

## ABSTRACT

PATEL, MUKUND KIRITBHAI. Disconnected, a C2H2 Zinc Finger Protein has a role in Appendage Formation and Gene Regulation in *Drosophila*. (Under the direction of James W. Mahaffey.)

The Homeotic/Hox proteins are evolutionarily conserved proteins that specify segment identity by regulating segment-specific batteries of target genes. One the biggest questions in developmental biology is how the Hox proteins direct such complex target gene regulation given that they have very little target specificity. One possible way to increase transcription specificity is through the use of cofactors. This has lead to extensive searches to identify potential Hox cofactors. Although several Hox cofactors have been identified, only Extradenticle has been well characterized. Previously, our lab identified paralogous redundant genes *disconnected* (*disco*) and *disco-related* (*disco-r*) as genetic cofactors for the *Hox* genes *Deformed* and *Sex combs reduced*, which are required for proper *Drosophila* embryonic head patterning. *disco* and *disco-r* encode C2H2 zinc finger transcription factors that are conserved among edysozoa, lophotrochozoa, and some deuterosomes. Here, we present a molecular characterization of Disco as a transcription factor. We determined that Disco is able to bind to sequences that contain TGACA at the core *in vitro*. Furthermore, we characterized the interaction of Disco with the corepressor CtBP. we also examine the significance of certain amino acid differences in the DNA-binding domain between ecdysozoa/lophothrochozoa and deuterosomes, concluding that one cysteine is required for complete Disco function.

In addition to the molecular characterization, we also describe Disco's role during development of the adult *Drosophila* demonstrating that Disco is part of the appendage

development network. These data are supported from studies in *Tribolium castaneum* done by Lisa Robertson and Nathaniel Grubbs.

Several redundant C2H2 zinc finger genes have been identified in *Drosophila* including *disco* and *disco-r* that have roles in segmental patterning. Based on this we undertook a genome wide study to identify other novel, redundant C2H2 zinc finger genes that also function during segmental patterning. The results of this study are presented including identification two genes, *CG5249* and *CG11798*. Lastly, the appendix contains data from a microarray experiment that examined transcriptional changes due to ectopic activation of *Disco*.

Disconnected, a C2H2 Zinc Finger Protein has a role in Appendage  
Formation and Gene Regulation in Drosophila.

By

Mukund Kiritbhai Patel

A dissertation submitted to the Graduate Faculty of  
North Carolina State University  
in partial fulfillment of the  
requirements for the Degree of  
Doctor of Philosophy

Genetics

Raleigh, North Carolina

2007

APPROVED BY:

---

Dr. James W. Mahaffey  
Chair of Advisory Committee

---

Dr. Laura Mathies  
Committee Member, Genetics

---

Dr. Pat Estes  
Committee Member, Genetics

---

Dr. Steve Spiker  
Committee Member, Genetics

## **BIOGRAPHY**

Mukund Patel was born in Pij, India, but has lived in North Carolina since the age of 7.

He proudly attended the University of North Carolina-Chapel Hill where he earned his

B.S. in Biology and minor in Chemistry.

## **ACKNOWLEDGEMENTS**

I would like to thank Department of Genetics for the opportunity they have provided me. I especially would like to thank my advisor Jim Mahaffey for allowing me to join his lab and the advice and guidance he has given me. I would also like to thank the members of my committee Pat Estes, Laura Mathies, and Steve Spiker for their guidance. Lastly, I thank all present and previous members of the Mahaffey Lab who have assisted me along the way.

## TABLE OF CONTENTS

	Page
LIST OF TABLES.....	vii
LIST OF FIGURES.....	viii
LIST OF ABBREVIATIONS.....	x
CHAPTER ONE: General Introduction.....	1
References.....	6
CHAPTER TWO: Drosophila Embryonic Development and the Role of HOM-C Proteins and Cofactors.....	8
Overview of Drosophila embryonic development.....	9
Establishment of the metameric Drosophila embryo.....	11
The <i>Homeotic</i> genes and specification of identity along the A-P axis.....	15
Expression of <i>Dfd</i> and mutant phenotypes.....	18
Expression of <i>Scr</i> and mutant phenotypes.....	19
The homeodomain and cofactors.....	21
Collaborative interactions vs. Cooperative interactions.....	24
Extradenticle as a cofactor.....	24
Teashirt as a cofactor.....	30
Disco and Disco-r as potential Hox cofactors.....	34
References.....	38
CHAPTER THREE: The Appendage Role of Insect <i>disco</i> Genes and Possible Implications on the Evolution of the Maggot Larval Form.....	45
Abstract.....	48
Introduction.....	49
Materials and Methods.....	50
Results.....	55
The <i>disco</i> genes are widely expressed in ventral imaginal discs of Drosophila.....	55
Cells lacking the <i>disco</i> genes do not contribute to the ventral appendages.....	57
Ectopic Disco transforms dorsal appendages to ventral identity.....	63
The appendage role of <i>disco</i> is conserved in the beetle, <i>Tribolium</i> .....	71
The <i>disco</i> genes operate as appendage factors during Drosophila embryogenesis.....	78

Discussion.....	78
<i>disco</i> as an appendage factor.....	78
The <i>Drosophila</i> embryo and appendage factors.....	82
Acknowledgements.....	85
Supplementary Material.....	86
References.....	90
CHAPTER FOUR: Functional Analysis of the Disconnected Protein.....	95
Introduction.....	96
Materials and Methods.....	97
Results.....	102
Significance of alanine to threonine change.....	103
Potentially different isoforms of the Disco protein.....	104
DNA binding by the zinc finger region of the Disco protein.....	109
<i>In vivo</i> significance of Disco binding site.....	114
Disco interacts with the corepressor CtBP.....	117
Discussion.....	123
References.....	129
CHAPTER FIVE: Identification of novel C2H2 Zinc Finger Genes in Epidermal Pattern Formation.....	132
Introduction.....	133
Materials and Methods.....	137
Results.....	140
Identification and expression patterns of potentially duplicated Genes.....	140
Highly conserved genes between <i>Anopheles</i> and <i>Drosophila</i> .....	144
Expression pattern of <i>CG5249 (Blimp-1)</i> and RNAi phenotypes..	144
Expression pattern of <i>CG11798</i> and RNAi phenotypes.....	146
Discussion.....	150
References.....	156
CHAPTER SIX: Summary and Future Work.....	161
Summary.....	162
Future work for Disco's role in gene regulation.....	164
<i>Distal-less</i> as a potential direct target of Disco.....	166
Further evidence to support that Disco as appendage factor during embryogenesis.....	168
Other potential protein interactions with Disco.....	169
Involvement of appendage genes in cardiac development.....	171
The role and function of Disco-r.....	172
Combinatorial expression of Disco with other Hox proteins.....	173
References.....	175

APPENDIX.....	180
Materials and Methods.....	181
Results.....	182
References.....	189

## LIST OF TABLES

CHAPTER THREE	
Table 1.....	61
CHAPTER FOUR	
Table 1: Hatch rate on <i>UAS-Cys102/Cyo X prd-GAL4/prd-GAL4</i> .....	107
CHAPTER FIVE	
Table 1: Primers used for PCR.....	139
Table 2: Expression of duplicated genes.....	142
Table 3: Expression patterns of highly conserved C2H2 zinc finger genes between <i>Drosophila melanogaster</i> and <i>Anopheles gambiae</i> .....	145
APPENDIX	
Table 1: Transcriptional Changes in 6-8 hr Ectopically Expressing Disco embryos.....	184
Table 2: Transcriptional Changes in 7-13 hr Ectopically Expressing Disco embryos.....	186

## LIST OF FIGURES

### CHAPTER TWO

Figure 1: Division of the <i>Drosophila</i> embryo into segments, parasegments, and compartments.....	12
Figure 2: The segmentation hierarchy.....	13
Figure 3: Expression of the <i>Drosophila HOM-C</i> genes along the anterior-posterior axis.....	16
Figure 4: The <i>Drosophila</i> larval head.....	20

### CHAPTER THREE

Figure 1: Expression of <i>disco</i> and <i>disco-r</i> in third instar imaginal discs.....	56
Figure 2: Relationship between <i>disco</i> and other appendage factors in the <i>Drosophila</i> wandering third instar discs.....	58
Figure 3: Clonal analysis of <i>disco</i> genes.....	62
Figure 4: Effects of ectopic, <i>dpp</i> -driven <i>disco</i> expression.....	65
Figure 5: Expression of leg determining genes in <i>dpp&gt;disco</i> transformed wing.....	69
Figure 6: Expression of <i>Tribolium disco</i> genes.....	73
Figure 7: <i>Tribolium</i> RNAi analysis.....	76
Figure 8: Role of the <i>disco</i> genes during appendage development in embryos.....	79
Supplemental Figure 1.....	86
Supplemental Figure 2: Map of <i>Df(1)ED7355</i> .....	88
Supplemental Figure 3: Expression of wing determining genes in the transformed wings.....	89

### CHAPTER FOUR

Figure 1: Disco is an evolutionarily conserved protein.....	105
---	-----

Figure 2: Ectopic expression of the different isoforms of Disco.....	108
Figure 3: Disco represses <i>rpr</i> expression in the maxillary segment.....	111
Figure 4: <i>lacZ</i> expression in deficiency and ectopic Disco embryos.....	112
Figure 5: Disco binds to the site TGACA within the <i>rpr4S3</i> enhancer region.....	115
Figure 6: The Disco binding site is necessary for Disco to bind DNA.....	116
Figure 7: Expression of <i>lacZ</i> from the <i>rpr4S3</i> reporter with Disco sites mutated.....	118
Figure 8: CtBP is only capable of interacting with Disco when the PLDFS site is present.....	120
Figure 9: Ectopic expression of UAS-Disco CtBP <sup>-</sup> and protein expression levels.....	122
 CHAPTER FIVE	
Figure 1: Alignment of <i>CG5249</i> to the Anopheles genome and Drosophila embryonic expression.....	148
Figure 2: Alignment of <i>CG11798</i> to the Anopheles genome and Drosophila embryonic expression.....	149
Figure 3: <i>Df(2R)XTE-58</i> and RNAi phenotypes of <i>CG11798</i> compared to wild type.....	151

## List of Abbreviations

<b>Abbreviation</b>	<b>Full Name</b>
<i>abd-A</i>	<i>abdominal-A</i>
<i>Abd-B</i>	<i>Abdominal-B</i>
<i>Antp</i>	<i>Antennapedia</i>
<i>ap</i>	<i>apontic</i>
<i>apt</i>	<i>apterous</i>
<i>bi</i>	<i>bifid</i>
<i>btd</i>	<i>buttonhead</i>
<i>cnc</i>	<i>cap n' collar</i>
<i>CtBP</i>	<i>C-terminal Binding Protein</i>
<i>dac</i>	<i>dachshund</i>
<i>Dfd</i>	<i>Deformed</i>
<i>dip1</i>	<i>disco interacting protein 1</i>
<i>dip2</i>	<i>disco interacting protein 2</i>
<i>disco</i>	<i>disconnected</i>
<i>disco-r</i>	<i>disco-related</i>
<i>Dll</i>	<i>Distal-less</i>
<i>dpp</i>	<i>decapentaplegic</i>
<i>exd</i>	<i>extradenticle</i>
<i>fkf</i>	<i>fork head</i>
<i>HOM-C</i>	<i>Homeotic Complex</i>
<i>hth</i>	<i>homothorax</i>
<i>lab</i>	<i>labial</i>
<i>lin</i>	<i>lines</i>
<i>mod</i>	<i>modulo</i>
<i>pb</i>	<i>proboscipedia</i>
<i>pnr</i>	<i>pannier</i>
<i>rpr</i>	<i>reaper</i>
<i>sal</i>	<i>spalt</i>
<i>Scr</i>	<i>Sex combs reduced</i>
<i>sd</i>	<i>scalloped</i>
<i>sn</i>	<i>singed</i>
<i>ss</i>	<i>spineless</i>
<i>tio</i>	<i>tiptop</i>
<i>tsh</i>	<i>teashirt</i>
<i>Ubx</i>	<i>Ultrabithorax</i>
<i>wg</i>	<i>wingless</i>

# **CHAPTER ONE**

## **General Introduction**

## General Introduction

Understanding the evolution of animal body plans is of major interest to biologists because of the enormous amount of diversity. For example the Arthropod lineage alone has millions of extant species that are morphologically very diverse. The arthropod body is essentially a series of repeated segments often have differing morphological characteristics, usually in the form of an appendage extending from that segment. These appendages come in many shapes and forms contributing to the diversity that we observe. Understanding the developmental processes that lead to segment identity has been the focus of many researchers. Much of what is known about segment identity comes from the study of *Homeotic/Hox* genes, in part because loss of *Hox* expression during development leads to segmental transformations suggesting that *Hox* genes control the identity of segments (McGinnis and Krumlauf, 1992).

The *Hox* genes were first identified in *Drosophila* but are now known to be in nearly all animals and in most cases they are colinearly organized on the chromosome with regards to the anterior posterior body axis (McGinnis and Krumlauf, 1992). *Hox* genes control segment identity by regulating specific target genes. Expression of a *Hox* gene in an inappropriate location leads to a different appendage being formed in that segment. A famous example of this is the gain-of-function mutations in *Antennapedia* that lead to legs being formed where the antennae should be (Denell et al., 1981). They are also referred to as “selector” genes that function by regulating downstream “realizators” which ultimately direct morphogenesis (Garcia-Bellido, 1977). Realizators affect cell behaviors such as cell shape, adhesion, apoptosis, etc., which will determine

tissue or organ shape (Garcia-Bellido, 1977).

Unfortunately, the way *Hox* genes function is not as simple as it is stated. Since they have such dramatic effects, each Hox protein is thought to regulate a variety of target genes; however, very few direct *in vivo* targets have been identified for any given *Hox* gene. Additionally, on top of the few *in vivo* targets identified, Hox proteins have very weak *in vitro* binding specificities (Ekker et al., 1994). So the question arises, how can these few genes produce the morphological diversity that we observe? One potential way is through the use of cofactors, but again few have been identified.

There are several lines of evidence suggesting that cofactors play a key role in Hox function. Many of the downstream targets of Hox regulation are not realizators themselves, but molecules that indirectly control realizators (Hombria and Lovegrove, 2003). Many of these genes appear to be part of transcription factor networks and signaling pathways whose combined inputs control realizator action. The understanding of these networks is the key to understanding Hox function. Another key observation is that even though in some cases *Hox* genes are expressed throughout a segment, Hox targets are regulated in a subset of cells within the segment, indicating that interaction between the Hox protein and other proteins is necessary to provide the positional specificity (Hombria and Lovegrove, 2003). Also, mutations in certain genes lead to misregulation of Hox target genes even though the *Hox* gene is still expressed normally, implying that the mutant genes may encode cofactors. Furthermore, loss-of-function mutations in potential cofactors produce phenotypes similar to loss-of-function mutations in *Hox* genes (de Zulueta et al., 1994; Mahaffey et al., 2001; Peifer and Wieschaus,

1990). Often potential cofactors are referred to as “genetic” cofactors because they act in parallel with the Hox proteins to regulate target gene expression but no direct interaction between the potential cofactor and the Hox protein has been demonstrated.

Two such potential cofactors that have been identified in our lab are the genes *disconnected* (*disco*) and *disco-related* (*disco-r*) (referred together as the *disco* genes). The *disco* genes encode for C2H2 zinc finger transcription factors, and they are genetic cofactors for the *Hox* genes *Deformed* and *Sex combs reduced* as well as possibly others during embryonic *Drosophila* development (Mahaffey et al., 2001).

The objective of my research was to characterize how *disco* functions and also investigate the role *disco* plays in adult *Drosophila* limb development. One question that arose was what is the DNA binding site for Disco? It is known that Disco can bind to DNA, though, no DNA binding site has been identified (Lee et al., 1999). Protein-protein interactions are important for the function of almost any given protein, so another question was what protein-protein interactions are important for Disco function? In addition to Disco’s role in embryonic development, it also has a role in adult development. Evidence for a role in adult development comes from expression in imaginal discs (Lee et al., 1991) and the enhancer trap line C50.1S1 which has been used as a marker in several other limb development studies (Bellen et al., 1989; Bishop et al., 1999). However, there has been no description of what happens to adult development if the *disco* genes are mutated. Lastly, due to its redundancy with *disco-r*, *disco*’s role in embryonic pattern formation had been overlooked, and we wanted to ask are there other genes that might be involved in embryonic pattern formation that might have been

overlooked due to redundancy similar to *disco* and *disco-r*? Using *Drosophila melanogaster* as a model organism we tried to answer these questions.

The next chapter will give an overview of *Drosophila* embryonic development and the *Hox* genes focusing primarily on *Dfd* and *Scr* since they are the two primary genes with which the *disco* genes act as genetic cofactors. It will also focus on the *disco* genes as genetic cofactors and other cofactors including *extradenticle* and *teashirt*. The following chapter extends what is currently known about *disco* by describing the important role that the *disco* genes play in appendage development. The chapter after that examines the DNA binding properties of Disco and protein-protein interaction with C-terminal Binding Protein (CtBP). The subsequent chapter will describe a study that attempts to identify novel genes that may be involved in embryonic pattern formation which may have been overlooked due to redundancy. The concluding remarks will focus on summarizing the work, presenting possible ideas for how Disco functions, and discussing ideas for future work. By studying Disco and how it functions we hope to advance our understanding of how the *Hox* proteins and their cofactors function to specify segment identity.

## REFERENCES

- Bellen, H. J., O'Kane, C. J., Wilson, C., Grossniklaus, U., Pearson, R. K. and Gehring, W. J.** (1989). P-element-mediated enhancer detection: a versatile method to study development in *Drosophila*. *Genes Dev* **3**, 1288-300.
- Bishop, S. A., Klein, T., Arias, A. M. and Couso, J. P.** (1999). Composite signalling from Serrate and Delta establishes leg segments in *Drosophila* through Notch. *Development* **126**, 2993-3003.
- de Zulueta, P., Alexandre, E., Jacq, B. and Kerridge, S.** (1994). Homeotic complex and teashirt genes co-operate to establish trunk segmental identities in *Drosophila*. *Development* **120**, 2287-96.
- Denell, R. E., Hummels, K. R., Wakimoto, B. T. and Kaufman, T. C.** (1981). Developmental studies of lethality associated with the antennapedia gene complex in *Drosophila melanogaster*. *Dev Biol* **81**, 43-50.
- Ekker, S. C., Jackson, D. G., von Kessler, D. P., Sun, B. I., Young, K. E. and Beachy, P. A.** (1994). The degree of variation in DNA sequence recognition among four *Drosophila* homeotic proteins. *EMBO J* **13**, 3551-60.
- Garcia-Bellido, A.** (1977). Homeotic and atavic mutations in insects. *Amer. Zoo.* **17**, 613-629.
- Hombria, J. C. and Lovegrove, B.** (2003). Beyond homeosis--HOX function in morphogenesis and organogenesis. *Differentiation* **71**, 461-76.
- Lee, K. J., Freeman, M. and Steller, H.** (1991). Expression of the disconnected gene during development of *Drosophila melanogaster*. *EMBO J* **10**, 817-26.
- Lee, K. J., Mukhopadhyay, M., Pelka, P., Campos, A. R. and Steller, H.** (1999). Autoregulation of the *Drosophila* disconnected gene in the developing visual system. *Dev Biol* **214**, 385-98.
- Mahaffey, J. W., Griswold, C. M. and Cao, Q. M.** (2001). The *Drosophila* genes disconnected and disco-related are redundant with respect to larval head development and accumulation of mRNAs from deformed target genes. *Genetics* **157**, 225-36.
- McGinnis, W. and Krumlauf, R.** (1992). Homeobox genes and axial patterning. *Cell* **68**, 283-302.

**Peifer, M. and Wieschaus, E.** (1990). Mutations in the *Drosophila* gene extradenticle affect the way specific homeo domain proteins regulate segmental identity. *Genes Dev* **4**, 1209-23.

## CHAPTER TWO

# Drosophila Embryonic Development and the Role of HOM-C Proteins and Cofactors

## **Overview of *Drosophila* embryonic development**

*Drosophila* embryonic development begins after fertilization, whereupon the embryo undergoes seven synchronous nuclear divisions to form the syncytial blastoderm, at which point most of the nuclei start migrating towards the periphery of the egg (Campos-Ortega, 1997). Some of these nuclei will migrate to the posterior of the egg to eventually form the germ cells. After several more nuclear divisions, membrane furrows form around individual nuclei at the periphery of the egg to form the cellular blastoderm. Gastrulation, the process of forming germ layers (groups of cells that will form the ectoderm, mesoderm, and endoderm), begins with the invagination of the mesodermal primordia along the ventral midline of the embryo to create the ventral furrow, and at about this point the cephalic furrow is also created at about 67% egg length (0% being the posterior end of the embryo) (Campos-Ortega, 1997). The establishment of the cephalic furrow divides the early embryo into roughly two regions which consist of the procephalon and the germ band. The germ band will give rise to the metameric regions of the embryo and is defined as the primordia of the three gnathal segments (mandibular, maxillary, and labial), three thoracic segments (pro-, meso-, and metathorax, also referred to as t1-t3), and nine abdominal segments (a1-a9) (Campos-Ortega, 1997). The segments themselves can be further subdivided with each segment containing epidermal, neural, and mesodermal layers.

The germ band undergoes the process of elongation and retraction. Germ band elongation begins immediately after ventral furrow formation and coincides with parts of the gastrulation process. During germ band elongation, the cells within the germ band

divide and extend around the posterior end of the embryo such that the caudal end of the embryo is positioned dorsoanteriorly. The germ band is fully extended at about 5 hrs and 20 mins post fertilization and is extended to about 75% egg length (Campos-Ortega, 1997). Shortly afterwards, the gnathal segments are clearly visible in the ventro-lateral region of the head as are the parasegmental furrows within the thoracic and abdominal segments dividing the embryo into metameric units (Campos-Ortega, 1997). After the germ band has extended, the germ band will retract restoring the normal anatomical positions of the embryo, i.e. the caudal end of the embryo will be located at the posterior of the embryo rather than dorsoanteriorly, and germ band segmentation becomes very evident (Campos-Ortega, 1997).

