty. Science *321*, 1314.

Huang, F. *et al.* (1991). The emergence of sense organs in the wing disc of Drosophila. Development *111*, 1087-1095.

Jafar-Nejad, H. *et al.* (2003). Senseless acts as a binary switch during sensory organ precursor selection. Genes Dev *17*, 2966-2978.

Jafar-Nejad, H. *et al.* (2006). Senseless and Daughterless confer neuronal identity to epithelial cells in the Drosophila wing margin. Development *133*, 1683-1692.

Jarman, A. P. *et al.* (1994). Atonal is the proneural gene for Drosophila photoreceptors. Nature *369*, 398-400.

Jennings, B. *et al.* (1994). The Notch signalling pathway is required for Enhancer of split bHLH protein expression during neurogenesis in the Drosophila embryo. Development *120*, 3537-3548.

Jennings, V. *et al.* (1995). Role of Notch and achaete-scute complex in the expression of Enhancer of split bHLH proteins. Development *121(11)*, 3745-3752.

Klambt, C. *et al.* (1989). Closely related transcripts encoded by the neurogenic gene complex enhancer of split of Drosophila melanogaster. EMBO J *8*, 203-210.

Klein, T. *et al.* (1997). An intrinsic dominant negative activity of serrate that is modulated during wing development in Drosophila. Dev Biol *189*, 123-134.

Knust, E. *et al.* (1992). Seven genes of the Enhancer of split complex of Drosophila melanogaster encode helix-loop-helix proteins. Genetics *132*, 505-518.

Kosman, D. *et al.* (2004). Multiplex detection of RNA expression in Drosophila embryos. Science *305*, 846.

Kramatschek, B., and Campos-Ortega, J. A. (1994). Neuroectodermal transcription of the Drosophila neurogenic genes E(spl) and HLH-m5 is regulated by proneural genes. Development *120*, 815-826.

Kunisch, M. *et al.* (1994). Lateral inhibition mediated by the Drosophila neurogenic gene delta is enhanced by proneural proteins. Proc Natl Acad Sci U S A *91*, 10139-10143.

Lai, E. C. *et al.* (2000). The enhancer of split complex of Drosophila includes four Notch-regulated members of the bearded gene family. Development *127*, 3441-3455.

Lai, E. C. *et al.* (2001). Drosophila neuralized is a ubiquitin ligase that promotes the internalization and degradation of delta. Dev Cell *1*, 783-794.

Le Borgne, R., and Schweisguth, F. (2003). Unequal segregation of Neuralized biases Notch activation during asymmetric cell division. Dev Cell *5*, 139-148.

Lee, T. *et al.* (2000). Essential roles of Drosophila RhoA in the regulation of neuroblast proliferation and dendritic but not axonal morphogenesis. Neuron *25*, 307-316.

Lehmann, R. *et al.* (1981). Mutations of early neurogenesis inDrosophila. Development Genes and Evolution *190*, 226-229.

Lehmann, R. *et al.* (1983). On the phenotype and development of mutants of early neurogenesis inDrosophila melanogaster. Development Genes and Evolution *192*, 62-74.

Leviten, M. W. *et al.* (1997). The Drosophila gene Bearded encodes a novel small protein and shares 3' UTR sequence motifs with multiple Enhancer of split complex genes. Development *124*, 4039-4051.

Leyns, L. *et al.* (1989). Two different sets of cis elements regulate scute to establish two different sensory patterns. Development Genes and Evolution *198*, 227-232.

Li, Y., and Baker, N. E. (2004). The roles of cis-inactivation by Notch ligands and of neuralized during eye and bristle patterning in Drosophila. BMC Dev Biol *4*, 5.

Markstein, M. *et al.* (2008). Exploiting position effects and the gypsy retrovirus insulator to engineer precisely expressed transgenes. Nat Genet *40*, 476-483.

Mercado-Pimentel, M. E. *et al.* (2002). Affinity purification of GST fusion proteins for immunohistochemical studies of gene expression. Protein Expr Purif *26*, 260-265.

Meyerowitz, E. M. (1999). Plants, animals and the logic of development. Trends Cell Biol *9*, M65-8.

Micchelli, C. A. *et al.* (1997). The function and regulation of cut expression on the wing margin of Drosophila: Notch, Wingless and a dominant negative role for Delta and Serrate. Development *124*, 1485-1495.

Miller, S. W. *et al.* (2009). Complex interplay of three transcription factors in controlling the tormogen differentiation program of Drosophila mechanoreceptors. Dev Biol *329*, 386-399.

Muller, B. *et al.* (2003). Conversion of an extracellular Dpp/BMP morphogen gradient into an inverse transcriptional gradient. Cell *113*, 221-233.

Murre, C. *et al.* (1989). Interactions between heterologous helix-loop-helix proteins generate complexes that bind specifically to a common DNA sequence. Cell *58*, 537-544.

Nam, Y. *et al.* (2006). Structural basis for cooperativity in recruitment of MAML coactivators to Notch transcription complexes. Cell *124*, 973-983.

Nellesen, D. T. *et al.* (1999). Discrete enhancer elements mediate selective responsiveness of enhancer of split complex genes to common transcriptional activators. Dev Biol *213*, 33-53.

Nolo, R. *et al.* (2000). Senseless, a Zn finger transcription factor, is necessary and sufficient for sensory organ development in Drosophila. Cell *102*, 349-362.

O'Neill, J. W., and Bier, E. (1994). Double-label in situ hybridization using biotin and digoxigenin-tagged RNA probes. Biotechniques *17*, 870, 874-5.

Park, M. *et al.* (1998). Mesodermal cell fate decisions in Drosophila are under the control of the lineage genes numb, Notch, and sanpodo. Mech Dev *75*, 117-126.

Parks, A. L. *et al.* (1997). The dynamics of neurogenic signalling underlying bristle development in Drosophila melanogaster. Mech Dev *63*, 61-74.

Pi, H. *et al.* (2004). phyllopod is a target gene of proneural proteins in Drosophila external sensory organ development. Proc Natl Acad Sci U S A *101*, 8378-8383.

Pinaudeau, C. *et al.* (2000). Stage of specification of the spinal cord and tectal projections from cortical grafts. Eur J Neurosci *12*, 2486-2496.

Pitsouli, C., and Delidakis, C. (2005). The interplay between DSL proteins and ubiquitin ligases in Notch signaling. Development *132*, 4041-4050.

Rebeiz, M., and Posakony, J. W. (2004). GenePalette: a universal software tool for genome sequence visualization and analysis. Dev Biol *271*, 431-438.

Rebeiz, M. *et al.* (2002). SCORE: a computational approach to the identification of cis-regulatory modules and target genes in whole-genome sequence data. Site clustering over random expectation. Proc Natl Acad Sci U S A *99*, 9888-9893.

Reeves, N., and Posakony, J. W. (2005). Genetic programs activated by proneural proteins in the developing Drosophila PNS. Dev Cell *8*, 413-425.

Rodriguez, I. *et al.* (1990). Competence to develop sensory organs is temporally and spatially regulated in Drosophila epidermal primordia. EMBO J *9*, 3583-3592.

Romani, S. *et al.* (1989). Expression of achaete and scute genes in Drosophila imaginal discs and their function in sensory organ development. Genes Dev *3*, 997-1007.

Rong, Y. S., and Golic, K. G. (2001). A targeted gene knockout in Drosophila. Genetics *157*, 1307-1312.

Rubin, G. M., and Spradling, A. C. (1982). Genetic transformation of Drosophila with transposable element vectors. Science *218*, 348-353.

Ruiz-Gomez, M., and Modolell, J. (1987). Deletion analysis of the achaete-scute locus of Drosophila melanogaster. Genes Dev *1*, 1238-1246.

Sakamoto, K. *et al.* (2002). Intracellular cell-autonomous association of Notch and its ligands: a novel mechanism of Notch signal modification. Dev Biol *241*, 313-326.

Sandelin, A. *et al.* (2004). JASPAR: an open-access database for eukaryotic transcription factor binding profiles. Nucleic Acids Res *32*, D91-4.

Seugnet, L. *et al.* (1997). Transcriptional regulation of Notch and Delta: requirement for neuroblast segregation in Drosophila. Development *124*, 2015-2025.

Simpson, P. (1990). Lateral inhibition and the development of the sensory bristles of the adult peripheral nervous system of Drosophila. Development *109*, 509-519.

Singson, A. *et al.* (1994). Direct downstream targets of proneural activators in the imaginal disc include genes involved in lateral inhibitory signaling. Genes Dev *8*, 2058-2071.

Skeath, J. B., and Carroll, S. B. (1991). Regulation of achaete-scute gene expression and sensory organ pattern formation in the Drosophila wing. Genes Dev *5*, 984-995.

Thummel, C. S. *et al.* (1988). Vectors for Drosophila P-element-mediated transformation and tissue culture transfection. Gene *74*, 445-456.

Tsuda, L. *et al.* (2006). An NRSF/REST-like repressor downstream of Ebi/SMRTER/Su(H) regulates eye development in Drosophila. EMBO J *25*, 3191-3202.

Usui, K., and Kimura, K. (1993). Sequential emergence of the evenly spaced microchaetes on the notum of Drosophila. Development Genes and Evolution *203*, 151-158.

Van Doren, M. *et al.* (1991). The Drosophila extramacrochaetae protein antagonizes sequence-specific DNA binding by daughterless/achaete-scute protein complexes. Development *113*, 245-255.

Van Doren, M. *et al.* (1992). Spatial regulation of proneural gene activity: auto- and cross-activation of achaete is antagonized by extramacrochaetae. Genes Dev *6*, 2592-2605.

Venken, K. J. *et al.* (2006). P[acman]: a BAC transgenic platform for targeted insertion of large DNA fragments in D. melanogaster. Science *314*, 1747-1751.

Warming, S. *et al.* (2005). Simple and highly efficient BAC recombineering using galK selection. Nucleic Acids Res *33*, e36.

Xiong, N. *et al.* (2002). Redundant and unique roles of two enhancer elements in the TCRgamma locus in gene regulation and gammadelta T cell development. Immunity *16*, 453-463.

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

Yao, L. C. *et al.* (2008). Multiple modular promoter elements drive graded brinker expression in response to the Dpp morphogen gradient. Development *135*, 2183-2192.

Yeh, E. *et al.* (2001). Neuralized functions as an E3 ubiquitin ligase during Drosophila development. Curr Biol *11*, 1675-1679.

Yeh, E. *et al.* (2000). Neuralized functions cell autonomously to regulate Drosophila sense organ development. EMBO J *19*, 4827-4837.

Zaffran, S., and Frasch, M. (2000). Barbu: an E(spl) m4/m(alpha)-related gene that antagonizes Notch signaling and is required for the establishment of ommatidial polarity. Development *127*, 1115-1130.

Chapter Two, in full, is a manuscript in preparation for publication: Cell non-autonomous auto-regulation of *neuralized* during *Drosophila* mechanosensory organ precursor lateral inhibition. Miller, S. W. and Posakony, J. W. The dissertation author was the primary researcher and author. James Posakony provided critical comments and oversight through the course of this investigation. Nick Reeves cloned the recombinant Chn construct, and Joseph Fontana generated the *E(spl)m4, E(spl)m*α double deletion lines.

CHAPTER THREE:

Complex interplay of three transcription factors in controlling the tormogen differentiation program of *Drosophila* mechanoreceptors

Developmental Biology 329 (2009) 386–399

Contents lists available at ScienceDirect

Developmental Biology

journal homepage: www.elsevier.com/developmentalbiology

Genomes & Developmental Control

Complex interplay of three transcription factors in controlling the tormogen differentiation program of *Drosophila* mechanoreceptors

Steven W. Miller [a], Tomer Avidor-Reiss [b,c,1], Andrey Polyanovsky [d], James W. Posakony [a,*]

[a] Division of Biological Sciences, Section of Cell and Developmental Biology, Mail Code 0349, UC San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA
[b] Howard Hughes Medical Institute and Division of Biological Sciences, University of California San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA
[c] Department of Neurosciences, University of California San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA
[d] Sechenov Institute, St. Petersburg, Russia

ARTICLE INFO

Article history:
Received for publication 10 December 2008
Revised 18 January 2009
Accepted 3 February 2009
Available online 20 February 2009

Keywords:
Sox15
Su(H)
Vvl
Tormogen
Socket cell
Notch signaling

ABSTRACT

We have investigated the expression and function of the Sox15 transcription factor during the development of the external mechanosensory organs of *Drosophila*. We find that *Sox15* is expressed specifically in the socket cell, and have identified the transcriptional cis-regulatory module that controls this activity. We show that Suppressor of Hairless [Su(H)] and the POU-domain factor Ventral veins lacking (Vvl) bind conserved sites in this enhancer and provide critical regulatory input. In particular, we find that Vvl contributes to the activation of the enhancer following relief of Su(H)-mediated default repression by the Notch signaling event that specifies the socket cell fate. Loss of *Sox15* gene activity was found to severely impair the electrophysiological function of mechanosensory organs, due to both cell-autonomous and cell-non-autonomous effects on the differentiation of post-mitotic cells in the bristle lineage. Lastly, we find that simultaneous loss of both *Sox15* and the autoregulatory activity of *Su(H)* reveals an important role for these factors in inhibiting transcription of the Pax family gene *shaven* in the socket cell, which serves to prevent inappropriate expression of the shaft differentiation program. Our results indicate that the later phases of socket cell differentiation are controlled by multiple transcription factors in a collaborative, and not hierarchical, manner.

Introduction

Metazoan development depends fundamentally on the capacity of a handful of cell-cell signaling pathways to effect an extremely diverse array of conditional cell fate decisions. This capacity is based in turn on the ability of each pathway to activate distinct subsets of its target gene repertoire in different contexts, which is achieved through the integrative properties of the associated transcriptional cis-regulatory modules (Barolo and Posakony, 2002).

The Notch (N) signaling pathway is particularly well suited to binary cell fate choices, in which two possible cell fates are partitioned among two or more adjacent or nearby cells (Lai, 2004; Bray, 2006). The architecture and function of the N pathway have long been studied in the context of mechanosensory organ development in *Drosophila*, a setting in which the pathway is used repeatedly from the selection of primary precursor cells to the specification of post-mitotic cell fates in the ensuing lineage (Hartenstein and Posakony, 1990; Posakony, 1994). The first step takes place in small groups of cells with neural cell fate potential called proneural clusters (PNCs). Within each PNC, one cell becomes stably committed to the sensory organ precursor (SOP) fate and inhibits all of its neighbors from adopting the same fate, in a N-mediated signaling process called lateral inhibition. The inhibited cells adopt the alternative, epidermal cell fate. The SOP then divides, initiating a fixed cell lineage that ultimately generates the four terminally differentiated cells comprising the mechanosensory organ found in the adult fly: tormogen (socket cell), trichogen (shaft cell), thecogen (sheath cell), and a bipolar neuron (Figs. 1A, B). The socket and shaft cells are sister cells in this lineage (resulting from the division of the pIIa secondary precursor), as are the sheath cell and neuron (progeny of the pIIIb tertiary precursor) (Hartenstein and Posakony, 1989; Gho et al., 1999). When these cell pairs are born, they are each bipotent: both sisters can adopt either of the two possible fates. Asymmetric N signaling is responsible for insuring that the two cells adopt different fates. Thus, one cell in each pair receives and responds to a N-mediated signal from its sister and is accordingly directed to the N-responsive fate (socket or sheath). The remaining sister adopts the N-non-responsive fate (shaft or neuron).

Following these cell fate specification events, each cell executes a complex program of differentiation that confers the distinctive structural and physiological properties that constitute its unique contribution to the construction of a functional mechanosensory organ (Hartenstein and Posakony, 1989). For example, the trichogen builds the long microtubule-based shaft structure that functions as a very sensitive receptor of mechanical stimuli; this cell degenerates

* Corresponding author. Fax: +1 858 822 3021.
E-mail address: jposakony@ucsd.edu (J.W. Posakony).
[1] Present address: Department of Cell Biology, Harvard Medical School, Boston, MA 02115, USA.

0012-1606/$ – see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.ydbio.2009.02.009

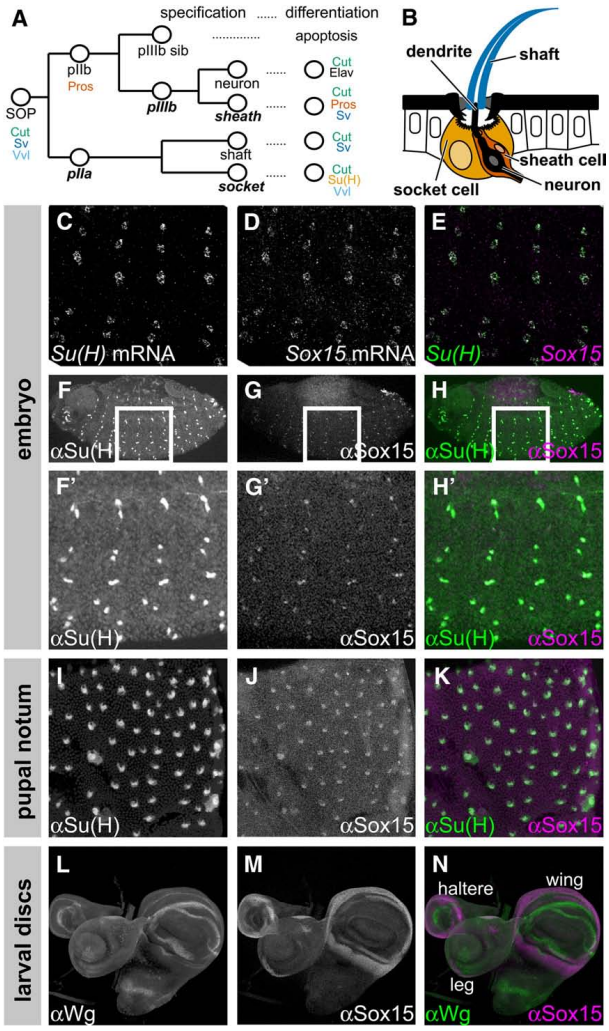


Fig. 1. *Sox15* is expressed specifically in socket cells of external sensory organs. (A) Diagram of the *Drosophila* mechanosensory bristle lineage, indicating protein expression markers characteristic of specific cells. Expression of the three SOP markers and of the pIIb marker Pros is maintained in all of their progeny cells during the mitotic phase of the lineage. As post-mitotic cells differentiate (24–36 hours APF), protein expression profiles refine to those indicated at right. Notch-dependent cell fates are shown in bold italics. (B) Cross-section diagram of adult mechanosensory bristle organ [adapted from Walker et al. (Walker et al., 2000)], showing the shaft, neuron, sheath, and socket cells. Also indicated is the dendrite of the neuron. (C–E) *Su(H)* mRNA (C, green in E) and *Sox15* mRNA (D, magenta in E) appear in the same cells in the late embryo. (F–N) Pattern of Sox15 protein accumulation (G, G', J, M, and magenta in H, H', K, N) compared with that of Su(H) (F, F', I, and green in H, H', K) in socket cells of the embryo and the pupal notum, and compared with the pattern of Wg protein (L, green in N) in wing, haltere, and leg imaginal discs of late third-instar larvae.

by the adult stage. The tormogen tightly surrounds the other cells with a cytoplasmic sheath, and produces the cuticular socket structure that surrounds the base of the shaft. The socket cell is also responsible for maintaining a distinct (high K^+, low Na^+) ionic environment in the endolymph, the fluid in the lymph cavity inside the organ. The composition of the endolymph is essential to the proper functioning of the mechanosensitive ion channels in the membrane of the neuron, and thus for mechanotransduction itself (Jarman, 2002).

The fundamental role played by N signaling in the development of *Drosophila* mechanosensory organs is founded on context-dependent transcriptional responses to activation of the N receptor. Suppressor of Hairless [Su(H)], the transducing transcription factor for the N pathway, is by default a repressor (Barolo et al., 2000b; Morel and Schweisguth,

2000; Furriols and Bray, 2001). In the absence of signaling, the adaptor protein Hairless (H) binds to Su(H) and recruits the co-repressors Groucho (Gro) and C-terminal Binding Protein (CtBP), thus conferring repressive activity on Su(H) bound to its targets (Morel et al., 2001; Barolo et al., 2002). N activation leads to proteolytic cleavage of the receptor's intracellular domain (NICD) and its translocation to the nucleus, where it and the co-activator protein Mastermind (Mam) form a complex that replaces the H/Gro/CtBP complex on Su(H), converting Su(H) from a repressor to an activator (Bray, 2006). Which subset of N pathway target genes is activated following any given signaling event is determined by "local activators" — transcription factors that cooperate with Su(H) to activate only those targets with the appropriate binding sites in their N-responsive cis-regulatory modules (Nellesen et al., 1999; Cooper et al., 2000; Barolo and Posakony, 2002). Different local activators are expressed in different developmental contexts. Thus, a full understanding of a particular cell's distinctive response to N signaling requires knowledge of the full set of pathway target genes activated by the signal in that cell; identification of each target's responding cis-regulatory module; and identification of the local activators that contribute to the context-specific activation of each module. The goal is to elucidate the cell's "specification-differentiation interface"; in other words, to understand how the same cell fate specification signal elicits a distinct differentiative outcome in each setting.

Our previous work has characterized one highly specific transcriptional response of the socket cell to the N signaling event that specifies its fate (Barolo et al., 2000b). This cell uniquely expresses very high levels of Su(H), far higher than that required to transduce a N signal. This is the result of the cell type-specific activation of the autoregulatory socket enhancer (ASE) downstream of the *Su(H)* gene. This autoactivation loop, and the elevated Su(H) levels it generates, are not required for specification of the socket cell fate, nor for many aspects of socket cell differentiation, including the normal formation of the socket cuticular structure. Instead, sensory organs in which the socket cell lacks ASE activity exhibit severe deficits in the electrophysiological properties that underlie mechanotransduction. This finding suggested that different components of the socket cell differentiation program are under separate regulatory control.

In the present study we show that the Sox-family transcription factor Sox15 is likewise expressed specifically in the socket cell of external sensory organs and we identify the cis-regulatory module that directs this expression. We provide evidence that this module is a direct target of Su(H) and the N pathway, and that N signaling acts principally to relieve default repression by Su(H), leaving to local activators the task of activating the module in the socket cell. We show that one such local activator is the POU-domain transcription factor Ventral veins lacking (Vvl). By analyzing flies carrying deletion mutations within the *Sox15* locus, we define an important role for Sox15 in the socket cell differentiation program. We find that Sox15, like Su(H), is required for the proper electrophysiological function of the sensory organ. Important phenotypic differences between *Sox15* and *Su(H)* ASE mutants, however, imply that the two factors control at least partially non-overlapping components of the socket program. Finally, we demonstrate that Sox15 and Su(H) collaborate to inhibit socket cell expression of *shaven (sv)*, which encodes a Pax family transcription factor that is a high-level regulator of the shaft differentiation program (Kavaler et al., 1999). Thus, actively preventing execution of the alternative (sister cell) program, which helps insure the robustness of cell fate commitment, is a critical response to N signaling in the mechanosensory organ lineage.

Materials and methods

Fly lines

KG09145 flies were a gift of Hugo Bellen; *Ubx-FLP* flies were a gift of Y. N. Jan; vvl^{H599} *FRT80B* flies were a gift of Adi Salzberg; *Ubi-nGFP FRT80B* and sequencing strain flies were obtained from the Bloomington Stock Center (Stock #5630 and #2057, respectively); the $Su(H)^{AR9}/Su(H)^{SF8}$; *Su(H)RC-ΔASE* stock has been described previously (Barolo et al., 2000b).

Sox15 *mutant alleles*

New mutant alleles of *Sox15* described in this study were created by mobilizing the P transposon insertion *KG09145* (Roseman et al., 1995; Bellen et al., 2004) by crossing this strain to flies carrying a transposase source (Cooley et al., 1988). F_1 progeny were crossed to a *CyO* balancer and stocks were generated from $y^- w^-$, *CyO* F_2 progeny. Stocks were screened for deletions by PCR, first in pools of ten, and then as individual lines from positive pools. DNA was amplified from individual positive lines, and PCR products were sequenced to determine exact deletion breakpoints.

Mosaic analysis

$vvl^{-/-}$ clones were generated utilizing the FLP/FRT system (Golic and Lindquist, 1989; Golic, 1991; Xu and Rubin, 1993). Males of the genotype *y w Ubx-FLP/Y;* vvl^{H599} *FRT80B/+* were crossed either to *y w Ubx-FLP; ASE-nGFP; Ubi-nGFP FRT80B* or to *y w Ubx-FLP; Sox7.5>nGFP; Ubi-nGFP FRT80B* females. Flies were reared at 25 °C to induce clones.

Electrophysiology

Recordings from adult anterior notopleural bristles were performed as described previously (Kernan et al., 1994; Walker et al., 2000). For each genotype, eight bristles were analyzed. Statistical significance was determined using a two-sample *t* test.

DNA-binding assays

The *vvl*-PA coding region was amplified from genomic DNA and cloned into the pGEX-5X-2 vector. GST-Vvl fusion protein was purified as described (Bailey and Posakony, 1995), and 6XHis-Su(H) as described (Janknecht et al., 1991). Electrophoretic mobility shift assays (EMSAs) were performed as described previously (Bailey and Posakony, 1995; Barolo et al., 2000b). Wild-type Probe 1 consists of the sequence TTACAAGTAATATTTACATTTTTCCCATGCTAA; the Vvl site mutant version is TTACAAGTAATA*GGG*ACC*G*TTTTCCCATGCTAA, the Su(H) site mutant version is TTACAAGTAATATTTACATTTTTGCCATGCTAA, and the Vvl/Su(H) double site mutant is TTACAAGTAATA*GG-GACCG*TTTTGCCATGCTAA. Wild-type Probe 2 consists of the sequence GCAGTCGACCATTTACATATTTACGTTTAC; the Vvl site mutant version is GCAGTCGACCA*GGGACCG*ATTTACGTTTAC.