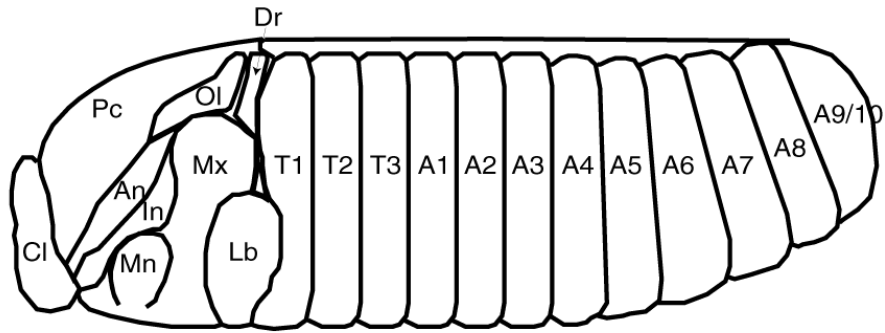
During germ band retraction, dorsal closure and head involution also take place. As the germ band retracts, the amnioserosa, an extraembryonic epithelial tissue, is exposed along the dorsal edge of the germ band. During dorsal closure, the dorsal epidermis on either side extends towards the amnioserosa and fuses at the dorsal midline which leads to the displacement of the amnioserosa into the embryo (Campos-Ortega, 1997). During head involution, several complex morphogenetic movements take place to form the atrium, the larval mouth. These include movement of the hypopharyngeal (pre-mandibular) lobes migrating into the stomodeum to form the floor of the pharynx, fusion of the labial lobes along the ventral midline to form the salivary duct and ventral border of the atrium, and also fusion of the mandibular and maxillary segments from which some cells will form the lateral wall of the atrium (Campos-Ortega, 1997; Chadwick and McGinnis, 1987). After dorsal closure and head involution, morphogenesis is essentially

complete and the epidermis will secrete cuticle to form the *Drosophila* larva.

### **Establishment of the metameric *Drosophila* embryo**

Segmentation is the process of dividing the body into repeated units, and the segments can be similar to each other or different. As was mentioned above, the initial division of the embryo at germ band extended stage is described as parasegmental. Parasegments are the same size as segments, but are out of register by one compartment. The segments consist of anterior and posterior compartments, thus the posterior of a segment is the anterior of a parasegment (Fig. 1) (Martinez-Arias and Lawrence, 1985). The usefulness of parasegments becomes evident when describing the expression and function of many regulatory genes, whose expression is often not segmental but parasegmental or transitions from parasegmental to segmental (Martinez-Arias et al., 1987; Martinez-Arias and Lawrence, 1985).

Formation of the segments of the *Drosophila* embryo involves genes that can be divided into four classes: the maternal, gap, pair-rule and segment polarity genes (Anderson and Nusslein-Volhard, 1984; Nusslein-Volhard et al., 1987; Nusslein-Volhard and Wieschaus, 1980). These genes act in a hierarchical manner to establish the segmented *Drosophila* embryo (Fig. 2). The maternal genes are at the top of the segmentation hierarchy and they will establish the anterior-posterior axis of the embryo. The maternal genes encode mRNAs and proteins that are loaded into the egg before fertilization. Two examples of maternal patterning genes are *bicoid* and *nanos*. *bicoid* is necessary for patterning the anterior end of the embryo and *nanos* is required for patterning the posterior (Berleth et al., 1988; Hulskamp et al., 1989).



Segment	Mn		Mx		Lb		T1	T2	T3	A1	A2	A3	A4	A5	A6	A7	A8	A9		
Compartments	A	P	A	P	A	P	A	P	A	P	A	P	A	P	A	P	A	P	A	P
Parasegments	0		1		2		3	4	5	6	7	8	9	10	11	12	13	14		

This figure was reprinted from Robertson, L.K. and Mahaffey, J.W. (2005). Insect Homeobox Genes and Development-Lessons from *Drosophila* and Beyond. In *Comprehensive Insect Science, Volume One--Reproduction and Development*, (eds. L.I. Gilbert, K. Iatrou, and S. Gill), pp. 247-303, Elsevier Limited, London, UK.

**Figure1. Division of the *Drosophila* embryo into Segments, Parasegments, and Compartments.**

Parasegments are the same width as segments but out of register by one compartment. Abbreviations: Cl-clypeolabrum, Pc-procephalic, An-antennal, In-intercalary, Ol-optic lobe, Dr-dorsal ridge, Mn-mandibular, Mx-maxillary, Lb-labial, T#-thoracic, A#-abdominal

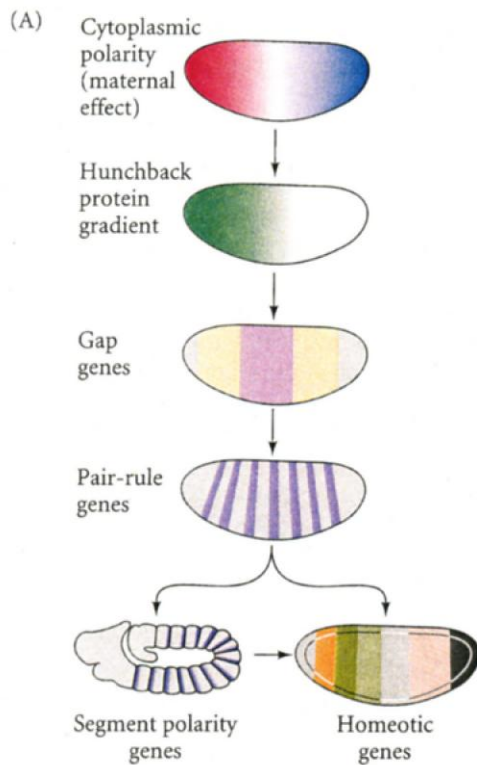


Figure reprinted from Gilbert, SF. 2003. *Developmental Biology*. Sunderland, Mass.: Sinauer Associates.

**Figure 2. The Segmentation Hierarchy.**

General diagram of the *Drosophila* segmentation hierarchy. (A) The segmentation hierarchy begins with the maternal effect genes which help establish the anterior-posterior axis. The maternal effect genes establish the gap genes which define broad regions of the embryo. The gap genes regulate the pair rule genes which subdivide the embryo into segment wide units. The action of these genes leads to the establishment of the segment polarity genes. Lastly the *homeotic* genes will give identity to each segment.

The gap genes are next in the segmentation hierarchy and are the targets of the anterior-posterior patterning system. Gap genes define broad regions of segmentation. In gap mutants a group of adjacent segments is missing (Knipple et al., 1985; Nusslein-Volhard and Wieschaus, 1980). Gap genes are expressed along the anterior-posterior axis in discrete domains, and their expression is limited by maternal factors and other gap genes.

The gap gene products assist in establishing the expression of the pair-rule genes. The pair-rule genes are expressed in seven stripes along the anterior-posterior axis and divide the embryo into segment wide divisions (DiNardo et al., 1985; Lawrence et al., 1987; Macdonald et al., 1986). Proper expression of the pair-rule genes relies not only on the gap gene products but also on the products of pair-rule genes themselves (Carroll and Scott, 1986; Frasch and Levine, 1987). Mutations in the pair-rule genes lead to loss of alternating segments (Nusslein-Volhard and Wieschaus, 1980). The combined expression of the pair rule genes *fushi tarazu* and *even skipped* define the initial fourteen parasegments (Lawrence et al., 1987).

The segment polarity genes are at the bottom of the segmentation hierarchy. Mutations in segment polarity genes affect anterior-posterior polarity within each segment. Segment polarity mutants have the normal number of segments but each segment has part of its normal pattern deleted and replaced by a mirror image duplication of the remaining pattern (Nusslein-Volhard and Wieschaus, 1980). Expression of the segment polarity genes is regulated by the pair-rule genes, which, like the pair-rule and gap genes, are regulated by each other. The expression of the gap, pair-rule, and

segment polarity genes help define the expression of the Homeotic genes, which will specify the identity of the gnathal and trunk segments.

### **The Homeotic genes and specification of Identity along the AP axis**

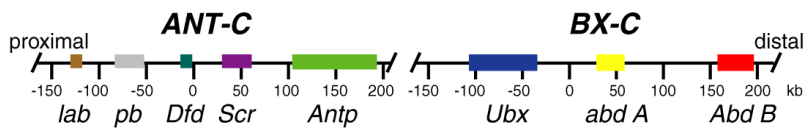
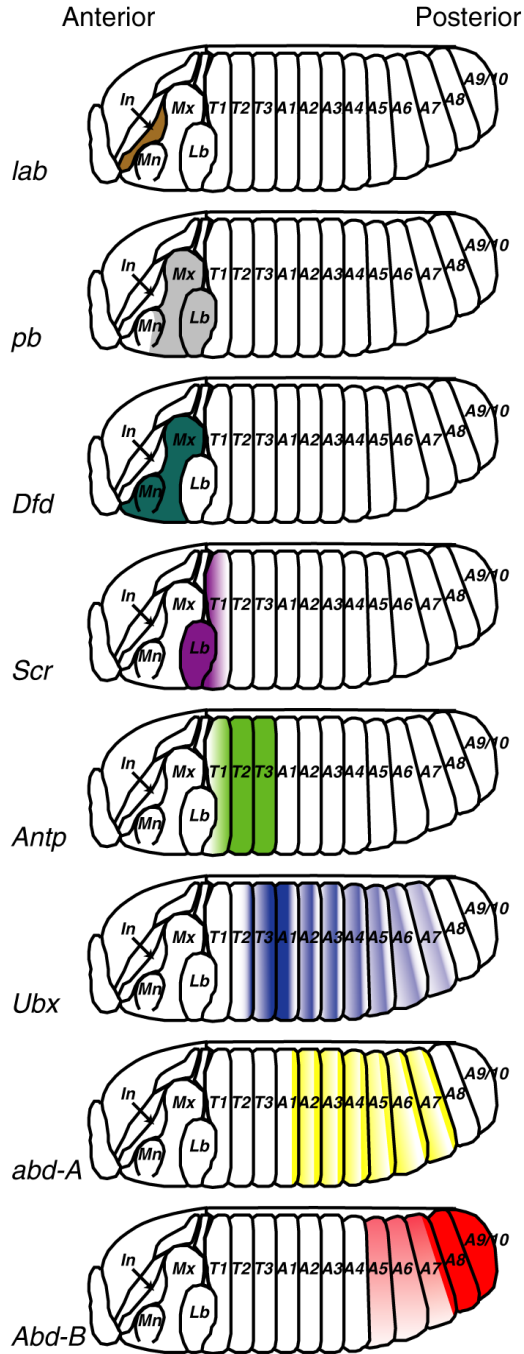
The *Homeotic* genes were first identified in *Drosophila*, but they are now known to be in many if not all animals, and they are now generally referred to as the *Hox* genes. In *Drosophila* the *Homeotic* genes are collectively referred to as the *Homeotic Complex (HOM-C)* genes. The eight fly *HOM-C/Hox* genes are *labial (lab)*, *proboscipedia (pb)*, *Deformed (Dfd)*, *Sex combs reduced (Scr)*, *Antennapedia (Antp)*, *Ultrabithorax (Ubx)*, *abdominal-A (abd-A)*, and *Abdominal-B (Abd-B)*. The genes are actually broken in two complexes located on the right arm of the third chromosome separated by approximately 7.5 Mb. The two complexes are called the Antennapedia complex (which contains the genes *lab*, *pb*, *Dfd*, *Scr*, and *Antp*) and the Bithorax complex (which contains the genes *Ubx*, *abd-A*, and *Abd-B*). *lab*, *pb*, and *Dfd* are expressed in the head segments while *Antp*, *Ubx*, *abd-A*, and *Abd-B* are expressed within the trunk. *Scr* is unique in that it is the only *Hox* gene that is expressed in the head and the trunk. *Hox* genes are expressed in a manner from anterior to posterior such that expression is colinear with their arrangement on the chromosome in most animals (Kaufman et al., 1990). This is referred to as colinearity. Expression of the *HOM-C* genes compared to their location on the chromosome in *Drosophila* along the anterior posterior axis of the embryo can be seen in Figure 3.

Some *Hox* genes are expressed in overlapping domains, while others are mutually exclusive. In general most *Hox* genes interact with each other to maintain their proper

**Figure 3. Expression the Drosophila HOM-C genes along the anterior-posterior axis.**

Expression of the HOM-C genes along the anterior-posterior axis corresponds to their arrangement along the right arm of the third chromosome. This is referred to as colinearity. Abbreviations: *lab*-labial, *pb*-proboscipedia, *Dfd*-Deformed, *Scr*-Sex combs reduced, *Antp*-Antennapedia, *Ubx*-Ultrabithorax, *abd-A*-abdominal-A, *Abd-B*-Abdominal-B.

This figure was reprinted from Robertson, L.K. and Mahaffey, J.W. (2005). Insect Hom-C Genes and Development-Lessons from Drosophila and Beyond. In Comprehensive Insect Science, Volume One--Reproduction and Development, (eds. L.I. Gilbert, K. Iatrou, and S. Gill), pp. 247-303, Elsevier Limited, London, UK.



domains of expression and often the more posterior gene will suppress the expression or function of the more anterior gene, an idea referred to as posterior prevalence (Harding et al., 1985; Struhl and White, 1985). In some cases loss of *Hox* gene expression in a posterior segment leads to expression of the more anterior *Hox* gene within that segment, which in turn changes the segment identity to the more anterior segment.

The *Hox* genes encode homeodomain-containing proteins. The homeodomain is a highly conserved 60 amino acid motif with three alpha helices of which residues 20-50 form a helix-turn-helix motif that is important for DNA recognition (Frazee et al., 2002; Qian et al., 1989). In addition to the three helices, homeoproteins also contain a conserved amino terminal flanking sequence. The Hox proteins are considered master transcriptional regulators that specify segmental identity by regulating specific target genes. Of particular importance to the studies of *disco* and *disco-r* are the *Hox* genes *Dfd* and *Scr*, thus their expression and mutant phenotypes will be examined more closely. Mutants for *disco* and *disco-r* are missing mouth structures that are specified by *Dfd* and *Scr*.

### **Expression of *Dfd* and mutant phenotypes**

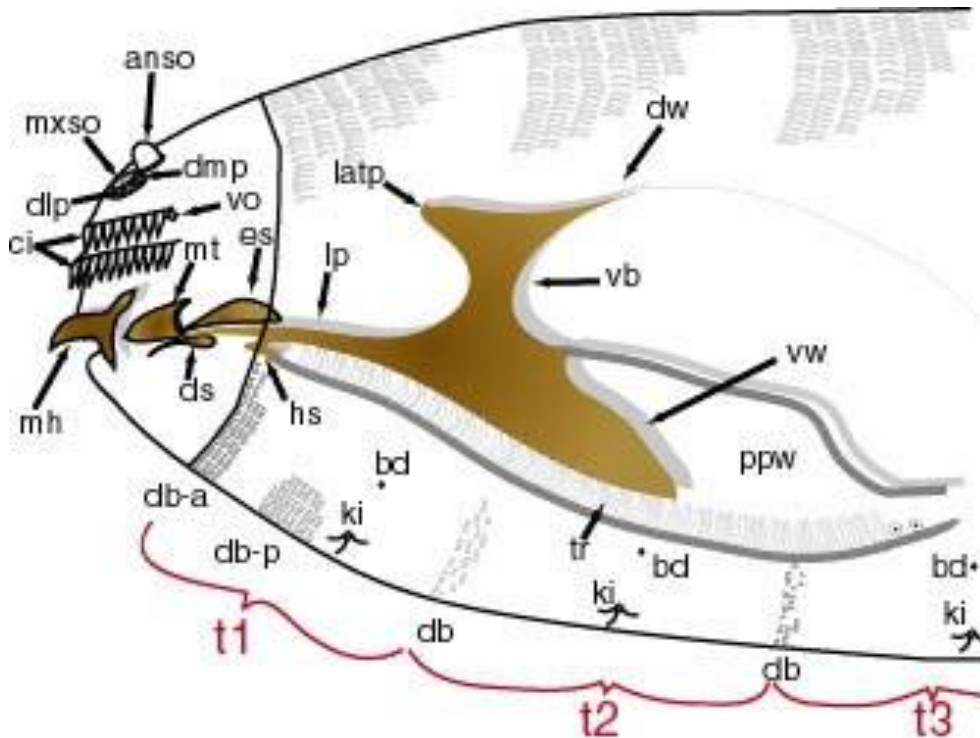
*Dfd* is the first *Hox* gene to be expressed during embryogenesis with transcripts detected around 65% egg length on the dorsal side and 70% on the ventral side, and as gastrulation occurs it is expressed in a circumferential stripe that encompasses the cephalic furrow. During germ band extended stage, transcripts are detected in the ectoderm of the mandibular and maxillary lobes and transiently in the hypopharyngeal lobes. When the germ band is retracted expression is restricted to the mandibular lobe

and throughout the maxillary lobe except for cells within the antero-lateral region and cells between the optic lobe and the dorsal ridge (Chadwick and McGinnis, 1987; Mahaffey et al., 1989).

Loss-of-function *Dfd* mutants have defects in larval head development. In particular, null mutations lead to defects in the cephalopharyngeal skeleton and some sense organs. The cephalopharyngeal skeleton consists of many cuticular structures including the mouth hooks, the median tooth, the H-piece, and the cephalopharyngeal plates (Fig. 4). In *Dfd* embryos, the mouth hooks and H-piece are missing (Merrill et al., 1987; Regulski et al., 1987). Also, in *Dfd* embryos the labial lobes do not migrate to the midline but do eventually fuse. Likewise, the maxillary lobes fail to fuse with either the mandibular or procephalic lobes, and they fail to internalize during head involution. The maxillary sense organ forms missing some sensory papillae but is separated from the antennal sense organ, and the maxillary cirri do not form (Merrill et al., 1987; Regulski et al., 1987).

### **Expression of *Scr* and mutant phenotypes**

Expression of *Scr* transcripts is first seen during the blastoderm stage as a band three to four cells wide in the dorso-lateral region adjacent but posterior to the expression of *Dfd*. Once the germ band is completely extended, expression can be seen in both the labial segment and the first thoracic segment. In the labial segment, *Scr* is expressed in the ectoderm and in the first thoracic segment expression is limited to the mesodermal layer and the anterior portion of the ectoderm. As the germ band retracts, expression within the labial segment remains the same, but expression in the first thoracic segment



This figure was reprinted from Robertson, L.K. and Mahaffey, J.W. (2005). Insect Hom-C Genes and Development-Lessons from *Drosophila* and Beyond. In *Comprehensive Insect Science, Volume One--Reproduction and Development*, (eds. L.I. Gilbert, K. Iatrou, and S. Gill), pp. 247-303, Elsevier Limited, London, UK.

#### Figure 4. The *Drosophila* Larval Head.

The *Drosophila* larval head contains many identifiable structures. The abbreviations for these structures and which segments they arise from are as follows: the acron produces the *dw*-dorsal wing, *latp*-latticed piece (dorsal bridge), *vb*-vertical bridge, and *es*-epipharyngeal sclerite; the labral produces the *mt*-median tooth and *anso*-antennal sense organ; the antennal produces the *ppw*-posterior pharyngeal wall, the *dlp*-dorsal-lateral papillae of the mxso; the intercalary produces the *lp*-lateral process; the mandibular produces the *vw*-ventral wing, *tr*-T-ribs, *mh*-mouth hooks, *ci*-cirri; the *mxso*-maxillary sense organ is produced by the mandibular and maxillary segments; the maxillary segment produces the *dmp*-dorsal-medial papillae of the mxso, the *vo*-ventral organ, *ds*-dental sclerite, the *hs*-hyostomal sclerite (H-piece); the labial segment produces the labial sense organ (not shown) and part of the H-piece; the thoracic segments produce *db*-denticle belts, *ki*-Keilin's organ, *bd*-ventral kolbchen (black dot organs); the first thoracic segment (t1) produces two denticle belts *db-a*-denticle belt anterior and *db-p*-denticle belt posterior. Of particular importance to these studies are the structures produced by the mandibular, maxillary, labial, and T1 segments whose identities are controlled by the HOM-C genes *Dfd* and *Scr*.

spreads through out most of the ectodermal layer. Scr protein distribution is similar to RNA transcript distribution except it seems that expression is delayed by 2-3 hours (Mahaffey and Kaufman, 1987; Martinez-Arias et al., 1987). Scr can also be detected in the midgut visceral mesoderm and salivary gland primordia (Reuter and Scott, 1990).

Loss of Scr leads to a transformation of the first thoracic segment to second thoracic identity. In the first thoracic segment, the beard denticles are lost and the denticles that remain resemble those of the second thoracic segment (Mahaffey and Kaufman, 1987; Pattatucci et al., 1991). In addition to transformation of the first thoracic segment, the labial lobes fail to fuse and enter the stomodeum leading to the sensory organs produced by the labial segments being external (Pederson et al., 1996). Loss of Scr also leads to loss of structures such as the bridge of the H-piece and also the salivary glands (Panzer et al., 1992).

### **The Homeodomain and cofactors**

As was mentioned previously, all *HOM-C/Hox* genes encode proteins containing a highly conserved 60 amino acid motif called the homeodomain that is capable of binding DNA and is found in many other transcription factors (Gehring et al., 1994). A model for how Hox proteins control cell fate is that they regulate the transcription of different sets of target genes; however, *in vitro* most HOM-C proteins bind very similar sites with similar affinities (Ekker et al., 1994). Thus, how HOM-C proteins regulate specific target genes *in vivo* still remains an important unanswered question.

The three dimensional structure of the homeodomain complexed with DNA has been resolved. The 3D structure indicates that the homeodomain consist of a three

helical regions where helix I is preceded by a flexible N-terminal arm and followed by a loop that separates it from helix II, and helix II and III together form a DNA binding helix-turn-helix motif (Gehring et al., 1994). The DNA contacts are mediated by the third helix which sits in the major groove. Overall the 'core' DNA sequence recognized by HOM-C proteins is 5' T<sub>1</sub>N<sub>2</sub>A<sub>3</sub>T<sub>4</sub>(G/T)<sub>5</sub>(G/A)<sub>6</sub> 3' where N is variable (Ekker et al., 1994). The residues 47, 50, 51, and 54 of the third helix contact positions 3-6 of the recognition sequence. Positions 1 and 2 of the sequence are recognized by residues 3 and 5 of the N-terminal arm. Interestingly, nearly all the residues found at these positions are conserved. All HOM-C homeodomains at position 50 have a glutamine and at position 54 have a methionine. Nearly all homeodomains have an asparagine at position 51 and seven of the eight homeodomains have an isoleucine at position 47 (Mann, 1995). Thus there is not enough variation at these positions to provide the specificity required to regulate different target genes.

Experiments have demonstrated that the N-terminal arm and the C-terminal tail also have roles in specificity (Chan and Mann, 1993; Gibson et al., 1990; Lin and McGinnis, 1992; Zeng et al., 1993). Of particular importance is a motif N-terminal to the homeodomain often called the YPWM motif. This particular motif is important because of its interaction with the Hox cofactor Extradenticle, which is discussed below. So how can one resolve how HOM-C proteins direct complex target gene regulation with such low DNA binding specificity? One possibility is that these slight DNA binding specificities *in vitro* actually have quite dramatic effects *in vivo*. However, such drastic differences in specificity which are seen *in vivo* more likely arise through interactions

with cofactors. In fact, no known Hox target gene's entire expression is regulated by a Hox protein alone (Hombria and Lovegrove, 2003).

There are two proposed models for cofactor interaction; the co-selective binding model and the widespread binding model (Biggin and McGinnis, 1997). In the co-selective model, Hox proteins require a cofactor to direct them to the target regulatory site. In the widespread binding model Hox proteins are able to bind to many sites but are able to regulate transcription only if the cofactor is present nearby. It is likely that *in vivo* a combination of these two models is occurring.

In order to be considered a potential Hox cofactor, certain requirements have to be met. First, null mutations of the cofactor should reproduce some aspect of the phenotype of the *Hox* mutant. Second, absence of the cofactor would lead to alteration or absence of Hox target gene expression. Third, the cofactor would act in parallel with the Hox protein, i.e., the cofactor is neither a target of the Hox protein nor regulates the Hox protein (Peifer and Wieschaus, 1990). Potential cofactors that meet these requirements include *extradenticle (exd)*, *cap n' collar (cnc)*, *apontic (apt)*, *lines (lin)*, *buttonhead (btd)*, *teashirt (tsh)*, *disconnected (disco)*, and *disco-related (disco-r)*. Some of these are known to be direct acting cofactors, capable of interacting with several HOM-C proteins, (such as Exd); however, others work specifically with a few HOM-C proteins (such as Tsh) (de Zulueta et al., 1994; Mann and Chan, 1996; Rauskolb and Wieschaus, 1994; Taghli-Lamalle et al., 2007). The genes *disconnected (disco)* and *disco-related (disco-r)*, which are the focus of my research, are proposed to encode potential cofactors for the HOM-C proteins Deformed and Sex combs reduced.

## **Collaborative interactions vs. Cooperative interactions**

The Hox-cofactor models presented above imply a direct interaction between the Hox protein and the proposed cofactor, often termed a cooperative interaction. However, in some instances there may not be a direct interaction between two proteins, but both are necessary to properly regulate transcription. This type of situation is taken into consideration in collaborative interactions (Walsh and Carroll, 2007). In this model Hox proteins do not have to cooperatively bind with cofactors, but it is the identity of the collaborative protein and/or topology of cis-regulatory elements that plays a critical role in gene regulation.

Regulation of *spalt (sal)* in the wing and haltere provides a good example of collaborative interactions. In the wing two Smad proteins, Mad and Medea, activate *sal* expression, but in the haltere Mad, Medea, and the Hox protein Ubx work together to repress *sal* expression. There is no direct interaction between Mad and Medea with Ubx, suggesting that some other factors might influence the repression of *sal* expression (Walsh and Carroll, 2007). It was shown that the spacing between the Ubx binding site and the Mad/Medea binding site played a key role in the regulation of *sal* expression. If the spacing of these sites is altered the proteins no longer collaborate to repress *sal* expression. This is an example of lack of requirement for direct interaction to regulate gene expression.

## **Extradenticle as a cofactor**

Out of all the proposed cofactors for HOM-C proteins, Exd has been the best characterized. Exd, itself, is a homeodomain protein that shares 71% identity with the

human proto-oncoprotein Pbx1 (Rauskolb and Wieschaus, 1994). *exd* is expressed throughout development, and mutations in *Exd* cause homeotic transformations without changing the expression of the *HOM-C* genes themselves (Rauskolb et al., 1993). Likewise, it appears that the *HOM-C* genes do not regulate the expression of *exd*, indicating that *exd* and the *Homeotic* genes work in parallel pathways to determine segment identity (Rauskolb et al., 1993). Several experiments have demonstrated the requirement of *Exd* to properly regulate known targets of several *Homeotic* genes. These include regulation of parasegmental stripe 8 of *wingless* (*wg*), which is controlled by the *Homeotic* gene *abd-A*; expression of *teashirt* in parasegmental stripes 5/6 and 8, which is influenced by *Antp*; and expression of *decapentaplegic* (*dpp*) in the embryonic midgut visceral mesoderm of parasegments 7 and 8, which are regulated by *Ubx* and *Abd-A* (Capovilla et al., 1994; Immergluck et al., 1990; Rauskolb et al., 1993). Expression patterns of *wg*, *tsh*, and *dpp* are all altered in the embryonic midgut in *exd* mutants. *Exd*'s role as a cofactor can also be seen in experiments dealing with the regulation of *fork head* (*fkh*) and *labial* (*lab*). In the following paragraphs, *Exd* requirement for *dpp*, *fkh*, and *lab* expression will be discussed in more detail.

*dpp*, a TGF $\beta$  family member, is initially expressed in the dorsal part of the embryo during the syncytial blastoderm; as the germ band elongates it is expressed in the region that will give rise to the dorsal epidermis, and even later during embryogenesis expression in the ectoderm changes into two stripes running along the anterior-posterior axis of the embryo (Padgett et al., 1987). Expression is also detected in the visceral mesoderm (St Johnston and Gelbart, 1987). The expression in the visceral mesoderm is

particularly important here because it is dependent on Ubx and Abd-A, and requires Exd. In the visceral mesoderm *dpp* is positively regulated by Ubx in parasegment 7, while Abd-A negatively regulates *dpp* in more posterior parasegments (Immergluck et al., 1990; Reuter and Scott, 1990). In *exd* mutants, ectopic *dpp* is seen in the visceral mesoderm anterior to parasegment 7, while some expression is lost in parasegment 7. Several  $\beta$ -galactosidase reporter constructs have been made that mimic the expression of *dpp* in the visceral mesoderm. These include the *dpp4000lacZ*, *dpp674lacZ*, and *dpp303lacZ* reporters. Each one is a subfragment of the larger fragment, i.e. *dpp303* is a subfragment of the *dpp674* enhancer. Expression of the *dpp303lacZ* is first seen in parasegment 7 at germ band extended stages, and in *exd* mutants some Ubx expressing cells within parasegment 7 are incapable of activating expression of the *dpp303lacZ* reporter (Rauskolb and Wieschaus, 1994). This evidence pointed towards a possible interaction between Exd and Ubx. Chan et al. (1994) went on to demonstrate that Exd interaction with Ubx increased Ubx binding to the *dpp80* enhancer (a subfragment of the *dpp303* enhancer) 6-30 fold (Chan et al., 1994). Furthermore this interaction is dependent on residues within the Ubx homeodomain and the C terminal tail (Chan et al., 1994; Johnson et al., 1995). However, there have been conflicting results from other investigators in regards to the findings by Chan et al. (1994).

Manak et al. (1994) also worked with the *dpp* enhancer region but they were not able to demonstrate Exd binding to the *dpp80* enhancer region. Manak et al. (1994) did show a footprint for this region, but binding by Exd. Although Manak et al. (1994) were not able to show that Exd binding to *dpp80*, they did demonstrate that another region of

the *dpp* enhancer did contain a binding site that was involved in regulating *dpp* expression in gastric caeca and PS7. van Dijk and Murre (1994) demonstrated that Exd interaction with Ubx required a region N-terminal to the Exd homeodomain, however, this region is absent from the Exd protein used in the experiments done by Chan et al. (1994). Although there are discrepancies in these experiments there have been other experiments that demonstrate Exd's role as a cofactor.

Exd's role as a Hox cofactor was further demonstrated from studies of the *Hox* gene, *labial*. *labial* is expressed in the neural and epidermal cells of the intercalary segment, a discrete loop within the midgut, and in progenitor sensory cells within the clypeolabrum, thoracic segments, and tail (Diederich et al., 1989). *Hoxb-1* is the vertebrate ortholog of *labial* whose autoregulatory elements have been well characterized. *Hoxb-1* has three related DNA elements (termed repeats 1-3) that are necessary and sufficient for the autoregulation of *Hoxb-1* in rhombomere 4 of the mouse hindbrain, and this autoregulation is dependent on Exd (Popperl et al., 1995). Chan et al. (1996) were able to reproduce a similar endogenous *lab* expression pattern by placing 3 copies of repeat 3 from *Hoxb-1* in a *lacZ* reporter in *Drosophila* called *3Xrpt3-lacZ*. However, there was expression in the gastric caeca, pouches that secrete digestive enzymes, and several other discrepancies that are not present in normal *lab* expression. In *lab<sup>-</sup>* embryos, expression of the *3Xrpt3-lacZ* reporter in the head and midgut was absent, but expression in the gastric caeca remained. In *exd<sup>-</sup>* embryos, *lacZ* expression was also absent indicating that Lab and Exd combined to regulate *labial* expression. Further experiments demonstrated that the *3Xrpt3* is weakly bound by Exd and not at all

by Lab; however, cooperatively they can both bind. Deletion of the YPWM motif greatly reduced cooperative binding between the two. Interestingly, this deletion allowed Exd and Lab to bind to *3Xrpt3* independently. Thus it seems that, in this instance, the YPWM motif inhibits binding of Lab to DNA, but Exd interaction with this motif would allow binding (Chan et al., 1996).

An interaction of Exd with the YPWM motif has been shown with several other Hox proteins (Chang et al., 1995; Johnson et al., 1995). In some cases though, such as with Ubx, the YPWM motif is important for the interaction with Exd but is not sufficient for this interaction (Johnson et al., 1995). In fact, Johnson et al. (1995) demonstrated that the Ubx homedomain and the C-terminal tail were important for interaction with Exd supporting the results from Chan et al. (1994). Recently a study has even shown that the YPWM motif in Abd-A is not necessary for Exd recruitment, DNA binding or target gene selection, but, is required for proper regulation of some Abd-A target genes, perhaps suggesting Exd independent roles for the YPWM motif (Merabet et al., 2003).

Another example of Exd acting as a cofactor with a Hox protein comes from *fork head* (*fkh*) enhancer studies. *fkh* is expressed throughout most of the gut and in the salivary gland primordium (Weigel et al., 1989). Expression of *fkh* in the salivary gland placode of the labial segment in particular is regulated by the Hox protein Sex combs reduced (*Scr*). Expression in the salivary gland placode has been shown to be controlled by several elements 10kb upstream from the coding sequence (Zhou et al., 2001). Ryoo et al. (1999) were able to identify a 37bp element called *fkh*[250] which, when placed in front of a *lacZ* reporter, can drive expression in parasegment 2, similar to endogenous *fkh*

expression (Ryoo and Mann, 1999). The *fkh*[250] element contains a binding site similar to the canonical Hox/Exd site 5'-TGATNNATNN-3' (Chan et al., 1997). Exd is not capable of binding this fragment on its own, and even though Scr, Antp, Ubx, and Abd-A all have similar homeodomains, only Scr makes a stable DNA bound heterodimer with Exd. Likewise, only ectopic expression of Scr was able to expand expression of *lacZ* anteriorly and posteriorly, similar to endogenous *fkh* expression when Scr is ectopically expressed (Ryoo and Mann, 1999). Thus this element cannot be activated by other highly similar Hox proteins.

Two models could explain how Exd acts as a cofactor with the Hox proteins (Ryoo and Mann, 1999). The first is that Exd enhances DNA binding specificity of the Hox proteins *in vivo*. The other possibility is that Exd changes the regulatory properties of the Hox proteins. Evidence for the first model comes mostly from *in vitro* studies. Since the Hox/Exd heterodimer binds to a 10bp binding site rather than a 6bp recognition site by Hox alone, these sites would occur less frequently throughout the genome, increasing specificity (Mann and Chan, 1996). Along similar lines, Hox proteins that normally bind similar sites tend to bind different sequences as Hox/Exd heterodimers (Chan et al., 1994; Mann and Chan, 1996). Thus Exd may influence the binding of one Hox protein over another at a particular binding site. This is the case with Scr at the *fkh* enhancer. Another example of this can be seen when the two central base pairs of the Hox/Exd site 5'-TGATNNATNN-3' are either GG or TA. Repeat 3 of *Hoxb1* contains the Hox/Exd site and in the *3Xrpt3* reporter, if the central base pairs of the site, GG, are switched to TA a change in reporter expression from *labial* to *Dfd* in

*vivo* results. *In vitro*, Lab/Exd complexes are still able to bind to the site with TA, but do so 3-8 fold less efficiently than Dfd/Exd heterodimers (Chan et al., 1997). Evidence for the second model comes from experiments comparing the activity of a Dfd monomer binding site to a Dfd/Exd binding site (Li et al., 1999). Alone, Dfd is capable of binding the Dfd monomer binding site, but it is not sufficient for transcriptional activation of reporter genes. In comparison, when Dfd/Exd complexes bind to Dfd/Exd binding sites, transcriptional activation occurs. Dfd binding affinity is the same for the Dfd monomer binding site and the Dfd/Exd composite site, indicating that in this particular case Exd is required to control the transcriptional activity of Dfd and not its binding specificity. Although there is more evidence supporting the first model, Exd function is probably a combination of the two and is probably very context specific.

### **Teashirt as a cofactor**

*teashirt* was first identified in a *p*-element screen for embryonic segmentation defects (Fasano et al., 1991). The gene is located on the second chromosome and encodes for a C2H2 zinc finger transcription factor that is initially expressed in a central ring of cells spanning 45% to 60% egg length. By the end of cellularization a transient stripe pattern appears, and at germ band extension expression is in the ectodermal cells of parasegments 3-13. This expression continues through later stages of embryogenesis where expression is seen from anterior prothoracic to anterior eighth abdominal segments. Expression can also be seen in the CNS of the trunk and around the second midgut constriction of the visceral mesoderm. Homozygous null mutants of *tsh* are lethal. Null embryos are reported to have 1) homeotic transformations of the ventral trunk

segments towards head-like segments, which are characterized by disorganized denticle belts that are reduced in size and number 2) dorsal and ventral portions of the prothorax are partially deleted 3) sclerotization of the anal opening, and 4) the presence of sclerotized material within the trunk, which resembles the sclerotized material found within the cephalopharyngeal skeleton. Defects are also seen in the neuronal clusters where the ventral group of neurons in the trunk are deleted or disrupted, and axons are also misrouted towards the head.

Although the phenotypes seen in mutants corresponds well with the expression pattern, one might expect more severe defects given the expression pattern. The reason that more severe phenotypes are not observed, is that *tsh* is partially redundant with the gene *tiptop* (*tio*) (Laugier et al., 2005). *tiptop* encodes a C2H2 zinc finger protein similar to *tsh* and has common and distinct domains of expression from *tsh*. Unlike *tsh*, *tio* mutants are viable and fertile. However, loss of *tio* and *tsh* produces a stronger phenotype than loss of *tsh* alone. In particular the ventral denticles are reduced and poorly differentiated, more so than in *tsh* mutants alone.

Roder et al. (1992) concluded that *tsh* works at the same hierarchical level as the Hox proteins, and work by de Zulueta et al. (1994) suggest that Tsh and HOM-C proteins regulate common sets of downstream target genes. Also, it has recently been demonstrated that Tsh can interact with the Hox proteins Scr, Antp, and Ubx (Taghli-Lamalle et al., 2007). *tsh* expression overlaps the expression of *Scr*, *Antp*, *Ubx*, *abd-A*, and *Abd-B*. Looking at different mutant combinations Roder et al. (1992) determined that *tsh* is not required for initiation of transcription of these genes, however, expression

of *Scr* is altered in *tsh* mutants. Similarly, loss of *Scr*, *Antp*, *Ubx*, *abd-A*, and *Abd-b* has no effect on the early expression of *tsh*, but during germ band retraction *tsh* is no longer detected in the posterior compartment of each trunk segment and is more strongly expressed in the anterior compartment. Based on these experiments Roder et al. (1992) concluded that *tsh* and the trunk *HOM-C* genes work together to specify trunk development while repressing head development.

*tsh*'s role as a potential *HOM-C* cofactor during development is probably best observed with the *HOM-C* gene *Scr*. *Scr* is expressed in and controls the identity of the labial and prothoracic segments, and in the prothoracic (T1) segment activity of *Scr* seems to rely on *Tsh*. When *Tsh* is absent, *Scr* cannot promote prothoracic identity and instead, represses trunk (denticle belt) development transforming the prothoracic segment into a labial-like identity. These results are supported by ectopic expression studies, which demonstrated that if *Scr* is ectopically expressed and *Tsh* is present, all thoracic segments have prothoracic identity (Gibson et al., 1990). Likewise, if *tsh* is ectopically expressed, the labial segment is transformed into prothoracic identity (Manfroid et al., 2004).

*tsh* has always been considered a genetic cofactor, but no one has shown a direct interaction with any of the Hox proteins until recently. It is now known that Teashirt can interact directly with *Scr* and *Antp*, and the interaction with *Scr* occurs through an acidic domain at N-terminal half of Teashirt (Taghli-Lamalle et al., 2007). If the acidic domain of Teashirt is removed, ectopic expression of Teashirt will no longer lead to a transformation of the labial segment to prothoracic. If the acidic domain mutant is

ectopically expressed in a *tsh*<sup>8</sup> background, all trunk defects are rescued except for prothoracic defects. This implies that the acidic domain is particularly important for prothoracic identity. Most importantly the experiments done by Taghli-Lamalle et al. (2007) show that Tsh helps Scr establish differences in gene expression along the anterior-posterior axis. This is demonstrated by the regulation of the gene *modulo* (*mod*).

*modulo* is a direct target of Tsh and has been a proposed target of Scr, because in *Scr* mutants no *mod* expression is seen in the labial segment (Alexandre et al., 1996). *Scr* positively regulates *mod* expression in the labial segment, and Tsh represses *mod* expression in the prothoracic segment. In *tsh*<sup>8</sup> mutants, both *Scr* and *mod* expression are observed in the ventral prothoracic segment. Ectopic expression of Tsh represses *mod* expression in the labial segment. The *mod-1050* regulatory fragment has been shown to reproduce a similar pattern of expression to *mod* when placed in front of a *lacZ* reporter (Alexandre et al., 1996). A subfragment of this regulatory region, *mod84* contains two binding sites for Tsh and also binding sites for Scr. Taghli-lamalle et al. (2007) were able to identify two Scr binding sites within the *mod84* fragment which are critical to *mod* expression. The presence of both the Tsh and Scr binding sites are necessary for expression of the *mod1050-lacZ* reporter gene *in vivo* (Taghli-Lamalle et al., 2007). Moreover, electromobility supershift assays showed that Tsh and Scr form complexes with the DNA on this regulatory element. Similar to endogenous *mod* expression, the *mod1050-lacZ* reporter is no longer repressed in the labial segment by the ectopic expression of the Tsh acidic domain mutant. However, the Tsh acidic domain mutant is still capable of binding to the *mod84* regulatory fragment as is Scr, suggesting that a

cooperative interaction between Tsh and Scr is required to properly regulate *mod* expression. Interestingly, the Tsh acidic domain contains a binding site for the corepressor C-terminal Binding Protein (CtBP) (see following chapters) and this site is necessary for interaction with CtBP (Manfroid et al., 2004). This information begs the question of what exactly is the importance of the interaction between Tsh and Scr. Does Tsh interact with Scr and change its regulatory effect on *mod* expression or is this repression achieved solely through Tsh interaction with CtBP?

There is still much work that needs to be done with Tsh to determine how Tsh along with the Hox proteins regulates target genes, but it appears that Tsh may function in a different manner than Exd. Exd/Hox proteins tend to bind to a composite site, where one half of the binding site binds Exd and the other the Hox protein. In contrast, Tsh and Scr regulate *mod* through separate binding sites. Since there are not many targets known for Tsh and Scr combined, this might be a context specific case. In support of this, Taghli-Lamalle et al. (2007) showed an interaction between Antp and Tsh. Ectopic expression of Tsh lacking the acidic domain rescues T2 defects, suggesting that Tsh and Antp function differently from Tsh and Scr, possibly through different domains.

### **Disco and Disco-r as potential Hox cofactors**

Our lab has primarily focused on the genes *disco* and *disco-r* (referred together as the *disco* genes) and their roles during embryonic Drosophila development. *disco* encodes a C2H2 zinc finger transcription factor whose expression can first be seen as a cap over the posterior pole of the blastoderm embryo; during cellularization of the blastoderm it is expressed within anterior regions as well, and by the time segmentation is

evident there is expression in the gnathal lobes, clypeolabrum, the antennal segment, the posterior region of the procephalic lobe, and in the thoracic segments and the visceral mesoderm (Lee et al., 1991). *disco* is also expressed within some of the imaginal discs, which will give rise adult structures.

In *disco* mutants the larval visual nerve, the Bolwig's nerve, fails to make proper connections with targets in the larval brain 50% of the time (Heilig et al., 1991; Steller et al., 1987). However, *disco* also has a role in embryonic gnathal development, which was not uncovered until both *disco* and *disco-r*, a gene with redundant function to *disco*, were knocked out simultaneously (Mahaffey et al., 2001). RNAi of *disco-r* in a *disco*<sup>1</sup> mutant background leads to embryonic lethality, and embryos are missing mouth structures that are derived from the gnathocephalic segments. More specifically, they are missing structures derived from the maxillary and mandibular segments specified by the *Hox* genes *Deformed (Dfd)* and labial structures specified by *Sex combs reduced (Scr)* (Merrill et al., 1987; Pattatucci et al., 1991; Pederson et al., 1996). These data along with data about changes in target gene regulation suggest that *disco* and *disco-r* are potential Hox cofactors.

In the absence of *disco/disco-r*, *Dfd* cannot properly regulate several target genes within the gnathal segments (Mahaffey et al., 2001). Loss of *disco/disco-r* does not affect the initiation of *Dfd* expression; likewise, loss of *Dfd* does not affect initiation of *disco/disco-r* expression, suggesting that the *disco* genes are potential cofactors for *Dfd* (Mahaffey et al., 2001). Direct interaction assays have been done between *Disco* and *Deformed*, however, no interaction has been detected. This may suggest collaborative

interactions between Disco and Deformed to regulate target gene expression (see Chapter Four).

In addition to loss-of-function studies, gain-of-function studies have also been conducted with *disco*. Ectopic expression of *disco* within the trunk represses dorsal development, which is indicated by the repression of a dorsal determinant gene *pannier* (Herranz and Morata, 2001; Robertson et al., 2004). Normal trunk specific structures were also affected by ectopic *disco* expression. For example, denticle formation was repressed, dorsal tracheal cells and oenocytes were missing, and the trunk peripheral nervous system was altered such that it may have a mixed gnathal/trunk identity (Robertson et al., 2002, 2004). In the opposite scenario, when *disco* and *disco-r* are absent there is often excessive dorsal development in the head indicated by the fusion of the labial lobe and sometimes the maxillary lobe to the dorsal ridge (Robertson et al., 2004).