Reporter transgene constructs

Various regions of the *Sox15* locus (see Fig. 2A) were amplified by PCR from genomic DNA using primers with 5' restriction site sequences added; these were cloned into either the pH-Stinger (nuclear eGFP) or pRed H-Stinger (nuclear DsRed) vectors (Barolo et al., 2000a; Barolo et al., 2004). The 7.5-kb enhancer fragment was amplified using the primer sequences GCCGGCGCGCCGATAGCCA-CCGTGCTCCGGATAATCGCTGC (Asc I site end) and TCCCCGCGGCA-TATGATCACGAACATCCACATCATCTGC (Sac II site end); the 1.3-kb enhancer fragment was amplified with the primer sequences TAACTCGAGCATATGATCACGAACATCCACATCATCTGC (Xho I site end) and TAAGCATGCTATATGCATTTCCAATCCAGCTTAGTCACG (Sph I site end). Other primer sequences are available upon request. The mutations described above under "DNA-binding assays" were introduced into the 1.3-kb enhancer fragment by overlap extension PCR mutagenesis (Ho et al., 1989). The wild-type version of the 1.3-kb fragment was amplified from the *D. melanogaster* iso-1 genome

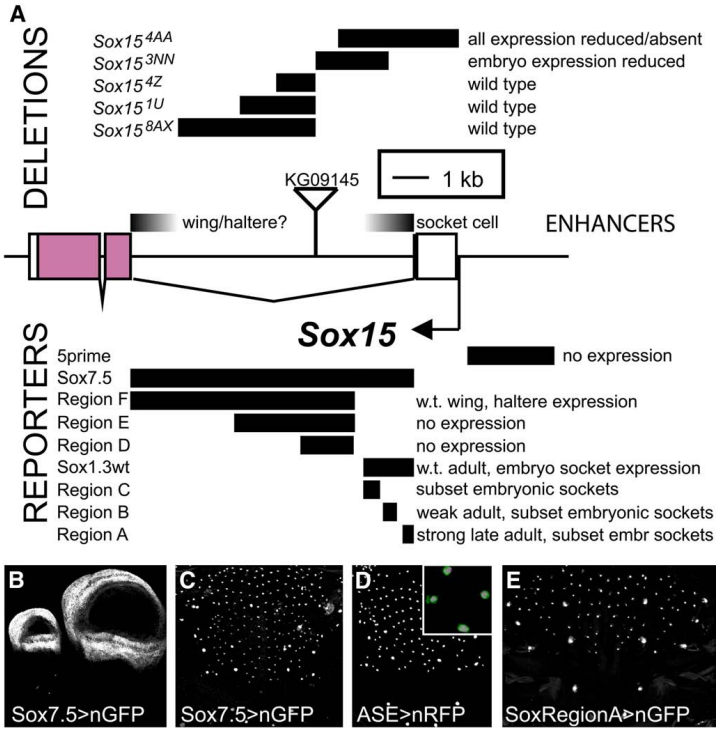


Fig. 2. Identification of the *Sox15* socket enhancer. (A) Schematic of the *Sox15* locus. Shown are the intron-exon structure, the location of the *KG09145* P-element transposon insertion, the extents and designations of the deletion mutations created by imprecise excision of the *KG09145* P element, the mutant phenotypes of the deletion alleles with respect to *Sox15* mRNA accumulation, the DNA fragments tested for enhancer activity in vivo, the nGFP expression pattern directed by each fragment, if any, and the deduced locations of the socket cell and wing/haltere enhancers. Note that a 1.3-kb intron fragment immediately adjacent to the first exon (Sox1.3) drives socket cell-specific reporter expression (see Fig. 4A). (B) Haltere (left) and wing (right) imaginal discs from late third-instar larvae, showing nGFP expression directed by the 7.5-kb enhancer fragment (Sox7.5; A); compare with Figs. 5C and B, respectively; see also Fig. 1M, N. (C) Pupal notum at 30 hours APF, also showing nGFP expression driven by Sox7.5. (D) 30-hour-APF notum showing nRFP expression driven by the *Su(H)* ASE. Inset is a high-magnification image of a 30-hour-APF notum from a fly carrying one copy each of *Sox7.5>nGFP* (green) and *ASE>nRFP* (magenta), demonstrating activity of both enhancers in the same nuclei. (E) Pharate adult notum showing nGFP expression driven by Region A (see A), a 380-bp fragment sufficient to direct robust reporter activity in socket cells by the time of eclosion.

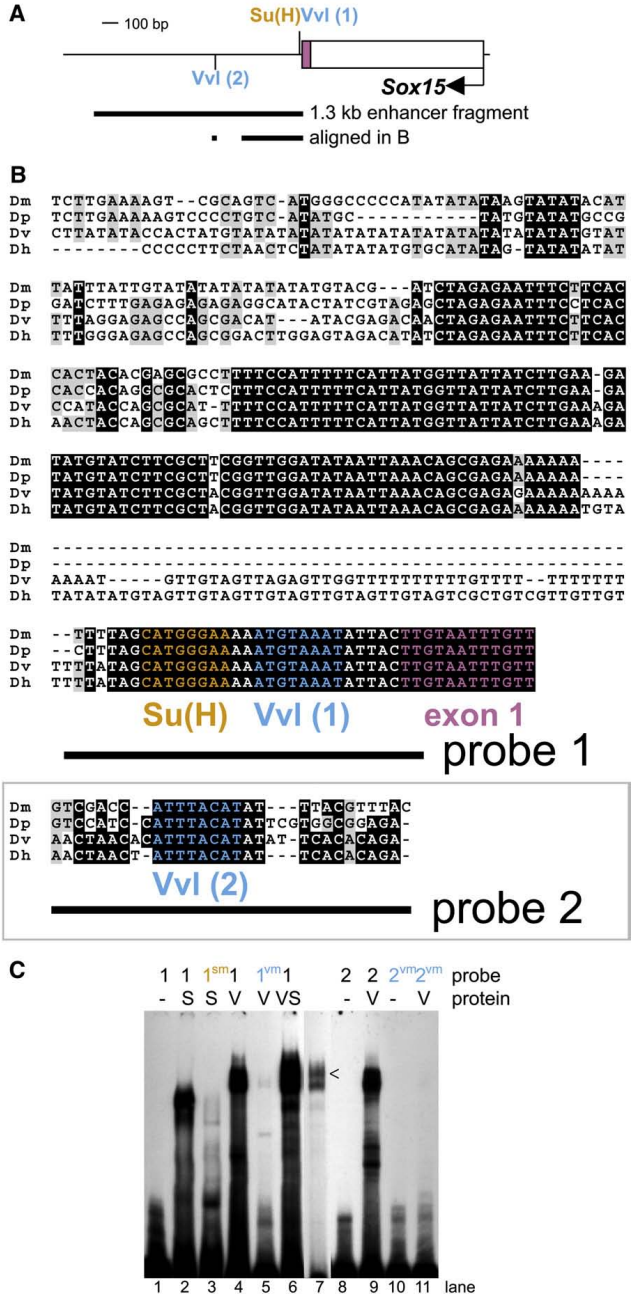
sequencing strain (Brizuela et al., 1994); both wild-type and mutant PCR products were sequenced and the results compared to the genome sequence in GenBank by BLAST search (Altschul et al., 1997) to detect unwanted point mutations. Reporter constructs were introduced into the germline of w^{1118} embryos using standard procedures (Rubin and Spradling, 1982).

Preparation of adult tissues

Wing and hinge tissue was removed from flies under CO_2 anesthesia and placed directly onto slides. Permount (Fisher) was then applied to coverslips and these were immediately placed on the tissue. For preparation of notum cuticle, adult flies were dehydrated overnight in methanol and then in ethanol. Flies were then carefully dissected in ethanol and dorsal nota were transferred into fresh ethanol. Ethanol was removed, and 1:2 CMC-10 mounting media (Masters Company, Wood Dale, IL) : lactic acid was added; tissue in this mixture was incubated, covered, overnight at 65 °C. Nota in suspension were removed with a pipette and transferred to slides.

In situ hybridization

A digoxygenin-labeled antisense RNA probe (Tautz and Pfeifle, 1989) to detect *Sox15* transcript in situ was made by transcribing with T7 RNA polymerase a linear *Sox15* pGEM-T (Promega) cDNA clone containing a *Sox15* PCR product from embryonic first-strand cDNA. *Su(H)* digoxygenin- and biotin-labeled antisense RNA probes (O'Neill and Bier, 1994) were generated from a HindIII-linearized *Su(H)* cDNA clone in pNB40 (Brown and Kafatos, 1988), transcribed with T7 RNA polymerase. A *sv* digoxygenin-labeled antisense RNA intron probe was constructed from genomic DNA PCR products covering 3.0 kb from intron 1 and 5.4 kb from intron 4 cloned into pGEM-T, linearized, and transcribed with T7 RNA polymerase. Fluorescent in situ hybridizations to embryos were performed as described (Kosman et al., 2004), using a biotin-labeled *Su(H)* probe and a digoxygenin-labeled *Sox15* probe. Imaginal disc, pupal notum, and adult abdomen in situ hybridizations were performed as described (O'Neill and Bier, 1994; Lai et al., 2000; Reeves and Posakony, 2005).


A
100 bp
Su(H)Vvl (1)
Vvl (2)
Sox15
1.3 kb enhancer fragment
aligned in B
B
Dm TCTTGAAAAGT--CGCAGTC-ATGGGCCCCCATATATATAAGTATATACAT
Dp TCTTGAAAAAGTCCCCTGTC-ATATGC----------TATGTATATGCCG
Dv CTTATATACCACTATGTATATATATATATATATATATATATATATGTAT
Dh --------CCCCCTTCTAACTCTATATATATGTGCATATAG-TATATATAT
Dm TATTTATTGTATATATATATATGTACG---ATCTAGAGAATTTCTTCAC
Dp GATCTTTGAGAGAGAGAGGCATACTATCGTAGAGCTAGAGAATTTCCTCAC
Dv TTTAGGAGAGCCAGCGACAT---ATACGAGAGCAACTAGAGAATTTCTTCAC
Dh TTTGGGAGAGCCAGCGGACTTGGAGTAGACATATCTAGAGAATTTCTTCAC
Dm CACTACACGAGGGCCTTTTCCATTTTTCATTATGGTTATTATCTTGAA-GA
Dp CACCACAGGCGGACTCTTTCCATTTTTCATTATGGTTATTATCTTGAA-GA
Dv CCATACCAGCGGCAT-TTTTCCATTTTTCATTATGGTTATTATCTTGAAAGA
Dh AACTAGCAGCGGCAGCTTTTCCATTTTTCATTATGGTTATTATCTTGAAAGA
Dm TATGTATCTTCGCTTCGGTTGGATATAATTAAACAGCGAGAAAAAAA----
Dp TATGTATCTTCGCTTCGGTTGGATATAATTAAACAGCGAGAAAAAA----
Dv TATGTATCTTCGCTACGGTTGGATATAATTAAACAGCGAGAGAAAAAAAA
Dh TATGTATCTTCGCTACGGTTGGATATAATTAAACAGCGAGAAAAAATGTA
Dm ----------
Dp ----------
Dv AAAAT-----GTTGTAGTTAGAGTTGGTTTTTTTTTGTTTT--TTTTTTT
Dh TATATATGTAGTTGTAGTTGTAGTTGTAGTTGTAGTCGCTGTCGTTGTTGT
Dm --TTTTAGCATGGGAAAAATGTAAATATTACTTGTAATTTGTT
Dp --CTTTAGCATGGGAAAAATGTAAATATTACTTGTAATTTGTT
Dv TTTTTAGCATGGGAAAAATGTAAATATTACTTGTAATTTGTT
Dh TTTTTAGCATGGGAAAAATGTAAATATTACTTGTAATTTGTT
Su(H) Vvl (1) exon 1
probe 1
Dm GTCGACG--ATTTACATAT---TTACGTTTAC
Dp GTCCATG-CATTTACATATTCGTGGCGGAGA-
Dv AACTAACACATTTACATATAT-TCACACAGA-
Dh AACTAAGT-ATTTACATAT---TCACACAGA-
Vvl (2)
probe 2
C
1 1 1sm1 1vm1 2 2 2vm2vm probe
- S S V V VS - V - V protein
1 2 3 4 5 6 7 8 9 10 11 lane

Antibody production

The *Sox15* protein coding sequence was amplified from embryo first-strand cDNA and cloned into pGEX-5X-2. The optimum conditions for purification were determined as described (Mercado-Pimentel et al., 2002). Since these consistently yielded some non-specific bands, purified protein was electrophoresed on a 10 cm × 10 cm SDS-PAGE gel and the location of the majority of the protein was determined by Coomassie staining of a thin lengthwise slice. A slice was then cut from the gel at this latitude and sent to Pocono Rabbit Farm and Laboratory for immunization of two guinea pigs. Upon receipt, antisera were preabsorbed against fixed embryos and tested for the anticipated staining pattern.

Immunohistochemistry

White pre-pupae were collected and aged to the appropriate developmental time at 25 °C. Late third-instar larvae or aged pupae were dissected in 1X PBS, fixed for 30 minutes in 1X PBS with 0.3% Triton X-100, and washed 5X for 10 minutes each in 1X PBS with 0.1% Triton X-100 (PBT). Tissue was then incubated with primary antibody overnight at 4 °C, followed by five washes in PBT, a 1-hour incubation with fluorescent secondary antibody, and five further PBT washes. After the last PBT wash, samples were suspended in 2.5% w/v DABCO, 50 mM Tris-HCl pH 8.0, 90% glycerol (Kosman et al., 2004) for fluorescence microscopy, or in 80% glycerol, 100 mM Tris-HCl pH 8.0 for light microscopy. Any tissue containing fluorescent reporters was kept in the dark from the point of fixation onward as much as possible.

Embryos were fixed and prepared for antibody staining as previously described (Kosman et al., 2004; Reeves and Posakony, 2005). They were then brought into 1X PBT and stained as described above for larval and pupal tissues.

Primary antibodies included mouse anti-Cut hybridoma supernatant (Developmental Studies Hybridoma Bank, University of Iowa), diluted 1:500; mouse anti-Prospero hybridoma supernatant (DSHB), diluted 1:20; rat anti-Elav hybridoma supernatant (DSHB), diluted 1:200; mouse anti-Wg hybridoma supernatant (DSHB), diluted 1:200; and rabbit anti-Su(H) (Santa Cruz Biotechnology), diluted 1:1000. Guinea pig anti-Sox15 antiserum was preabsorbed against embryos and used at either a 1:500 or 1:1000 dilution. All secondary antibodies were used at a 1:1000 dilution and included anti-mouse-HRP conjugate (Jackson Laboratories), anti-rat-Alexa647 conjugate, and anti-mouse-Alexa555 conjugate (Molecular Probes). HRP was detected with DAB as described (Goldstein and Fyrberg, 1994).

Microscopy

Fluorescence microscopy was performed on a Leica TCS SP2 confocal microscope equipped with Leica Confocal Software v2.5 (Leica Microsystems). Figure panels are composed of maximum projections of stacks taken along the apical-basal axis at 2 μm increments. Fluorophores were excited separately at 488 nm (GFP), 543 nm (RFP, Alexa 555), or 633 nm (Alexa 647).

Samples for scanning electron microscopy were collected and dehydrated overnight in isoamyl acetate as described (Bang et al., 1991). Imaging was performed at the Scripps Institution of Oceanography Unified Laboratory Facility on an FEI Quanta 600 instrument.

Gene diagrams

Gene diagrams in Figs. 2–4 and Fig. S4 were created using GenePalette (Rebeiz and Posakony, 2004) and edited in Adobe Illustrator.

Results

Sox15 *is expressed specifically in the socket cell of external sensory organs*

A previous survey of the expression patterns of *Sox* family genes in *Drosophila* revealed that transcripts from *Sox15* accumulate specifically in the developing peripheral nervous system (PNS) in the embryo, and it was suggested that this expression might be in the socket cells of external sensory organs (Cremazy et al., 2001) (Figs. 1A, B). We thus sought to define the cell-type specificity of *Sox15* expression in both the larval and adult PNSs. In the embryo, we detect *Sox15* transcript accumulation specifically in the cells that exhibit high levels of *Su(H)* transcript (Schweisguth and Posakony, 1992), indicating that *Sox15* is indeed expressed in socket cells of the larval PNS (Figs. 1C–E). The cell-type specificity of *Sox15* expression in the pupal notum at 30 hours APF was investigated by observing its response to experimentally induced cell fate changes (see Supplementary Fig. S1). This analysis shows explicitly that *Sox15* expression in the sensory organ lineage is triggered by activation of the N pathway and specification of the socket cell fate. Despite their co-expression in socket cells, we find that the temporal patterns of *Su(H)* and *Sox15* mRNA accumulation in the pupal notum differ markedly, in that high-level expression of *Sox15* is considerably delayed relative to that of *Su(H)* (see Supplementary Fig. S2).

We complemented this analysis of the pattern of *Sox15* transcript accumulation with a parallel examination of Sox15 protein expression. Immunofluorescence analysis utilizing a polyclonal antiserum raised against recombinant GST-Sox15 protein revealed expression specifically in the nuclei of socket cells in both the larval and adult PNSs, as indicated by co-localization with Su(H) immunoreactivity (Figs. 1F–K). Our antibody also revealed the unique pattern of Sox15 protein accumulation in the imaginal discs of late third-instar larvae (Fig. 1M). In wing and haltere discs, Sox15 localizes to the presumptive hinge region, while in leg discs the protein is found in discrete patches on opposite sides of the proximal leg primordium. These patterns of Sox15 protein expression in imaginal discs closely mimic the corresponding patterns of *Sox15* transcript accumulation (see Figs. 5A–C) (Cremazy et al., 2001). We note that Sox15 protein accumulation in the wing hinge primordium abuts the more distal zone of Wingless (Wg) protein expression (Figs. 1L–N), which may be suggestive of an important regulatory relationship.

Identification and analysis of the Sox15 socket enhancer

To begin to define how *Sox15* expression in the PNS is regulated at the transcriptional level, we sought to identify the *Sox15* socket cell enhancer module. We took two approaches to this goal: generation of deletion mutations within the *Sox15* locus, and analysis of reporter gene expression driven by non-coding genomic DNA fragments from the gene.

The first effort was aided by the availability of the *KG09145* line from the BDGP gene disruption collection (Roseman et al., 1995;

Fig. 3. Analysis of conserved sequence motifs in the *Sox15* socket enhancer reveals Su(H) and Vvl as potential regulatory inputs. (A) Schematic of the 1.3-kb socket enhancer fragment in relation to the first exon of *Sox15*, indicating the positions of two Vvl and one Su(H) binding motifs. The sequences aligned in B are also indicated. (B) Alignment of the region of high sequence conservation among *D. melanogaster* (Dm), *D. pseudoobscura* (Dp), *D. virilis* (Dv), and *D. hydei* (Dh) with the Su(H) site (orange) and Vvl site 1 (blue) highlighted. Also shown is an alignment of the region surrounding Vvl site 2 (blue) from the same species. Sequences included in oligonucleotide probes used in gel shift assays (see C) are indicated. (C) Electrophoretic mobility shift assays showing the ability of both His-tagged Su(H) (S) and a GST-Vvl fusion protein (V) to bind radiolabelled oligonucleotides from the *D. melanogaster* sequence (S: lanes 2,6,7; V: lanes 4,6,7,9), while failing to bind mutant oligonucleotides (S: lane 3; V: lanes 5,11). Lane 7 is a shorter exposure of lane 6, to indicate the slower mobility band (caret) seen when both His-Su(H) and GST-Vvl are allowed to bind probe 1.

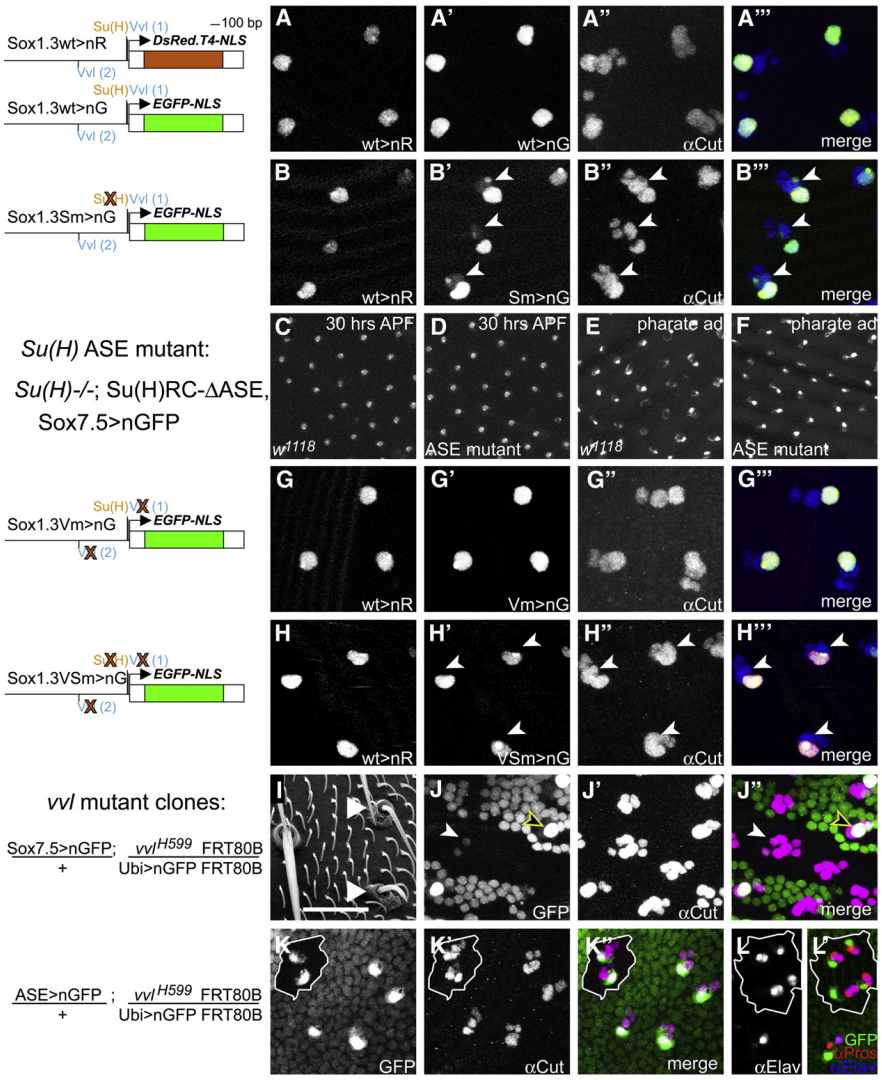


Fig. 4. Loss of Su(H) and Vvl binding in cis and in trans affects *Sox15* socket enhancer activity. (A–A‴) 30-hour-APF notum from a fly bearing one copy each of *Sox1.3wt>nRFP* (wild-type enhancer) and *Sox1.3wt>nGFP* (wild-type enhancer), showing the nRFP signal (A, red in A‴), the nGFP signal (A′, green in A‴), and Cut immunoreactivity (A″, blue in A‴). (B–B‴) 30-hour-APF notum from a fly bearing one copy each of *Sox1.3wt>nRFP* and *Sox1.3Sm>nGFP* [Su(H) site mutant enhancer], with the different channels displayed as in A–A‴. Arrowheads in B′–B‴ indicate positions of GFP-positive shaft cell nuclei. (C–F) nGFP expression directed by the Sox7.5 enhancer fragment in either a wild-type (w^{1118}; C, E) or *Su(H)* ASE mutant background [$Su(H)^{AR9}/Su(H)^{SF8}$; *Su(H)RC-ΔASE*; D, F; genotype shown at left], at either 30 hours APF (C, D) or pharate adult (E, F). (G–G‴) 30-hour-APF notum from a fly bearing one copy each of *Sox1.3wt>nRFP* and *Sox1.3Vm>nGFP* (Vvl site mutant enhancer), with the different channels displayed as in A–A‴. (H–H‴) 30-hour-APF notum from a fly bearing one copy each of *Sox1.3wt>nRFP* and *Sox1.3VSm>nGFP* [Vvl site-Su(H) site double mutant enhancer], with the different channels displayed as in A–A‴. Arrowheads in H′-H‴ indicate positions of GFP-negative shaft cell nuclei (compare with B–B‴). (I) Scanning electron micrograph of notum region of a fly bearing $vvl^{-/-}$ clones. Arrowheads indicate mutant bristles. (J–J″) 36-hour-APF notum from a fly bearing one copy of *Sox7.5>nGFP* and $vvl^{-/-}$ clones (genotype shown at left; mutant tissue marked by the absence of low-level ubiquitous nGFP), showing the nGFP signal from the reporter gene (J, green in J″) and Cut immunoreactivity (J′, magenta in J″). Filled arrowhead in J and J″ indicates *Sox7.5>nGFP* expression within the $vvl^{-/-}$ mutant territory; yellow-edged open arrowhead in the same panels indicates expression of the reporter in adjacent $vvl^{+/-}$ tissue for comparison. (K–L′) 30-hour-APF nota from flies bearing one copy of *ASE>nGFP* and $vvl^{-/-}$ clones (genotype shown at left; mutant tissue marked as in J and J″ and outlined in white). Compare level of *ASE>nGFP* expression (K, green in K″ and L′) within the $vvl^{-/-}$ mutant territory with that in adjacent $vvl^{+/-}$ tissue; sensory organ cells marked by Cut immunoreactivity (K′, magenta in K″). (L–L′) $vvl^{-/-}$ tissue at 30 hours APF shows two cells positive for the neuron marker Elav in most positions (L, blue in L′), and three cells positive for Pros immunoreactivity (red in L′), of which two are the Elav-positive cells.