Even with all this information, little is known about Disco's role in gene regulation and adult *Drosophila* development. Expression of *disco* in the imaginal discs has implicated that *disco* might play a role in adult *Drosophila* development. However, this expression has not been well characterized and neither has the role *disco* plays in the development of adult appendages. One of the primary goals of my research is to address *disco*'s role in adult appendage development by providing a more thorough analysis of *disco* expression in the imaginal discs and also presenting data from loss- and gain-of-function studies.

Another goal of this research is to understand Disco's molecular function. Many

studies have been done with C2H2 zinc finger transcription factors and most of these transcription factors are capable of binding DNA. In fact some scientists have suggested that the DNA binding site of any given C2H2 zinc finger protein might be predictable (Meng et al., 2007; Thiesen and Bach, 1991; Thiesen and Bach, 1993). Disco being a C2H2 zinc finger transcription factor should be capable of binding DNA, and indeed Lee et al. (1999) demonstrated that Disco can, though no DNA binding site was identified (Lee et al., 1999). Identifying a binding site will assist in our understanding and identification of genes that are directly regulated by Disco. Along the same lines, it is important to identify potential protein-protein interactions that might affect Disco function. One protein-protein interaction that we identified was with the corepressor CtBP, and this interaction may have some affect on Disco function. Overall, the work presented here will contribute to the understanding of pattern formation.

## REFERENCES

- Alexandre, E., Graba, Y., Fasano, L., Gallet, A., Perrin, L., De Zulueta, P., Pradel, J., Kerridge, S. and Jacq, B.** (1996). The Drosophila teashirt homeotic protein is a DNA-binding protein and modulator, a HOM-C regulated modifier of variegation, is a likely candidate for being a direct target gene. *Mech Dev* **59**, 191-204.
- Anderson, K. V. and Nusslein-Volhard, C.** (1984). Information for the dorsal--ventral pattern of the Drosophila embryo is stored as maternal mRNA. *Nature* **311**, 223-7.
- Berleth, T., Burri, M., Thoma, G., Bopp, D., Richstein, S., Frigerio, G., Noll, M. and Nusslein-Volhard, C.** (1988). The role of localization of bicoid RNA in organizing the anterior pattern of the Drosophila embryo. *EMBO J* **7**, 1749-56.
- Biggin, M. D. and McGinnis, W.** (1997). Regulation of segmentation and segmental identity by Drosophila homeoproteins: the role of DNA binding in functional activity and specificity. *Development* **124**, 4425-33.
- Campos-Ortega, J. A. a. H., V.** (1997). The embryonic development of Drosophila melanogaster. Berlin; New York: Springer.
- Capovilla, M., Brandt, M. and Botas, J.** (1994). Direct regulation of decapentaplegic by Ultrabithorax and its role in Drosophila midgut morphogenesis. *Cell* **76**, 461-75.
- Carroll, S. B. and Scott, M. P.** (1986). Zygotically active genes that affect the spatial expression of the fushi tarazu segmentation gene during early Drosophila embryogenesis. *Cell* **45**, 113-26.
- Chadwick, R. and McGinnis, W.** (1987). Temporal and spatial distribution of transcripts from the Deformed gene of Drosophila. *EMBO J* **6**, 779-789.
- Chan, S. K., Jaffe, L., Capovilla, M., Botas, J. and Mann, R. S.** (1994). The DNA binding specificity of Ultrabithorax is modulated by cooperative interactions with extradenticle, another homeoprotein. *Cell* **78**, 603-15.
- Chan, S. K. and Mann, R. S.** (1993). The segment identity functions of Ultrabithorax are contained within its homeo domain and carboxy-terminal sequences. *Genes Dev* **7**, 796-811.
- Chan, S. K., Popperl, H., Krumlauf, R. and Mann, R. S.** (1996). An extradenticle-induced conformational change in a HOX protein overcomes an inhibitory function of the conserved hexapeptide motif. *EMBO J* **15**, 2476-87.

**Chan, S. K., Ryoo, H. D., Gould, A., Krumlauf, R. and Mann, R. S.** (1997). Switching the in vivo specificity of a minimal Hox-responsive element. *Development* **124**, 2007-14.

**Chang, C. P., Shen, W. F., Rozenfeld, S., Lawrence, H. J., Largman, C. and Cleary, M. L.** (1995). Pbx proteins display hexapeptide-dependent cooperative DNA binding with a subset of Hox proteins. *Genes Dev* **9**, 663-74.

**de Zulueta, P., Alexandre, E., Jacq, B. and Kerridge, S.** (1994). Homeotic complex and teashirt genes co-operate to establish trunk segmental identities in *Drosophila*. *Development* **120**, 2287-96.

**Diederich, R. J., Merrill, V. K., Pultz, M. A. and Kaufman, T. C.** (1989). Isolation, structure, and expression of labial, a homeotic gene of the Antennapedia Complex involved in *Drosophila* head development. *Genes Dev* **3**, 399-414.

**DiNardo, S., Kuner, J. M., Theis, J. and O'Farrell, P. H.** (1985). Development of embryonic pattern in *D. melanogaster* as revealed by accumulation of the nuclear engrailed protein. *Cell* **43**, 59-69.

**Ekker, S. C., Jackson, D. G., von Kessler, D. P., Sun, B. I., Young, K. E. and Beachy, P. A.** (1994). The degree of variation in DNA sequence recognition among four *Drosophila* homeotic proteins. *EMBO J* **13**, 3551-60.

**Fasano, L., Roder, L., Core, N., Alexandre, E., Vola, C., Jacq, B. and Kerridge, S.** (1991). The gene teashirt is required for the development of *Drosophila* embryonic trunk segments and encodes a protein with widely spaced zinc finger motifs. *Cell* **64**, 63-79.

**Frasch, M. and Levine, M.** (1987). Complementary patterns of even-skipped and fushi tarazu expression involve their differential regulation by a common set of segmentation genes in *Drosophila*. *Genes Dev* **1**, 981-95.

**Frazee, R. W., Taylor, J. A. and Tullius, T. D.** (2002). Interchange of DNA-binding modes in the deformed and ultrabithorax homeodomains: a structural role for the N-terminal arm. *J Mol Biol* **323**, 665-83.

**Gehring, W. J., Qian, Y. Q., Billeter, M., Furukubo-Tokunaga, K., Schier, A. F., Resendez-Perez, D., Affolter, M., Otting, G. and Wuthrich, K.** (1994). Homeodomain-DNA recognition. *Cell* **78**, 211-23.

**Gibson, G., Schier, A., LeMotte, P. and Gehring, W. J.** (1990). The specificities of Sex combs reduced and Antennapedia are defined by a distinct portion of each protein that includes the homeodomain. *Cell* **62**, 1087-103.

- Harding, K., Wedeen, C., McGinnis, W. and Levine, M.** (1985). Spatially regulated expression of homeotic genes in *Drosophila*. *Science* **229**, 1236-42.
- Heilig, J. S., Freeman, M., Lavery, T., Lee, K. J., Campos, A. R., Rubin, G. M. and Steller, H.** (1991). Isolation and characterization of the disconnected gene of *Drosophila melanogaster*. *EMBO J* **10**, 809-15.
- Herranz, H. and Morata, G.** (2001). The functions of pannier during *Drosophila* embryogenesis. *Development* **128**, 4837-46.
- Hombria, J. C. and Lovegrove, B.** (2003). Beyond homeosis--HOX function in morphogenesis and organogenesis. *Differentiation* **71**, 461-76.
- Hulskamp, M., Schroder, C., Pfeifle, C., Jackle, H. and Tautz, D.** (1989). Posterior segmentation of the *Drosophila* embryo in the absence of a maternal posterior organizer gene. *Nature* **338**, 629-32.
- Immergluck, K., Lawrence, P. A. and Bienz, M.** (1990). Induction across germ layers in *Drosophila* mediated by a genetic cascade. *Cell* **62**, 261-8.
- Johnson, F. B., Parker, E. and Krasnow, M. A.** (1995). Extradenticle protein is a selective cofactor for the *Drosophila* homeotics: role of the homeodomain and YPWM amino acid motif in the interaction. *Proc Natl Acad Sci U S A* **92**, 739-43.
- Kaufman, T. C., Seeger, M. A. and Olsen, G.** (1990). Molecular and genetic organization of the antennapedia gene complex of *Drosophila melanogaster*. *Adv Genet* **27**, 309-62.
- Knipple, D. C., Seifert, E., Rosenberg, U. B., Preiss, A. and Jackle, H.** (1985). Spatial and temporal patterns of Kruppel gene expression in early *Drosophila* embryos. *Nature* **317**, 40-4.
- Laugier, E., Yang, Z., Fasano, L., Kerridge, S. and Vola, C.** (2005). A critical role of teashirt for patterning the ventral epidermis is masked by ectopic expression of tiptop, a paralog of teashirt in *Drosophila*. *Dev Biol* **283**, 446-58.
- Lawrence, P. A., Johnston, P., Macdonald, P. and Struhl, G.** (1987). Borders of parasegments in *Drosophila* embryos are delimited by the fushi tarazu and even-skipped genes. *Nature* **328**, 440-2.
- Lee, K. J., Freeman, M. and Steller, H.** (1991). Expression of the disconnected gene during development of *Drosophila melanogaster*. *EMBO J* **10**, 817-26.

- Lee, K. J., Mukhopadhyay, M., Pelka, P., Campos, A. R. and Steller, H.** (1999). Autoregulation of the *Drosophila* disconnected gene in the developing visual system. *Dev Biol* **214**, 385-98.
- Li, X., Murre, C. and McGinnis, W.** (1999). Activity regulation of a Hox protein and a role for the homeodomain in inhibiting transcriptional activation. *EMBO J* **18**, 198-211.
- Lin, L. and McGinnis, W.** (1992). Mapping functional specificity in the Dfd and Ubx homeo domains. *Genes Dev* **6**, 1071-81.
- Macdonald, P. M., Ingham, P. and Struhl, G.** (1986). Isolation, structure, and expression of even-skipped: a second pair-rule gene of *Drosophila* containing a homeo box. *Cell* **47**, 721-34.
- Mahaffey, J. W., Diederich, R. J. and Kaufman, T. C.** (1989). Novel patterns of homeotic protein accumulation in the head of the *Drosophila* embryo. *Development* **105**, 167-74.
- Mahaffey, J. W., Griswold, C. M. and Cao, Q. M.** (2001). The *Drosophila* genes disconnected and disco-related are redundant with respect to larval head development and accumulation of mRNAs from deformed target genes. *Genetics* **157**, 225-36.
- Mahaffey, J. W. and Kaufman, T. C.** (1987). Distribution of the Sex combs reduced gene products in *Drosophila melanogaster*. *Genetics* **117**, 51-60.
- Manfroid, I., Caubit, X., Kerridge, S. and Fasano, L.** (2004). Three putative murine Teashirt orthologues specify trunk structures in *Drosophila* in the same way as the *Drosophila* teashirt gene. *Development* **131**, 1065-73.
- Manak, J. R., Mathies, L. D. Scott M. P.** (1994). Regulation of a *decapentaplegic* midgut enhancer by Homeotic proteins. *Development* **120**, 3605-3619.
- Mann, R. S.** (1995). The specificity of homeotic gene function. *Bioessays* **17**, 855-63.
- Mann, R. S. and Chan, S. K.** (1996). Extra specificity from extradenticle: the partnership between HOX and PBX/EXD homeodomain proteins. *Trends Genet* **12**, 258-62.
- Martinez-Arias, A., Ingham, P. W., Scott, M. P. and Akam, M. E.** (1987). The spatial and temporal deployment of Dfd and Scr transcripts throughout development of *Drosophila*. *Development* **100**, 673-83.
- Martinez-Arias, A. and Lawrence, P. A.** (1985). Parasegments and compartments in the *Drosophila* embryo. *Nature* **313**, 639-42.

- Meng, X., Thibodeau-Beganny, S., Jiang, T., Joung, J. K. and Wolfe, S. A.** (2007). Profiling the DNA-binding specificities of engineered Cys2His2 zinc finger domains using a rapid cell-based method. *Nucleic Acids Res* **35**, e81.
- Merabet, S., Kambris, Z., Capovilla, M., Berenger, H., Pradel, J. and Graba, Y.** (2003). The hexapeptide and linker regions of the AbdA Hox protein regulate its activating and repressive functions. *Dev Cell* **4**, 761-8.
- Merrill, V. K., Turner, F. R. and Kaufman, T. C.** (1987). A genetic and developmental analysis of mutations in the Deformed locus in *Drosophila melanogaster*. *Dev Biol* **122**, 379-95.
- Nusslein-Volhard, C., Frohnhofer, H. G. and Lehmann, R.** (1987). Determination of anteroposterior polarity in *Drosophila*. *Science* **238**, 1675-81.
- Nusslein-Volhard, C. and Wieschaus, E.** (1980). Mutations affecting segment number and polarity in *Drosophila*. *Nature* **287**, 795-801.
- Padgett, R. W., St Johnston, R. D. and Gelbart, W. M.** (1987). A transcript from a *Drosophila* pattern gene predicts a protein homologous to the transforming growth factor-beta family. *Nature* **325**, 81-4.
- Panzer, S., Weigel, D. and Beckendorf, S. K.** (1992). Organogenesis in *Drosophila melanogaster*: embryonic salivary gland determination is controlled by homeotic and dorsoventral patterning genes. *Development* **114**, 49-57.
- Pattatucci, A. M., Otteson, D. C. and Kaufman, T. C.** (1991). A functional and structural analysis of the Sex combs reduced locus of *Drosophila melanogaster*. *Genetics* **129**, 423-41.
- Pederson, J. D., Kiehart, D. P. and Mahaffey, J. W.** (1996). The role of HOM-C genes in segmental transformations: reexamination of the *Drosophila* Sex combs reduced embryonic phenotype. *Dev Biol* **180**, 131-42.
- Peifer, M. and Wieschaus, E.** (1990). Mutations in the *Drosophila* gene extradenticle affect the way specific homeo domain proteins regulate segmental identity. *Genes Dev* **4**, 1209-23.
- Popperl, H., Bienz, M., Studer, M., Chan, S. K., Aparicio, S., Brenner, S., Mann, R. S. and Krumlauf, R.** (1995). Segmental expression of Hoxb-1 is controlled by a highly conserved autoregulatory loop dependent upon exd/pbx. *Cell* **81**, 1031-42.

- Qian, Y. Q., Billeter, M., Otting, G., Muller, M., Gehring, W. J. and Wuthrich, K.** (1989). The structure of the Antennapedia homeodomain determined by NMR spectroscopy in solution: comparison with prokaryotic repressors. *Cell* **59**, 573-80.
- Rauskolb, C., Peifer, M. and Wieschaus, E.** (1993). extradenticle, a regulator of homeotic gene activity, is a homolog of the homeobox-containing human proto-oncogene pbx1. *Cell* **74**, 1101-12.
- Rauskolb, C. and Wieschaus, E.** (1994). Coordinate regulation of downstream genes by extradenticle and the homeotic selector proteins. *EMBO J* **13**, 3561-9.
- Regulski, M., McGinnis, N., Chadwick, R. and McGinnis, W.** (1987). Developmental and molecular analysis of Deformed; a homeotic gene controlling Drosophila head development. *EMBO J* **6**, 767-777.
- Reuter, R. and Scott, M. P.** (1990). Expression and function of the homoeotic genes Antennapedia and Sex combs reduced in the embryonic midgut of Drosophila. *Development* **109**, 289-303.
- Robertson, L. K., Bowling, D. B., Mahaffey, J. P., Imiolczyk, B. and Mahaffey, J. W.** (2004). An interactive network of zinc-finger proteins contributes to regionalization of the Drosophila embryo and establishes the domains of HOM-C protein function. *Development* **131**, 2781-9.
- Robertson, L. K., Dey, B. K., Campos, A. R. and Mahaffey, J. W.** (2002). Expression of the drosophila gene disconnected using the UAS/GAL4 system. *Genesis* **34**, 103-6.
- Roder, L., Vola, C. and Kerridge, S.** (1992). The role of the teashirt gene in trunk segmental identity in Drosophila. *Development* **115**, 1017-33.
- Ryoo, H. D. and Mann, R. S.** (1999). The control of trunk Hox specificity and activity by Extradenticle. *Genes Dev* **13**, 1704-16.
- St Johnston, R. D. and Gelbart, W. M.** (1987). Decapentaplegic transcripts are localized along the dorsal-ventral axis of the Drosophila embryo. *EMBO J* **6**, 2785-91.
- Steller, H., Fischbach, K. F. and Rubin, G. M.** (1987). Disconnected: a locus required for neuronal pathway formation in the visual system of Drosophila. *Cell* **50**, 1139-53.
- Struhl, G. and White, R. A.** (1985). Regulation of the Ultrabithorax gene of Drosophila by other bithorax complex genes. *Cell* **43**, 507-19.

**Taghli-Lamalle, O., Gallet, A., Leroy, F., Malapert, P., Vola, C., Kerridge, S. and Fasano, L.** (2007). Direct interaction between Teashirt and Sex combs reduced proteins, via Tsh's acidic domain, is essential for specifying the identity of the prothorax in *Drosophila*. *Dev Biol* **307**, 142-51.

**Thiesen, H. J. and Bach, C.** (1991). Determination of DNA binding specificities of mutated zinc finger domains. *FEBS Lett* **283**, 23-6.

**Thiesen, H. J. and Bach, C.** (1993). DNA recognition of C2H2 zinc-finger proteins. Evidence for a zinc-finger-specific DNA recognition code. *Ann N Y Acad Sci* **684**, 246-9.

**van Dijk, M. A. and Murre C.** (1994). *extradenticle* raises the DNA binding specificity of homeotic selector gene products. *Cell* **78**, 617-624.

**Walsh, C. M. and Carroll, S. B.** (2007). Collaboration between Smads and a Hox protein in target gene repression. *Development* **134**, 3585-92.

**Weigel, D., Jurgens, G., Kuttner, F., Seifert, E. and Jackle, H.** (1989). The homeotic gene fork head encodes a nuclear protein and is expressed in the terminal regions of the *Drosophila* embryo. *Cell* **57**, 645-58.

**Zeng, W., Andrew, D. J., Mathies, L. D., Horner, M. A. and Scott, M. P.** (1993). Ectopic expression and function of the Antp and Scr homeotic genes: the N terminus of the homeodomain is critical to functional specificity. *Development* **118**, 339-52.

**Zhou, B., Bagri, A. and Beckendorf, S. K.** (2001). Salivary gland determination in *Drosophila*: a salivary-specific, fork head enhancer integrates spatial pattern and allows fork head autoregulation. *Dev Biol* **237**, 54-67.

## **CHAPTER THREE**

**The appendage role of insect *disco* genes and possible implications on the evolution of the maggot larval form**

The following chapter was published in *Developmental Biology*. 2007. 309(1): 56-59.  
The *Tribolium* studies from this chapter were done by Lisa Robertson and Nathaniel Grubbs. The clones were generated by Laila Farzana.

**The appendage role of insect *disco* genes and possible implications on the  
evolution of the maggot larval form**

Mukund Patel, Laila Farzana, Lisa K. Robertson<sup>1</sup>, Jennifer Hutchinson, Nathaniel  
Grubbs, Mark N. Shepherd, and James W. Mahaffey\*

Department of Genetics  
Campus Box 7614  
North Carolina State University  
Raleigh NC, 27695-7614

<sup>1</sup>Current address:

Cato Research

4364 South Alston Avenue

Durham, NC 27713-2280

\*Author for correspondence

email: jim\_mahaffey@ncsu.edu

voice: 919-515-5791

fax: 919-515-3355

Running title: appendages genes and pattern formation

## **ABSTRACT**

Though initially identified as necessary for neural migration, *Disconnected* and its partially redundant paralog, *Disco*-related are required for proper head segment identity during *Drosophila* embryogenesis. Here, we present evidence that these genes are also required for proper ventral appendage development during development of the adult fly, where they specify medial to distal appendage development. Cells lacking the *disco* genes cannot contribute to the medial and distal portions of ventral appendages. Further, ectopic *disco* transforms dorsal appendages toward ventral fates; in wing discs, the medial and distal leg development pathways are activated. Interestingly, this appendage role is conserved in the red flour beetle, *Tribolium* (where legs develop during embryogenesis), yet in the beetle we found no evidence for a head segmentation role. The lack of an embryonic head specification role in *Tribolium* could be interpreted as a loss of the head segmentation function in *Tribolium* or gain of this function during evolution of flies. However, we suggest an alternative explanation. We propose that the *disco* genes always function as appendage factors, but their appendage nature is masked during *Drosophila* embryogenesis due to the reduction of limb fields in the maggot style *Drosophila* larva.

**Keywords:** *Drosophila*, segment identity, zinc finger, appendage, limb, pattern formation

## INTRODUCTION

The arthropod body is composed of a series of repeated segments along the anterior/posterior body axis, and though sometimes similar in appearance, different segments often have differing morphological characteristics. The differences in segment morphology are often evident as variations in the form of appendages extending from each segment. The wormlike *Drosophila* larva lacks appendages except for small sensory organs that are thought to be the remnants of larval appendages. Adult appendages arise from the imaginal discs, which are blocks of cells set-aside during embryogenesis. The genetic hierarchy governing appendage development has been extensively studied in *Drosophila* (see reviews by (Kojima, 2004; Morata, 2001; Panganiban, 2000). Many factors have been identified that have proximal to distal domains of expression and control of appendage development. Some factors have roles in all appendages, while others are specific for a particular class of appendage. Comparative studies in other insects indicate that much of this process has been conserved (for examples, see Abzhanov and Kaufman, 2000; Angelini and Kaufman, 2005a; Angelini and Kaufman, 2005b; Beermann et al., 2004; Jockusch et al., 2004; Prpic et al., 2001; Williams and Nagy, 2001).