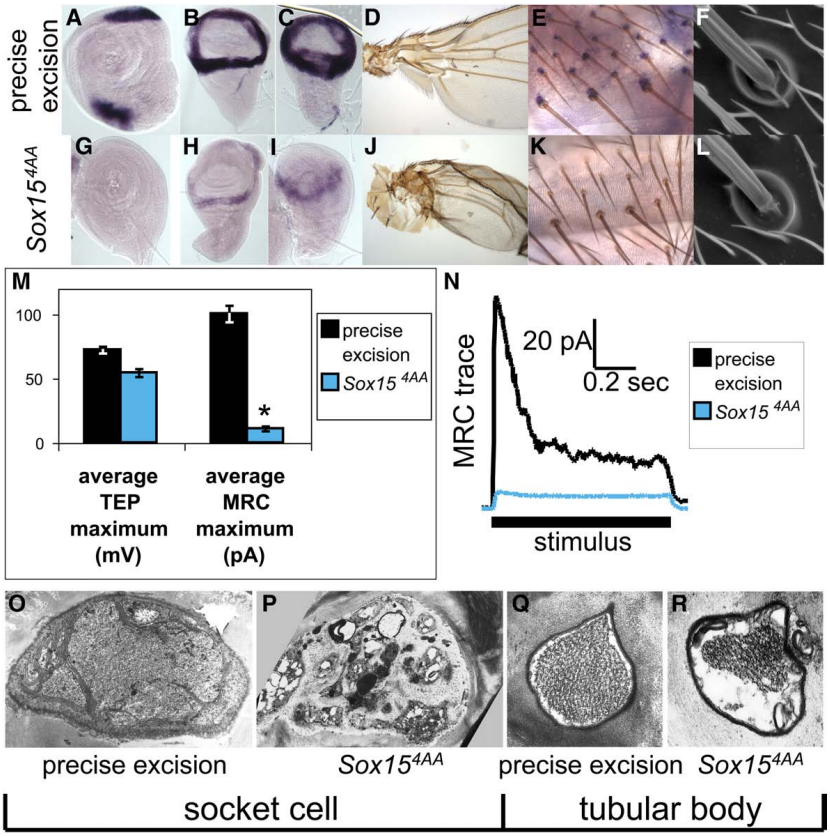


Fig. 5. *Sox15* loss-of-function phenotypes. (A–F, O, Q) Tissue from wild-type (homozygous precise excision) flies. (G–L, P, R) Tissue from $Sox15^{4AA}$ homozygous flies. (A–C, G–I) Pattern of *Sox15* mRNA accumulation in imaginal discs of late third-instar larvae; see also (Cremazy et al., 2001). (A, G) Leg discs. (B, H) Wing discs. (C, I) Haltere discs. (D, J) Proximal wing and body wall. (E, K) *Sox15* mRNA accumulation in socket cells of mechanosensory bristles on the pharate adult abdomen. (F, L) Scanning electron micrographs of external cuticular structures of representative adult mechanosensory bristles. (M) Average transepithelial potential (TEP) and mechanoreceptor current (MRC) recorded from precise excision (black) and $Sox15^{4AA}$ (blue) adult flies. Asterisk indicates statistically significant difference from the precise excision value ($p<0.001$). (N) Representative MRC traces recorded from the same genotypes. (O–R) Transmission electron micrograph sections through the adult socket cell (O, P) and through the tubular body of the neuronal dendrite (Q, R).

Bellen et al., 2004), which contains a P-element insertion within the 7.5-kb first intron of *Sox15* (Fig. 2A). Five new mutant alleles of *Sox15* were created through imprecise excision of this element, resulting in a series of deletions internal to the *Sox15* locus (Fig. 2A). Only two of these alleles, $Sox15^{4AA}$ and $Sox15^{3NN}$, display altered accumulation of *Sox15* mRNA. $Sox15^{4AA}$ is a deletion that removes part of the intron proximal to the first exon, as well as the first exon itself (-21 to +3078), while the $Sox15^{3NN}$ deletion endpoints (+1812 to +3645) are contained entirely within the intron (which extends from +1124 to +8360). $Sox15^{4AA}$ homozygotes have lost expression of the gene in leg imaginal discs as well as in the socket cell, and expression in wing and haltere discs, while not eliminated, is dramatically reduced (see Figs. 5G–I, K). This phenotype, which affects multiple tissues, is likely due in part to the deletion of the transcription Initiator sequence of *Sox15* (Juven-Gershon et al., 2008), but it could also be the result of deleting one or more cis-regulatory modules (see below). $Sox15^{3NN}$ homozygous embryos have lost most socket cell expression, while adult socket expression appears normal (see Supplementary Fig. S3), suggesting that the embryonic and adult tormogen regulatory elements are at least to some degree distinct.

We next tested the capacity of a number of genomic DNA fragments from the *Sox15* locus to direct GFP reporter gene expression in the socket cell. A fragment comprising the entire 7.5-kb intron (Sox7.5; Fig. 2A) drives reporter expression in late third-instar wing and haltere imaginal discs (Fig. 2B), in a pattern recapitulating that of transcript accumulation from the endogenous gene (see Figs. 5B, C). This fragment also directs reporter activity in single cells in both the embryonic and adult PNSs; these are confirmed to be socket cells by the co-expression of the socket-specific *Su(H)* reporter *ASE>nRFP* (Figs. 2C, D). A shorter fragment containing the intronic 1.3 kb proximal to exon 1 (Sox1.3wt; Fig. 2A) is likewise sufficient to recapitulate this full socket cell expression pattern, while even smaller subfragments (Regions A-C, Fig. 2A) are capable of directing expression in only subsets of embryonic and adult socket cells (Fig.

2E). The 5.5 kb of the first intron distal to exon 1 (Region F, Fig. 2A) directs reporter gene expression only in wing and haltere discs (data not shown). *Sox15*$^{\Delta AX}$, the imprecise excision allele bearing the largest deletion within the intron, has its distal endpoint 1.2 kb from exon 2 (Fig. 2A), and flies homozygous for this allele show no obvious defect in endogenous *Sox15* mRNA accumulation (see Fig. S3; data not shown). Thus, the combined evidence from deletion mutations of *Sox15* and from enhancer fragment activity localizes the wing/haltere and socket cell regulatory sequences to opposite ends of the gene's large first intron, with the former presumed to be in the 1.2 kb distal and the latter in the 1.3 kb proximal to exon 1 (Fig. 2A).

Since the 1.3-kb proximal enhancer region is the smallest fragment tested capable of recapitulating the full socket cell expression dynamics of *Sox15* mRNA, we examined this region (Fig. 3A) for conserved sequence elements that might function as binding sites for important regulatory factors. This analysis revealed two occurrences of the sequence ATGTAAAT (Figs. 3A, B), which is a single-base-pair mismatch to the strong binding site ATGCAAAT for the POU-domain transcription factor Vvl (Certel et al., 1996; Ma et al., 2000). In the adult mechanosensory bristle lineage, Vvl is detectable in all cells from the SOP to the post-mitotic cell types, at least initially (Inbal et al., 2003). By 42 hours APF, however, expression persists only in the socket cell, and mosaic analysis has revealed differentiation defects in both the socket and shaft cuticular structures in *vvl* mutant sensory organs (see Fig. 4I) (Inbal et al., 2003). We find that a purified GST-Vvl fusion protein binds specifically in vitro to each of the Vvl motifs in the *Sox15* socket enhancer (Fig. 3C). Six bp distal to Vvl(1), the more proximal of these sites, is the conserved sequence CATGGGAA (Figs. 3A, B), previously shown to be bound strongly by Su(H) in vitro (Nellesen et al., 1999). Indeed we find that this sequence in the *Sox15* enhancer is bound specifically by purified His-tagged Su(H) (Fig. 3C). With both Vvl and Su(H) as strong candidates for key regulatory factors, we next investigated their possible roles in the operation of the *Sox15* socket cell module.

Sox15 *is a direct target of the N pathway in the socket cell*

Interfering with Su(H) regulation of N target genes often results in two concurrent defects in gene expression (Barolo and Posakony, 2002). One effect is loss of gene activation in the N signal-receiving cell due to loss of the Su(H)/NICD/Mam activation complex at the enhancer (Bailey and Posakony, 1995; Lecourtois and Schweisguth, 1995). The second phenotype is de-repression of N target genes in the N signal-sending cell because of the inability of the Su(H)/H/Gro/CtBP repression complex to be recruited to the enhancer (Barolo et al., 2000b; Morel and Schweisguth, 2000; Castro et al., 2005; Koelzer and Klein, 2006). Su(H) binding in vitro and in vivo can be abolished by mutating the YRTGDGAA motif to YRTGDCAA (Fig. 3C, lane 3) (Bailey and Posakony, 1995). When applied to the 1.3-kb socket enhancer fragment, this single-base-pair mutation results in weak ectopic activation of reporter gene expression in a neighboring large Cut-expressing nucleus, that of the shaft cell, by 30 hours APF (Figs. 4A, B). Such de-repression in the N signal-sending shaft cell indicates a role for default repression of *Sox15* in this cell by Su(H) (Barolo and Posakony, 2002). While we can detect ectopic activation of the *Sox15* socket enhancer in the shaft cell when its Su(H) site is mutated, we fail to detect any effect on activation of the reporter gene throughout the life of the socket cell (Fig. 4B and data not shown). Moreover, when the socket cell lacks the function of the *Su(H)* ASE, there is no effect on either wild-type 7.5-kb reporter gene expression (Figs. 4C–F) or accumulation of endogenous *Sox15* transcript in this cell (data not shown). Collectively, these data indicate that the principal direct effect of normal N signaling on *Sox15* socket enhancer activity is the relief of default repression in the socket cell. This allows activators other than Su(H), one or more of which would be present in both the socket and shaft cells, to initiate transcription of *Sox15* only in the appropriate cell.

Vvl activates Sox15 *enhancer activity in both the socket and shaft cells*

The presence of Vvl binding sites in the 1.3-kb *Sox15* socket cell enhancer, combined with the Vvl expression pattern and the phenotype of *vvl* mutant bristles (Inbal et al., 2003) (see Fig. 4I), suggested that Vvl might function as one of the activators of the enhancer. Mutating the two identified Vvl sites (see Figs. 3A, B) in a manner that abolishes binding in vitro (see Fig. 3C), however, does not result in loss of reporter gene expression in the socket cell (Fig. 4G). When these sites are mutated in combination with the Su(H) site mutation, though, we find that ectopic reporter expression in the shaft cell is eliminated (Fig. 4H), suggesting a role for Vvl as an activator in the shaft cell at a minimum. The lack of an effect on reporter activity in the socket cell could indicate that Vvl is not required for activation in this cell, that Vvl could be working through additional sequence motifs in the socket enhancer region, or that Vvl functions, at least in part, indirectly in the activation of the socket module. To help distinguish among these possibilities, we examined the expression of a wild-type reporter gene in *vvl* mutant clones. First, however, we further investigated the effects of the loss of *vvl* function on the sensory organ lineage (Inbal et al., 2003).

vvl mutant bristles on the adult notum have smaller, deformed shafts and disorganized socket structures (Fig. 4I). At 42 hours APF, supernumerary cells expressing Cut protein are detected at these positions, suggesting that one or more cells in the lineage may undergo inappropriate divisions (Inbal et al., 2003). At 30 hours APF, we find that most *vvl* mutant bristle positions display five Cut-positive nuclei (Figs. 4K, L). To determine which cells might be dividing inappropriately, we stained wild-type and mutant tissue with anti-Prospero (Pros), which identifies the sheath cell, and anti-Elav, which labels the neuron. Wild-type positions contain one Pros-positive, Elav-negative cell (sheath) and one Elav-positive, Pros-negative cell (neuron). *vvl* mutant positions, however, often have three cells expressing Pros, and in most of these positions two of the Pros-positive cells also express Elav (Fig. 4L). By contrast, loss of *vvl* function appears not to affect expression of *ASE>nGFP*, as the GFP level and pattern are indistinguishable from that in wild-type territories (Figs. 4K, L), and reporter activity is detected in a single large Cut-positive nucleus at 30 hours APF (Fig. 4K). The normal expression of *ASE>nGFP* in *vvl* mutant territory indicates that *vvl* function is not required for the high levels of *Su(H)* expression in the socket cell, and suggests that loss of *vvl* activity does not affect specification of the socket fate, as *ASE>nGFP* is one of the earliest and most specific markers of this event.

The smaller size of the socket and shaft structures in adult external sensory organs within *vvl* mutant clones suggests that the external cells divide inappropriately; we confirm this by detecting two nuclei expressing *ASE>nGFP* in *vvl* clones around the time of eclosion (data not shown). By 36 hours APF, some *vvl* mutant microchaete positions contain up to eight Cut-positive nuclei, implying the occurrence of additional divisions of one or more normally post-mitotic cells in the bristle lineage. The observation that in these positions none of the nuclei are as large as either the wild-type socket cell or shaft cell nuclei may indicate that indeed the external cells are dividing at this time, and/or that they have failed to undergo the normal rounds of endoreplication.

It is only at 36 hours APF, 19 hours after it comes on in the wild-type positions outside of *vvl* clone boundaries, that we can first detect any expression of the *Sox7.5>nGFP* reporter in *vvl* mutant territory. Fig. 4J shows a mutant position at which weak activity of the *Sox7.5>nGFP* reporter is observed in two small nuclei; compare reporter gene expression in wild-type (open arrowhead) and mutant (filled arrowhead) cells. The lack of an expression delay in the case of the *ASE>nGFP* reporter suggests that Vvl is required for proper regulation of *Sox15* expression. However, since the *Sox15* reporter gene still becomes expressed in the socket cell, albeit with a long

delay, Vvl is likely not the only factor contributing to the activation of *Sox15* in this cell.

Sox15 *activity in the socket cell is required for normal mechanosensory organ function*

Because of the specificity and timing of its expression in the socket cell (see Fig. 1 and Supplementary Figs. S1 and S2), *Sox15* is well positioned as a potential regulator of differentiative gene expression subsequent to N signaling-dependent cell fate specification. The imprecise excision allele $Sox15^{4AA}$ proved to be a crucial reagent for analyzing the function of *Sox15* in the socket cell. Most of the endogenous expression pattern is affected in $Sox15^{4AA}$ homozygous flies; no transcript is detected in late third-instar leg imaginal discs and adult socket cells, and only weak expression is detected in wing and haltere discs (Fig. 5; compare A,B,C,E with G,H,I,K, respectively). The reduction of expression in these latter tissues results in a strong mutant phenotype in the adult wing and haltere, mainly affecting the formation of the hinge region (Figs. 5D, J). Such a phenotype is observed in a mis-expression mutant of another *Sox* gene, *Dichaete*, which Russell has proposed is possibly due to a dominant-negative effect on another endogenous *Sox* gene (Russell, 2000). Our observations would indicate that *Sox15* is that gene. The absence of leg disc expression in $Sox15^{4AA}$ homozygotes does not cause an obvious physical deformity, but results in reduced movement at the coxa-trochanter joint, preventing extension of the femur and interfering with the ability to walk (data not shown). The $Sox15^{4AA}$ deletion allele, in addition to removing the Initiator sequence at the gene's transcription start site, also removes the first exon and the proximal first intron sequence that contains the socket cell regulatory information, as indicated above. For these reasons $Sox15^{4AA}$ homozygotes fail to accumulate either *Sox 15* transcript or Sox15 protein in the socket cells of developing notum bristles at 30 hours APF (data not shown); likewise, they lack detectable *Sox15* transcript in the socket cells of abdominal bristles in pharate adults (Figs. 5E, K). Despite this lack of expression, the socket fate is properly specified, and by scanning electron microscopy we could detect no defect in the cuticular structures elaborated by the external cells (Figs. 5F, L). This result suggests that *Sox15* may be required instead for proper functional differentiation of the socket and possibly other cells in the bristle organ. Indeed, we have reported earlier that *Su(H)* ASE mutants, while displaying apparently normal specification of the socket cell fate, exhibit dramatic reductions of both the trans-epithelial potential (TEP) and the mechanoreceptor current (MRC) of adult mechanosensory organs (Barolo et al., 2000b; Walker et al., 2000). We find that $Sox15^{4AA}$ flies display only a mild TEP phenotype but show a strong MRC defect (Figs. 5M, N). Interestingly, this particular type of TEP/MRC phenotype is most often observed with mutations affecting neuronal, as opposed to socket cell, function (Avidor-Reiss et al., 2004).

We further investigated the $Sox15^{4AA}$ electrophysiological defect by looking for abnormalities in socket cell and/or neuronal internal structure using transmission electron microscopy. Consistent with the physiological phenotype, we find a reduction in the quantity of microtubules in the tubular body of the ciliary dendrite of the neuron (Figs. 5Q, R), which could account for the MRC defect. In addition to this neuronal abnormality, the socket cell exhibits signs of cell death. The cell loses its characteristic apical membrane folds, and the cytoplasm contains multiple large vesicles lacking electron density (Figs. 5O, P). Indeed, the results of studies presented in the Supplementary Material (see Fig. S4) suggest that the *Sox15* mutant socket cell undergoes necrosis (not apoptosis).

Sox15 *and* Su(H) *function together to repress* sv *expression and the shaft differentiation program in the socket cell*

Data presented above and previously (Barolo et al., 2000b) support the conclusion that both *Su(H)* ASE function and *Sox15* play critical roles in the differentiation, but not the fate specification, of the socket cell. Loss of either of these activities causes severe defects in the sensory organ's electrophysiological function. Given the similarities in mutant phenotype, we sought to examine the effect of loss of both *Sox15* and *Su(H)* ASE activity on the socket cell (Fig. 6). We find that removing the function of both of these factors results in a more

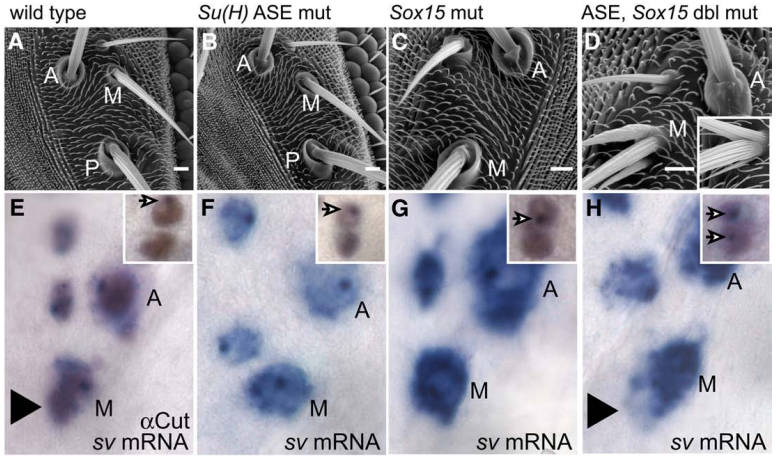


Fig. 6. Socket cell phenotype of the *Su(H)* ASE-*Sox15* double mutant. (A–D) Scanning electron micrographs of adult cuticle and (E–H) in situ hybridization detection of *sv* mRNA at 36 hours APF using a cDNA probe, highlighting orbital bristles of the following genotypes: (A, E) Wild type (w^{1118}). (B, F) $Su(H)^{ARS}/Su(H)^{SF8}$; *Su(H)RC-ΔASE*. (C, G) $Sox15^{4AA}$. (D,H) $Su(H)^{ARS}$ $Sox15^{4AA}/Su(H)^{SF8}$ $Sox15^{4AA}$; *Su(H)RC-ΔASE*. Inset in D shows bristle organ of a fly in which *sv* is misexpressed specifically in the socket cell (genotype *ASE-GAL4/+; UAS-sv/+*). A: anterior orbital; M: middle orbital; P: posterior orbital. Large arrowhead in E and H points to the position of the socket cell as indicated by Cut immunoreactivity shown in E. Insets in E–H are thoracic microchaete socket-shaft pairs stained with anti-Cut (brown) and a *sv* intron in situ hybridization probe to detect nascent transcript (purple nuclear dots, arrows).

dramatic socket differentiation defect than loss of either alone. The majority of double-mutant bristle positions display a convex, rather than a concave, socket-shaft cuticular interface (Figs. 6A–D). Some of these bulbous socket structures extend one or more small cuticular projections, while others extend one or two distinctly shaft-like structures (Fig. 6D). This latter phenotype is the most severe, and is primarily restricted to the medial orbital macrochaetes. It is reminiscent of the effect of mis-expressing in the socket cell the Pax transcription factor Shaven (Sv, formerly D-Pax2), a high-level regulator of shaft cell differentiation (see inset in Fig. 6D) (Kavaler et al., 1999). Sv is expressed in the bristle lineage starting late in SOP development and persisting through the specification of each of the post-mitotic cell types. Once specified, however, the socket cell and neuron fail to maintain Sv expression, and by 32 hours APF it is undetectable in the socket (Kavaler et al., 1999). We thus sought to examine whether the *Su(H)* ASE-*Sox15* double mutant phenotype we observe could result in part from the maintenance of *sv* transcription in the socket cell. Using RNA in situ hybridization probes generated from *sv* coding sequence, we often detect a cloud of transcript accumulation around the socket cell nuclei of orbital macrochaetes only in *Su(H)* ASE-*Sox15* double mutants at 36 hours APF (Figs. 6E–H). We confirmed late socket cell transcription of *sv* in the microchaete field using RNA in situ hybridization probes generated from intron sequence to detect nascent transcript (insets in Figs. 6E–H). These data indicate that *Su(H)* and *Sox15* act together early in socket cell differentiation to inhibit the maintenance of *sv* expression and hence to prevent inappropriate execution of the shaft differentiation program in this cell.

Discussion

Distinct transcriptional regulation of Sox15 *and the* Su(H) *ASE in the socket cell*

After *Su(H)* (Barolo et al., 2000b), *Sox15* is the second transcription factor gene known to be activated specifically in the postmitotic socket cell of the *Drosophila* external sensory organ lineage. Three observations reported here indicate that although both genes come to be expressed at high levels in this cell, the underlying regulatory logic may be quite different (Fig. 7).

The first is the distinct dynamics of ASE-stimulated *Su(H)* transcription versus *Sox15* expression. *Su(H)* is immediately activated at high levels following the specification of the socket cell, due at least in part to the establishment of an autoregulatory loop working through the ASE. *Sox15* expression, however, exhibits a significant delay between socket cell specification and the time peak levels of transcript accumulation are achieved.

The second observation concerns the role played by Vvl in the activation of the *Sox15* socket enhancer and the *Su(H)* ASE. Conserved within the ASE lies a motif, CATAAAT, that might act as a weak Vvl binding site (not shown) (Certel et al., 1996), suggesting the possibility that Vvl could play a part in the high-level activation of *Su(H)* in the socket cell. However, this appears not to be the case, since *ASE-GFP* is activated within the same temporal window, and just as strongly, in *vvl* mutant clones as in neighboring wild-type tissue. By contrast, while *Sox7.5>GFP* is also activated in *vvl* mutant sensory organs, there is a substantial delay in this expression, which is often not detectable until the socket cell has begun to divide aberrantly. At this time, neighboring wild-type sensory organs are already strongly expressing *Sox7.5>GFP*. Vvl thus appears to be one factor present in the socket cell that is necessary for the full activation of *Sox15*, but not of *Su(H)*.

Finally, there is the observed role of N-activated Su(H) in contributing to the transcriptional activation of the *Sox15* socket enhancer versus the *Su(H)* ASE. A major difference between the two genes is made apparent by the contrasting effects on reporter gene expression of mutating the high-affinity Su(H) site(s) in their respective socket cell enhancers. In the case of the *Su(H)* ASE, mutation of the Su(H) sites causes a strong reduction in socket cell activity at early times, along with ectopic activity in the shaft cell; by the adult stage, the mutant enhancer is inactive (Barolo et al., 2000b). Thus, N-activated Su(H) contributes critically to the transcriptional activation of the *Su(H)* ASE. The Su(H)-site-mutant *Sox15* enhancer, on the other hand, shows no apparent diminution of its socket cell

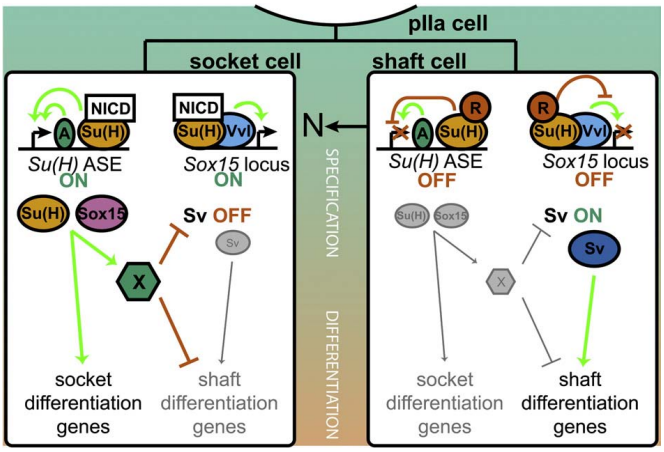


Fig. 7. Summary model of the collaborative roles of *Su(H)* and *Sox15* as N pathway targets required for socket cell differentiation. In the N signal-resistant shaft cell (right), Su(H) in its default repression mode keeps both the *Su(H)* ASE and the *Sox15* socket enhancer off (in the latter case, despite the presence of Vvl). The result is that the Pax family factor Sv accumulates to high levels in the shaft cell and directs its differentiation program. Successful activation of the N pathway in the socket cell (left) relieves Su(H)-mediated repression of the two enhancers in this cell. The initially low basal level of Su(H), now complexed with NICD and Mam, contributes critically to the rapid (auto)activation of the *Su(H)* ASE, but acts permissively on the *Sox15* enhancer, allowing Vvl and other factors to activate *Sox15* gradually. Su(H) and Sox15 then function collaboratively (not hierarchically) to activate target genes required primarily for the later (physiological) phase of socket cell differentiation. They also collaborate to shut off continued transcription of *sv*, probably indirectly through the action of a hypothetical repressor X. This inactivates the shaft differentiation program in the socket cell.

activity early (when it also drives ectopic expression in the shaft cell), and remains fully active in the pharate adult. In the case of *Sox15*, then, activation of Su(H) by the N signaling event appears to serve only the purpose of relieving Su(H)-mediated default repression; activation of the enhancer is evidently accomplished entirely through the action of other factors such as Vvl. This distinction in the role of N signaling in enhancer activation has been referred to as "Notch instructive" [*Su(H)* ASE] versus "Notch permissive" (*Sox15* socket enhancer) (Bray and Furriols, 2001).