Previously, we presented evidence that the redundant genes *disconnected* (*disco*) and *disco-related* (*disco-r*) (together referred to as the *disco* genes below) function in parallel with the head Hox genes to specify identity in the larval head segments during *Drosophila* embryogenesis (Mahaffey, 2005; Mahaffey et al., 2001; Robertson et al., 2004). In this manuscript we describe our investigation into the role of the *disco* genes

during development of the adult fly. Earlier work (Lee et al., 1991) demonstrated that *disco* is expressed in some of the imaginal discs. Furthermore, *lacZ* from the enhancer trap line C50.1S1 (reported to reside near *disco*) has been used as a marker for studies of leg joint formation (Bishop et al., 1999; Mirth and Akam, 2002). Below we demonstrate that the *disco* genes play a prominent role in appendage specification where they are necessary in order for cells to contribute to the medial and distal portions of the ventral appendages. Highlighting the substantial role these genes play during appendage development, ectopic expression transforms dorsal appendages to ventral fate; for example, ectopic activation of *disco* transforms the wings towards legs. We also demonstrate that the appendage specification role is conserved in the beetle, *Tribolium castaneum*. However, we found no evidence for an embryonic head specification role in *Tribolium*. We discuss this apparent discrepancy between appendage and segmentation roles for *disco* genes, and propose that, in actuality, there may only be an appendage role. We offer that the reduction of appendage fields during evolution of the derived, limbless maggot larval form may be the reason these genes to appear to be regional head specification factors in the embryo.

## **MATERIALS AND METHODS**

### ***Drosophila* and *Tribolium* stocks and culture**

Flies were reared on standard cornmeal-agar-molasses medium. Deficiency *Df(1)ED7355* (14A8-14B7) (FBst0008899) was generated by the DrosDel Project (Ryder et al., 2004) and was obtained from the Bloomington Stock center. *Tribolium* (GA1 wild type) were reared at 30°C on whole-wheat flour supplemented with 5% yeast powder (Berghammer

et al., 1999).

### **Induction of *UAS-disco***

We induced ectopic expression of *disco* with the *UAS-disco* lines described in (Robertson et al., 2002) at 17°C, 25°C and 29°C. We tried several imaginal disc drivers, *P{GawB}E132*, (Halder et al., 1995), *P{GAL4}klu<sup>G410</sup>* (Klein and Campos-Ortega, 1997), *P{GAL4-dpp.blk1}* (Stahling-Hampton et al., 1994) and *P{GawB-ΔKE}Bx<sup>MS1096-KE</sup>* (Capdevila and Guerrero, 1994). Only *P{GawB-ΔKE}Bx<sup>MS1096-KE</sup>* and *P{GAL4-dpp.blk1}* supported development to produce some pharate adults that could be dissected for examination. Even with these drivers, most died before pupation.

### **Cloning of *Tc-disco***

The *Drosophila disco* sequence was compared to *Tribolium* genome trace files available at NCBI (<http://www.ncbi.nlm.nih.gov/>) using BLAST (Altschul et al., 1997).

Overlapping regions were assembled to obtain the contiguous genomic region using Biolign 2.0.9 and the CAP contig assembly program (Huang, 1992). ORF Finder (<http://www.ncbi.nlm.nih.gov/gorf/gorf.html>) was used to identify the predicted open reading frame. For in situ probes and RNAi studies, primers were designed to amplify the ORF as a DNA template for RNA transcription. Genomic DNA was isolated from *Tribolium* pupae with the Qiagen DNeasy Tissue Kit. Two sets of primers were necessary to amplify overlapping regions of the *Tc-disco* ORF. The first pair amplified a 2220 bp region: Forward—5' ATGTCACCTAACCATCGCC 3'; Reverse—5' GTAATGCGTTTTACGCCGA 3'. The second pair amplified a 2177 bp region: Forward—5' AGACGTTTTGCGACAAAGG 3'; Reverse—5'

TCACGAACTCTCCGAACTCTT 3'. The resulting PCR products were cloned into Promega's pGEM-T Easy vector. Both clones were digested with XbaI and SacII, and the appropriate, adjoining fragments were isolated and ligated together yielding a 2931 bp *Tc-disco* ORF. Sense and anti-sense transcripts were generated using T7 or SP6 polymerase (Promega).

### **RNAi Injections**

Injections were performed as described for parental RNAi (Bucher et al., 2002). Approximately 0.2 ul of dsRNA (0.5 ug/ul to 3 ug/ul) was injected into each pupa. Injected females were allowed to complete pupation and were mated to wild type males. Larval offspring and unhatched eggs were collected. To control for the mechanical effects of injection or effects coming from the injection buffer, control injections were performed with buffer only. To control for effects which may arise from activation of the RNAi pathway, but not specific to *Tc-disco*, we also performed control injections using dsRNA prepared from a portion of the *Drosophila giant* coding region. Unhatched, or newly hatched *Tribolium* larvae were prepared for cuticle examination essentially as described for *Drosophila* (Pederson et al., 1996). Efficacy of RNAi was assayed by in situ hybridization with the *Tc-disco* probe to embryos collected from injected females.

### **In situ localizations of mRNA and protein**

*Drosophila* imaginal tissues were dissected from larvae and pupae in Phosphate buffered saline and placed into the standard in situ fixative. Digoxigenin labeled antisense RNAs were prepared and in situ hybridizations were done essentially as in Tautz and Pfeifle (1989). For *Tribolium*, the partial *Tc-disco* ORF comprising the first PCR generated

clone described above was used as a template. Overnight hybridization was done at 55°C. Probes for *disco*, *disco-r*, and *Dll* mRNAs were described in Mahaffey et al., (2001). The *wingless* clone was kindly provided by Dr. Amy Bejsovec (Duke University). Other probes were obtained from *Drosophila* genomic DNA using PCR. The primers used to generate clones were as follows: *bifid*, AGATACGACGTCCAGGAGCTG (forward) and TGCCGCTCTTGGTGATGA (reverse); *apterous*, TTGGTACTCGCCGATGCT (forward) and CAAGTTAAGTGGCGGTGTGC (reverse); *scalloped*, CTATGTGTTTGAGGTGGCGG (forward) and GCTGAACTAAAGTCGGTT (reverse). Immunological detection of proteins followed the protocol described in Pederson et al. (1996).

Imaginal discs for fluorescent in situ hybridization were dissected and fixed as mentioned above. Probes for fluorescent in situ hybridization were made essentially as described with different haptens. The hapten used for *disco* probes was digoxigenin and for *teashirt* probes was biotin. The protocol used for hybridization is essentially described in Kosman et. al. (2004), however, there was no treatment with xylenes or Proteinase K for imaginal discs. For detection of *disco* a sheep-anti DIG HRP (Roche) antibody (1:400) was used and for *teashirt* a mouse anti-Biotin (1:400) was used as a primary then a Goat anti-mouse HRP (1:250) as a secondary. All antibodies were preabsorbed against fixed *Drosophila* embryos. In both cases the Cy3/Flourescein tyramide signal amplification kit from Perkin Elmer was used. For *disco* the tyramide reaction was for 25 min and for *teashirt* was 5 min in imaginal discs. For embryos both tyramide reactions were 15 min. Detection of bound Dachshund antibodies was with Goat anti-mouse

labeled with fluorescein. Images were obtained using the Zeiss Pascal confocal microscope, and then subsequently processed using the 3D imaging software Velocity LE (Improvision) and Adobe Photoshop.

### **Mosaic Analysis of Drosophila**

We used the FLP/FRT recombination system (Xu and Rubin, 1993). *Df(1)ED7355* (FBab0030966) females were crossed with  $w^{1118}$ , *sn3*, *P{neoFRT}19A* (FBst0001740) males to generate recombinants between the deficiency and the *19A FRT* site. We generated both *sn* marked and non-marked *Df(1) ED7355*, *FRT* bearing chromosomes, permitting us to mark in separate experiments either the homozygous deficient cells or the twinned wild type cells with the *sn* bristle marker. The recombinant females were crossed with *P{neoFRT}19A*, *P{tubP-GAL80}LL1*, *P{hsFLP}1*,  $w^*$  (FBst0005132) males. To increase the rate of mosaicism,  $w^{1118}$ , *sn3*, *P{neoFRT}19A* and *P{neoFRT}19A*, *P {tubP-GAL80}LL1*, *P{hsFLP}1*,  $w^*$  females were crossed with  $w^{1118}$ ; *P {70FLP}10* (FBst0006938) males that carry *FLPase* on second chromosome. Male progeny were mated with the deficiency-FRT recombinant females. Females were allowed to lay eggs for either 4 or 24 hours followed by heat shock for 4 hours by submerging vials in a 37°C water bath. Resulting female progeny of appropriate genotype were analyzed.

To mark the clones with Green Fluorescent Protein (GFP), we generated a second chromosome that carried *UAS-GFP*, *arm-Gal4* and *hsFLP* by following the standard recombination methods. This chromosome was homozygosed in *sn*-, *Df(1)ED7355*, *FRT* background as well as in the *P{neoFRT}19A*, *P{tubP-GAL80}LL1*, *P{hsFLP}1*,  $w^*$  flies. Females of the first genotype were mated with males of the latter genotype to create

mosaic flies. Heat shock treatment was the same as described before.

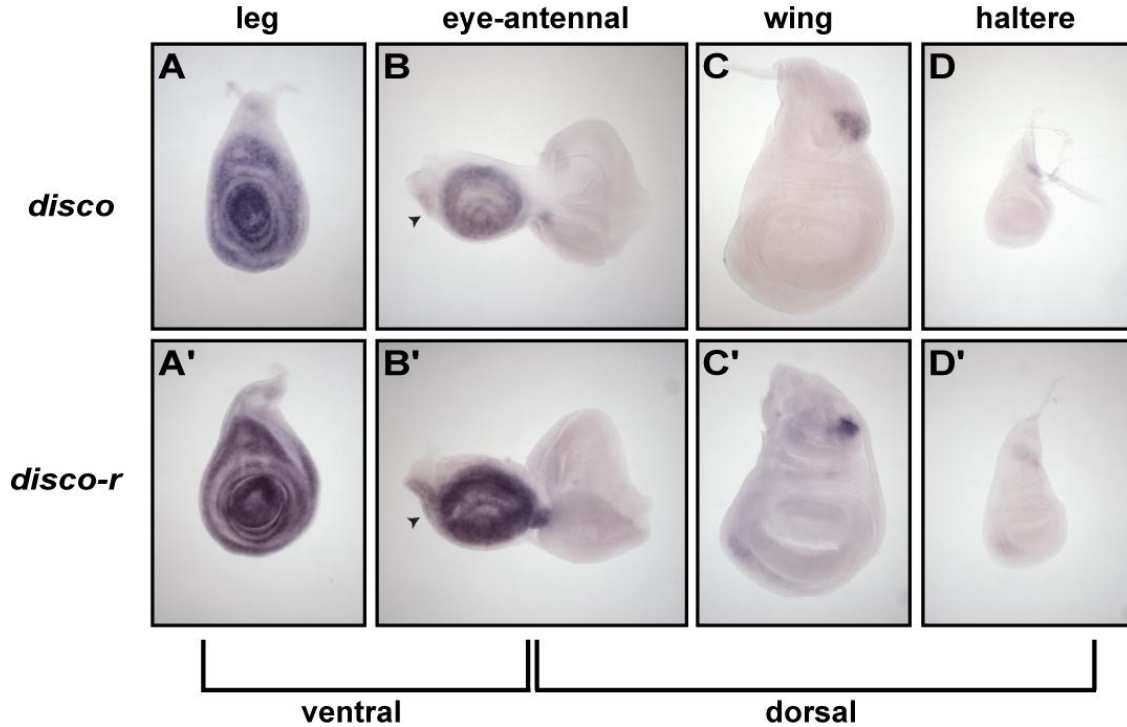
## RESULTS

### **The *disco* genes are widely expressed in ventral imaginal discs of *Drosophila***

Prior work mentioned that *disco* is expressed in many of the imaginal discs (Lee et al., 1991). Below, we extend the description of *disco* and include *disco-r* expression in the imaginal discs. In early third instar larvae, the *disco* genes are expressed in the ventral imaginal discs (Fig. 1): the antennal and maxillary palp regions of the eye-antennal discs, the labial discs (data not shown), and in the leg discs. In addition, they are expressed in a small patch of cells in the future scutellum of the wing discs, in a similar position in the haltere discs, and weakly in an anterior region of the eye disc. We did not detect any difference in the spatial expression of the two genes, though *disco-r* appeared to be more abundant in the discs which is opposite to what we previously observed in embryos (Mahaffey et al., 2001)

To help define the spatial and temporal aspects of expression, we compared *disco* with two other known appendage factors, *teashirt* (*tsh*) and *dachshund* (*dac*). *tsh* is required for proper development of the trunk segments during embryogenesis (de Zulueta et al., 1994; Fasano et al., 1991; Roder et al., 1992) and of the body wall and proximal regions of ventral discs during development of the adult (Erkner et al., 1999; Wu and Cohen, 2000). In accord with this, *tsh* is expressed in the outer ring of the leg discs (Fig. 2A,E'). *disco* mRNAs were co-expressed in many of the *tsh* expressing cells of the leg discs, except in the dorsal (stalk) region (Fig. 2E,E',E"). In the medial and distal portions of the discs (where *tsh* is not expressed) there was uniform expression of the *disco* genes.

Figure 1



**Figure 1. Expression of *disco* and *disco-r* in third instar imaginal discs.**

The top panels show expression of *disco* and the lower panels *disco-r* (denoted by '). The *disco* genes were expressed throughout most of the leg discs (A and A'), and the antennal and maxillary palp (arrowhead) portion of the eye-antennal discs (B and B'), and there was weak expression in the anterior-ventral fold of the eye discs. In the wing (C and C') the *disco* genes were expressed in a small region of the scutellum. Expression was observed in an equivalent region of the haltere disc (D and D'). The legs and antennal and maxillary portion of the eye-antennal disc are ventral appendages while the wing and haltere are dorsal appendage discs.

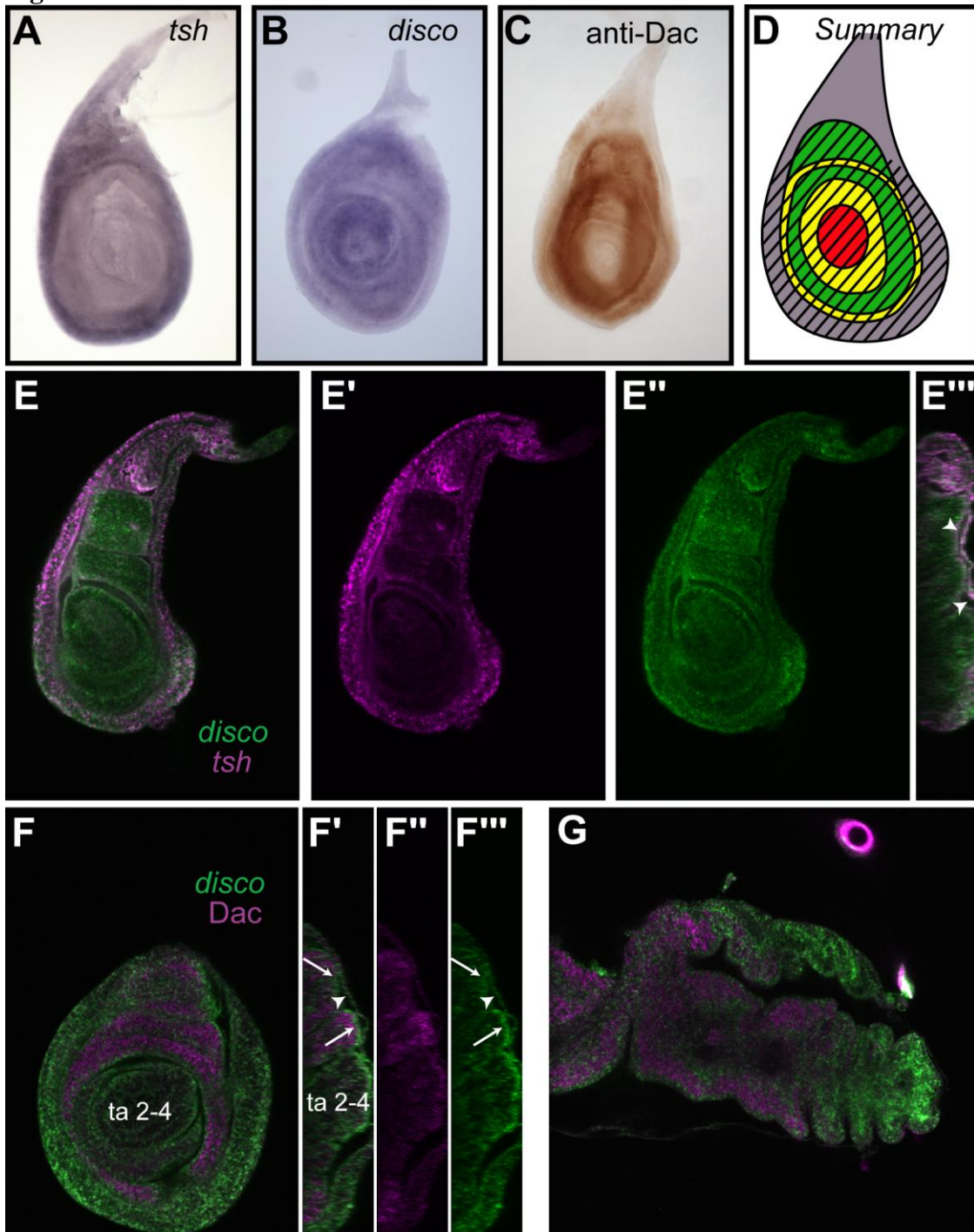
*dac* is required for medial appendage development (Mardon et al., 1994). We used an antibody to detect Dac, while detecting *disco* mRNA. Since Dac is a nuclear protein and the nuclei of the columnar disc cells are basally located, Dac staining appeared below the cytoplasmic staining of the *disco* mRNA. However, z-section reconstructions (Fig. 2F-G) permitted unequivocal demonstration that Dac positive cells also expressed *disco*. As above, we present data from the leg discs as an example. The distribution of Dac protein was entirely contained within the *disco* domain. In the dorsal region, the expression boundaries are nearly congruent. However, ventrally the *disco* mRNA was more proximally expressed, in the region where *disco* and *tsh* overlap (Fig. 2 F). Furthermore, while Dac accumulated only in the medial portion of the discs, *disco* expression extended through the distal region (FIG. 2F,G), thus overlapping the *Distal-less (Dll)* domain (data not shown). This distal expression was more pronounced during eversion of the leg (Fig.2G). We also note that both *disco* genes are weakly expressed near the joints of completely extended legs, while the C50.1S1 enhancer trap is strongly expressed making it a good marker for studies of joint formation (See supplemental data Fig. 1) In sum, the *disco* genes are widely expressed throughout the majority of the leg disc cells, indicating these genes may have a prominent role during development of these tissues.

### **Cells lacking the *disco* genes do not contribute to the ventral appendages**

Embryos homozygous for loss of both *disco* genes die during late embryogenesis, so mosaic analysis, generating small clones of homozygous mutant cells in a heterozygous background, is needed to assess the role of these genes during later stages. However,

**Figure 2.** Relationship between *disco* and other appendage factors in the *Drosophila* wandering third instar discs. (A) *tsh*, (B) *disco*, (C) Dac. (D) A summary of our interpretation of gene expression in the leg discs; *tsh* (violet), Dac (green), *Dll* (red), Dac + *Dll* (yellow) and *disco* (striped region). (E-E''') Fluorescent in situ detection of *disco* (green) and *tsh* (magenta) mRNAs. E, composite of both; E', *tsh*; E'', *disco*; and E''', a z-section of the stacked images. White arrowheads point to the peripodial membrane. F-F''', Dac protein (magenta) and *disco* mRNA (green). F, composite XY section; F' composite YZ section. Arrows point to the apical location of the *disco* mRNA, while the nuclear Dac protein is more basally located. F'', Dac; F''' *disco* mRNA. (G) Extending leg labeled as in F.

Figure 2



redundancy requires that we eliminate both genes, and our attempts to generate double mutants have been unsuccessful so far. Therefore, we took advantage of a small deficiency generated by the DrosDel project (Ryder et al., 2004), *Df(1)ED7355* which removes both *disco* and *disco-r* and only a few neighboring genes (See Supplemental Fig.2). We used this deficiency with the FLP/FRT system (Xu and Rubin, 1993). We generated an FRT-bearing chromosome containing *Df(1)ED7355* and with the *sn* bristle marker to mark homozygous deficiency clones in adults (see materials and methods for more details). Furthermore, *Df(1)ED7355* contained the mini-white marker, so dark red sectors would mark homozygous deficiency clones in the eyes, while their “twinned” homozygous normal cells would be white. Using this technique, we were able to generate homozygous *Df(1)ED7355* clones in many tissues (Table 1A). *sn* bristles appeared in the dorsal and ventral body wall of the thorax, in the abdomen, and in the dorsal and ventral head capsule (see thoracic example in FIG. 3A,B). Dark red sectors accompanied by white “twinned” regions were observed in the eyes (FIG. 3C). In contrast, we never observed *sn*-marked bristles in the antennae, maxillary palps or proboscises (Table 1A). In the legs, we did find three flies each with one or two *sn* bristles in the coxa (i.e. the proximal-most portion of the legs), but we never observed *sn* bristles in the trochanter or femur (Table 1A). The occasional clone in the coxa could reflect that the *disco* genes are expressed only in part of the proximal regions of the legs, or that these were late-occurring recombinants arising after *disco* function. As a control, we used *sn* to mark bristles in cells homozygous for the non-deficiency chromosome (see Materials and Methods), resulting in *sn* bristles in all regions of the fly (Table 1B).

**TABLE 1.** In the upper table, homozygous *Df(1)ED7355* clones were marked with *sn* bristles. In the lower table, *sn* bristles mark the homozygous non-*Df(1)ED7355* clones. In both, eye clones are marked by combinations of mini-white. The three clones listed as coxa/trochanter in table A were all in the coxa. The reduced number of clones in the dorsal and ventral head capsule (reduced when compared to the reciprocal cross, Table 1B) could indicate that the *disco* genes play some role in specification of this region. In both tables, the top row shows absolute numbers from the experiment, and the bottom row shows the percent of total flies scored. We did note that even where homozygous *Df(1)ED7355* cells survived, non-deficiency *sn* clones were often larger, possibly indicating there was somewhat of a growth disadvantage to the homozygous deficiency cells. This may have accounted for the lower numbers of total clones identified when marking the deficiency cells. Abbreviations: Ant=Antenna; Max=Maxillary Palps; Prob=Proboscis; Troch=Trochanter; Abd=Abdomen.