Functions of Sox15 *and* Su(H) *in socket cell and sensory organ differentiation*

Our investigation of the loss-of-function phenotype of *Sox15* has revealed that, like *Su(H)*, it has an important role in controlling the socket cell differentiation program (Fig. 7). Comparison of the phenotypic effects of losing *Sox15* function, *Su(H)* function, or both, suggests an incomplete overlap in the target gene batteries regulated by the two factors. Loss of either *Sox15* or *Su(H)* ASE activity causes a serious defect in mechanosensory organ function. The lack of *Su(H)* ASE activity confers the more severe phenotype, including significant reductions of both TEP and MRC. The TEP defect signifies an inability of the socket cell to establish the receptor lymph cavity itself, the proper ionic composition of the receptor lymph, or a combination of the two. The genes required for these events have yet to be identified, but it is likely that Su(H) plays a role in regulating their expression in the socket cell. Sox15, on the other hand, does not appear to share this role, based on the apparent lack of a major TEP defect in *Sox15* mutants. Instead, Sox15 appears to regulate targets that contribute to socket cell viability. Without these target factors, the cell eventually becomes necrotic. In addition, the principal physiological phenotype of *Sox15* mutants is the MRC defect, which is also conferred by loss of *Su(H)* ASE function. Loss of MRC is indicative of a failure in neuronal function, yet both *Sox 15* and the *Su(H)* ASE are active specifically in the socket cell. This apparent paradox indicates an important role for the socket cell as a support cell for the mechanosensory neuron. To date three proteins — Sox15 (this paper), Su(H) (Barolo et al., 2000b), and the cytochrome P450 Cyp303a1 (Willingham and Keil, 2004) — expressed in and required specifically for socket cell differentiation appear to contribute to neuronal function in mechanosensation. Given that the socket cell envelops the other cells of the sensory organ as it develops, the socket may be intimately involved in their normal differentiation and in the establishment of structural and functional connectivity between them. Defects in these processes could readily manifest themselves in an MRC phenotype. Thus, the abnormal microtubule bundling in the sensory dendrite in *Sox15* mutants may very well be the result of a defect in the socket cell's ability to contribute as it should to the neuron's normal development. It is unclear at this point if the dendrite defect is due to a failure to activate Sox15-dependent target genes directly involved in the socket cell's support function, or if it is an indirect consequence of the degeneration of the socket cell.

Inhibition of the sister cell differentiation program is one consequence of N-mediated cell fate specification in the bristle lineage

Previous studies have established that both daughters of the pIIa secondary precursor division are bipotent cells that can adopt either the shaft or socket cell fate (Bang and Posakony, 1992). Asymmetric N signaling specifies that the posterior daughter expresses only the signal-dependent socket fate and the anterior daughter only the signal-independent shaft fate. Correspondingly, our investigation of socket cell fate specification has largely focused on its positive aspects; i.e., those ways in which the N signaling event promotes the socket cell from the "default" (signal-independent) shaft fate to the alternative fate, triggering its execution of the distinctive socket differentiation program. We have shown here that socket cell-specific activation of *Sox15* expression is an important component of this program. But the present study has also revealed the other side of the coin, by showing that the N signaling event also results in the activation of a mechanism for suppressing in the socket cell the capacity to execute the shaft differentiation program (Fig. 7). We have shown that this suppression mechanism involves the combined action of *Sox15* and *Su(H)* in inhibiting transcription of the *sv* gene, which encodes a Pax transcription factor that is a high-level activator of the shaft differentiation program (Kavaler et al., 1999). Without this inhibition, the socket cell generates both socket and shaft cuticular structures. It is clear, then, that much of the network circuitry necessary for the execution of the shaft differentiation program remains intact in the socket cell even after its fate has been specified. Our results show that robust N-mediated cell fate specification in the mechanosensory bristle lineage involves not only promoting the signal-dependent fate, but also actively inhibiting the alternative program.

It is likely that at least Su(H)'s role in inhibiting *sv* expression in the socket cell is indirect, and occurs via an as yet unidentified repressor (Fig. 7). An attractive candidate for this factor X would be one or more basic helix-loop-helix (bHLH) repressors encoded in the *Enhancer of split* Complex [E(spl)-C]. Multiple E(spl)-C bHLH repressor genes are activated directly by Su(H) in response to N signaling in a variety of developmental contexts (Bailey and Posakony, 1995; Lecourtois and Schweisguth, 1995). Consistent with this possibility, we have observed (data not shown) that socket cell-specific overexpression of E(spl)m7-VP16, a form of the E(spl)m7 bHLH repressor that has been converted to a strong activator, phenocopies the ectopic-shaft effect of *sv* overexpression in the same cell (see Fig. 6D, inset).

Separable regulation of early and late phases of the socket cell differentiation program

The results of this and earlier studies (Barolo et al., 2000b) afford us a glimpse of the regulatory architecture of the socket differentiation program, which is set in motion by the N signaling event that specifies the socket cell fate. It seems useful to distinguish two broad phases of this program, which no doubt overlap each other in time and are also very likely to share at least some components of the regulatory network. These two phases might be referred to as the earlier "morphogenetic" and the later "physiological" subdivisions of the socket program. The distinction is prompted by our observations of the phenotypes conferred by loss of the two socket cell-specific transcription factor activities identified so far, Su(H) and Sox15. In both cases, we find that many characteristic aspects of the socket's cellular differentiation proceed completely normally, most notably the construction of the complex socket cuticular structure that surrounds the shaft structure (morphogenesis). By contrast, loss of Su(H) or Sox15 function in the socket cell results in major deficits in the electrophysiological capacity of the sensory organ (physiological differentiation). As described above, the specifics of these deficits differ for *Su(H)* versus *Sox15* mutants, and include distinctive cell-autonomous defects in the socket cell and defects in other cells likely due to the failure of some aspects of the socket cell's support function. But the phenotypic commonalities (emphasizing the physiological and not the morphogenetic) are striking nonetheless. It is perhaps reasonable to speculate that transcription factors like Su(H) and Sox15 that are activated for the first time in the sensory organ lineage specifically in the socket cell will tend to function primarily in the later physiological phase of the differentiative program. By contrast, we may expect that the earlier morphogenetic phase is controlled primarily by factors first expressed earlier in the lineage, at least in the pIIa precursor cell and perhaps in the SOP. Vvl exemplifies this notion: It is first expressed in the SOP, and loss of its activity causes visible defects in the socket cuticular structure, as well as aberrations in the mitotic status of the normally postmitotic socket cell. Investigation of the roles of additional transcriptional regulators in

directing the socket differentiation program will test the viability of this broad conceptual framework.

Overall, our comparison of the roles of Sox15, Su(H), and Vvl in controlling aspects of the socket differentiation program indicates that they function largely in parallel, and collaboratively, rather than in a hierarchical fashion. This may suggest that the socket program will prove to be characterized by an ensemble of such parallel regulatory inputs that collectively direct the complex differentiation of the cell. It is perhaps useful to note that this picture contrasts already with what is known about the control of the shaft differentiation program, which is dominated by the function of Sv as a high-level regulator (Kavaler et al., 1999). Whether this reflects some important difference in how the differentiative programs of N-responsive versus N-non-responsive cell types are controlled will become clearer as we learn more about the gene regulatory network that underlies mechanosensory organ development.

Acknowledgments

We dedicate this paper to the memory of Danny L. Brower, undergraduate advisor, role model, and friend to S.W.M. We are indebted to Jamy Peng for making the GST-Sox15 construct, to Tammie Stone for prepping His-Su(H) protein, and to Sui Zhang for critical steps in Sox15 antibody production and characterization. We would like to thank Scott Barolo, Nick Reeves, and Mark Rebeiz for useful discussions throughout this project, members of the Posakony lab for critical comments on the manuscript, and two anonymous reviewers for their helpful suggestions. S.W.M. received support from NIH predoctoral training grant GM07240. This work was supported by Grant 07-04-01127 from the Russian Foundation for Basic Research to A.P. and by NIH grants GM062279 and GM046993 to J.W.P.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ydbio.2009.02.009.

References

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25, 3389–3402.

Avidor-Reiss, T., Maer, A.M., Koundakjian, E., Polyanovsky, A., Keil, T., Subramaniam, S., Zuker, C.S., 2004. Decoding cilia function: defining specialized genes required for compartmentalized cilia biogenesis. Cell 117, 527–539.

Bailey, A.M., Posakony, J.W., 1995. Suppressor of Hairless directly activates transcription of *Enhancer of split* Complex genes in response to Notch receptor activity. Genes Dev. 9, 2609–2622.

Bang, A.G., Posakony, J.W., 1992. The *Drosophila* gene *Hairless* encodes a novel basic protein that controls alternative cell fates in adult sensory organ development. Genes Dev. 6, 1752–1769.

Bang, A.G., Hartenstein, V., Posakony, J.W., 1991. *Hairless* is required for the development of adult sensory organ precursor cells in *Drosophila*. Development 111, 89–104.

Barolo, S., Posakony, J.W., 2002. Three habits of highly effective signaling pathways: principles of transcriptional control by developmental cell signaling. Genes Dev. 16, 1167–1181.

Barolo, S., Carver, L.A., Posakony, J.W., 2000a. GFP and β-galactosidase transformation vectors for promoter/enhancer analysis in *Drosophila*. Biotechniques 29, 726–732.

Barolo, S., Walker, R.G., Polyanovsky, A.D., Freschi, G., Keil, T., Posakony, J.W., 2000b. A Notch-independent activity of *Suppressor of Hairless* is required for normal mechanoreceptor physiology. Cell 103, 957–969.

Barolo, S., Stone, T., Bang, A.G., Posakony, J.W., 2002. Default repression and Notch signaling: Hairless acts as an adaptor to recruit the corepressors Groucho and dCtBP to Suppressor of Hairless. Genes Dev. 16, 1964–1976.

Barolo, S., Castro, B., Posakony, J.W., 2004. New *Drosophila* transgenic reporters: insulated P-element vectors expressing fast-maturing RFP. Biotechniques 36, 436–442.

Bellen, H., Levis, R., Liao, G., He, Y., Carlson, J., Tsang, G., Evans-Holm, M., Hiesinger, P., Schulze, K., Rubin, G., Hoskins, R., Spradling, A., 2004. The BDGP gene disruption project: single transposon insertions associated with 40% of Drosophila genes. Genetics 167, 761–781.

Bray, S.J., 2006. Notch signalling: a simple pathway becomes complex. Nat. Rev., Mol. Cell Biol. 7, 678–689.

Bray, S., Furriols, M., 2001. Notch pathway: making sense of Suppressor of Hairless. Curr. Biol. 11, R217–R221.

Brizuela, B.J., Elfring, L., Ballard, J., Tamkun, J.W., Kennison, J.A., 1994. Genetic analysis of the *brahma* gene of *Drosophila melanogaster* and polytene chromosome subdivisions 72AB. Genetics 137, 803–813.

Brown, N.H., Kafatos, F.C., 1988. Functional cDNA libraries from *Drosophila* embryos. J. Mol. Biol. 203, 425–437.

Castro, B., Barolo, S., Bailey, A.M., Posakony, J.W., 2005. Lateral inhibition in proneural clusters: Cis-regulatory logic and default repression by Suppressor of Hairless. Development 132, 3333–3344.

Certel, K., Anderson, M.G., Shrigley, R.J., Johnson, W.A., 1996. Distinct variant DNA-binding sites determine cell-specific autoregulated expression of the *Drosophila* POU domain transcription factor Drifter in midline glia or trachea. Mol. Cell. Biol. 16, 1813–1823.

Cooley, L., Kelley, R., Spradling, A., 1988. Insertional mutagenesis of the *Drosophila* genome with single P elements. Science 239, 1121–1128.

Cooper, M.T., Tyler, D.M., Furriols, M., Chalkiadaki, A., Delidakis, C., Bray, S., 2000. Spatially restricted factors cooperate with Notch in the regulation of *Enhancer of split* genes. Dev. Biol. 221, 390–403.

Cremazy, F., Berta, P., Girard, F., 2001. Genome-wide analysis of Sox genes in *Drosophila melanogaster*. Mech. Dev. 109, 371–375.

Furriols, M., Bray, S., 2001. A model Notch response element detects Suppressor of Hairless-dependent molecular switch. Curr. Biol. 11, 60–64.

Gho, M., Bellaïche, Y., Schweisguth, F., 1999. Revisiting the *Drosophila* microchaete lineage: a novel intrinsically asymmetric cell division generates a glial cell. Development 126, 3573–3584.

Goldstein, L.S.B., Fyrberg, E.A. (Eds.), 1994. *Drosophila melanogaster*: Practical Uses in Cell and Molecular Biology. Methods Cell Biol. 44. Academic Press, San Diego.

Golic, K.G., 1991. Site-specific recombination between homologous chromosomes in *Drosophila*. Science 252, 958–961.

Golic, K., Lindquist, S., 1989. The FLP recombinase of yeast catalyzes site-specific recombination in the *Drosophila* genome. Cell 59, 499–509.

Hartenstein, V., Posakony, J.W., 1989. Development of adult sensilla on the wing and notum of *Drosophila melanogaster*. Development 107, 389–405.

Hartenstein, V., Posakony, J.W., 1990. A dual function of the *Notch* gene in *Drosophila* sensillum development. Dev. Biol. 142, 13–30.

Ho, S.N., Hunt, H.D., Horton, R.M., Pullen, J.K., Pease, L.R., 1989. Site-directed mutagenesis by overlap extension using the polymerase chain reaction. Gene 77, 51–59.

Inbal, A., Levanon, D., Salzberg, A., 2003. Multiple roles for *u-turn/ventral veinless* in the development of *Drosophila* PNS. Development 130, 2467–2478.

Janknecht, R., De Martynoff, G., Lou, J., Hipskind, R.A., Nordheim, A., Stunnenberg, H.G., 1991. Rapid and efficient purification of native histidine-tagged protein expressed by recombinant vaccinia virus. Proc. Natl. Acad. Sci. U. S. A. 88, 8972–8976.

Jarman, A.P., 2002. Studies of mechanosensation using the fly. Hum. Mol. Genet. 11, 1215–1218.

Juven-Gershon, T., Hsu, J.Y., Theisen, J.W., Kadonaga, J.T., 2008. The RNA polymerase II core promoter - the gateway to transcription. Curr. Opin. Cell Biol. 20, 253–259.

Kavaler, J., Fu, W., Duan, H., Noll, M., Posakony, J.W., 1999. An essential role for the *Drosophila Pax2* homolog in the differentiation of adult sensory organs. Development 126, 2261–2272.

Kernan, M., Cowan, D., Zuker, C., 1994. Genetic dissection of mechanosensory transduction: Mechanoreception-defective mutations of *Drosophila*. Neuron 12, 1195–1206.

Koelzer, S., Klein, T., 2006. Regulation of expression of Vg and establishment of the dorsoventral compartment boundary in the wing imaginal disc by Suppressor of Hairless. Dev. Biol. 289, 77–90.

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

Lai, E.C., 2004. Notch signaling: control of cell communication and cell fate. Development 131, 965–973.

Lai, E.C., Bodner, R., Posakony, J.W., 2000. The *Enhancer of split* Complex of *Drosophila* includes four Notch-regulated members of the Bearded gene family. Development 127, 3441–3455.

Lecourtois, M., Schweisguth, F., 1995. The neurogenic Suppressor of Hairless DNA-binding protein mediates the transcriptional activation of the *Enhancer of split* Complex genes triggered by Notch signaling. Genes Dev. 9, 2598–2608.

Ma, Y., Certel, K., Gao, Y., Niemitz, E., Mosher, J., Mukherjee, A., Mutsuddi, M., Huseinovic, N., Crews, S.T., Johnson, W.A., Nambu, J.R., 2000. Functional interactions between *Drosophila* bHLH/PAS, Sox, and POU transcription factors regulate CNS midline expression of the *slit* gene. J. Neurosci. 20, 4596–4605.

Mercado-Pimentel, M.E., Jordan, N.C., Aisemberg, G., 2002. Affinity purification of GST fusion proteins for immunohistochemical studies of gene expression. Protein Expr. Purif. 26, 260–265.

Morel, V., Schweisguth, F., 2000. Repression by Suppressor of Hairless and activation by Notch are required to define a single row of *single-minded* expressing cells in the *Drosophila* embryo. Genes Dev. 14, 377–388.

Morel, V., Lecourtois, M., Massiani, O., Maier, D., Preiss, A., Schweisguth, F., 2001. Transcriptional repression by Suppressor of Hairless involves the binding of a Hairless-dCtBP complex in *Drosophila*. Curr. Biol. 11, 789–792.

Nellesen, D.T., Lai, E.C., Posakony, J.W., 1999. Discrete enhancer elements mediate selective responsiveness of *Enhancer of split* Complex genes to common transcriptional activators. Dev. Biol. 213, 33–53.

O'Neill, J.W., Bier, E., 1994. Double-label *in situ* hybridization using biotin and digoxigenin-tagged RNA probes. Biotechniques 17, 870, 874–875.

Posakony, J.W., 1994. Nature versus nurture: asymmetric cell divisions in Drosophila bristle development. Cell 76, 415–418.
Rebeiz, M., Posakony, J.W., 2004. GenePalette: A universal software tool for genome sequence visualization and analysis. Dev. Biol. 271, 431–438.
Reeves, N., Posakony, J.W., 2005. Genetic programs activated by proneural proteins in the developing *Drosophila* PNS. Dev. Cell 8, 413–425.
Roseman, R.R., Johnson, E.A., Rodesch, C.K., Bjerke, M., Nagoshi, R.N., Geyer, P.K., 1995. A *P* element containing *suppressor of Hairy-wing* binding regions has novel properties for mutagenesis in *Drosophila melanogaster*. Genetics 141, 1061–1074.
Rubin, G.M., Spradling, A.C., 1982. Genetic transformation of *Drosophila* with transposable element vectors. Science 218, 348–353.
Russell, S., 2000. The *Drosophila* dominant wing mutation *Dichaete* results from ectopic expression of a Sox-domain gene. Mol. Gen. Genet. 263, 690–701.
Schweisguth, F., Posakony, J.W., 1992. *Suppressor of Hairless*, the Drosophila homolog of the mouse recombination signal-binding protein gene, controls sensory organ cell fates. Cell 69, 1199–1212.
Tautz, D., Pfeifle, C., 1989. A non-radioactive *in situ* hybridization method for the localization of specific RNAs in *Drosophila* embryos reveals translational control of the segmentation gene *hunchback*. Chromosoma 98, 81–85.
Walker, R.G., Willingham, A.T., Zuker, C.S., 2000. A *Drosophila* mechanosensory transduction channel. Science 287, 2229–2234.
Willingham, A.T., Keil, T., 2004. A tissue specific cytochrome P450 required for the structure and function of *Drosophila* sensory organs. Mech. Dev. 121, 1289–1297.
Xu, T., Rubin, G.M., 1993. Analysis of genetic mosaics in developing and adult *Drosophila* tissues. Development 117, 1223–1237.

Chapter Three, in full, is a reprint of the material as it appears in Developmental Biology 2009. Miller, S.W., Avidor-Reiss, T., Polyanovsky, A., and Posakony, J. W., Complex interplay of three transcription factors in controlling the tormogen differentiation program of Drosophila mechanoreceptors. *Dev Biol*. 2009 May 15;**329**(2):386-399. The dissertation author was the primary investigator and author of this paper. James W. Posakony provided critical oversight and discussion throughout this work. Tomer Avidor-Reiss performed the electrophysiological studies, and Andrey Polyanovsky performed the transmission electron microscopy analysis.

CHAPTER FOUR:

Auto-regulation, proneural synergy, and mutual inhibitory interactions shape expression of *shaven* and influence the tormogen and tricogen differentiation programs

Abstract

The challenge for developmental biology is to link information about cell fate determination mechanisms with changes in gene expression between different cell fates. Based upon its phenotype, the Pax2 transcription factor, Shaven, is believed to be a high-level regulator of the differentiation program of shaft cell of *Drosophila melanogaster* mechanoreceptors. Explaining the logic of Sv expression has long been challenging because its pattern within the sensory organ lineage does not follow a pattern consistent with the pattern of N signaling, the major fate determining mechanism. In this study we present an initial examination of *sv* cis-regulatory control. We demonstrate the activity of an enhancer that recapitulates the expression of Sv in the lineage, and present evidence that auto-regulation is one of the major activating forces. Lastly, we present evidence that the major trans-regulators in the sister socket and shaft cells have an antagonistic relationship, perhaps contributing to the ultimate specificity of expression in the post-mitotic cells.

Introduction

"Set it, and forget it" for developmental programming

Two of the most important themes in establishing cell fate during the development of an organism are the activities of extrinsic and autonomous determinants. Autonomous determinants convey a particular outcome to all cells that contain that determinant; while extrinsic determinants come from the surrounding environment to enact changes in all cells that respond to that signal. Both types of information are used multiple times throughout the development of the *Drosophila* mechanosensory organ lineage. Such a strategy allows cells to proceed down a particular developmental path--following the orders of the autonomous determinants—unabated unless they receive information otherwise from an extrinsic source. The proneural proteins direct all cells that continue to express them down a path toward sensory organ precursor (SOP) development (Ruiz-Gomez and Modolell, 1987; Romani et al., 1989; Jimenez and Campos-Ortega, 1990); how unfortunate that so many of them receive a signal from the Notch pathway that prevent them from continuing this route, save a select few (Cabrera, 1990; Simpson, 1990).

These successful SOPs discover a new chest of autonomous determinants (see Chapter One), only to divide and distribute their bounty to their daughters. Both of these daughters receive Senseless, which is necessary to become an external cell, but only in conjunction with a N signal (Jafar-Nejad et al., 2006). One daughter alone, however, awakens to find it has been given Numb (Knoblich et al., 1995), an intrinsic protein that thwarts the efforts of Notch (Frise et al., 1996; Guo et al., 1996)and allows this daughter to express Prospero, conferring an internal cell fate (Reddy and Rodrigues, 1999). The other daughter, helpless against the N signal emanating from the Prospero-expressing cell has no choice but to bow to the whims of Senseless and follow the path to external cell production.

Shaven as a high-level regulator in a post-mitotic differentiation program

Once restricted to the external cell fate, this daughter of the SOP, termed pIIa, must now divide as well and produce the two terminally differentiated cells, the socket and the shaft cells. The pIIa cells makes new Numb protein (M. Rebeiz, S. Miller, and J. Posakony,

unpublished data) and gives it to one of the daughters, preventing receipt of a N signal. The other cell faithfully responds to its N instruction and becomes the socket cell. The socket cell activates high levels of Ventral veins lacking (Vvl) (Inbal et al., 2003), Suppressor of Hairless [Su(H)] (Gho et al., 1996; Barolo et al., 2000b), and Sox15 proteins. If the socket cell fails to activate these latter two, it begins to take on characteristics of the shaft cell, including the expression of the gene *shaven* (Miller et al., 2009) (*sv)*. Sv is an autonomous determinant of the shaft differentiation program (Kavaler et al., 1999). Without Sv expression, the cell fails to embody the shaft fate, and can even take on socket cell characteristics. If the socket cell expresses Sv, it will likewise build shaft-like projections, as though it thinks it should be a shaft cell (S. Barolo and J. Posakony, unpublished results). Given that Sv is both necessary and sufficient to promote activation of the shaft cell differentiation program, it is considered an autonomous, high-level regulator of this program.

Shaven expression and Notch signaling

To merely state that Sv is expressed in the shaft cell does not do justice to the intricacies of its expression pattern. Sv protein and transcript is first detected in the SOP, and continues to be expressed until the ultimate birth of all the terminally differentiated daughter cells (Kavaler et al., 1999; Figure 4.1A). As these cells differentiate, however, two cells, the shaft and the internal sheath cell, upregulate Sv while the others lose their expression. This pattern of expression is particularly puzzling because it is completely inconsistent with N activity in these two cells, and thus hypotheses for explaining how Sv becomes activated in these cells are not straightforward; while the shaft cell is a N-independent cell fate, the sheath is N-dependent. If N is used to activate Sv expression in the sheath, how is it prevented from being activated in the socket cell, the other N-dependent post-mitotic fate? Similarly, if N is used to repress Sv expression in the socket cell, how does it still allow Sv expression in the sheath? The simplest model is that Sv expression in these cells truly is N-independent, but at the end of the day, everything in the bristle lineage is linked to N signaling eventually. Discovering how this unique

expression pattern is generated, then, will likely reveal undiscovered players in gene regulation in the mechanosensory organ lineage.

In this study we take a cursory look at the regulation of *sv* expression in the shaft cell. We describe the activity of an enhancer that directs expression in the mechanosensory organ lineage that parallels Sv protein accumulation quite well. We also deduce some of the key regulators of Sv expression in the lineage by analyzing conservation and function of binding sites for known factors upon the activity of this enhancer. Lastly, we expand upon previous work on the socket differentiation program and present evidence that builds upon the theme of an antagonistic relationship between the differentiation programs in these two cells.

Materials and Methods

Fly stocks

ASE>GAL4 has been used previously (Miller et al., 2009), svWT>GAL4 was created as described below, UAS-Sv was constructed by Josh Kavaler and has been reported (Kavaler et al., 1999), UAS-Sox15 was cloned by Par Towb and Jamy Peng as described below, UAS-m7 and UAS-m7::VP16 have both been reported previously (Jimenez and Ish-Horowicz, 1997; Ligoxygakis et al., 1999).

Transgenic constructs

The svWT enhancer was amplified from sequencing strain genomic DNA as a 2 kb fragment with Asc I and Sac II restriction sites incorporated into the primer sequences and cloned into pH-Stinger (Barolo et al., 2000a). Mutations were generated by the overlap extension PCR approach (Ho et al., 1989) using the svWT fragment as template. All constructs were sequenced as pGEM clones before incorporating into pH-Stinger and injecting. All svRmut pH-Stinger transgenic flies were used to amplify the MCS insert and verify

sequence given the lack of reporter expression: in all cases the sequence was as desired. The svWT>GAL4 construct was digested from the pH-Stinger clone using the same restriction enzymes and cloned into H-GAL4, which is described by Castro *et al* (Castro et al., 2005). UAS-Sox15 was generated by amplifying Sox15 coding sequence from w^{1118} first strand cDNA and cloning it into pUAST through EcoR I + Xho I linked ends.

Antibody generation

Guinea pig anti-Sv antibody was generated by Pocono Rabbit Farm and Laboratories using GST-Sv protein as an immunogen. GST-Sv protein was produced identically as has been described for Sox15 antibody production (Miller et al., 2009). Two guinea pigs and two rabbits were immunized. Stains shown in this work were using the first bleed from guinea pig 1997 because it was awesome.

Immunohistochemistry

All tissue preparation prior to antibody incubation was performed as described (Miller et al., 2009). Rabbit anti-Su(H) was purchased

from Santa Cruz Biotechnology and used at 1:1000. Guinea pig anti-Sv was used at 1:2000. Rabbit anti-Sox15 has been described previously (Miller et al., 2009), and was used at 1:100 – 1:300.

Adult tissue preparation

Adult cuticle for phenotypic photography was prepared as described (Miller et al., 2009).

Gene diagrams

Gene diagram and alignment was generated using the latest version of GenePalette, http://www.genepalette.org (Rebeiz and Posakony, 2004), and images were edited in Adobe Illustrator.

Microscopy

All confocal images were collected as described (Miller et al., 2009).

Primers

Sequences of all oligonucleotides used in this work are available upon request.

Results

A 2 kb region upstream of sv *can recapitulate all aspects of Sv protein accumulation in the PNS lineage*

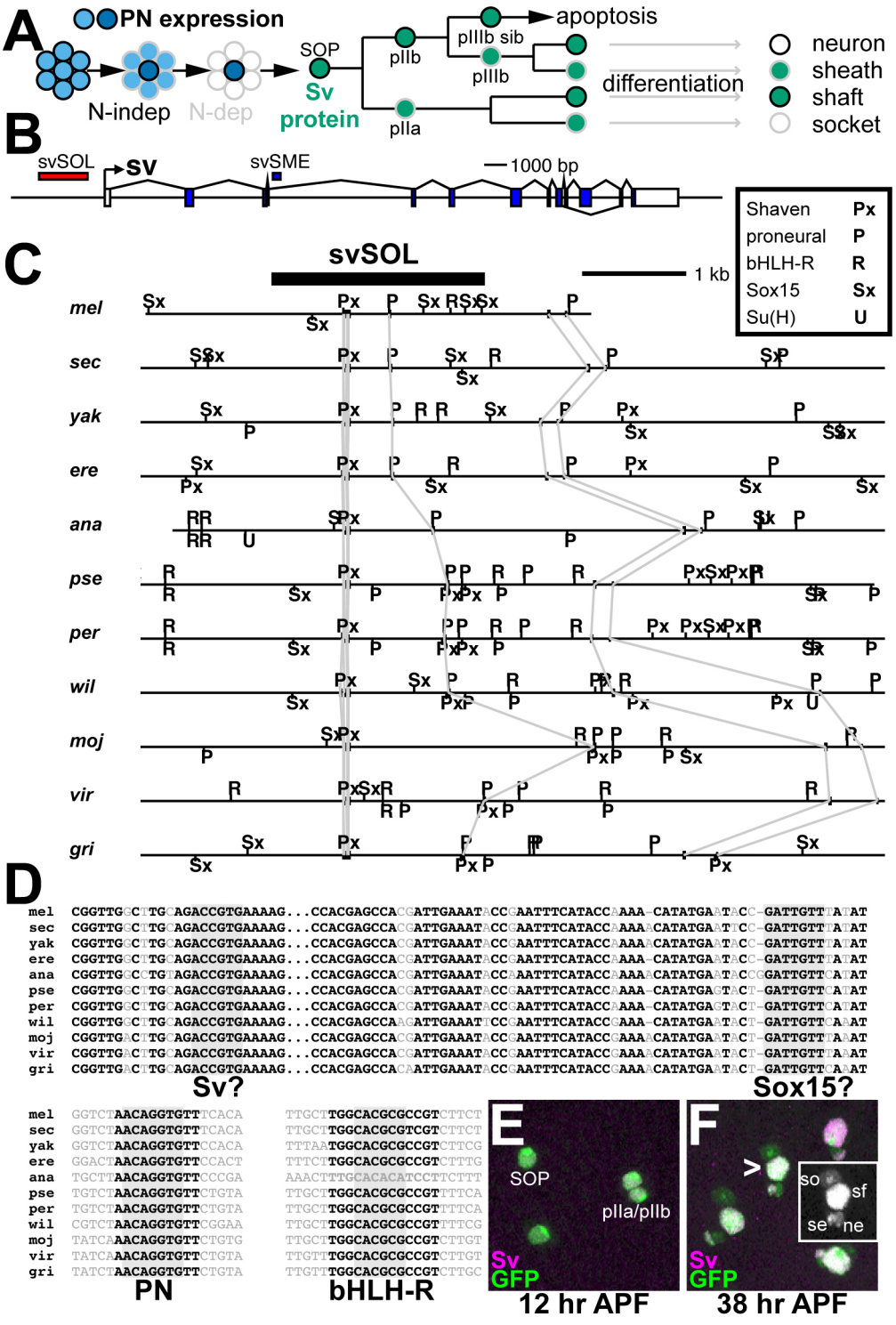
Some of the first classes identified mutations in the *sv* locus caused two distinct phenotypes—one affecting eye development, the other affecting the mechanosensory bristle pattern—that were later proposed to be the result of disruptions of separate enhancers for the same locus (Fu et al., 1998). Indeed, an enhancer directing expression in the eye has been described that was not noted to give expression in the bristle lineage (Flores et al., 2000). This enhancer is located in an intron precisely in the region of lesions giving eye phenotypes. Deletions which cause bristle defects alone map to the region 5' to the promoter of *sv*, implying such a location for the mechanosensory lineage enhancer. Consistent with these mapping data, we find bristle lineage activity in this region (S. Barolo and J. Posakony, unpublished data; Figure 4.1B). Intriguingly, the enhancer activity largely recapitulates the expression of endogenous Sv protein (Figure 4.1E,F): reporter activity first becomes activated in SOPs, and expression is mostly uniform throughout the lineage past 24 hrs APF;

between 24 and 36 hrs APF, GFP expression is slowly lost from the socket and neuron until eventually reporter expression is only detected in the shaft and sheath. While slightly delayed relative to endogenous Sv protein, these results at minimum suggest that the dynamics of Sv protein accumulation is based in a transcriptional, as opposed to degradating/maintaining, mechanism.

The sv2kb region contains conserved binding sites for the proneural proteins and Sv itself that are required for proper regulation of the enhancer

We chose to examine the pattern of sequence conservation across the twelve sequenced *Drosophila* genomes to find clues for important regulatory sequences. This region appears to be lacking from the annotated sequence of *Drosophila simulans*, and throughout the remaining eleven species there is a general dearth of conservation (Figure 4.1C). There is a proneural consensus sequence that is conserved in all eleven, as well as a bHLH-repressor binding site conserved in ten of the eleven species, with *Drosophila ananassae* as the exception. Finally, though much of this 2 kb region shows little

Figure 4.1: Identification of and conservation within a sheath/shaft enhancer upstream of *sv*.
(A) Diagram of Sv protein expression in the *Drosophila* mechanosensory organ lineage. A developmental series progressing from the proneural cluster at the left to the terminally differentiated cell types at the right. Proneural protein expression is indicated by blue, and Sv protein by green. Notch-dependent fates are outlined in grey and Notch-independent fates in black. Note that ultimately Sv accumulates in one N-dependent fate, the sheath, and one N-independent fate, the shaft. (B) Diagram of the *sv* locus, noting the locations of the two identified enhancers. svSME expresses in the eye (Noll *et al*., 1998), and svSOL (SOL = sensory organ lineage) is described here. (C) Alignment of the region including and surrounding the svSOL enhancer, noting the location of binding sites as indicated in the legend and text. Sequence blocks identical for at least 9 bp in all eleven genomes are connected by grey lines; those within the svSOL boundaries are explicitly shown in D. (D) Alignment of identical sequences from the eleven genomes in C. Identities are in bold, positions with mismatch are in regular, grey text. Putative binding sites for Sv, proneural, proteins, and bHLH-Rs are indicated by a grey background box, as is a highly conserved sequence that mismatches a Sox15 consensus. This mismatch site was not analyzed. (E) Expression of svSOL (a.k.a. svWT) at 12 hrs APF is coincident with Sv protein expression (magenta) in SOPs and newly divided SOPs. (F) At 38 hrs APF, GFP expression has been mostly downregulated in socket and neuron, while upregulated in sheath and shaft. The majority of positions at this stage have no detectable Sv protein in socket and neuron (magenta). Arrowhead points to position shown in inset, which shows only the GFP channel and indicates the cellular identities.


A
PN expression
N-indep
N-dep
SOP
Sv protein
pIIb
pIIa
pIIIb sib
pIIIb
apoptosis
differentiation
neuron
sheath
shaft
socket
B
svSOL
sv
svSME
1000 bp
Shaven Px
proneural P
bHLH-R R
Sox15 Sx
Su(H) U
C
svSOL
1 kb
mel
sec
yak
ere
ana
pse
per
wil
moj
vir
gri
D
Sv?
Sox15?
PN
bHLH-R
E
SOP
pIIa/pIIb
Sv GFP
12 hr APF
F
so
sf
ne
Sv GFP
38 hr APF

conservation, a tight cluster of 230 bp that is highly conserved above background, with three blocks that contain >10 bp of identity across all eleven genomes (Figure 4.1D). Within one of these identical blocks is the sequence ACCGTG, which we have found can be bound by Sv protein in vitro (T. Stone and J. Posakony, unpublished data).

Most notably, we do not find binding sites for Su(H) within this 2 kb sequence, nor do we find conserved instances of the sequence AACAATR, which can be bound by Sox15 protein in vitro (J. Peng, N. Reeves, S. Barolo, and J. Posakony, unpublished data). In the socket cell Su(H) contributes to its own auto-activation, therefore a model utilizing direct binding by Su(H) for repression in the socket cell would be less favorable. While no putative Sox15 binding sites are conserved, there are instances in all 11 genomes so we chose assess the consequences of eliminating these sequences.

Mutation of the putative Sv binding results in strong loss of expression in both the shaft and sheath cells (Figure 4.2C). Moreover, the expression levels among all four post-mitotic cells remain fairly uniform, in sharp contrast to the expression of the svWT enhancer. The invariant expression is maintained well into 40 – 44 hrs APF, long

after both the svWT enhancer and Sv protein expression has been downregulated in the socket cell and neuron. Expression during the lineage prior to the differentiation of the post-mitotic cells is unaltered (data not shown), indicating the requirement for other sequences for this earlier activity of *sv* expression. These data suggest this single site is capable of providing both shaft/sheath activation as well as socket/neuron repression during the differentiation of these cell types.

Given that the putative Sv binding site has apparently no effect upon expression during the lineage, the strongly conserved proneural binding site seemed like a very attractive candidate. Mutation of this site resulted in no reduction in reporter expression either during the lineage or post-mitotically, inconsistent with such a hypothesis (Figure 4.2B). However, when we combine mutation of this sequence with the mutation of the putative Sv binding site we observe nearly complete loss of expression at all stages of mechanosensory organ development (Figure 4.2D), suggesting the lineage phase of *sv* expression is coordinated by a synergistic activity of proneural proteins and Sv itself.

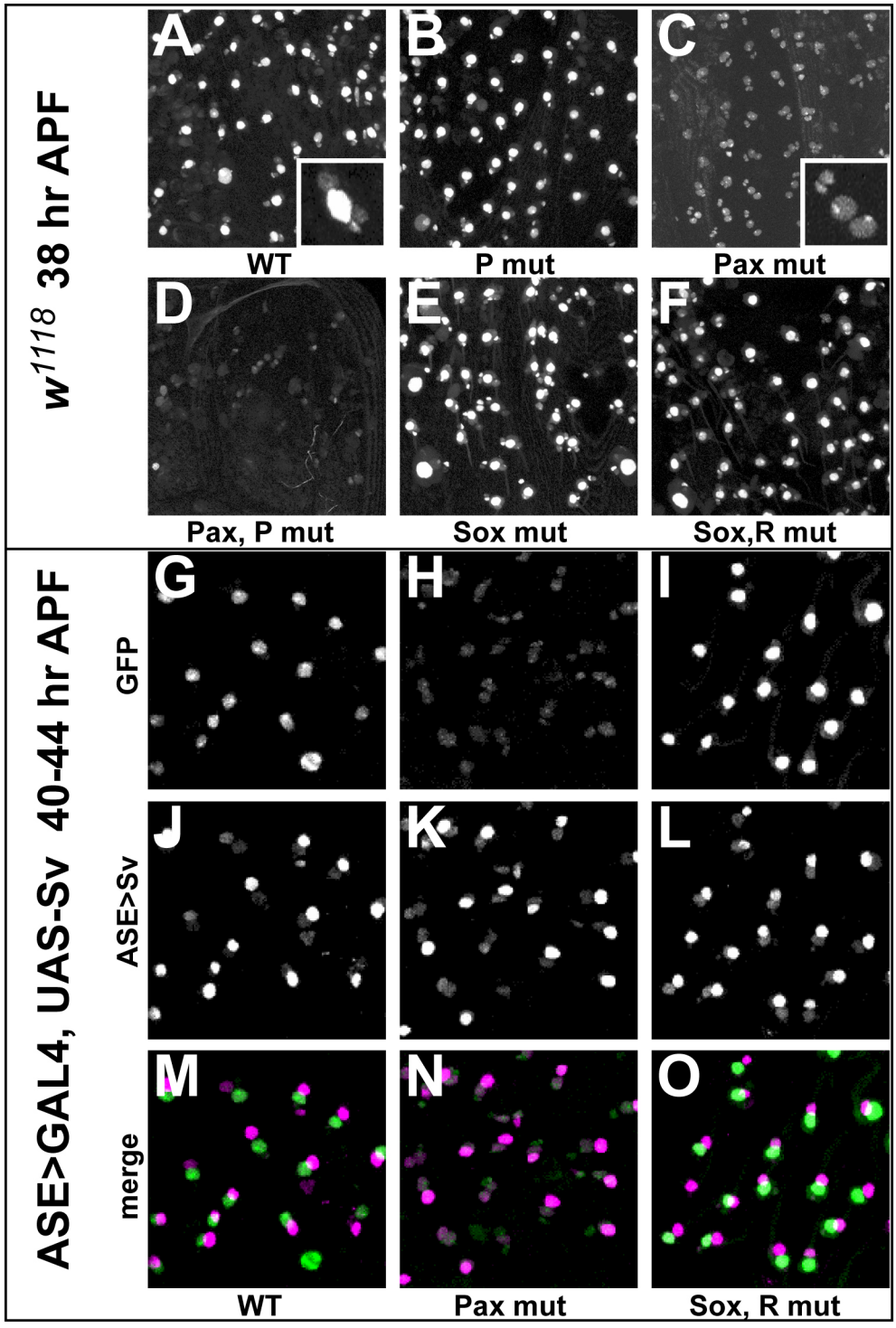
The other sequences that could be hypothesized to play a role in the repression of *sv* expression in the socket cell all appeared to

have no effect upon reporter expression. Loss of all the Sox15 sites matching the AACAATR consensus failed to result in either loss of expression or ectopic reporter activity (Figure 4.2E). Mutation of the bHLH-R sites actually resulted in complete loss of reporter expression (data not shown). In spite of all reporter lines with this single mutation giving the same result, this would appear to be an artifact since this same DNA fragment is perfectly capable of directing GAL4 expression (data not shown), and when placed in combination with the Sox15 site mutations gives the same phenotype as the Sox15 site mutant alone (Figure 4.2F). In addition to giving ectopic activity in the socket cell, repression via the bHLH-Rs could also be operating much earlier, in the non-SOP cells of the proneural cluster. However, examining expression of the svSoxRmut>GFP reporter at 12 hrs APF fails to show any ectopic activity in these cells as well (data not shown).

Ectopic socket expression of Sv is incapable of activating reporter expression

These data for Sox15 and bHLH-R binding sites are complicated by the potential for Sv auto-regulatory activity suggested

Figure 4.2: *sv* expression in the lineage is controlled by auto-regulatory activity.
(A – F) GFP expression upon mutation of svSOL at 38 hrs APF. (A) svWT>GFP, with inset showing a single microchaete position. (B) svPmut>GFP, carrying mutation of the proneural binding site, has no effect upon expression. (C) svPaxmut>GFP, mutating the putative Sv binding site, results in loss of high level expression in shaft and sheath, and failure to completely downregulate expression in socket and neuron. (D) svPaxPmut>GFP, mutation of both Sv and proneural binding sites leaves only weak residual reporter activity. (E) svSoxmut>GFP, mutating all the Sox15 binding sites (AACAATR sequences) does not affect reporter expression. (F) Similarly, in svSoxRmut>GFP, where both the bHLH-R and the Sox15 binding sites are mutated, no effect upon reporter expression is observed. (G – O) Analyzing the effect of reporter expression in an ASE>GAL4, UAS-Sv background, where high levels of Sv are ectopically expressed in the socket cell. (G, J, M) svWT>GFP. (H, K, N) svPaxmut>GFP. (I, L, O) svSoxRmut>GFP. (G – I) GFP expression. (J – L) Sv protein. (M – O) Merged images of Sv (magenta) and GFP (green) expression. Note the complete complementary pattern of the high level socket expression of Sv and GFP expression, which fails to become ectopically activated in the socket.


w1118 38 hr APF
A
B
C
WT
P mut
Pax mut
D
E
F
Pax, P mut
Sox mut
Sox,R mut
ASE>GAL4, UAS-Sv 40-44 hr APF
GFP
G
H
I
ASE>Sv
J
K
L
merge
M
N
O
WT
Pax mut
Sox, R mut

above; if the dominant activator in the post-mitotic cells is Sv itself, we may not be able to observe de-repression in cells which do not normally express Sv. To address this possibility, we examined the activity of our reporters in flies which are ectopically expressing high levels of Sv in the socket cell through the activity of ASE>GAL4. These conditions are sufficient to induce the socket cell to generate ectopic shaft structures and should thus activate in the socket cell components of the shaft differentiation program, including a *sv* reporter. However, under these conditions we fail to even observe ectopic expression of svWT>GFP by 40 – 44 hrs APF, in spite of the high levels of Sv protein (Figure 4.2G). Similar results were obtained for svSoxmut>GFP (data not shown) and svSoxRmut>GFP (Figure 4.2I). The Sv site mutant reporter still maintains its uniform expression through all four post-mitotic cells, and does not appear to increase GFP expression in the socket cell (Figure 4.2H). In total, our reporter data support a model for activation of *sv* expression in the lineage through combinatorial inputs of Sv auto-regulation and proneural proteins followed by Sv auto-regulation alone in shaft and sheath cells. Our results are also notably inconsistent with a model for repression of

sv expression in the socket cell mediated by Sox15, bHLH-Rs, or Su(H); instead repression in the socket and neuron may be also mediated by Sv auto-regulation, although in the form of repression (see Discussion).

Reciprocal negative regulation between Sv, Su(H), and Sox15

Our previous work has demonstrated a combinatorial activity of Su(H) and Sox15 for repressing *sv* expression in the socket cell (Miller et al., 2009). We also know that high-level Sv expression in the socket cell results in the ectopic production of shaft structures in the socket cell. However, this phenotype does not resemble a socket-to-shaft fate conversion, but rather the independent incorporation of two differentiation programs in the same cell. An ASE>GAL4, UAS-Sv socket cell still generates a cuticular structure that envelopes the shaft cell; the shaft-like projections coming from the socket cell emerge from this cuticular ring, on either side of the bristle. While this phenotype suggests a failure to interrupt the normal socket cell differentiation, we sought to confirm this by examining expression of both Su(H) and Sox15. Under these conditions, we find that expression for both of

these factors is severely compromised (Figure 4.3A-D), suggesting that ectopic expression of Sv and the subsequent shaft differentiation program induction does negatively affect normal socket cell development.

With the strong consequences of ectopic Sv expression in the socket cell, we asked if the reciprocal is true by inducing ectopic expression of socket-specific factors in the shaft cell using svWT>GAL4. With this driver, UAS-Su(H) results in a strong double-socket phenotype (Figure 4.3E,F). This phenotype is most consistent with ectopic Su(H) activity in the shaft cell overcoming Hairless-mediated repression, resulting in a conversion of this fate to a socket cell. A subset of the activity of Su(H) in the socket cell is likely mediated by bHLH-Rs, and even though we have yet to find evidence for direct regulation of Sv expression by bHLH-Rs, they may participate in the socket cell repression of shaft cell expressing genes. Consistent with this model, ASE:>GAL4 UAS-HLHm7::VP16 flies can exhibit a low level of ectopic shaft structures similar to UAS-Sv (S. Barolo and J. Posakony, unpublished data). Indeed, when we drive ectopic HLHm7 in the shaft cell with svWT>GAL4, we not only observe bristle positions

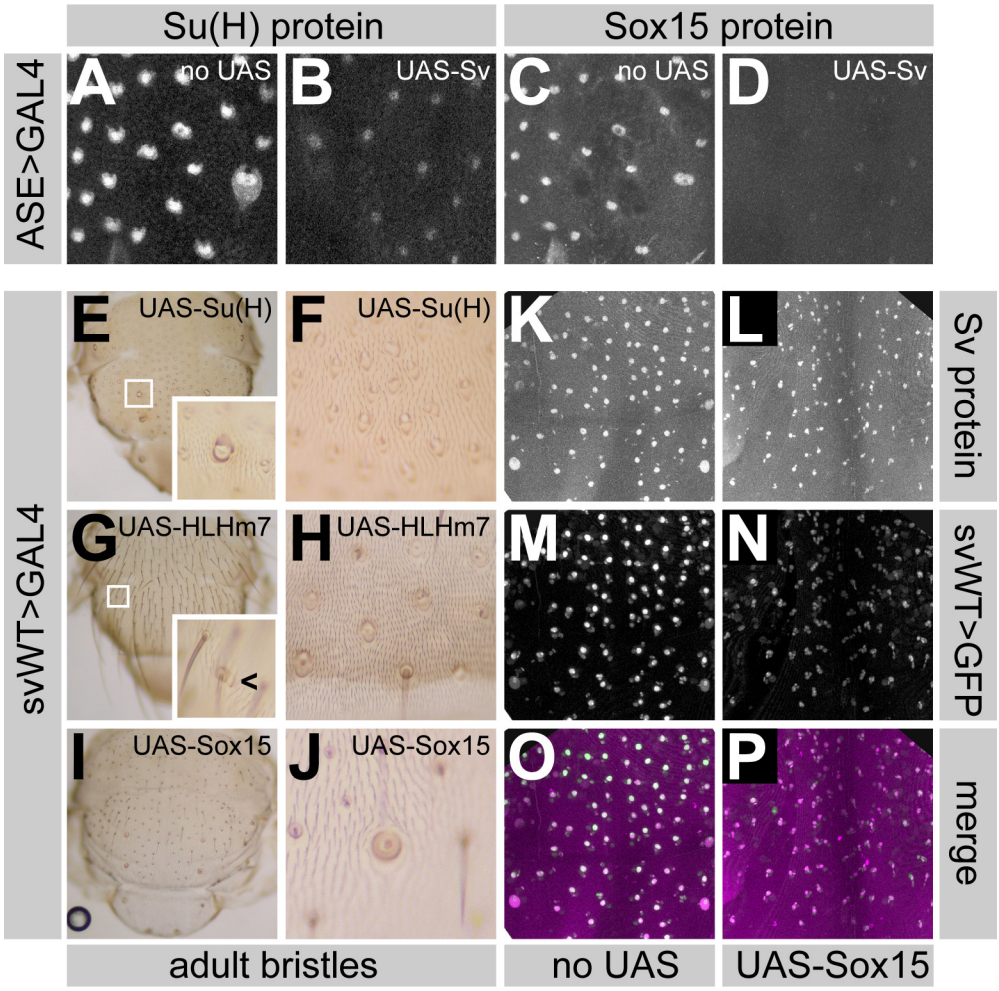


Figure 4.3: Mutual inhibition of the activities of Su(H), Sox15, and Sv. (A – D) Expression of either Su(H) (A, B) or Sox15 (C, D) protein in response to ectopic expression of Sv in the socket cell (B, D) using ASE>GAL4. (E – J) Cuticular phenotypes upon ectopic shaft expression using svWT>GAL4 of Su(H) (E, F); HLHm7 (G, H); or Sox15 (I, J) Insets in E and G are higher magnifications of the boxed regions in the same figure. Arrowhead in G points to ectopic socket cuticle. H shows abdominal bristle positions. F and J are higher magnifications of positions in E and G. (K – P) Effects upon Sv protein levels (K, L, magenta in O, P) or svWT>GFP expression (M, N, green in O, P) upon ectopic expression of Sox15 (L, N, P) using svWT>GAL4.

with double sockets, but also positions with reduced shafts and many positions unaffected (the strongest phenotype is observed on the abdominal bristles, rather than the notum; Figure 4.3G, H). Overall, this phenotype contrasts strongly with HLHm7 expression driven by neur>GAL4, which causes balding due to loss of SOPs. This difference is quite striking even though the enhancers regulating both drivers begin their activity in the SOP, and suggests that the primary activity of svWT>GAL4 acts either predominantly in the shaft cell or only very late in SOP development. Lastly, with ectopic Sox15 expression we detect a strong shaft-reduction phenotype, and fail to detect positions with double sockets (Figure 4.3I, J). This phenotype is consistent with a role for Sox15 primarily in the regulation of differentiation rather than fate. Moreover, despite the lack of evidence for Sox15 binding the *sv* tr/th enhancer, both Sv protein and svWT>GFP appear reduced when high levels of Sox15 are driven in the socket cell (Figure 4.3K-P).