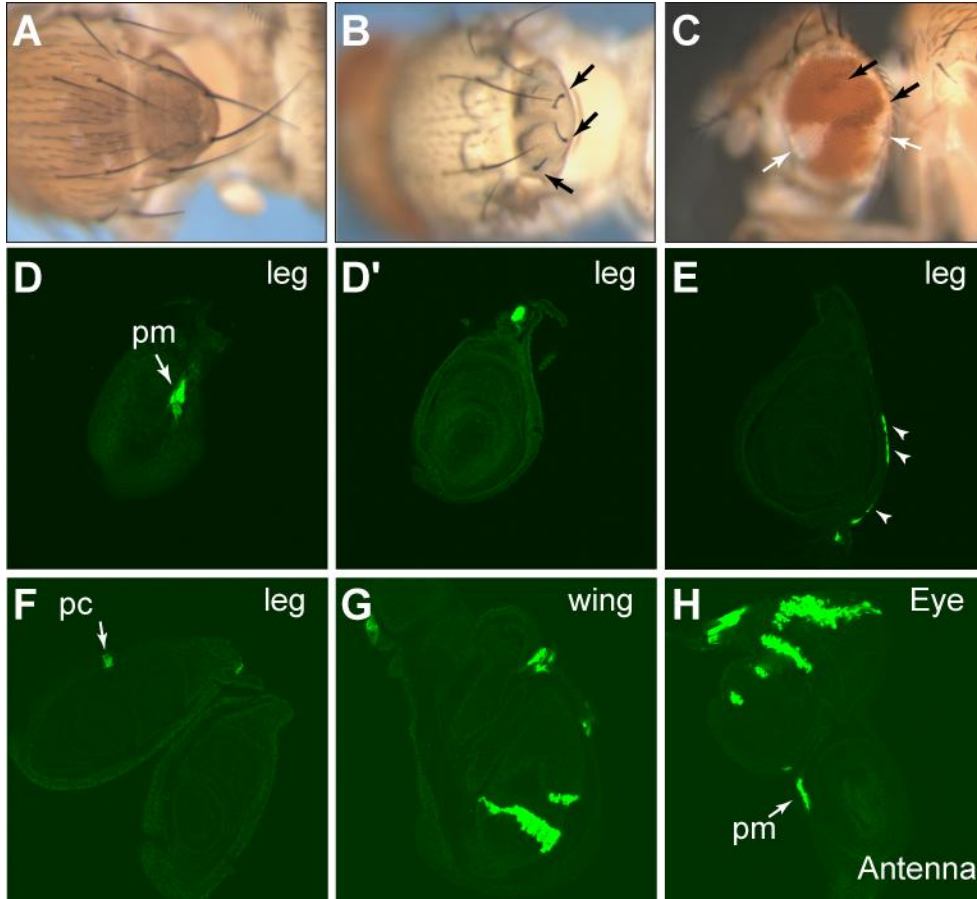
**A. *sn, Df(1)ED7355, FRT/FM7h x FLP, FRT/y; +/-FLP***

Fly #	Eye	Dorsal Head	Ventral Head	Ant	Max	Prob	Femur	Coxa/Troch	Thorax	Abd
208	188	39	16	0	0	0	0	3	83	134
%	90	18	8	0	0	0	0	1	40	64

**B. *Df(1)ED7355, FRT/FM7h x sn, FRT/y; +/-FLP***

Fly #	Eye	Dorsal Head	Ventral Head	Ant	Max	Prob	Femur	Coxa/Troch	Thorax	Abd
91	88	66	55	31	25	20	38	36	75	80
%	96	72	60	34	27	21	41	39	82	88

**Figure 3**



**Figure 3.** Clonal analysis of *disco* genes. (A) scutellar bristles on the thorax of a wild type *Drosophila* adult. (B) *sn* marked bristles developing in a clone of cells that are homozygous for *Df(1)ED7355*. (C) Clones of cells in the eye. Homozygous wild type cells are white, while those homozygous for *Df(1)ED7355* are dark red. (D-H) GFP marked clones in imaginal discs. (D and D') Two different Z sections from the same leg disc. (D) A clone in the peripodial membrane (pm) while D' shows there are no clones in the columnar cells of the disc proper. (E) Another leg disc with clones in cells along the outer edge of the disc (arrowheads). These appear to be hematocytes, and are not disc-specific cells. (F) Paired first thoracic leg discs. Note the clone (pc) in the columnar cells within the proximal region of the disc. In contrast to the leg discs, marked cells appear in many regions of the dorsal imaginal discs. (G) A wing disc, (H) An eye-antennal disc with several clones in the eye, a dorsal structure, but only in the peripodial membrane (pm) of the antennae, a ventral disc.

Since *sn* only marked bristle cells, we established a variation of the MARCM system (Lee and Luo, 2001) so that we could mark homozygous *Df(1)ED7355* clonal cells with green fluorescent protein (GFP) (see Materials and Methods for details). In agreement with the *sn* marker, we were able to generate numerous GFP-positive clones in the eyes, abdomen, thoracic and head capsule of adult flies (data not shown). The only GFP-positive clones in the legs, antenna, maxillary palps or proboscis were in neurons that ran down the center of the appendages (data not shown). We also examined clone positions in third instar larval imaginal discs. Overall, clones were prevalent in the dorsal imaginal discs (wing and eye, Fig. 3G,H respectively). GFP-positive cells were present in all parts of these discs. By contrast, GFP-positive cells were rare in ventral discs, and they were not observed in the medial and distal portions of the ventral discs (Fig. 3D',E,F). Marked clones were observed in the peripodial membranes (Fig. 3D,H), and occasionally in the proximal portions of the disc proper (Fig. 3F), but never in the medial or distal regions. We also detected what appeared to be GFP-labeled hemocytes/macrophage in and around the leg discs (Fig.3E). The GFP-tagged data agrees with our *sn* marked clonal results, and we conclude that homozygous *Df(1)ED7355* clones can survive in many regions, but not where the *disco* genes are expressed. We suspect that clonal cells lacking the *disco* genes were generated, but that they cannot survive or are removed from the medial and distal portions of the ventral appendages.

### **Ectopic Disco transforms dorsal appendages to ventral identity**

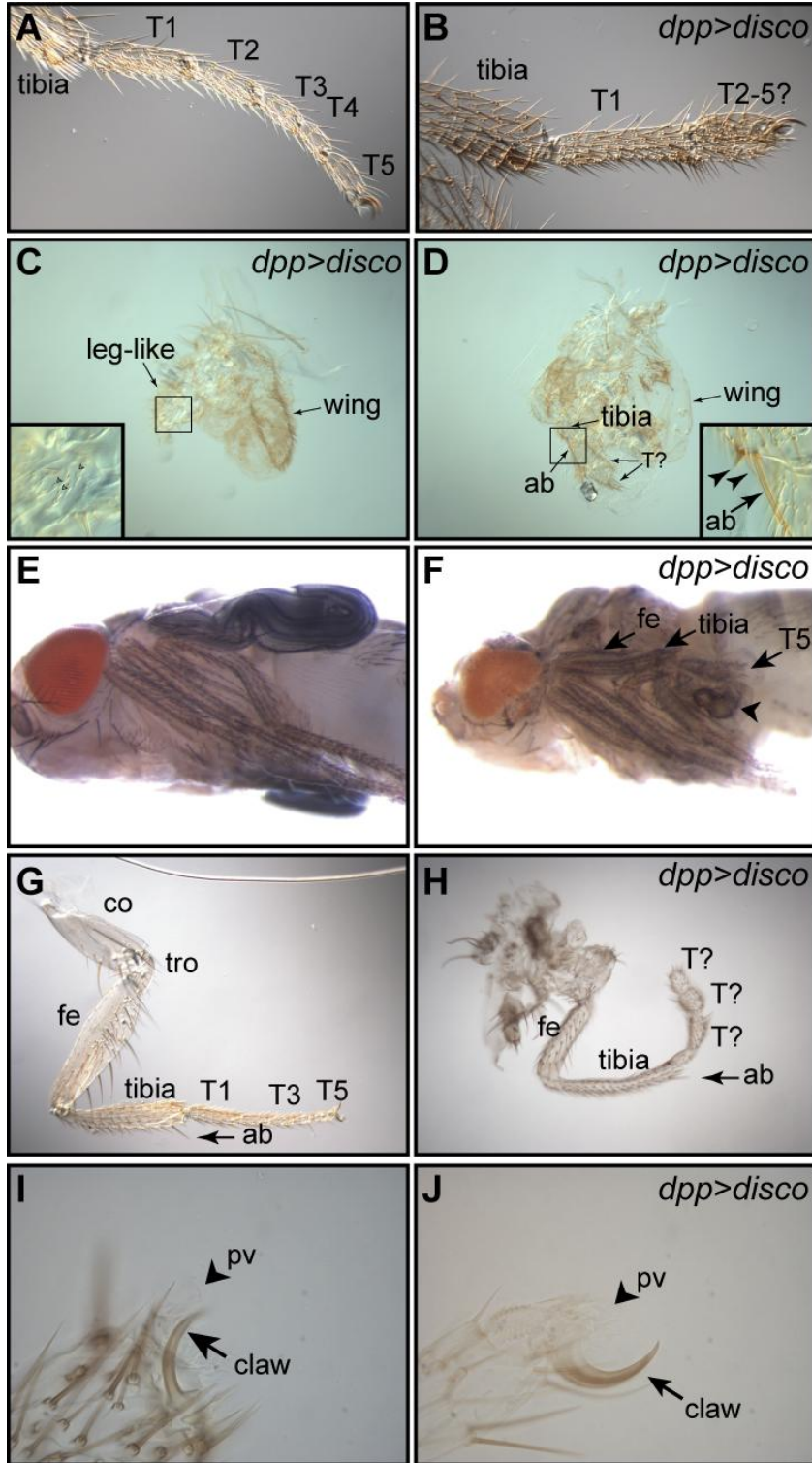
That cells lacking the *disco* genes were not found in the ventral appendages could imply that either these genes are required to establish this region of the appendage, or that they

are required for cells to remain viable in this region. To distinguish between these possibilities, we ectopically expressed *disco* (Robertson et al., 2002) using the UAS/Gal4 system (Brand and Perrimon, 1993). All flies ectopically expressing *disco* died prior to eclosing as adults, and most died during larval stages. However, when they did survive to form pupae, we dissected them from the pupal case to examine the consequences of ectopic *disco* expression. Using the *P{GAL4-dpp.blk1}* driver, we noted that most ventral appendages (antenna, mouthparts and legs) were nearly normal. The only defect noted was that the number of tarsal segments in the legs was sometimes reduced (Fig. 4A,B). It was not clear if this was caused by fusion or deletion of the tarsal subsegments.

In contrast to the limited effect on ventral appendages, ectopic *disco* had an impressive effect on dorsal appendages; for example, wings and halteres were transformed into leg-like appendages (Fig. 4C-H), and antenna-like structures occasionally developed in the eyes. The transformation was sensitive to growth temperature. No larvae survived to form pupae at 29°, and most did not form pupae even at 25°C. For those that did form pupae at 25°C the wings were in various states of transformation toward legs. About equal numbers were of the weak and moderate phenotype (Fig. 4 C, D respectively). Though the extent of transformation did vary, in all cases, what should have been wing contained easily recognized leg tissues. Bracted bristles, a characteristic of legs, were found on all of the transformed wings. In many cases tarsal segments were present, and often we could identify apical bristles and spurs, specialized bristles found on second thoracic legs, indicating that the leg tissue had appropriate second thoracic segment identity. About 20% of the transformed wings

**Figure 4.** Effects of ectopic, *dpp*-driven *disco* expression. Except for panel C, all panels on the left are of wild type flies and those on the right (along with C are of *dpp*-driven *disco*. Normal *Drosophila* legs (for example see panel G) are composed of six segments – coxa, trochanter, femur, tibia, and tarsi (divided into five subsegments) and the claws and pulvilli are found on the distal pretarsus. In *dpp>disco* pharate adults, the tarsal subsegments of the normal legs are altered (compare panels A and B). It is not clear whether this represents a fusion or loss of sub-segments, occasionally only two subsegments were present. (C) An example of a weaker wing-to-leg transformation. Bracted bristles (see insert, a few are marked by arrowheads) indicate leg identity. Often distinctions between leg segments are visible, though individual segment identities are unclear. (D) An example of a moderately transformed wing. In addition to bracted bristles, specific leg segments are visible. Often, as is the case here, an apical bristle (ab) and spurs (arrowheads) are present (see insert) permitting identification of distinct leg segments (for example, the apical bristle and spurs are on the tibia). T?=tarsal segments (E) A wild-type fly just before eclosing that was dissected from the pupal case. Note the wing is not inflated, which occurs after eclosion. Also note the three folded legs. (F) A *dpp>disco* pharate adult that was dissected from the pupal case. Note the absence of the wing, and the fourth leg above the three normal legs. The femur (fe), tibia and tarsal segments of the ectopic leg are labeled. The arrowhead points to the haltere, which is also transformed toward a leg-like identity. (G) Higher magnification of a second thoracic wild type leg, which consists of the coxa (Co), trochanter (tro), femur (fe), tibia, and the five tarsal segments (T1-T5). The apical bristle (ab) is also marked. (H) Higher magnification of the transformed wing from a *dpp>disco* fly. The femur and tibia are clearly identified based on the presence of specific bristles, and several tarsal segments are present. Normal femurs have both bracted and non-bracted bristles, which was also the case for the femur-like segment of the transformed wings. Additionally, there are distinguishing bristles at the distal end of the normal second thoracic tibia – stout bristles referred to as spurs and two longer bristles, the apical (closest to the spurs) and preapical bristle. The presence of these bristles on the transformed wings indicated second thoracic identity, as would be expected since wings are second thoracic structures. (I) In a wild-type fly the pretarsal segment contains the terminal claws (arrow) and the pulvilli (arrowhead, and out of the focal plane). (J) In the *dpp>disco* transformed wing, one of the terminal claws is clearly visible (arrow) as is a pulvillus (arrowhead).

Figure 4



developed as well-formed leg-like structures (Fig. 4 E-J). In these cases the identity of tissues in the proximal wing was difficult to determine, though it was possibly a combination of wing and leg tissues. We could not rule out the presence of some coxa- and/or trochanter-like material. Clear femur, tibia, tarsal and pretarsal segments were formed, and the pretarsal segment contained at least one claw-like bristle and a pulvillus (Fig. 4J). Well-formed joints separated the femur and tibia and the tibia from the terminal region. The  $P\{GawB-\Delta KE\}Bx^{MS1096-KE}$  driver also generated leg-like tissues in the wing (data not shown). Regions of bracted bristles were always observed and occasionally a bristle resembling the apical/preapical bristle was observed. In some cases, regions of bracted bristles were separated by what might have been vestiges of leg segmentation.

Ectopic expression has been used to study other limb patterning genes (Chen et al., 1997; Duncan et al., 1998; Estella et al., 2003; Gorfinkiel et al., 1997) and several have been proposed to transform the wings into leg tissues. We examined *spineless*, *homothorax (hth)*, *dac*, and *tsh* using the *dpp*-Gal4 driver (data not shown). All of those we tested, and those described in the literature, had rather modest transformations compared with what we observed resulting from ectopic *disco*. From these observations, we conclude that *disco* is a potent member of the ventral appendage network, capable of transforming dorsal appendages toward ventral fate.

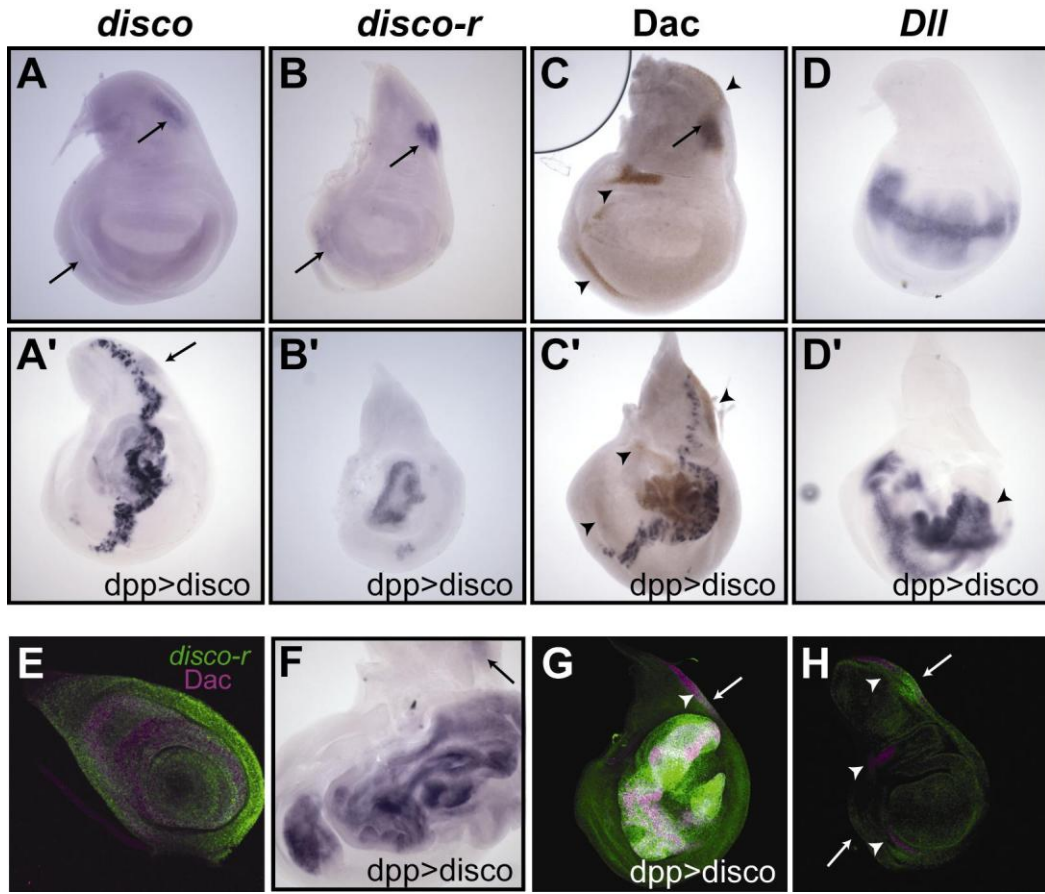
That such complete legs formed from the wing imaginal disc was surprising considering  $P\{GAL4-dpp.blk1\}$  activates expression only in a narrow line of cells along the anterior/posterior compartment boundary of the wing discs (Staebling-Hampton et al., 1994). To investigate how limited *disco* expression could cause such a complete wing-to-

leg transformation, we examined expression of several leg and wing factors following ectopic activation of *disco* in the wing discs. As expected with the *dpp* driver (Staepling-Hampton et al., 1994), *dpp*-driven *disco* mRNA accumulated to high levels along the A/P border (Fig. 5A,A'). However, we noted that ectopic *disco* (Fig. 5A') and *disco-r* (Fig. 5B',F,G) mRNAs accumulated anterior to this stripe, in the anterior wing blade region. Prior cell lineage tracing studies (Weigmann and Cohen, 1999) indicated that these anterior wing blade cells arise as daughters of cells initially in the *dpp* stripe. These cells cease expressing *dpp* as they progress anteriorly. Lee et al. (1999) demonstrated that *disco* can autoactivate, so perhaps the continued ectopic *disco* and *disco-r* expression may be due to autoactivation of the endogenous genes. This region continues to grow producing a large number of cells with high levels of ectopic expression of the endogenous *disco* genes (Fig. 5F,G). *Dac* (Fig. 5C,C'), required for medial appendage development (Beermann et al., 2004; Mardon et al., 1994), and *Distal-less (Dll)* (Fig. 5D,D') required for distal appendage (Cohen et al., 1989; Cohen and Jurgens, 1989) were also activated in these cells. As the transformed region continued to expand, a region composed of cells with high levels of *disco* (data not shown), *disco-r* (Fig. 5F), *Dac* (Fig. 5G) and *Dll* expression arose (data not shown). We also examined expression of *hth* and *tsh*, both encoding proximal appendage factors, and neither was induced by ectopic *disco* (data not shown). This indicates that activation of *disco* does not just induce leg, but it induces the portions of the leg normally controlled by *disco*, the medial and distal regions.

Since dramatic changes were observed in the expression of the *disco* genes and *Dac*, we followed the expression of these genes in more detail, using *disco-r* as a probe so

**Figure 5.** Expression of leg determining genes in *dpp>disco* transformed wing. (Note, except for *Dll*, expression in normal leg discs are shown in Figure 1 and 2.) Top row, normal wings; second row, *dpp*-driven *disco* wing discs (Denoted by '), third row examining *disco-r* and Dac in more detail. (A,A') *disco*. In normal wings *disco* is expressed in a small region of the wing disc that will give rise to a portion of the scutellum. In the transformed wings, *disco* mRNA accumulates in the *dpp* stripe along the anterior/posterior border, as would be expected for this driver. However, note the weaker staining in what should be the wing blade region, anterior to the *dpp* stripe. These cells are the daughters of cells initiating in the *dpp* stripe (Weigmann and Cohen, 1999). As they divide, they move away from the anterior/posterior border and cease expressing *dpp*. However, it appears that *disco* expression continues. (B, B') *disco-r* expression mirrors that of *disco* except that the gene is not activated by *dpp*-Gal4, hence no stripe of staining along the anterior/posterior border. However, *disco-r* mRNA accumulated in the anterior portion of the wing blade region, in daughter cells that have moved from the *dpp*-driven stripe. (C,C') Staining to detect *disco* mRNA and Dac protein demonstrated that Dac was ectopically activated in the region of the wing blade where there was ectopic expression of the endogenous *disco* and *disco-r* genes (also see G below). (D,D') *Dll* mRNA accumulation was enhanced in the region of *dpp>disco* expression. The region of ectopic *Dll* activation is marked with an arrowhead. (E) *disco-r* (green) and Dac (magenta) in a wild type leg disc. Dac and *disco-r* appear distinct only because Dac is localized to the nucleus (see text) (F) *disco-r* mRNA localization in a *dpp*-driven *disco* transformed wing from a late third instar larva. Note the extensive proliferation that has occurred in the *disco-r*-expressing wing blade region. Additional growth is also observed in the region expressing *disco-r* below the wing blade. This disc was photographed at the same magnification as those above. (G) Dac and *disco-r* in a transformed wing disc from a somewhat younger third instar larva. Note that ectopic Dac protein was completely within the *disco-r* region. (H) Dac and *disco-r* in a wild type wing disc. In all panels, Arrows point to the regions of *disco* or *disco-r* mRNA accumulation; arrowheads point to normal Dac accumulation.

Figure 5



that the *dpp*-driven expression of *disco* would not complicate the analysis (Fig. 5E-H). It appeared that the Dac-positive cells completely overlapped the *disco-r* expressing cells, though many cells expressed *disco-r* but not Dac. And though there was no way to determine whether or not the discs we examined could have formed pharate adults, in all cases the expression patterns were similar.