Discussion

The quandary of the sv *expression pattern and N signaling*

Understanding the *sv* expression pattern has long been puzzling due to its failure to track with the pattern of N signaling in the lineage (Kavaler et al., 1999). Using a series of mutant reporters, we have uncovered a significant clue to the *sv* expression pattern: auto-regulatory activity. Sv binding first cooperates with proneural binding during the elaboration of the bristle lineage after SOP specification until the production of the post-mitotic cells, where Sv plays an essential role in activation in the shaft and sheath cells.

The lack of effect of mutating the proneural binding site alone puts *sv* in welcome company with a number of other SOP genes. These include *neur*, *sens*, and *phyl*, all genes which have both SOP expression and strongly conserved proneural binding sites yet these enhancers all appear to require the activities of additional factors to cooperate with the proneurals to give their SOP and lineage expression. Sv protein appears to accumulate in late SOPs, unlike the other genes (Kavaler et al., 1999). It is likely, then, that *sv* expression in the SOP is initiated well downstream of the activation of the SOP

program and could therefore be imagined to become activated in a manner independently of the proneurals themselves. Nevertheless, it is indeed satisfying to observe that the two major cognate conserved binding sites provide the vast majority of enhancer activity.

Man cannot live by transcriptional regulation alone

In examining these conserved sequences, we have also presented preliminary evidence that Sv might also be contributing to its own repression in the socket cell and neuron. We observe that mutation of this sequence not only causes the loss of high-level expression in the shaft and sheath, but also elevates the levels in the other two post-mitotic cells to create uniform levels among all four cells which persist longer than the levels in the socket cell and neuron driven by svWT>GFP. With this result we can no longer sustain the simple hypothesis of Sv auto-activation, but rather auto-activation/repression activity regulated by a likely post-translational mechanism. Such a hypothesis is not new; Sv contains an octomer amino acid sequence that has been shown to bind a member of the

Groucho family of transcriptional co-repressors that is able to abrogate transactivation activity (Cai et al., 2003).

If Groucho is binding to and converting Sv protein into a transcriptional repressor in the socket and neuron, it may perhaps help to explain the failure to activate reporter activity in the presence of high levels of Sv protein in the socket cell. Even though such levels exist and can readily be observed by antibody stain, if conditions in the socket cell are more amenable to the repressive mode of Sv we may not be able to observe the activation of svWT>GFP or perhaps even other shaft targets in the socket cell. A mutational analysis of Sv coding sequence in the context of this assay may both illuminate the repressive sequence in Sv as well as confirm auto-activation of the svWT reporter.

The dueling swords of differentiation

Over the course of study of the regulation of both *Sox15* and *sv* expression in their respective cell types, we have apparently uncovered an important theme in the differentiation of post-mitotic cells: to incorporate not only an array of target genes to promote the

correct differentiation, but also a gene battery preventing incorrect differentiation. These two sets may actually be one in the same, but such a theme could have been assumed to be unnecessary. If the cell fate specification machinery is very robust, differentiation could be completely reliant upon the outcome of the specification event; in this scenario, differentiation could be altered only if the specification signal is also affected. The phenotypes of loss of *Su(H)* ASE and *Sox15* in the socket cell and *sv* in the shaft cell clearly elicit differentiation defects in spite of a correct specification event (Kavaler et al., 1999; Barolo et al., 2000b; Miller et al., 2009). Furthermore, loss of Su(H) and Sox15 in the socket cell results in ectopic *sv* expression, which from this study we find is a danger to the expression of both of these proteins. We can suggest that Sv expression would be particularly dangerous given its auto-regulatory activity. Also, we find that ectopic Sox15 is able to affect the shaft differentiation program without affecting the cell fate, likely in part by downregulating *sv* expression through an undetermined mechanism. It would appear that key components of both differentiation pathways, therefore, operate not

only to positively regulate differentiation of their respective cell types, but also prevent ectopic differentiation of the sister cell program.

References

Barolo, S. *et al.* (2000a). GFP and beta-galactosidase transformation vectors for promoter/enhancer analysis in Drosophila. Biotechniques *29*, 726, 728, 730, 732.

Barolo, S. *et al.* (2000b). A notch-independent activity of suppressor of hairless is required for normal mechanoreceptor physiology. Cell *103*, 957-969.

Cabrera, C. V. (1990). Lateral inhibition and cell fate during neurogenesis in Drosophila: the interactions between scute, Notch and Delta. Development *110*, 733-742.

Cai, Y. *et al.* (2003). Groucho suppresses Pax2 transactivation by inhibition of JNK-mediated phosphorylation. EMBO J *22*, 5522-5529.

Castro, B. *et al.* (2005). Lateral inhibition in proneural clusters: cis-regulatory logic and default repression by Suppressor of Hairless. Development *132*, 3333-3344.

Flores, G. V. *et al.* (2000). Combinatorial signaling in the specification of unique cell fates. Cell *103*, 75-85.

Frise, E. *et al.* (1996). The Drosophila Numb protein inhibits signaling of the Notch receptor during cell-cell interaction in sensory organ lineage. Proc Natl Acad Sci U S A *93*, 11925-11932.

Fu, W. *et al.* (1998). shaven and sparkling are mutations in separate enhancers of the Drosophila Pax2 homolog. Development *125*, 2943-2950.

Gho, M. *et al.* (1996). Subcellular localization of Suppressor of Hairless in Drosophila sense organ cells during Notch signalling. Development *122*, 1673-1682.

Guo, M. *et al.* (1996). Control of daughter cell fates during asymmetric division: interaction of Numb and Notch. Neuron *17*, 27-41.

Ho, S. N. *et al.* (1989). Site-directed mutagenesis by overlap extension using the polymerase chain reaction. Gene *77*, 51-59.

Inbal, A. *et al.* (2003). Multiple roles for u-turn/ventral veinless in the development of Drosophila PNS. Development *130*, 2467-2478.

Jafar-Nejad, H. *et al.* (2006). Senseless and Daughterless confer neuronal identity to epithelial cells in the Drosophila wing margin. Development *133*, 1683-1692.

Jimenez, F., and Campos-Ortega, J. A. (1990). Defective neuroblast commitment in mutants of the achaete-scute complex and adjacent genes of D. melanogaster. Neuron *5*, 81-89.

Jimenez, G., and Ish-Horowicz, D. (1997). A chimeric enhancer-of-split transcriptional activator drives neural development and achaete-scute expression. Mol Cell Biol *17*, 4355-4362.

Kavaler, J. *et al.* (1999). An essential role for the Drosophila Pax2 homolog in the differentiation of adult sensory organs. Development *126*, 2261-2272.

Knoblich, J. A. *et al.* (1995). Asymmetric segregation of Numb and Prospero during cell division. Nature *377*, 624-627.

Ligoxygakis, P. *et al.* (1999). Ectopic expression of individual E(spl) genes has differential effects on different cell fate decisions and underscores the biphasic requirement for notch activity in wing margin establishment in Drosophila. Development *126*, 2205-2214.

Miller, S. W. *et al.* (2009). Complex interplay of three transcription factors in controlling the tormogen differentiation program of Drosophila mechanoreceptors. Dev Biol *329*, 386-399.

Rebeiz, M., and Posakony, J. W. (2004). GenePalette: a universal software tool for genome sequence visualization and analysis. Dev Biol *271*, 431-438.

Reddy, G. V., and Rodrigues, V. (1999). Sibling cell fate in the Drosophila adult external sense organ lineage is specified by prospero function, which is regulated by Numb and Notch. Development *126*, 2083-2092.

Romani, S. *et al.* (1989). Expression of achaete and scute genes in Drosophila imaginal discs and their function in sensory organ development. Genes Dev *3*, 997-1007.

Ruiz-Gomez, M., and Modolell, J. (1987). Deletion analysis of the achaete-scute locus of Drosophila melanogaster. Genes Dev *1*, 1238-1246.

Simpson, P. (1990). Lateral inhibition and the development of the sensory bristles of the adult peripheral nervous system of Drosophila. Development *109*, 509-519.

Chapter Four, in full, is a manuscript in preparation for publication: Auto-regulation, proneural synergy, and mutual inhibitory interactions shape expression of *shaven* and influence the tormogen and tricogen differentiation programs. Miller, S. W. and Posakony, J. W. The dissertation author was the primary researcher and author. James Posakony provided critical comments and oversight through the course of this investigation. Scott Barolo originally located the enhancer, Jamy Peng generated the UAS-Sox15 construct, and Tammie Stone performed the random binding site selection that was used to deduce a Sv binding consensus.

CHAPTER FIVE:

pMemor: A highly versatile system for rapid enhancer transgenesis and beyond

Abstract

With the rapid production of voluminous genomic information, there is increasing demand for high-throughput validation methods. To this end we report the development and characterization of the Memor system for rapid enhancer transgenesis with theoretical applicability from *Drosophila* to vertebrates. Memor is a Minos transposon-based docking site containing a reporter and promoter present at the site, allowing the efficient integration of an enhancer in either orientation through ΦC31 integrase recombination mediated cassette exchange (RMCE). This site also contains the capacity to be completely replaced by Cre-mediated cassette, maximizing the utility of a single genomic location. In addition to demonstrating functionality and applicability of this system we report on previously unreported anomalies associated with RMCE.

Introduction

Advances in transgenic techniques have paved the way for applications a various as fine-scale gene regulatory analysis to gene therapy. *Drosophila melanogaster*, particularly through the use of transposable elements, has proven to be a valuable platform for technological development with applications for a variety of organisms. Indeed, while the simple P transposon continues to maintain its utility for most purposes, the last decade has witnessed a number of significant variations in terms of transposon type, structure, and genomic integration.

The simplest form of transposon, the empty P element, proved highly useful for integration of genomic DNA for mutant rescue (Rubin and Spradling, 1983; Hazelrigg et al., 1984). Incorporating a strong promoter and reporter gene (such as LacZ) enabled the discovery of a number of enhancer specificities by hijacking their function (Bellen et al., 1989). This basic reporter transposon has facilitated fine-scale analyses of regulation of specific enhancer elements, illuminating a number of developmental regulatory networks (Basler et al., 1989; Bowtell et al., 1989; Buttgereit et al., 1991; Schultz et al., 1991;

Wagner-Bernholz et al., 1991). These analyses have also revealed features of P element insertions: including insertion position effects, 5' UTR insertion bias, and a narrow range of species compatibility, all of which encouraged the implementation of improvements (reviewed in Ryder and Russell, 2003; Venken and Bellen, 2007).

One of the first improvements upon the standard P element was the addition of binding sites for the Su(Hw) insulator factor (gypsy elements). These elements were introduced in the Pelican/Stinger vectors, which also included the addition of fluorescent reporter genes as an alternative to LacZ (Barolo et al., 2000). Adding insulator elements cut down the influence of genomic location upon the reporter construct, perhaps due to an overall boost in the level of activation of the reporter construct (Markstein et al., 2008). While these constructs reduced the effect of position, they still required the injection of an maintenance of multiple independent insertions for due scientific rigor. To address this issue, Oberstein *et al* (Oberstein et al., 2005), first introduced the use of a genomic "docking site," a location in the genome harboring a pair of incompatible loxP sequences that can promote recombination with matching loxP sites on an exogenous

plasmid. This allowed mutant and wild-type constructs to be injected into the same genomic location, eliminating the influence of position effect differences and reducing the necessity of maintaining multiple stocks. However, Cre recombinase is able to interact with other cryptic loxP sequences, contributing to toxic effects and unintended recombination events (Siegal and Hartl, 1996). Concerns about these negative effects of Cre likely has contributed to the lack of rapid adoption of the Cre/lox docking site.

More recently, several groups have begun to report constructs using the ΦC31 integrase for *Drosophila* transgenesis (Groth et al., 2004; Bateman et al., 2006; Venken et al., 2006; Bischof et al., 2007; Venken et al., 2008; Venken et al., 2009). This bacteriophage integrase promotes the irreversible recombination between two sequences, commonly known as attB and attP, with high efficiency (Thorpe and Smith, 1998). Unlike, Cre, ΦC31 integrase also appears to interact with a minimal number of endogenous genomic sequences (Groth et al., 2004). As a further tool, Bischof has created several genomic insertions of the integrase gene, complete with regulatory elements from the *vasa* or *nanos* genes, resulting in germline-specific

expression (Bischof et al., 2007). These improvements have greatly improved the efficiency, as well as utility of the docking site technology and have indeed been rapidly adopted.

Despite these improvements for *Drosophila* transgenesis, P elements remain largely phylogenetically restricted (Rio et al., 1988), and for many systems cannot be utilized, even to introduce a docking site. Therefore, several groups have worked to characterize other transposable elements both from within *Drosophila* as well as from other insects to circumvent this difficulty. Some of these additional elements, including piggyBac, mariner, and Minos, display surprising phylogenetic promiscuity with their insertion specificity (Loukeris et al., 1995b; Lobo et al., 1999; Plasterk et al., 1999). Minos elements, in particular, have even been demonstrated to readily insert into vertebrate genomes (Klinakis et al., 2000). In addition, these transposons also lack the 5' UTR bias of the P element, relieving worry about unintended gene disruption upon insertion (Loukeris et al., 1995a).

While these developments in both transposon utility and genomic integration technology hold great promise for myriad

applications for the next decade and beyond. In the near future, however, the rapid production of genomic information from projects like ENCODE and modENCODE (2004; Celniker et al., 2009), demand high-throughput strategies for both integration as well as plasmid construction. There have certainly been improvements in cloning technology in the last few years, notably Gateway™ technology from Invitrogen and In-Fusion™ from Clontech, these commercial products come at a price that may dissuade labs with less diverse funding sources to tackle the hypotheses generated from this new-found wealth of genomic data.

With all of these developments in mind, here we report the design and implementation of the Memor docking site, a genomic location for the rapid testing of enhancer predictions. This docking site, contains both a GFP reporter and Hsp70 promoter already in the site, and enables the rapid integration of enhancer fragments via a ΦC31 integrase-mediated cassette exchange. These enhancer fragments can be cloned using T-A cloning, not requiring additional reagents beyond standard laboratory fare. Included in the design are pseudo-loxP sequence that enable the complete replacement of all elements

within the docking site, expanding its use beyond enhancer studies. Lastly, we report on the use of hairpin constructs as markers for transgenesis, which both reduce the size of the docking site and enable its use in organisms that lack mutant strains and therefore are incompatible with systems such as mini-w+ rescue, as has been used in *Drosophila melanogaster*.

Materials and Methods

Fly stocks

vas-ΦC31int flies were generously donated by Konrad Basler (Bischof et al., 2007). dpp>GAL4, GMR>GAL4, and act5C>GAL4 flies were obtained from Bloomington. ASE>GAL4 was generated by Scott Barolo and has been reported previously (Miller et al., 2009). *w; noc^{Sco}/Cyo, MiT w+* flies have been reported previously (Metaxakis et al., 2005), and were generously donated by Hugo Bellen.

Construction of pSWU variants

pSWU: A series of 60 bp overlapping oligonucleotides was used to synthetically construct a PCR product containing SCP, *white* intron, UTR, and surrounding restriction enzyme sequences in a standard PCR reaction. This product was T-A clone into pGEM-T (Promega). The insert was then cut out with a Not I + Mlu I double digest and ligated back into pGEM-T that had been cut with PspOM I + Mlu I, destroying the PspOM I sequence.

pSWUwhite and pSWUgarnet: Primers designed to amplify ~500 bp fragments from *white* and *garnet* with 5' compatible restriction

sites were used to generate products to clone in inverted orientation on either side of the *white* intron, with the antisense orientation immediately downstream of the SCP. The GMR enhancer was amplified from genomic DNA collected from GMR>GAL4 flies using primers with 5' compatible restriction site and cloned upstream of the SCP.

pSWUyellow: Left and right hairpin arms for *yellow* were generated and cloned as above. A fragment containing the *yellow* body enhancer (Geyer and Corces, 1987) and the endogenous promoter was amplified and cloned into pSWU in a manner that removed the SCP.

pHWU and pHWUwhite: Primers against Hsp70 were designed with restriction sequences compatible with pSWU, and a product was amplified from an Hsp70 pGEM-T clone (below) and cloned into pSWU with the SCP excised to create pHWU. The complete hairpin construct was cut from pSWUwhite and cloned into pHWU to create pHWUwhite.

Construction of pBASE starter vectors and cloning of pMemor variations

pBASEa and pBASEb: Three fragments, DNA1a, DNA1b, and DNA2, were synthetically generated using a series of 60 bp overlapping oligonucleotides in a standard PCR reaction. The only difference between DNA1a and DNA1b is the loxP variant, either lox66 in DNA1a or lox72 in DNA1b. DNA1 variants were T-A cloned into pGEM-T. Positive clones were excised with a Not I + Mlu I digest and ligated back into pGEM-T cut with PspOM I + Mlu I to create pMLPa and pMLPb (MLP = Minos LTR, lox, attP). 5X UAS sequence was amplified from pUAST and combined with DNA2 using overlap extension (Ho et al., 1989). This fragment was digested with Bss HII, and ligated into pMLP variants that had been digested with Mlu I to create pBASEa and pBASEb.

pAw+HG, pAWHR, pAYHG, and pAGSG: SCP was synthetically created using overlapping oligonucleotides in a standard PCR reaction which was then T-A cloned into pGEM-T. Hsp70, *white*, and eGFP were amplified from pH-Stinger; DsRed was amplified from pH-RedStinger; these products were then T-A cloned into pGEM-T.

Promoters were cloned into the pBASEa with a PspOM I + Avr II digest, and reporters with a Avr II + Spe I digest. Hairpin constructs from pSWU and the *white* pGEM clone were excised and cloned into pBASEa with a Bgl II + Mlu I digest.

pBWHR and pBWSG: A similar strategy for cloning as for pAWHR and pAGSG as described above was used for these constructs.

pBhWHG: The Hsp70-*white* hairpin from pHWUwhite was cloned into pBASEbHG as described above.

pBLw+HG: Primers were designed with incorporated new loxP sequences to amplify the 5X UAS 3' attP region in pBASE. This fragment was T-A cloned into pGEM-T. This clone was digested with BamH I + PspOM I (creating two small fragments) and ligated via triple ligation into pBw+HG that had been digested with Bgl II + PspOM I to create pLBASEbHG. *white* was then cloned into this vector as described above.

Cloning Memor-compatible enhancer fragments

All enhancer fragments were generated by adding the sequence 5'-CGGGTGCCAGGGCGTGCCCTTGGGCTCCCCGGGCGCGTACG-3' to the 5' ends of both primers, and amplifying fragments using their respective pH-Stinger clones as template. PCR products were T-A cloned into pGEM-T and sequence verified before prepping with the GenElute Maxi Kit (Sigma) and injecting.

Injection of pMemor docking site starter lines using Minos transposase mRNA

Minos transposase mRNA was generated using the mMessage mMachine T7 kit (Ambion) from the pBlue(SK)MimRNA plasmid as described (Pavlopoulos et al., 2004). Plasmid was generously donated by Charalambos Savakis. Injection cocktail contained 500 ng/μL Memor plasmid and 100 ng/μL transposase mRNA as suggested (Metaxakis et al., 2005).

Mapping Memor genomic insertion locations by inverse PCR

Single fly preps and iPCR were performed using the protocol by Gloor and Engels found online

(http://engels.genetics.wisc.edu/flyDNA.html) and previously described (Gloor et al., 1993). With much trial and error, we found best results were obtained using the Mbo I enzyme. Primer sequences are available upon request.

Minos transposition using endogenous transposase source

Transposition was performed as previously described (Metaxakis et al., 2005). AWHR lines were crossed to *w*$^{+}$*; noc*^{Sco}*/Cyo MiT* virgins. After two days of laying, the adults were transferred to new vials. Upon removal of adults, larvae were heat-shocked for 1 hour at 37 °C everyday until pharate adults were visible in the vial. 70 AWHR/Cyo MiT males were crossed in pairs to Canton S virgins. Straight-winged flies were scored for an eye color different from the starter line and new eye colors were collected, balanced, and mapped by iPCR as described above.

ΦC31 integrase-mediated enhancer integration

Different Memor starter lines were crossed into the X-linked vas-ΦC31 integrase line on 2A and maintained as stable stocks. Embryos

from these lines were injected with enhancer-attB pGEM-T clones at a concentration of 500 ng/μL using standard injection procedures. Surviving injectees were crossed to balancer lines and screened for loss of transgenesis marker (See Supplemental Figure 5.4).

Cre/lox docking site exchange

Embryos from starting Memor docking site flies were injected with 500 ng/μL replacement construct and 300 ng/μL pMLS104 helper plasmid as described (Oberstein et al., 2005). Surviving injectees were crossed to appropriate balancer lines and screened for white eyes.

Results and Discussion

Design and strategy of the Memor system

The theme we employed in the development of a new enhancer transgenesis system was the use of different "cassettes" to gather information—or spy—on the genome. For this reason, we named our system "Memor," after the Italian name for the character Soundwave of the Transformers™ cartoon franchise, who spies on his enemies with the aid of small robot workers who dock into his body in the form of stereo cassette tapes.

Our goal was to design a system from scratch which contained a number of key features: 1) a docking site within a transposon with more random genomic insertion than P elements that could insert into a variety of genomes; 2) the presence of both the reporter and promoter already in the docking site to eliminate the need to clone an enhancer into a reporter vector; 3) the ability of each integration location of the docking site to be tested for its permissibility of reporter expression; 4) the integration of the enhancer via a ΦC31 integrase-mediated cassette exchange, allowing both orientations of enhancer to be tested; 5) the use of hairpins as a marker for docking site integration

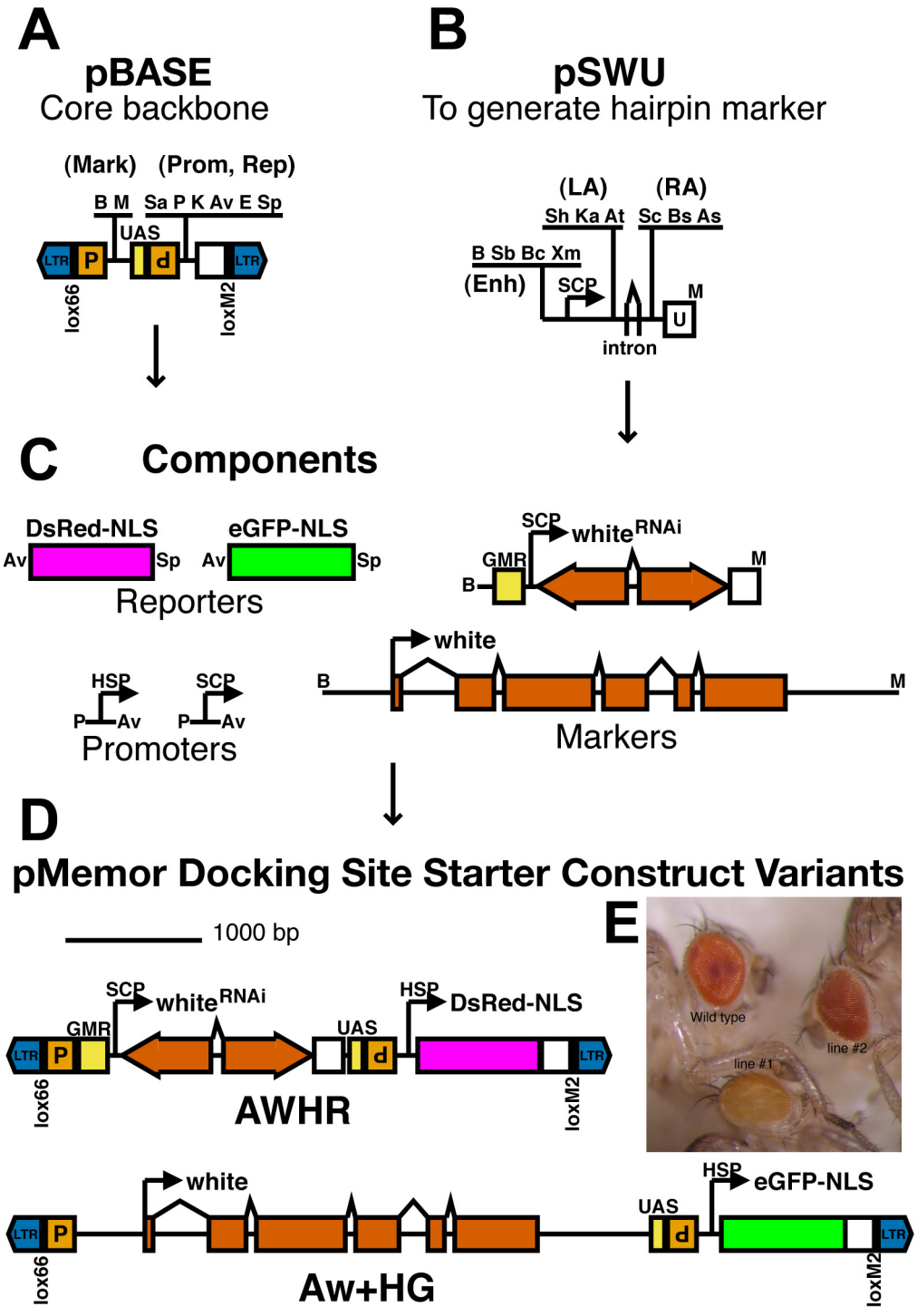
instead of mini-w+, to decrease both the size of the initial plasmid as well as the size of the cassette exchanged during enhancer integration; 6) the presence of pseudo-loxP sequences to allow the entire docking site to be changed at will to incorporate different reporter genes, promoters, transgenesis markers, or docking sites of completely different design to allow standardization of a minimal number of docking sites for nearly every transgenic need; and 7) the use of common and compatible restriction enzyme sites to allow changes to the original docking site to be performed rapidly in vitro. To fulfill these initial aims, we created three starting vectors: pBASEa, pBASEb, and pSWU (Figure 5.1A,B).

Both pBASE variants contain the same architecture: a Minos transposon with two non-intercompatible pseudo-loxP sequences, containing inverted 200bp ΦC31 integrase attP sequences flanking 5X tandem UAS sequences, the 3' UTR sequence downstream of GFP from pH-Stinger (Barolo et al., 2000), and a number of common, compatible restriction enzyme sites for cloning in the remaining elements for the desired docking site. The only difference between pBASEa and pBASEb is in the identity of the 5' loxP site: in pBASEa it

is lox66, while in pBASEb it is lox72; the 3' loxP site in both is loxM2. lox66 and lox72 both contain mutations in the inverted repeat arms of the loxP sequence, while loxM2 contains mutation of the central core. The nature of the recombination of lox66 and lox72 (they cannot mediate recombination of homotypic, only heterotypic sites) results in end products that contain either a wild-type loxP site or a site with mutations in both arms severely limiting any further recombination with this site (Oberstein et al., 2005). We chose to have lox66 in the starting docking site construct so that upon Cre-mediated docking site replacement only a single loxP sequence would remain the genome, allowing additional loxP sequences to be utilized elsewhere for some of the current technologies that also use these sequences.

Our original design was to make use of short hairpin RNA constructs as markers for transgenesis, for reasons described above. With versatility in mind, we generated the pSWU vector for easy generation of hairpin constructs to knock-down transcripts for a wide number of genes. pSWU contains the "Super Core Promoter" (SCP) developed by Jim Kadonaga's group (Juven-Gershon et al., 2006), the small second intron from the *white* gene, and a 3' UTR sequence

Figure 5.1: The Memor system for enhancer transgenesis.
(A – D) DNA constructs to generate pMemor docking site starter constructs. (A) The backbone pBASE is a Minos transposon with numerous restriction enzyme sequences, a 3'UTR sequence for reporters, inverted ΦC31 integrase attP sequences, and 5X UAS sequences, with flanking pseudo-loxP sites. The marker for transgenesis is cloned upstream of the 5X UAS, and the promoter and reporter are cloned between the attP cassette and the 3'UTR. (B) The vector pSWU is used to construct any desired hairpin construct for use as a marker in the Memor system. The core contains the Super Core Promoter (SCP), the short second intron from the *white* gene, and a 3' UTR sequence. Hairpin arms are cloned into inverted flanking positions around the intron and any enhancer can be cloned upstream of the SCP to tightly control hairpin specificity. (C) Different reporters, promoters, and transgenesis markers used to generate the main constructs used in this study. (D) Schematics of the two initial starting docking sites created, AWHR and Aw+HG. (E) Eye phenotypes resulting from transformation of the AWHR docking site (line 2, right, and line 1, bottom) into Canton S flies (top left). Restriction site abbreviations throughout the figure: B, Bgl II; M, Mlu I; Sa, Sac I; P, PspOM I; K, Kpn I; Av, Avr II; E, EcoR I; Sp, Spe I; Sb, Sbf I; Bc, Bcl I; Xm, Xma I; Sh, Sph I; Ka, Kas I; At, Aat II; Sc, Sac II; Bs, BsiW I; As, Asc I.


A
pBASE
Core backbone
(Mark) (Prom, Rep)
B M Sa P K Av E Sp
UAS
LTR P d LTR
lox66
loxM2
B
pSWU
To generate hairpin marker
(LA) (RA)
Sh Ka At Sc Bs As
B Sb Bc Xm
(Enh) SCP
M
U
intron
C Components
DsRed-NLS
Av Sp
eGFP-NLS
Av Sp
Reporters
SCP white RNAi
GMR
B
M
HSP
P Av
SCP
P Av
Promoters
white
B
M
Markers
D
pMemor Docking Site Starter Construct Variants
1000 bp
SCP white RNAi
GMR
HSP DsRed-NLS
UAS
LTR P d LTR
lox66
loxM2
AWHR
white
HSP eGFP-NLS
UAS
LTR P d LTR
lox66
loxM2
Aw+HG
E
Wild type
line #1
line #2

downstream of DsRed from the pH-RedStinger reporter vector. With this basic vector, identical sequences are cloned in an inverted manner on either side of the *w* intron, promoting splicing of the hairpin and avoiding toxic effects of having a directly repeated sequences as has been reporter and as we have experienced (data not shown). The hairpin construct can have specificity for any desired cell type by cloning any cognate enhancer upstream of the SCP. The entire hairpin construct is excised from pSWU as a Bgl II/Mlu I restriction fragment that is cloned into either pBASE variant and is inserted between the inverted attP sequences, upstream of the 5X UAS sites. The first hairpins we constructed included two constructs using the GMR enhancer that drove hairpins for the *white* and *garnet* genes, and a *yellow* hairpin under control of its own promoter and body enhancer. These hairpins were cloned into the pBASEa vector that also contained either DsRed.T4 or eGFP from pH-RedStinger or pH-Stinger, respectively, and either SCP or the Hsp70 promoter from pH-Stinger to create the first pMemor clones: pAWHR (pBASEa with the *white* hairpin and Hsp70 promoter upstream of DsRed.T4), pAYHG (*yellow*

hairpin and Hsp70 promoter upstream of eGFP), and pAGSG (*garnet* hairpin and SCP upstream of eGFP).

Docking site transformation

To transform these new docking sites into flies, we coinjected with Minos transposase mRNA into Canton S flies and screened for the phenotypes predicted for each of the respective hairpin constructs. While we were unable to detect any phenotypes from either the pAYHG or pAGSG constructs (and therefore unable to identify any insertion lines), we succeeded in finding two transformant lines for pAWHR (Figure 5.1D, E). Line #1 is a lethal insertion that mapped by inverse PCR to the coding sequence of *Gmd*, as verified by sequencing from both ends of the element; from line #2 we were only able to collect sequence from the 3' end, which indicated insertion 3.2 kb downstream of the coding sequence of the gene *CG31668*. Subsequent attempts to verify the 5' end have failed and we reason integration may have compromised the Minos LTR on this side of the element (see below).

Table 5.1: Memor Docking Site Starter Line Mapping Information.

Type	Line #	Location	Description	Background Activity	GAL4 Response	Enhancers Tested	Cre/lox Replacement?
Aw+HG	62	2L:10763100	in intron of CG6495				
Aw+HG	62	2L:10763100	in intron of CG6495				
Aw+HG	90	2L:11525751	in intron of CG31705				
Aw+HG	58	2L:13341666	in between CG16970 and CG16826				
Aw+HG	41	2L:16761724	in between CG13278 and Mhc	small # CNS	ActGAL4	neur4DWT, Sox1.3WT, svWT	
Aw+HG	96	2L:18738860	in between CG15167 and CG10348				
Aw+HG	14	2L:18981327	in intron of CG10641				
Aw+HG	55	2L:2001344	in intron of CG33543				
Aw+HG	84	2L:20019887	in between CG13966 and CG10659 (downstream of CG10659)				
Aw+HG	45	2L:20722164	in between CG1864 (Hr38) and CG9316				
Aw+HG	1	2L:22128523	in coding of Cyclin K				
Aw+HG	48	2L:22315487	in intron of CG17018				
Aw+HG	64	2L:3029889	in coding of CG17219				
Aw+HG	63	2L:3029890	in coding of CG17219				
Aw+HG	63	2L:3029890	in coding of CG17219				
Aw+HG	4	2L:677739	in intron of ds				
Aw+HG	70	2R:15210238	in intron of Rgk1				
Aw+HG	70	2R:15210238	in intron of Rgk1				
Aw+HG	15	2R:16656811	in between CG34202 and CG13443				
Aw+HG	59	2R:18635400	in intron of CG3536				
Aw+HG	32	2R:1938651	in 5' UTR of CG14593				

Table 5.1 Continued:

Type	Line #	Location	Description	Background Activity	GAL4 Response	Enhancers Tested	Cre/lox Replacement?
Aw+HG	33	2R:1938651	in 5' UTR of CG14593				
Aw+HG	39	2R:19503472	5' of CG4019	small # CNS	ActGAL4	neur4DWT, Sox1.3WT, svWT	BhWHG, BLw+HG
Aw+HG	19	2R:21036543	in between Tkr and CG34038				
Aw+HG	52	2R:7544767	in intron of Roc2				
Aw+HG	52	2R:7544767	in intron of Roc2				
Aw+HG	6c	3L:11269371	in between CG6175 and scyl (downstream of both) (same as 7a2, 7a1)				
Aw+HG	7a1	3L:11269371	in between CG6175 and scyl (downstream of both) (same as 7a2, 6c)				
Aw+HG	7a2	3L:11269372	in between CG6175 and scyl (downstream of both)				
Aw+HG	23	3L:13469008	in between CG10089 and stv	small # CNS	ActGAL4		
Aw+HG	88	3L:16906867	in intron of Lmpt				
Aw+HG	92D	3L:17495687	in intron of Oatp74D				
Aw+HG	83	3L:1823903	in between GV1 and Cpr62Ba				
Aw+HG	49	3L:18836515	3' of CG32027, in intron of Indy	many CNS	ActGAL4		
Aw+HG	47	3L:18836528	3' of CG32027- in intron of Indy (same as 49)				
Aw+HG	27	3L:19592915	in coding of Taf6				
Aw+HG	51	3L:20143679	in 3' UTR of CG7323				
Aw+HG	16	3L:20150533	in intron of CG7323				
Aw+HG	77b	3L:2669380	in between CG16976 and CG12187				
Aw+HG	31	3L:3327149	in intron of CG11526				
Aw+HG	37d	3L:3327149	in intron of CG11526				
Aw+HG	85	3L:3437143	in intron of eIF5B				
Aw+HG	53	3L:5603187	3' of mthl2				

Table 5.1 Continued:

Type	Line #	Location	Description	Background Activity	GAL4 Response	Enhancers Tested	Cre/lox Replacement?
Aw+HG	53	3L:5603187	3' of mthl2				
Aw+HG	12	3L:6353840	3' of CG32398	small # CNS	ActGAL4	neur4DWT, Sox1.3WT, svWT	BWHR, BWSG
Aw+HG	86	3L:673768	in between Rev1 and CG17129				
Aw+HG	77a	3L:6941950	in between CG14821 and CG10077				
Aw+HG	60	3L:9536718	in coding of CG12525				
Aw+HG	9	3R:11141505	in between CG14866 and Tm2				
Aw+HG	79	3R:13171950	in intron of beat-Iia				
Aw+HG	78	3R:1333447	in between CG31547 and Nmdar1				
Aw+HG	36	3R:1334679	in 3' UTR of Nmdar1	subset CNS	ActGAL4		
Aw+HG	38	3R:1334680	in 3' UTR of Nmdar1 (same as 36)				
Aw+HG	61	3R:14240483	upstream of CG31122				
Aw+HG	26	3R:16382877	in intron of att-ORFA				
Aw+HG	35	3R:17023688	in intron of SNF4Agamma	not detected	ActGAL4		
Aw+HG	11	3R:20183759	3' of CG18528	small # CNS	ActGAL4	neur4DWT, Sox1.3WT, svWT	BWHR, BWSG
Aw+HG	43	3R:21685486	in between CG5886 and CG5890				
Aw+HG	30	3R:25724207	in intron of CG31038	not detected	ActGAL4		
Aw+HG	81	3R:25906700	in intron of sima				
Aw+HG	6b	3R:27670321	in 3' UTR of RhoGAP100F				
Aw+HG	7c	3R:27670324	in 3' UTR of RhoGAP100F				
Aw+HG	71	3R:293553	in intron of CG31523				
Aw+HG	20	3R:3033206	in the 3' UTR of CG1288	not detected	ActGAL4		
Aw+HG	94	3R:4662054	in intron of pyd				

Table 5.1 Continued:

Type	Line #	Location	Description	Background Activity	GAL4 Response	Enhancers Tested	Cre/lox Replacement?
Aw+HG	8	3R:5306327	in between CG16779 and CG8147				
Aw+HG	93	3R:5822510	in coding of Gr85a (immediately after the start ATG)				
Aw+HG	93	3R:5822510	in coding of Gr85a (immediately after the start ATG)				
Aw+HG	68	3R:6199286	in intron of Rfx				
Aw+HG	46	3R:9267836	in between yrt and CG14372				
Aw+HG	28	3R:9927160	in intron of CG9918				
Aw+HG	57	4:297073	upstream of Nfl				
Aw+HG	21	4:739501	in intron of ey				
Aw+HG	13	unlocalized	in mgd1 transposable element				
Aw+HG	66	X:11562100	in intron of Ptp10D				
Aw+HG	66	X:11562100	in intron of Ptp10D				
Aw+HG	17	X:11817888	in intron of CG2471				
Aw+HG	54	X:11896214	in 5' UTR of fw				
Aw+HG	3	X:12318684	in between CG32652 and Pkcdelta				
Aw+HG	75	X:1254640	in between png and CG14770				
Aw+HG	75	X:1254640	in between png and CG14770				
Aw+HG	44	X:2973702	5' of CG4116 and kirre	small subset CNS	ActGAL4		
Aw+HG	40	X:3719063	in intron of ec	not detected	ActGAL4		
Aw+HG	67	X:6887035	in intron of inx7				
Aw+HG	87	X:9196067	in coding of lz				
AWHR	AWHR line 9	2L:16348982	in coding of chif				
AWHR	AWHR line 4	2L:16608410	in coding of CG4631				

Table 5.1 Continued:

Type	Line #	Location	Description	Background Activity	GAL4 Response	Enhancers Tested	Cre/lox Replacement?
AWHR	AWHR line 2	2L:2200661	in between CG34172 and CG31668	epithelial, semi-ubiquitous	yes (Act5C, ASE, dpp)	neur4DWT, Sox1.3WT, svWT, GFP	
AWHR	AWHR line 1	2L:4976284	in coding of CG8890, GDP-mannos 4,6-dehydratase	none	none (Act5C, ASE, dpp)	neur4DWT, Sox1.3WT, svWT, GFP	
AWHR	AWHR line 3b	2L:4982988	in intron of CG34126				
AWHR	AWHR line 5B	2R:12133269	in intron of Cng				
AWHR	AWHR line 5A	3L:16043263	in intron of th				
AWHR	AWHR line 6	chr2	in copia element: not localized				
AWHR	AWHR line 10a	same as line 1					

We hypothesized two reasons for the poor rate of transformation: either low efficiency of Minos transgenesis or ineptitude of the hairpin constructs. We chose to address these hypotheses by attempting to induce new pAWHR insertions by endogenous Minos transposase ("hopping") and creating an additional pMemor variant, pAw+HG, which instead of a hairpin contained the mini-*white* gene from pH-Stinger.

In a small-scale transposition experiment, we made use of the endogenous Minos transposase source (Metaxakis et al., 2005) to attempt to mobilize the original AWHR insertions. We were easily able to collect a range of eye phenotypes using AWHR line 1 to create new hops, collecting 10 new eye colors which ultimately led to 6 new insertions following mapping and unfortunate loss of unhealthy stocks. Despite this ease of production of new lines, we were unable to produce new hops using AWHR line 2 as the starting element. It is for this reason, combined with the inability to generate inverse PCR sequence for the 5' end, that we assume integration may have compromised the Minos LTR at this end of the construct. Although the ease of generating new insertions by hopping would seem to suggest

that the strength of the hairpin construct was not the reason for low transformation efficiency, when we injected the pAw+HG construct, we were burdened by an abundance of transformants of this docking site, collecting > 90 potential independent insertions. This result would tend to favor the hairpin as the culprit for the apparent low transformation efficiency (see summary of docking sites, Table 5.1).

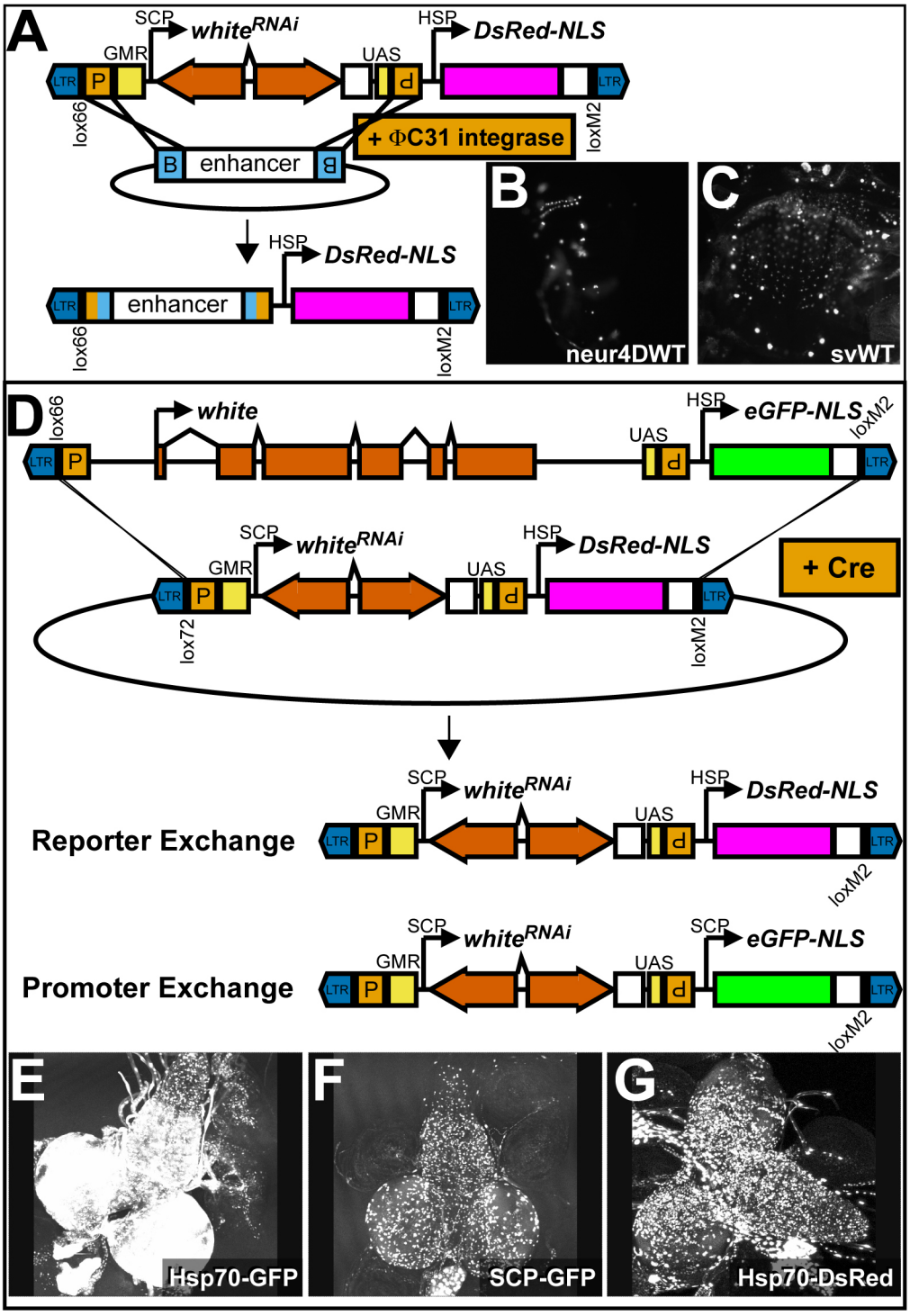
With these new docking site insertions in hand, we tested a number of them for both their background level of reporter expression due to genomic insertion location and their ability to respond to GAL4 expression. Looking at 12 Aw+HG lines and the two original AWHR lines in third instar larval imaginal discs, we find that in general most lines give no background reporter expression except for a small number of cells in the brain and eye imaginal disc (Table 5.2; Figure 5.3A-C). The exceptions among these was Aw+HG line 49, which produced expression in a large number of cells in the brain, and AWHR line 2, which was expressed in pharate adult epithelium in the abdomen. When we crossed the flies containing these docking sites to actin5C>GAL4, most lines gave expression in a all discs with stronger reporter activity in subregions of the brain and the notum region of the

wing imaginal disc, the latter of which may be representative of presumptive muscle cell identity. The exception among these lines was AWHR line 1, which not only responded to actin5C>GAL4, but also ASE>GAL4 and dpp>GAL4. While this line remains the only coding sequence insertion we have examined, we cannot confirm that this class of insertion is the reason for the lack of GAL4 response. Nevertheless, most lines appeared to respond well to GAL4 expression; we next sought to test the main purpose of this docking site: enhancer integration.

Enhancer cloning and integration

To maximize the efficiency of enhancer integration via ΦC31 integrase-mediated RMCE, we made use of the flies created by Konrad Basler's group which contain an endogenous source of germline-enriched integrase (Bischof et al., 2007). Enhancers to be integrated into the Memor docking sites were amplified with primers containing 40mer attB sequences at their 5' ends,T-A cloned, sequenced, and immediately injected (Figure 5.2A). As an enhancer test set, we included the enhancers examined in previous chapters,

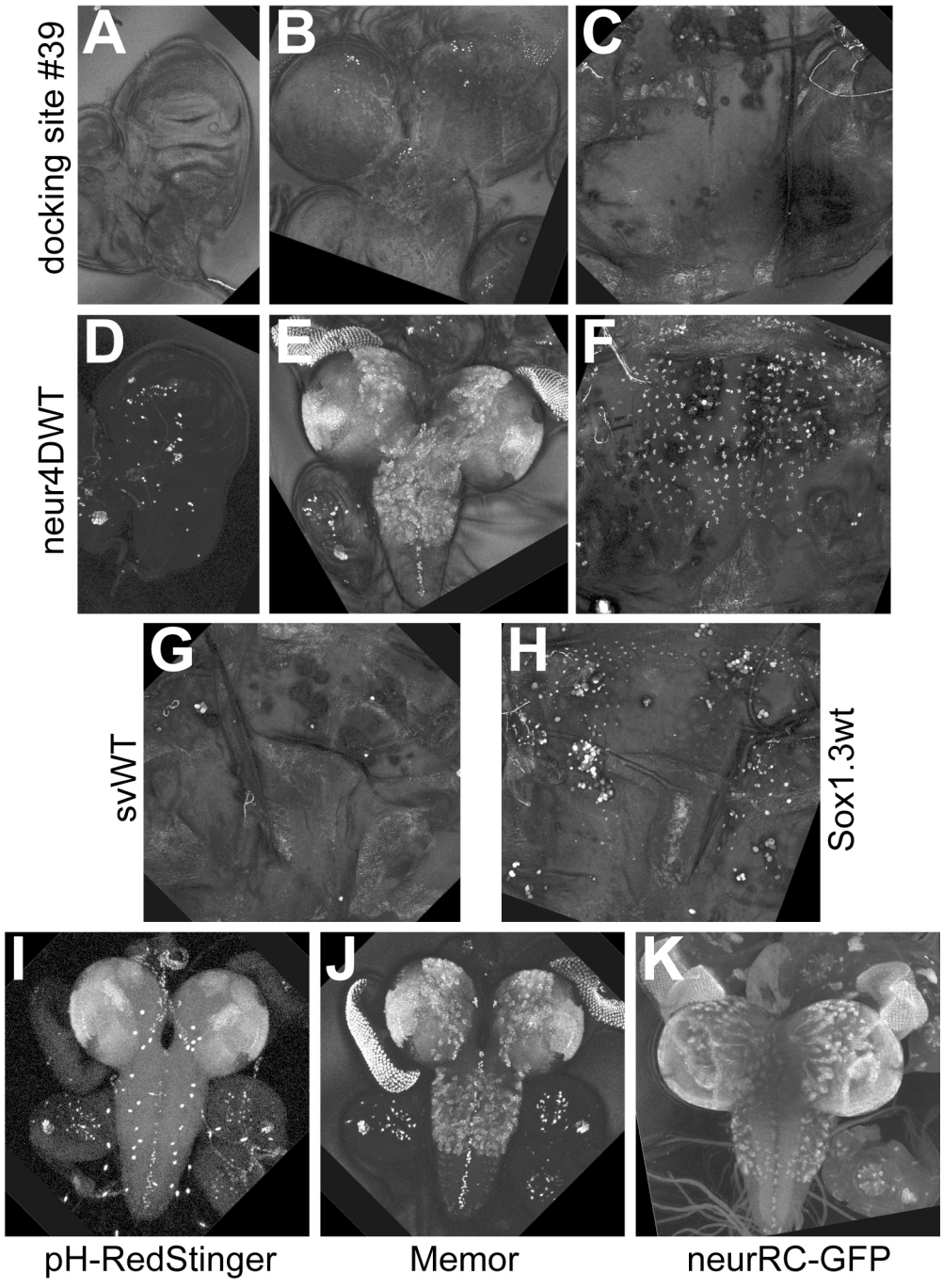
Figure 5.2: Recombination mediated cassette exchanges (RMCEs) at Memor docking sites.
(A – C) ΦC31 integrase RMCE of enhancers into the docking site. An enhancer with 40 bp inverted attB sequences is T-A cloned and injected into flies containing the docking site and endogenous integrase (see Supplemental Figure). The cassette containing the marker for transgenesis is exchanged for the enhancer, resulting in loss of the marker that can be screened for. (B) Expression of DsRed in third instar wing imaginal discs following integration of the neur4DWT enhancer into AWHR docking site line 2. (C) Pupal expression of DsRed in the microchaete field following integration of the svWT enhancer into AWHR docking site line 2. (D) Complete docking site replacement through Cre RMCE. Diagramed is the replacement of a docking site with a mini-w^+ marker and GFP for a docking site with a *w* hairpin marker and DsRed. Such a recombination can mediate any desired change at a given docking site. (E – G) Reporter expression in third instar larval brains in response to actin5C>GAL4 activity either at (E) Aw+HG docking site 11 prior to Cre RMCE or after RMCE replacement for (F) conversion of the promoter from Hsp70 to SCP or (G) conversion of the reporter from GFP to DsRed. Note the differences in expression level between Hsp70 and SCP.