We also examined expression of several wing-determining genes, *bifid* (*bi*), *apterous* (*ap*), *scalloped* (*sd*) and *wingless* (*wg*) (see supplemental data Fig. 3). Though *bi* and *ap* appeared unchanged, *sd* was reduced, though the distribution appeared normal, and *wg* was at normal levels but the normally closed line of expression along the dorsal wing blade boundary was broken. It was likely that this break was caused by the ectopic growth in the wing blade induced by *disco* expression. Overall, at this stage of development, wing-specific gene expression was not altered as much as we might have expected.

### **The appendage role of *disco* is conserved in the beetle, *Tribolium***

To investigate the role of *disco* homolog in another insect, we turned to the red flour beetle, *Tribolium castaneum*. Unlike the fly maggot larva, which lacks visible appendages, appendages are present on the head and thoracic segments of *Tribolium* larvae. These appendages develop during *Tribolium* embryogenesis. Furthermore, unlike the maggot fly larvae with internalized head segments, the beetle head segments with appendages remain external. These differences make comparisons between the fly and beetle informative.

Our analyses of the *T. castaneum* genome identified only a single *disco*-like gene

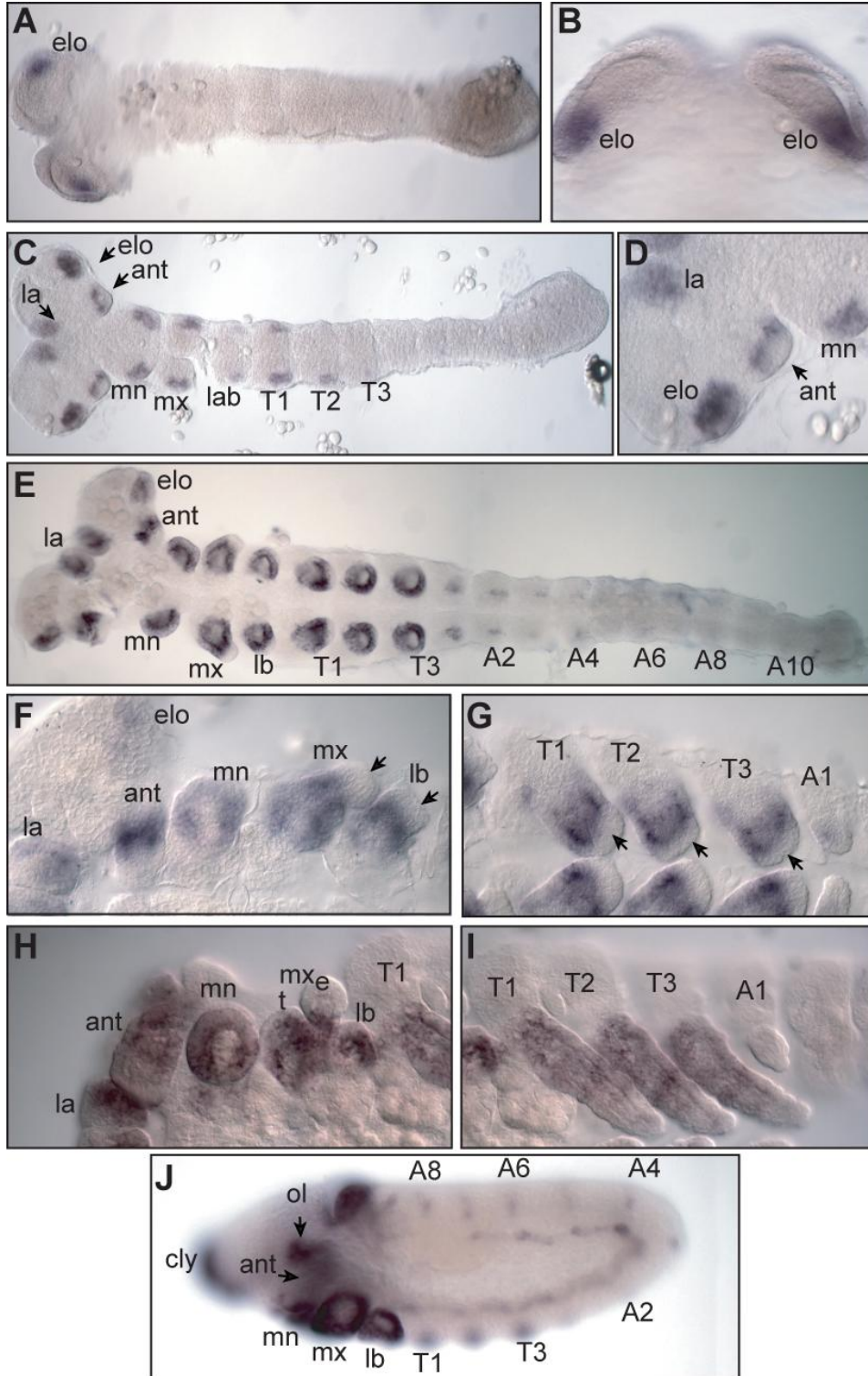
(referred to as *Tc-disco* below). The structure of the encoded Tc-Disco protein was more similar to the vertebrate homolog Basonuclin (Tseng and Green, 1992), containing three paired zinc finger domains, in contrast to the one and two pairs of Disco and Disco-r, respectively. Still, the N-terminal pair of zinc fingers, nearly identical in all animals examined, permitted unequivocal identification. We used a fragment from the largest *Tc-disco* exon as a probe to visualize mRNA distribution during *Tribolium* development (Fig. 6). *Tc-disco* transcripts were detected in the medial portions of all appendages, at all stages of embryogenesis. Other than the transient expression in the abdominal segments, *Tc-disco* mRNA was not detected in other regions of the body or head of the beetle.

There were obvious similarities between *disco* expression in *Tribolium* and *Drosophila*. Transient expression in the abdomen was observed during embryogenesis in both organisms. As appendages everted, concentric rings of stronger expression arose both in beetles and flies (for fly images see Supplemental Fig. 1). Yet, there were also differences. Both the *Tribolium* and *Drosophila disco* genes were expressed in the medial portions of the legs, but only in the fly did expression extend to the distal tip of the leg (for fly images see Supplemental Fig. 1). However, the most significant difference was that, in the beetle, only an appendage role was evident.

To investigate the effect of reducing *Tc-disco* function during *Tribolium* development, we used parental RNA interference (RNAi) (Bucher et al., 2002). Efficacy of RNAi was determined by in situ hybridization with the *Tc-disco* probe to fixed embryos collected from injected females. Eggs laid by injected females hatched, but the larvae were nearly immobile and unable to feed due to severe truncations of all

**Figure 6.** Expression of *Tribolium disco* genes. *Tc-disco* mRNA distribution during *Tribolium* development. Transcripts were first detected during early germ band elongation in the eye lobes (A,B), but as germ band elongation continued, transcripts accumulated in an anterior to posterior manner in the primordia of each head and trunk appendage—beginning first with the labral segment, then appearing in the antennal, mandibular, maxillary, labial, first thoracic (T1), second thoracic (T2), and the third thoracic (T3) segments (C-E). Expression was also detected in the region that will give rise to the pleuropodia. Weak, transient accumulation was also detected in the second through tenth abdominal segments (E). As the appendages began to evert, *Tc-disco* mRNA became limited to the medial portion of each appendage (F-I), with stripes of higher accumulation apparent in the antenna and legs (H,I). Transcripts were not detected in the body portion of the head or trunk segments or in the distal-most portions of the appendages. (J) *disco* expression in a germ band extended *Drosophila* embryo for comparison. Abbreviations: elo, eye lobe; la, labrum; ant, antenna; mn, mandible; mx, maxillary; lb, labial; T1-3, thoracic; A1-10 abdominal; cly, clypeolabrum; ol, optic lobe.

Figure 6

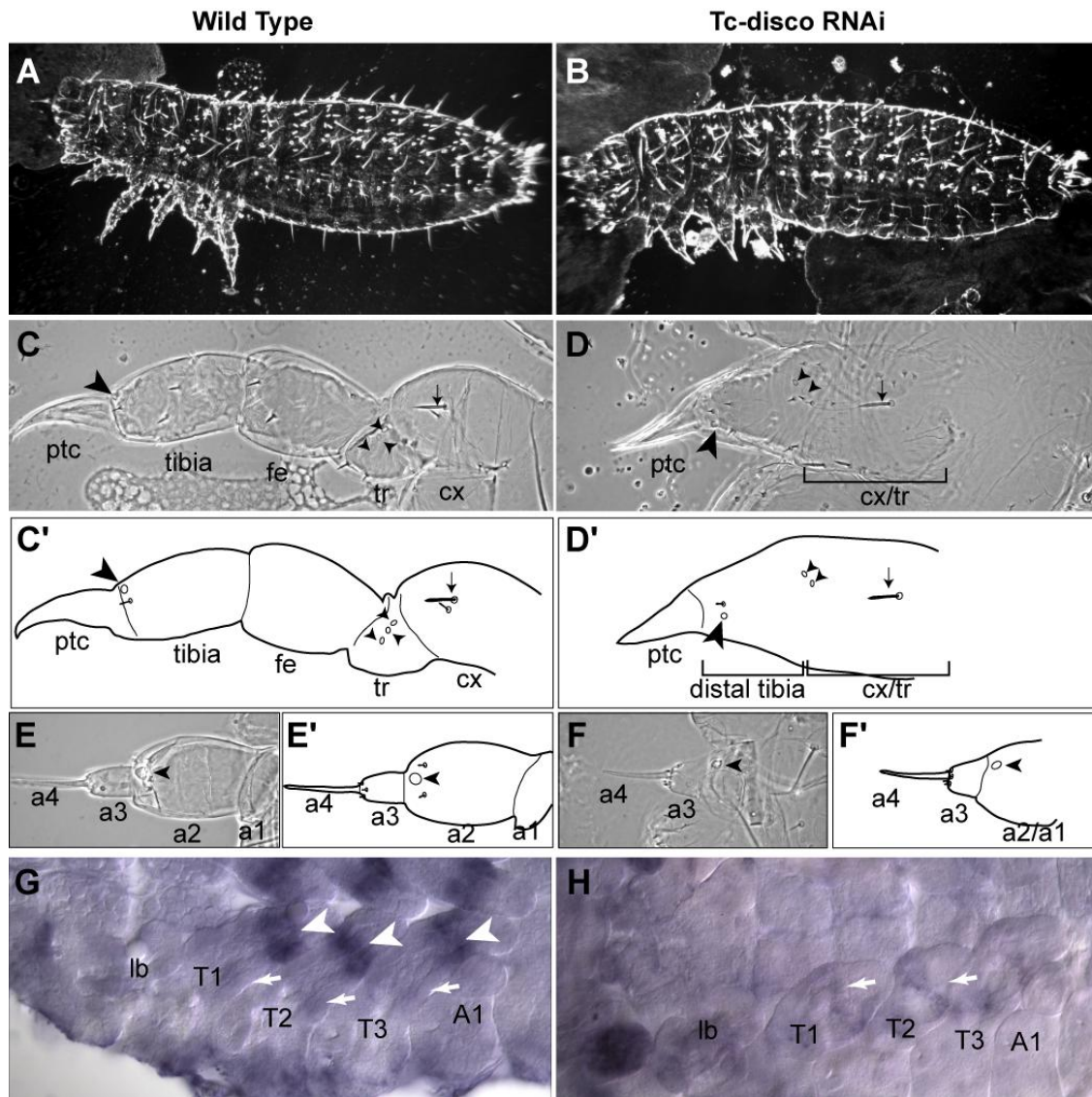


appendages (Fig. 7). Similar medial deletions were observed in all appendages. We highlight the alterations in the legs and antennae (Fig. 7C-F). In a typical collection of 115 larvae from injected females, 82 (71%) had the strong phenotype. The legs were markedly shorter due to loss of the medial regions (Fig. 7D) as were the antenna (Fig. 7F). The distal-most portion of the leg, the pretarsal claw, was present as were the sensory structures characteristic of the distal tibiotarsus, immediately above the claw. The proximal-most portion of the leg was more difficult to characterize, but remaining sensory organs and bristles indicated that the coxa and perhaps a small portion of the trochanter were present. Twenty-five larvae (22%) had a less severe phenotype with three clear leg segments, but lacking the trochanter. The remaining 8 larvae appeared to have normal larval legs. All RNAi defects were restricted to the regions of *Tc-disco* expression. Careful comparison between RNAi larvae and wild type larvae revealed no other differences, such as a more general disruption of head development as might have been expected from the role attributed to *disco* during *Drosophila* embryogenesis.

We also examined the *Tc-disco* RNAi phenotype on a molecular level using the *Tribolium dachshund* (*Tc-dac*) homolog as a marker for medial leg (Fig 7 G,H). At the fully contracted germ band stage, normal *Tribolium* legs have three regions of *Tc-dac* expression (Prpic et al., 2001), an intense ring around the middle of the leg, a weaker stained proximal band, and a spot at the junction of the leg and the body wall. In the strongest class of *Tc-disco* RNAi embryos, little or no *Tc-dac* mRNA was detected. It was not clear whether the slight staining that did remain was from the weaker proximal band or a small portion of the normally larger distal band. However, we can conclude that

**Figure 7.** Tribolium RNAi analysis. (A) A dark field image of a wild type Tribolium larval cuticle. (B) Similar image of a larva progeny of a female that was injected as a pupa with *Tc-disco* double stranded RNA. The normal Tribolium larval leg has, from proximal to distal, the coxa, trochanter, femur, tibiotarsus and pretarsal claw. *Tc-disco* RNAi legs (B, D,D') were markedly shorter than wild type (A,C,C') due to loss of the medial region. (C) A higher magnification of a wild type larval leg, and (C') a drawing that highlights specific leg segment markers. (D) High magnification view of an RNAi leg and (D') a drawing to highlight the changes. Specific sensory organs (arrow heads) and bristles (arrows) were used to determine the extent of the deletion. The pretarsal claw was present and well formed in the RNAi leg indicating that the distal-most portion of the leg was present. The large bristles of the coxa were also present. (Note the smaller bristle is out of the focal plane.) (E and E') wild type antenna. (F and F') antenna from RNAi embryos. Note the shorter structure of the antenna with reduced proximal to mid domain. (G) and (H) *Tc-dac* mRNA in wild type and *Tc-disco* RNAi developing legs, respectively. In (G) large arrowheads point to the major band of *Tc-dac* mRNA accumulation; smaller arrows denote lower intensity bands. In (H) the arrows point to the remaining reduced expression of *Tc-dac* after reducing *Tc-disco* via RNAi. Abbreviations: ptc, pretarsal claw; fe, femur; tr, trochanter; cx, coxa; a1-a4, antennal segments a1-a4; lb, labial segment; T1-T3, thoracic legs; A1, first abdominal segment with pleuropodia.

Figure 7



the leg deletion caused by reducing *Tc-disco* eliminates most, if not all, of the *Tc-dac* and medial regions of the Tribolium appendages.

### **The *disco* genes operate as appendage factors during *Drosophila* embryogenesis**

Though wormlike in appearance, *Drosophila* larvae are thought to have remnants of appendages in the thorax, the paired ventral Keilin's Organs (Angelini and Kaufman, 2005b; Bolinger and Boekhoff-Falk, 2005; Cohen et al., 1991). In addition, the cells that will give rise to the thoracic imaginal discs also arise from this region of the thorax (Bolinger and Boekhoff-Falk, 2005; Cohen et al., 1993; Cohen et al., 1991).

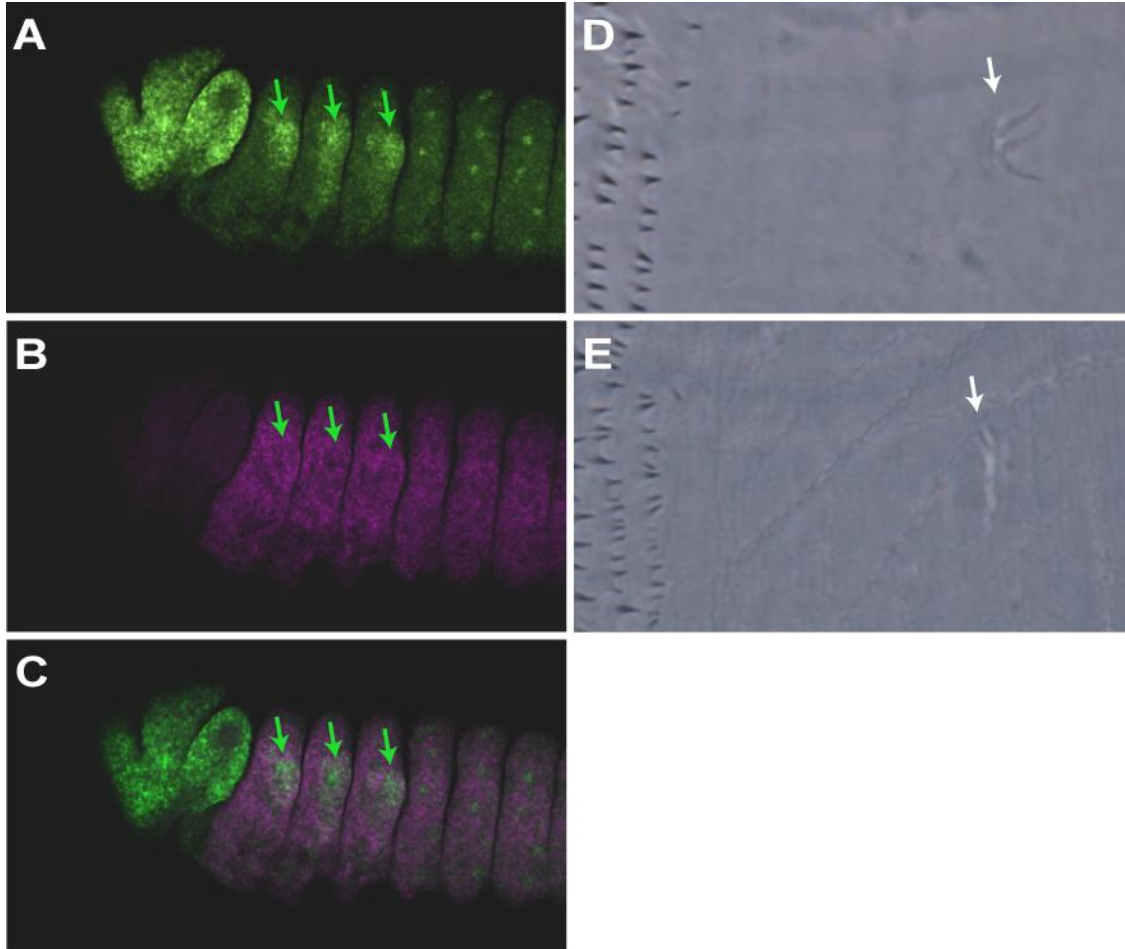
Development of the Keilin's Organs requires many of the appendage network genes. In the embryo, both *disco* genes are expressed in the region of the thorax where the Keilin's Organs appear (Fig. 8A-D), so we suspected these genes might also be required for their development. Indeed, it appeared that Keilin's Organ development was incomplete in *Df(1)ED7355* hemizygous embryos (Fig. 8E,F). An opening in the cuticle is present; however, the remainder of the organ is absent. Therefore, as in development of the adult appendages, the *disco* genes are necessary for development of the remnant of larval legs, the Keilin's Organs, indicating that, as in development of the adult, these genes have an appendage role.

## **DISCUSSION**

### ***disco* as an appendage factor**

The *Drosophila disco* gene was initially identified as required for proper neural migration (Steller et al., 1987), and our later work demonstrated that *disco* along with the paralog *disco-r* were required for embryonic pattern formation (Mahaffey et al., 2001; Robertson

**Figure 8**



**Figure 8.** Role of the *disco* genes during appendage development in embryos. (A-C) Comparison of *disco* (green) and *tsh* (magenta) mRNA distribution in stage 12 *Drosophila* embryos. Note the spots of *disco* accumulation in the thoracic segments (arrows). This expression marks the position where the Keilin's Organs and leg disc primordia are located. The *disco* genes are expressed in cells of both (Mahaffey et al., 2001). Note, further, that many of the central *disco*-positive cells are not expressing *tsh*. This arrangement of *disco* and *tsh* is similar to that in the third instar imaginal discs. (D,E) Keilin's organ phenotype in embryos lacking both *disco* genes. (D) Wild type third thoracic segment Keilin's Organ. Note the three bristles extending from the opening in the cuticle (arrows). (E) In embryos hemizygous for *Df(1)ED7355* the only visible sign of the Keilin's Organs are openings in the cuticle. There are no sockets or bristles. This may indicate that the outer regions of the organ develop normally, but that the inner components, those that would be more distal in an appendage scenario, are absent.

et al., 2004). Yet because mutations in *disco* affected migration of neurons during development of the adult, and work by Lee et al. (1991) indicated that *disco* was expressed in many of the imaginal discs, we suspected that the *disco* genes would have a role after embryogenesis. Furthermore, *disco-lacZ* from the enhancer trap line C50.1S1 has been used as a marker for leg joint formation (Bishop et al., 1999; Mirth and Akam, 2002). Here, we present evidence that *disco* genes are conserved members of the insect proximal/distal appendage specification network. In both *Drosophila* and *Tribolium*, these genes are expressed in the ventral appendages. Loss-of-function evidence from *Drosophila* and *Tribolium* indicates that the *disco* genes are required for ventral appendage development. Clonal analyses indicated that cells homozygous for *Df(1)ED7355* were lost from adult tissues that developed from *disco*-expressing regions of the ventral imaginal discs, though such cells were viable elsewhere. There is no evidence that any genes in *Df(1)ED7355* other than *disco* and *disco-r* affect viability or pattern formation. The proteins encoded by these other genes are unlikely to have such effects, and none are known to be expressed in a pattern similar to *disco* and *disco-r*. Therefore, it would be quite unlikely that one of these other genes would have an effect only in *disco* expressing regions while being viable elsewhere. Furthermore, our clonal analysis in *Drosophila* and RNAi experiments in *Tribolium* yielded complementary results. It is unlikely that two nonspecific processes could yield such similar results. Gain-of-function studies in *Drosophila* demonstrate that ectopic *disco* transforms dorsal appendages to ventral fates. Altogether, we feel this is strong evidence that the *disco* genes are ventral appendage factors. Though the C50.1S1 enhancer trap has been used by

others to mark the region of leg joint formation, we could not test for a direct role for the *disco* genes during leg joint formation, since the clones of *Df(1)ED7355* cells did not survive in the leg discs. However, we did not observe any indication of additional leg joints forming due to ectopic expression. The high levels of  $\beta$ -gal from the *disco-lacZ* enhancer trap C50.1S1 present in the leg joint regions may reflect the persistence of  $\beta$ -gal.