A
SCP
white^RNAi
HSP
DsRed-NLS
GMR
UAS
LTR
P
d
lox66
loxM2
+ ΦC31 integrase
B
enhancer
B
HSP
DsRed-NLS
enhancer
lox66
loxM2
B
neur4DWT
C
svWT
D
lox66
white
HSP
eGFP-NLS
loxM2
UAS
SCP
white^RNAi
GMR
HSP
DsRed-NLS
+ Cre
lox72
loxM2
Reporter Exchange
SCP
white^RNAi
GMR
HSP
DsRed-NLS
UAS
loxM2
Promoter Exchange
SCP
white^RNAi
GMR
SCP
eGFP-NLS
UAS
loxM2
E
Hsp70-GFP
F
SCP-GFP
G
Hsp70-DsRed

neur4DWT, 1.3WT from *Sox15* (Miller et al., 2009), and svWT. We selected four lines of the Aw+HG type from each arm of the second and third chromosomes that localized to intergenic regions and for which there was minimal background reporter activity, as well as the two AWHR lines, to use as the starting test sites for ΦC31 integration. For the two AWHR lines, we also included the GFP coding sequence flanked by 200 bp attB sequences which has been shown to integrate well using ΦC31 RMCE (Bateman et al., 2006).

Enhancer integration in to each of the two AWHR lines further revealed differences between these insertions. Even though AWHR line 1 was able to mobilize efficiently, none of the tested constructs were able to integrate into this docking site (Table 5.2). We find it interesting that not only is this docking site location the only line which has so far failed to integrate a DNA fragment via ΦC31-mediated RMCE, but also this is the only site which failed to show reporter activation with GAL4. In contrast, AWHR line 2, though unable to mobilize, was demonstrably functional for both GAL4 response and ΦC31 RMCE. We note that we experienced significant difficulty in general with all the injections in the Canton S background, from the

Figure 5.3: Enhancer activity in Memor docking site #39.
(A – C) Background reporter expression of docking site #39 in (A) a third instar wing imaginal disc, (B) a third instar larval brain, and (C) a 24 hr APF notum. (D – F) Reporter expression following ΦC31 RMCE with the neur4DWT enhancer in (D) wing imaginal disc, (E) a larval brain, and (F) a 24 hr APF notum. (G) A 24 hr APF notum upon FC31 RMCE with the svWT enhancer; no reporter expression is observed, contrary to expectation. (H) A 24 hr APF notum upon ΦC31 RMCE with the Sox1.3wt enhancer; in this case expected expression is observed. (I – K) In some contexts, Memor is more representative of endogenous expression than other reporter constructs. (I) A third instar larval brain expressing DsRed under control of the neur4DWT enhancer in a pH-RedStinger transgenic insertion. (J) GFP expression in the same brain with the same enhancer in the Memor docking site 39, showing significant difference from DsRed specificity (neur4DWT integrant line C). (K) A GFP-tagged rescue construct is much more similar to the Memor pattern than the Stinger pattern.


docking site #39
A
B
C
neur4DWT
D
E
F
svWT
G
H
Sox1.3wt
I
J
K
pH-RedStinger
Memor
neurRC-GFP

initial transformation with the docking site itself through the process of enhancer integration. Much of this we attributed to the general health of the CS stock in our hands; indeed, we have since seen improvement since obtaining a healthy Oregon R stock as our *w+* strain (data not shown).

We found that all four of the Aw+HG lines were produced greater efficiency of RMCE integration of enhancers and significantly better viability than the AWHR docking site lines (21% vs 4.8% embryo survival and 61% vs 36% fertility). The average integration efficiency across all lines was 30% for each fertile cross for the Aw+HG lines, with the lowest integration at 16% for line 11 and the highest at 34% for line 39.

As we created stocks for these integrants we noticed that several lines were not homozygous viable, even though the original docking site was. Furthermore, across all docking site lines, while we were able to find representative integrants to demonstrate that all enhancers were functional in this system, we were also able to find many lines that failed to give enhancer expression. This failure did not

correlate with docking site, orientation, enhancer, or viability (Table 5.4, Figure 5.3D-H).

For integrants which did produce reporter expression, we compared the pattern from Memor with the same enhancer cloned into the pH-Stinger vector by looking a the progeny of a cross between these two flies. We find that for both neur4DWT and Sox1.3wt, Memor enhancers give weaker overall expression than enhancers in pH-Stinger, which is likely due to the signal amplification associated with insulator sequences (Markstein et al., 2008) in pH-Stinger. In addition to this difference in signal levels, when we examined neur4DWT integrants, we find that reporter expression in the brain differs markedly between the two systems. Surprisingly, of the two, the Memor system presents a pattern that more accurately represents the endogenous *neur* pattern in this tissue, as seen with a GFP-tagged *neur* rescue construct (Figure 5.3I-K; see Chapter 2). While this specificity difference may also be attributable to the insulator sequences, we have not yet confirmed this assumption.

Docking site replacement

Table 5.2: ΦC31 RMCE into Memor docking sites.

Line	Transgenesis Marker	Location	Insert	Embryos Injected	# Crosses	# Fertile Crosses	# Vials With Integrant	Percentage
1	white-hairpin	2L:4976284	GFP	600	60	26	0	0
			neur4DWT	600	48	15	0	0
			svWT	600	41	13	0	0
			Sox1.3wt	600	38	13	0	0
2	white-hairpin	2L:2200661	GFP	600	11	6	1	17%
			neur4DWT	600	9	2	2	100%
			svWT	600	9	4	1	25%
			Sox1.3wt	600	14	4	1	25%
11	mini-w+	3R:20183759	neur4DWT	280	41	20	3	15%
			svWT	280	51	22	5	23%
			Sox1.3wt	280	46	31	4	13%
12	mini-w+	3L:6353840	neur4DWT	280	33	12	2	16%
			svWT	280	45	21	7	33%
			Sox1.3wt	280	49	28	10	36%
39	mini-w+	2R:19503472	neur4DWT	280	55	36	10	28%
			svWT	280	82	54	19	36%
			Sox1.3wt	280	86	55	21	38%
41	mini-w+	2L:16761724	neur4DWT	280	55	34	11	32%
			svWT	280	75	51	15	29%
			Sox1.3wt	280	83	62	20	32%

To assess whether the anomalies we have observed are inherent to the Memor docking site or specific to ΦC31-mediated RMCE, we used RMCE integration by Cre/lox to replace the mini-*white*-Hsp70-eGFP docking site architecture with either *white* hairpin-Hsp70-DsRed or *white* hairpin-SCP-eGFP (Figure 5.2D). Using lines 11 and 12 as the starting docking sites, we achieved replacement, as scored by the loss of the mini-w+ eye phenotype, with an overall average of 16% of fertile crosses (Table 5.4). We checked to see if these replacement constructs were functional by crossing either to Oregon R flies to check the *white* hairpin construct or to actin5C>GAL4 to test the promoter-reporter pair. Both replacement constructs into either docking site failed to give sufficient knock-down of *white* gene product to observe a phenotype (data not shown). Despite the lack of hairpin activity, all lines tested showed reporter activation in response to actin5C>GAL4. Moreover, we noticed a striking difference in the expression of GFP under the control of the SCP, as compared with the Hsp70-GFP construct in the initial starting line, with much strong expression in the latter (Figure 5.3E-G). Such a strong difference may be accounted for by the presence of GAGA sites in the Hsp70

Table 5.3: Memor Enhancer Integrant Orientation and Expression.

Docking Site	Location	Enhancer	Line	Determined Orientation	Homozygous Viable?	Expression?
11	3R:20183759	neur4DWT	A		no	
11	3R:20183759	neur4DWT	B		no	
11	3R:20183759	svWT	A		no	
11	3R:20183759	svWT	B		no	
11	3R:20183759	svWT	C		yes	
11	3R:20183759	Sox1.3wt	A		yes	
12	3L:6353840	neur4DWT	A	+	yes	yes
12	3L:6353840	svWT	A	+	yes	no (24 hr APF)
12	3L:6353840	svWT	B	+	yes	no (24 hr APF)
12	3L:6353840	svWT	D		no	
12	3L:6353840	svWT	E		yes	no (24 hr APF)
12	3L:6353840	Sox1.3wt	A	-	yes	no (24 hr APF)
12	3L:6353840	Sox1.3wt	B		no	
12	3L:6353840	Sox1.3wt	D		yes	
12	3L:6353840	Sox1.3wt	E		yes	yes
12	3L:6353840	Sox1.3wt	H	-	yes	no
12	3L:6353840	Sox1.3wt	I		yes	no
39	2R:19503472	neur4DWT	B		yes	
39	2R:19503472	neur4DWT	C	-	yes	yes
39	2R:19503472	neur4DWT	D		no	
39	2R:19503472	svWT	B		yes	
39	2R:19503472	svWT	C		no	
39	2R:19503472	svWT	D	-	yes	no (24 hr APF)
39	2R:19503472	svWT	F		no	
39	2R:19503472	svWT	G	+	yes	no (24 hr APF)
39	2R:19503472	Sox1.3wt	B		yes	
39	2R:19503472	Sox1.3wt	E	-	yes	
39	2R:19503472	Sox1.3wt	F	+	yes	yes
39	2R:19503472	Sox1.3wt	G		yes	yes
39	2R:19503472	Sox1.3wt	H		yes	yes
41	2L:16761724	neur4DWT	B	+		yes
41	2L:16761724	neur4DWT	C	+		yes
41	2L:16761724	neur4DWT	E	-	yes	

Table 5.3 Continued:

Docking Site	Location	Enhancer	Line	Determined Orientation	Homozygous Viable?	Expression?
41	2L:16761724	neur4DWT	F	+	yes	
41	2L:16761724	neur4DWT	G	+		no
41	2L:16761724	neur4DWT	H	-	yes	yes
41	2L:16761724	neur4DWT	I	-	yes	yes
41	2L:16761724	neur4DWT	J	-	yes	yes
41	2L:16761724	svWT	A	+		no (24 hr APF)
41	2L:16761724	svWT	B	+	yes	
41	2L:16761724	svWT	D	+		no (24 hr APF)
41	2L:16761724	svWT	G	-	yes	no (24 hr APF)
41	2L:16761724	svWT	J	-	yes	
41	2L:16761724	svWT	K	+	yes	no (24 hr APF)
41	2L:16761724	svWT	M	-	yes	no (24 hr APF)
41	2L:16761724	svWT	N	-	yes	no (24 hr APF)
41	2L:16761724	svWT	O	+	yes	
41	2L:16761724	Sox1.3wt	A	+	yes	yes
41	2L:16761724	Sox1.3wt	B	-	yes	yes
41	2L:16761724	Sox1.3wt	G	+	yes	
41	2L:16761724	Sox1.3wt	H	+		yes
41	2L:16761724	Sox1.3wt	I		yes	
41	2L:16761724	Sox1.3wt	J	-	yes	yes
41	2L:16761724	Sox1.3wt	K		yes	no (24 hr APF)
41	2L:16761724	Sox1.3wt	O			
41	2L:16761724	Sox1.3wt	S	+	yes	
2	2L:2200661	neur4DWT	A			yes
2	2L:2200661	svWT	A			yes

promoter sequences (J. Kadonaga, personal communication). These reporter data suggest that a reason for the lack of hairpin activity upon replacement, and perhaps also the reason for low transformation efficiency of AWHR initially, may be the SCP activity within the hairpin construct.

With our suspicion of the relative activity of the SCP, we replaced the SCP in pSWU with Hsp70 to create pHWU. We then cloned in the GMR enhancer and *white* arms to complete the hairpin construct, and then cloned this marker into pBASEb with an Hsp70-GFP pair to create BhWHG. We again used Cre/lox to replace the Aw+HG architecture with this construct at the docking site #39. We achieved similar Cre RMCE efficiency, around 22%. We examined the integrity of the recombination event by PCR, which showed no difference between the different replacements when assayed using primers at the ends of the construct and in the surrounding genomic region. When we examined the ability of the construct to knock-down *white* gene activity, we saw differences in phenotypic variation, ranging from dark red-brown eyes to orange eyes. These data demonstrate that variability comes from not only with ΦC31 RMCE, but also Cre

Table 5.4: Cre RMCE of Memor Docking Site Architecture.

Line	Location	Description	Replacement Construct Description	# Embryos Injected	# Crosses	# Fertile Crosses	# Vials With white-eyed fly	%
11	3R:20183759	Aw+HG	BWHR	600	198	124	20	16%
11	3R:20183759	Aw+HG	BWSG	600	130	69	7	10%
12	3L:6353840	Aw+HG	BWHR	600	154	90	13	14%
12	3L:6353840	Aw+HG	BWSG	600	115	52	13	25%
39	2R:19503472	Aw+HG	BhWHG	600	ND	54	12	22%
39	2R:19503472	A-neur4DWT-HG integrant #C	BLw+HG	600	ND	49	3	6%

RMCE. Importantly, since with this construct we were able to detect knock-down of *white* activity, these data favor the use of hairpin construct that contain Hsp70 as their promoter, rather than SCP.

Concluding remarks

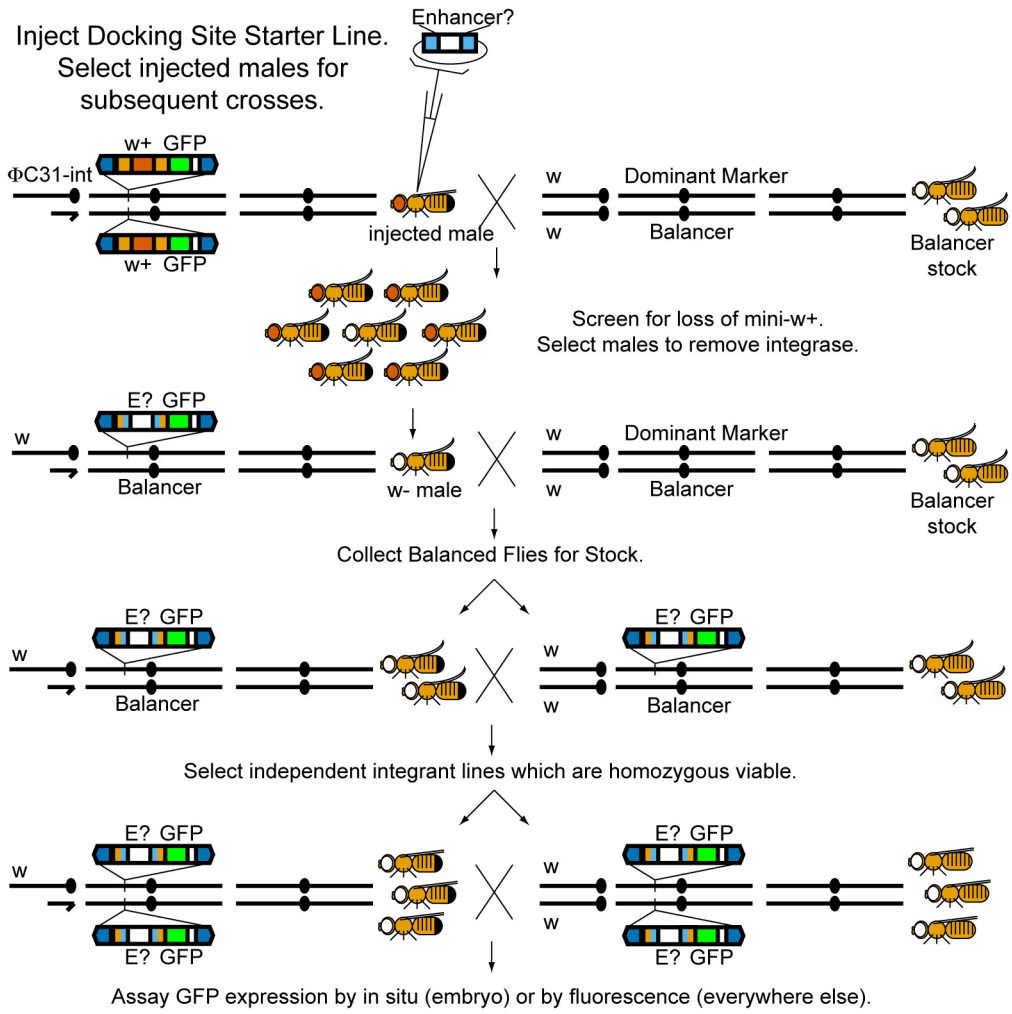
Here we have presented evidence for the functionality of a completely new strategy for enhancer transgenesis, and uncovered anticipated challenges not previously reported in the literature. We have described the first functional use of the SCP in the context of both a hairpin and reporter; while it does indeed show functionality, it appears to have much weaker activity than Hsp70 when used in the same construct. The original intent of SCP design was maximum compatibility with different types of enhancers (J. Kadonaga, personal communication), and this particular capability was not tested in our assays. We have also demonstrated the utility of a single docking site that can be completely replaced for additional purposes and uses, through the use of Cre upon pseudo-loxP sequences. We have shown that the presence of an enhancer-promoter pair, coupled with 5X UAS sequences provides the ability to test the quality of a docking site

before any construct is even integrated at that site. Indeed, as we have seen with AWHR line 1, the mere observation of transgenic marker activity provides no indication of docking site functionality.

The main confounding factor in our analyses is the currently unexplained variability in both viability and enhancer expression upon enhancer integration. One contributing factor affecting viability is the activity of the integrase on pseudo-attP sequences within the *Drosophila* genome, which could lead to lethal genomic rearrangements. Both viability and expression could also be subject to cellular responses to double strand breaks occurring in the exchange. It has been observed previously that P element insertions at the same nucleotide sequence can result in differences in viability correlated with changes in the size, type, and orientation of element insertion (Bellen et al., 1992), though this has not been explained mechanistically. We plan to explore minimizing the size change resulting from enhancer integration by making use of the smaller and apparently more versatile Hsp70 *white* hairpin construct to see if these problems can be rectified.

If these obvious flaws in the system can be fixed, our data suggest that the Memor system shows great promise for greater utility

beyond both enhancer transgenesis and *Drosophila*. With a minimal collection of defined docking sites, experimental variation can be greatly controlled from lab to lab. Moreover, the capacity to use the same docking site for RMCE (both Cre and ΦC31-mediated) as well as single attP construct integration further expands the applications for a single site. The potential uses extend beyond even flies, however, as both Minos elements and ΦC31 integrase have been demonstrated to function even in vertebrate genomes with reasonable efficiency (Thyagarajan et al., 2001; Drabek et al., 2003). Furthermore, using hairpins as a transgenic marker eliminates the need for both a mutant recipient strain (for which there often is not one for emerging model organisms) and for tools for transgenesis detection by fluorescence (which may be limiting for some labs).



Supplemental Figure 5.4: The crossing scheme for collecting RMCE integrants in the Memor system.
Males are selected from the injected embryos, which are homozygous for the docking site and contain an integrase source on the X chromosome. These males are crossed to a balancer stock lacking the transgenesis marker. Progeny from this cross are screened for presence of the balancer chromosome and loss of the marker. Such males are back-crossed again to the balancer line individually to check for homozygous viability and stock creation.

References

Barolo, S. *et al.* (2000). GFP and beta-galactosidase transformation vectors for promoter/enhancer analysis in Drosophila. Biotechniques *29*, 726, 728, 730, 732.

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

Bateman, J. R. *et al.* (2006). Site-specific transformation of Drosophila via phiC31 integrase-mediated cassette exchange. Genetics *173*, 769-777.

Bellen, H. J. *et al.* (1992). The Drosophila couch potato protein is expressed in nuclei of peripheral neuronal precursors and shows homology to RNA-binding proteins. Genes Dev *6*, 2125-2136.

Bellen, H. J. *et al.* (1989). P-element-mediated enhancer detection: a versatile method to study development in Drosophila. Genes Dev *3*, 1288-1300.

Bischof, J. *et al.* (2007). An optimized transgenesis system for Drosophila using germ-line-specific phiC31 integrases. Proc Natl Acad Sci U S A *104*, 3312-3317.

Bowtell, D. D. *et al.* (1989). Regulation of the complex pattern of sevenless expression in the developing Drosophila eye. Proc Natl Acad Sci U S A *86*, 6245-6249.

Buttgereit, D. *et al.* (1991). During Drosophila embryogenesis the beta 1 tubulin gene is specifically expressed in the nervous system and the apodemes. Mech Dev *33*, 107-118.

Celniker, S. E. *et al.* (2009). Unlocking the secrets of the genome. Nature *459*, 927-930.

Drabek, D. *et al.* (2003). Transposition of the Drosophila hydei Minos transposon in the mouse germ line. Genomics *81*, 108-111.

Geyer, P. K., and Corces, V. G. (1987). Separate regulatory elements are responsible for the complex pattern of tissue-specific and developmental transcription of the yellow locus in Drosophila melanogaster. Genes Dev *1*, 996-1004.

Gloor, G. B. *et al.* (1993). Type I repressors of P element mobility. Genetics *135*, 81-95.

Groth, A. C. *et al.* (2004). Construction of transgenic Drosophila by using the site-specific integrase from phage phiC31. Genetics *166*, 1775-1782.

Hazelrigg, T. *et al.* (1984). Transformation of white locus DNA in drosophila: dosage compensation, zeste interaction, and position effects. Cell *36*, 469-481.

Ho, S. N. *et al.* (1989). Site-directed mutagenesis by overlap extension using the polymerase chain reaction. Gene *77*, 51-59.

Juven-Gershon, T. *et al.* (2006). Rational design of a super core promoter that enhances gene expression. Nat Methods *3*, 917-922.

Klinakis, A. G. *et al.* (2000). Genome-wide insertional mutagenesis in human cells by the Drosophila mobile element Minos. EMBO Rep *1*, 416-421.

Lobo, N. *et al.* (1999). Transposition of the piggyBac element in embryos of Drosophila melanogaster, Aedes aegypti and Trichoplusia ni. Mol Gen Genet *261*, 803-810.

Loukeris, T. G. *et al.* (1995a). Introduction of the transposable element Minos into the germ line of Drosophila melanogaster. Proc Natl Acad Sci U S A *92*, 9485-9489.

Loukeris, T. G. *et al.* (1995b). Gene transfer into the medfly, Ceratitis capitata, with a Drosophila hydei transposable element. Science *270*, 2002-2005.

Markstein, M. *et al.* (2008). Exploiting position effects and the gypsy retrovirus insulator to engineer precisely expressed transgenes. Nat Genet *40*, 476-483.

Metaxakis, A. *et al.* (2005). Minos as a genetic and genomic tool in Drosophila melanogaster. Genetics *171*, 571-581.

Miller, S. W. *et al.* (2009). Complex interplay of three transcription factors in controlling the tormogen differentiation program of Drosophila mechanoreceptors. Dev Biol *329*, 386-399.

Oberstein, A. *et al.* (2005). Site-specific transgenesis by Cre-mediated recombination in Drosophila. Nat Methods *2*, 583-585.

Pavlopoulos, A. *et al.* (2004). Efficient transformation of the beetle Tribolium castaneum using the Minos transposable element: quantitative and qualitative analysis of genomic integration events. Genetics *167*, 737-746.

Plasterk, R. H. *et al.* (1999). Resident aliens: the Tc1/mariner superfamily of transposable elements. Trends Genet *15*, 326-332.

Rio, D. C. *et al.* (1988). Evidence for Drosophila P element transposase activity in mammalian cells and yeast. J Mol Biol *200*, 411-415.

Rubin, G. M., and Spradling, A. C. (1983). Vectors for P element-mediated gene transfer in Drosophila. Nucleic Acids Res *11*, 6341-6351.

Ryder, E., and Russell, S. (2003). Transposable elements as tools for genomics and genetics in Drosophila. Brief Funct Genomic Proteomic *2*, 57-71.

Schultz, J. R. *et al.* (1991). A muscle-specific intron enhancer required for rescue of indirect flight muscle and jump muscle function regulates Drosophila tropomyosin I gene expression. Mol Cell Biol *11*, 1901-1911.

Siegal, M. L., and Hartl, D. L. (1996). Transgene Coplacement and high efficiency site-specific recombination with the Cre/loxP system in Drosophila. Genetics *144*, 715-726.

(2004). The ENCODE (ENCyclopedia Of DNA Elements) Project. Science *306*, 636-640.

Thorpe, H. M., and Smith, M. C. (1998). In vitro site-specific integration of bacteriophage DNA catalyzed by a recombinase of the resolvase/invertase family. Proc Natl Acad Sci U S A *95*, 5505-5510.

Thyagarajan, B. *et al.* (2001). Site-specific genomic integration in mammalian cells mediated by phage phiC31 integrase. Mol Cell Biol *21*, 3926-3934.

Venken, K. J., and Bellen, H. J. (2007). Transgenesis upgrades for Drosophila melanogaster. Development *134*, 3571-3584.

Venken, K. J. *et al.* (2009). Versatile P[acman] BAC libraries for transgenesis studies in Drosophila melanogaster. Nat Methods

Venken, K. J. *et al.* (2006). P[acman]: a BAC transgenic platform for targeted insertion of large DNA fragments in D. melanogaster. Science *314*, 1747-1751.

Venken, K. J. *et al.* (2008). Recombineering-mediated tagging of Drosophila genomic constructs for in vivo localization and acute protein inactivation. Nucleic Acids Res *36*, e114.

Wagner-Bernholz, J. T. *et al.* (1991). Identification of target genes of the homeotic gene Antennapedia by enhancer detection. Genes Dev *5*, 2467-2480.

Chapter Five, in full, is a manuscript in preparation for publication: pMemor: A highly versatile system for rapid enhancer transgenesis and beyond. Miller, S. W. and Posakony, J. W. The dissertation author was the primary researcher and author. James Posakony provided key inputs into the initial designs of the system and provided crucial oversight throughout the study. Scott Barolo also contributed to construct design.