That cells lacking *disco* function either fail to proliferate or die when they are in the medial to distal portions of the appendage primordia implies that they are recognized as aberrant or inappropriately determined cells which are removed through autonomous actions or by their normal counterparts. Cell communication is important in establishing appendage regions (for example, Goto and Hayashi, 1999), and our observations demonstrate that there must be some form of communication between *disco*-expressing and non-expressing cells in the developing medial appendage region. Identifying the mechanism responsible is an important quest for the future.

Though we have not attempted to extensively address the regulatory interactions between the Disco proteins and other appendage factors, some insights are apparent. Tsh represses *disco* and *disco-r* during *Drosophila* embryogenesis (Robertson et al., 2004). That *disco* transcripts were not expressed throughout may indicate that Tsh represses the *disco* genes during appendage morphogenesis as well. In addition, that ectopic *disco* induced *dac* expression in the wing discs might indicate that *dac* is a target of Disco, though other explanations are certainly possible. At this time it is not known whether these regulatory events are direct.

### **The *Drosophila* embryo and appendage factors.**

During *Drosophila* embryogenesis, the *disco* genes act in parallel with the *hox* genes to establish proper segment identity in the head (Mahaffey, 2005; Mahaffey et al., 2001; Robertson et al., 2004). In this role, the *disco* genes are similar to the *teashirt* gene (*tsh*), which during embryogenesis encodes a trunk segment specification factor (de Zulueta et al., 1994; Fasano et al., 1991; Roder et al., 1992). Both the *disco* genes and *tsh* encode regionally expressed zinc finger transcription factors and, strictly from studies of *Drosophila* embryogenesis, they appear to establish zones along the anterior posterior axis of the embryo (*Disco*, head; *Tsh*, trunk) (Mahaffey, 2005; Robertson et al., 2004). Interaction between these two systems is evident in that *tsh* expression represses *disco* and *disco-r*, limiting their expression in the trunk segments (Robertson et al., 2004).

The newly discovered appendage role for the *disco* genes is intriguing in light of the segment specification function during *Drosophila* embryogenesis. As mentioned above, *tsh* also is required for proper proximal development of adult appendages. As we found with *disco*, the appendage role of *tsh* appears to be conserved, while the trunk segment specification role is found only during fly embryogenesis (Herke et al., 2005; Peterson et al., 1999). It is possible that the embryonic and appendage functions are distinct. If so, this would indicate that, for *disco*, either the head specification role was newly acquired in the fly lineage, or that it was lost in the beetle. Yet even during embryogenesis, we do find that the *disco* genes are required for the appendage primordium and for the Keilin's Organs, which are proposed to be remnants of larval appendages (Angelini and Kaufman, 2005b; Bolinger and Boekhoff-Falk, 2005; Cohen et

al., 1991). Therefore, even during *Drosophila* embryogenesis, *disco* is functioning as an appendage factor. But what about the expression in the head segments? To address this, we consider the differences between *Drosophila* and other insects as well as between the larval and adult forms of the fly. We suggest that, perhaps, *disco* is always an appendage factor, including during specification of the *Drosophila* larval head segments.

Most insects have well-formed appendages when they hatch as larva, but this is not the case for the worm-like larva of higher dipterans. In these insects visible appendages do not arise until the pupal stage when the adult body develops. Appendages arise from the imaginal discs, which are blocks of cells set-aside during embryogenesis. Certainly, reduction of the distal and medial appendage domains could account for the Keilin's Organs being derived from larval legs. Less obvious, but perhaps more significant in terms of novelty, are the changes that occurred to generate the internalized larval feeding apparatus of fly larvae. In *Drosophila* larvae the embryonic head segments are highly reduced and internalized, unlike most other insects (Diederich et al., 1991; Jurgens et al., 1986; Rogers and Kaufman, 1996). Perhaps, to form an internalized, multi-segmental feeding apparatus, the mouthpart appendages (mandibles, maxilla and labial palps) have been reduced, as occurred with the legs. However, instead of reducing the medial and distal portions of the appendage, perhaps in the head the proximal tissues were reduced so that the medial appendage domain, governed by *disco*, remains and is prominent in these head segments. In this regard, it is interesting to note that in less highly derived insects with external larval head appendages, homologs of *tsh* are expressed in the ventral portion of the head segments (Herke et al., 2005; Peterson et al.,

1999). If this model of the evolution of the *Drosophila* larval head is correct, then the *disco* genes would have analogous roles as appendage factors in the head and trunk segments, also uniting their roles in establishing appendages in the larva and adult.

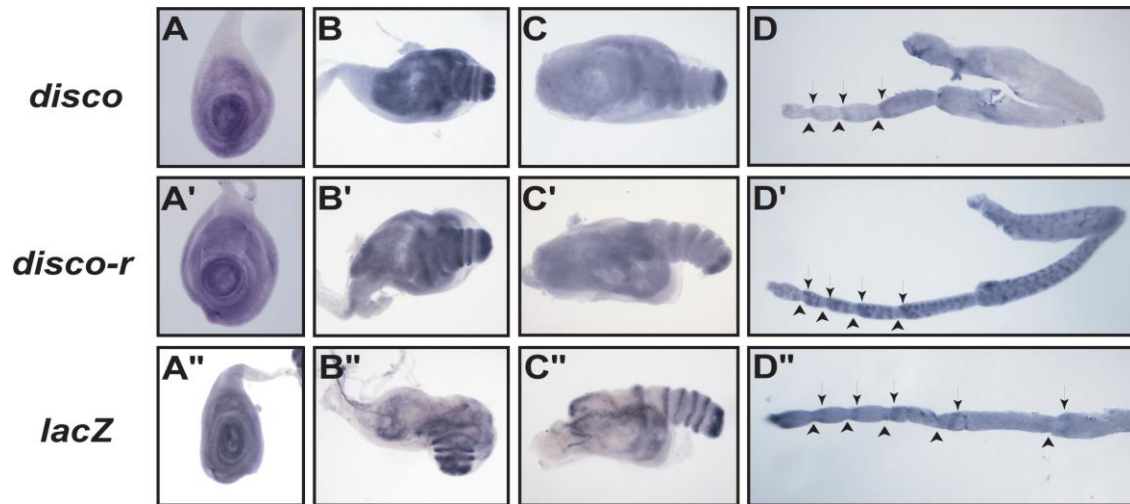
## ACKNOWLEDGEMENTS

We thank Dr. Amy Bejsovec (Duke University) for the *wingless* clone, Grace Boekhoff-Falk (University of Wisconsin) for the *dpp*-Gal4 flies and the Bloomington Indiana Drosophila Stock Center for *Df(1)ED7355* and other stocks used in this study. The Dachshund (Dac) antibody (Mardon et al., 1994) was obtained from the Developmental Studies Hybridoma Bank developed under the auspices of the NICHD and maintained by The University of Iowa, Department of Biological Sciences, Iowa City, IA 52242. We also thank Drs. Brown, Lorenzen and Shippy (Kansas State University) for advice on *Tribolium* and Drs. Ian Dworkin, Patricia Estes and Laura Mathies (Dept. of Genetics, NCSU) for helpful comments on the manuscript. LKR was supported by NIH pre-doctoral fellowship GM-08443-10. This work is supported by grant (IOB-0445540) from the Developmental Mechanisms Program of the National Science Foundation.

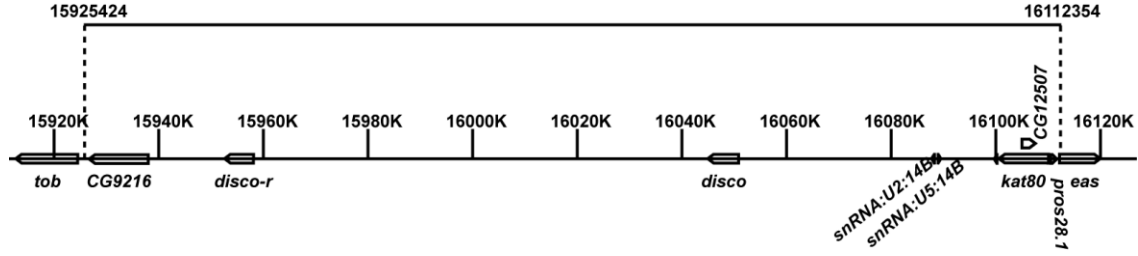
## SUPPLEMENTARY MATERIAL

**Supplemental Figure 1.** Expression of *disco* (A-D) and *disco-r* (A'-D') mRNAs in Ore-R third instar leg imaginal discs and everting prepupal legs compared to *lacZ* mRNA expression in C50-1S1 third instar leg imaginal disc and everting prepupal leg (A''-D''). In third instar leg imaginal discs *disco* expression was observed throughout the leg disc with rings of expression near the center of the disc (A). As the legs started to evert, strong concentric rings of expression were seen proximal to the constrictions that separate the leg segments (B). Later during leg eversion expression was much fainter and medial to the constrictions that will separate tarsals 1-5 (which are denoted by arrowheads) and the other leg segments; however, expression at the distal tip remained strong (C). In a fully everted leg, *disco* expression (arrow) was proximal to the constrictions that will form the joints (arrowhead) in the leg similar to the *lacZ* expression (D). *disco-r* expression was identical to *disco* expression in third instar leg imaginal discs and early everting legs (A'-C'); however, in addition to stripes of expression proximal to the constrictions that separate leg segments in a fully everted leg, there were spots of expression which appear to correspond to where each of the bristles would form (D'). *lacZ* expression in C50-1S1 was seen throughout the third instar leg disc with rings of expression around the center of the discs (A''). As the legs start to evert, *lacZ* expression was strong at the distal tip and became restricted to rings of expression of several cells wide proximal to the constrictions that separates tarsals 1-5 and the other segments of the leg (B'',C''). In a fully everted leg, *lacZ* expression was proximal to the constrictions that will form the joints in the leg (D''). Overall, it appears that the enhancer trap line has a similar expression pattern to *disco* and *disco-r* except for the strong rings of expression seen in the later stages of leg eversion which were not seen with *disco* or *disco-r*. This may be due to the persistence of *lacZ* mRNA or continued expression of the enhancer trap promoter. We note that previous studies have used a *disco* enhancer trap insertion, C50.1S1 as a marker for joint development in *Drosophila* legs (Bishop et al., 1999; Cohen et al., 1991; Couso and Bishop, 1998; Mirth and Akam, 2002). We have mapped this insertion (data available upon request) and found that it is closer to *disco-r* than *disco*, and though some aspects of *lacZ* accumulation mimic mRNA distribution from the *disco* genes, there are differences. The main difference is that *lacZ* mRNA is much more stable, so presents more the appearance of a role in joint formation. We are not able to address the role in joint development directly, because reducing *disco* function in either *Tribolium* or *Drosophila* prevents cells from contributing to the ventral appendages, so no joints can form. We do note that ectopic expression in the legs by the *dpp-Gal4* driver did not induce ectopic joints. In contrast, it eliminated joints in the tarsal region of the legs, though this is likely due to a change in developmental programming of this region as much as via an effect on joint development. It is also noteworthy that in the transformed wings, the legs that developed had appropriately positioned joints, giving rise to a well constructed leg. Further work will be required to ascertain the exact role of the *disco* genes in this later aspect of leg development.

Supplemental Figure 1

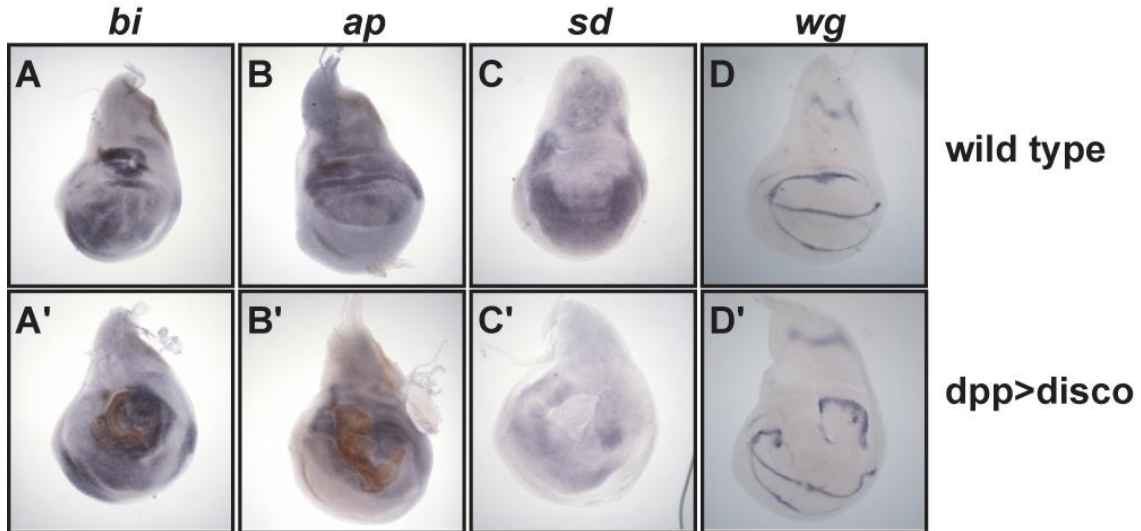


## Supplemental Figure 2



**Supplemental Figure 2.** Map of *Df(1)ED7355*. The breakpoints are shown at the top (15925424 to 16112354). Genes deleted or partially deleted in order from distal to proximal along the chromosome; *CG9216*, *disco-r*, *disco*, *snRNA:U2:14B*, *snRNA:U5:14B*, *CG12507*, *kat80* (katanin 80), *CG3415*, *Pros28.1* (Proteasome 28kd subunit 1) and *eas* (easily shocked). Note, none of these genes has been reported to be lethal when absent, except, together, *disco* and *disco-r*.

### Supplemental Figure 3



**Supplemental Figure 3.** Expression of wing determining genes in the transformed wings. Third row – normal wings; bottom row – *dpp>disco* transformed wings (denoted by '). (A,A') *bifid* expression (also known as *optimotor-blind*), (B,B') *apterous mRNA*. Both were co-stained with antibodies to *Dac* to facilitate identification of transformed wings. No change was noted in *bifid* or *apterous mRNA* distributions. (C,C') In contrast, *scalloped mRNA* was significantly reduced in the *dpp>disco* transformed wing. Surprisingly, the reduction in *scalloped mRNA* affected a broader region than we had detected ectopic *disco*, *disco-r*, *Dac* or *Dll*. This could indicate that a signaling mechanism was involved. (D,D') We also noted a change in the appearance of *wg mRNA* distribution. However, this likely was the result of morphological changes in the wing discs in response to ectopic *disco*.

## REFERENCES

- Abzhanov, A. and Kaufman, T. C.** (2000). Homologs of Drosophila appendage genes in the patterning of arthropod limbs. *Dev Biol* **227**, 673-89.
- Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D. J.** (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**, 3389-402.
- Angelini, D. R. and Kaufman, T. C.** (2005a). Comparative Developmental Genetics and the Evolution of Arthropod Body Parts. *Annu Rev Genet.* **39**, 95-119.
- Angelini, D. R. and Kaufman, T. C.** (2005b). Insect appendages and comparative ontogenetics. *Dev Biol.* **286**, 57-77.
- Beermann, A., Aranda, M. and Schroder, R.** (2004). The Sp8 zinc-finger transcription factor is involved in allometric growth of the limbs in the beetle *Tribolium castaneum*. *Development* **131**, 733-42.
- Berghammer, A., Bucher, G., Maderspacher, F. and Klingler, M.** (1999). A system to efficiently maintain embryonic lethal mutations in the flour beetle *Tribolium castaneum*. *Dev Genes Evol* **209**, 382-9.
- Bishop, S. A., Klein, T., Arias, A. M. and Couso, J. P.** (1999). Composite signalling from Serrate and Delta establishes leg segments in *Drosophila* through Notch. *Development* **126**, 2993-3003.
- Bolinger, R. A. and Boekhoff-Falk, G.** (2005). Distal-less functions in subdividing the *Drosophila* thoracic limb primordium. *Dev Dyn* **232**, 801-16.
- Brand, A. H. and Perrimon, N.** (1993). Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* **118**, 401-15.
- Bucher, G., Scholten, J. and Klingler, M.** (2002). Parental RNAi in *Tribolium* (Coleoptera). *Curr Biol* **12**, R85-6.
- Chen, R., Amoui, M., Zhang, Z. and Mardon, G.** (1997). Dachshund and eyes absent proteins form a complex and function synergistically to induce ectopic eye development in *Drosophila*. *Cell* **91**, 893-903.
- Cohen, B., Simcox, A. A. and Cohen, S. M.** (1993). Allocation of the thoracic imaginal primordia in the *Drosophila* embryo. *Development* **117**, 597-608.

- Cohen, B., Wimmer, E. A. and Cohen, S. M.** (1991). Early development of leg and wing primordia in the *Drosophila* embryo. *Mech Dev* **33**, 229-40.
- Cohen, S. M., Bronner, G., Kuttner, F., Jurgens, G. and Jackle, H.** (1989). Distal-less encodes a homoeodomain protein required for limb development in *Drosophila*. *Nature* **338**, 432-4.
- Cohen, S. M. and Jurgens, G.** (1989). Proximal-distal pattern formation in *Drosophila*: cell autonomous requirement for Distal-less gene activity in limb development. *EMBO J* **8**, 2045-2055.
- de Zulueta, P., Alexandre, E., Jacq, B. and Kerridge, S.** (1994). Homeotic complex and teashirt genes co-operate to establish trunk segmental identities in *Drosophila*. *Development* **120**, 2287-96.
- Diederich, R. J., Pattatucci, A. M. and Kaufman, T. C.** (1991). Developmental and evolutionary implications of labial, Deformed and engrailed expression in the *Drosophila* head. *Development* **113**, 273-81.
- Duncan, D. M., Burgess, E. A. and Duncan, I.** (1998). Control of distal antennal identity and tarsal development in *Drosophila* by spineless-aristopedia, a homolog of the mammalian dioxin receptor. *Genes Dev* **12**, 1290-303.
- Erkner, A., Gallet, A., Angelats, C., Fasano, L. and Kerridge, S.** (1999). The role of Teashirt in proximal leg development in *Drosophila*: ectopic Teashirt expression reveals different cell behaviours in ventral and dorsal domains. *Dev Biol* **215**, 221-32.
- Estella, C., Rieckhof, G., Calleja, M. and Morata, G.** (2003). The role of buttonhead and Sp1 in the development of the ventral imaginal discs of *Drosophila*. *Development* **130**, 5929-41.
- Fasano, L., Roder, L., Core, N., Alexandre, E., Vola, C., Jacq, B. and Kerridge, S.** (1991). The gene teashirt is required for the development of *Drosophila* embryonic trunk segments and encodes a protein with widely spaced zinc finger motifs. *Cell* **64**, 63-79.
- Gorfinkiel, N., Morata, G. and Guerrero, I.** (1997). The homeobox gene Distal-less induces ventral appendage development in *Drosophila*. *Genes Dev* **11**, 2259-71.
- Goto, S. and Hayashi, S.** (1999). Proximal to distal cell communication in the *Drosophila* leg provides a basis for an intercalary mechanism of limb patterning. *Development* **126**, 3407-13.
- Halder, G., Callaerts, P. and Gehring, W. J.** (1995). Induction of ectopic eyes by targeted expression of the eyeless gene in *Drosophila*. *Science* **267**, 1788-92.

- Herke, S. W., Serio, N. V. and Rogers, B. T.** (2005). Functional analyses of tiptop and Antennapedia in the embryonic development of *Oncopeltus fasciatus* suggests an evolutionary pathway from ground state to insect legs. *Development* **132**, 27-34.
- Huang, X.** (1992). A contig assembly program based on sensitive detection of fragment overlaps. *Genomics* **14**, 18-25.
- Jockusch, E. L., Williams, T. A. and Nagy, L. M.** (2004). The evolution of patterning of serially homologous appendages in insects. *Dev Genes Evol* **214**, 324-38.
- Jurgens, G., Lehmann, R., Schardin, M. and Nusslein-Volhard, C.** (1986). Segmental organisation of the head in the embryo of *Drosophila melanogaster*. A blastoderm fate map of the cuticle structures of the larval head. *Roux Arch. dev. Biol.* **195**, 359--377.
- Klein, T. and Campos-Ortega, J. A.** (1997). klumpfuss, a *Drosophila* gene encoding a member of the EGR family of transcription factors, is involved in bristle and leg development. *Development* **124**, 3123-34.
- Kojima, T.** (2004). The mechanism of *Drosophila* leg development along the proximodistal axis. *Dev Growth Differ* **46**, 115-29.
- Kosman, D., Mizutani, C. M., Lemons, D., Cox, W. G., McGinnis, W. and Bier, E.** (2004). Multiplex detection of RNA expression in *Drosophila* embryos. *Science* **305**, 846.
- Lee, K. J., Freeman, M. and Steller, H.** (1991). Expression of the disconnected gene during development of *Drosophila melanogaster*. *EMBO J* **10**, 817-26.
- Lee, K. J., Mukhopadhyay, M., Pelka, P., Campos, A. R. and Steller, H.** (1999). Autoregulation of the *Drosophila* disconnected gene in the developing visual system. *Dev Biol* **214**, 385-98.
- Lee, T. and Luo, L.** (2001). Mosaic analysis with a repressible cell marker (MARCM) for *Drosophila* neural development. *Trends Neurosci* **24**, 251-4.
- Mahaffey, J. W.** (2005). Assisting Hox proteins in controlling body form: are there new lessons from flies (and mammals)? *Curr Opin Genet Dev* **15**, 422-9.
- Mahaffey, J. W., Griswold, C. M. and Cao, Q. M.** (2001). The *Drosophila* genes disconnected and disco-related are redundant with respect to larval head development and accumulation of mRNAs from deformed target genes. *Genetics* **157**, 225-36.
- Mardon, G., Solomon, N. M. and Rubin, G. M.** (1994). dachshund encodes a nuclear protein required for normal eye and leg development in *Drosophila*. *Development* **120**, 3473-86.

- Mirth, C. and Akam, M.** (2002). Joint development in the *Drosophila* leg: cell movements and cell populations. *Dev Biol* **246**, 391-406.
- Morata, G.** (2001). How *Drosophila* appendages develop. *Nat Rev Mol Cell Biol* **2**, 89-97.
- Panganiban, G.** (2000). Distal-less function during *Drosophila* appendage and sense organ development. *Dev Dyn* **218**, 554-62.
- Pederson, J. D., Kiehart, D. P. and Mahaffey, J. W.** (1996). The role of HOM-C genes in segmental transformations: reexamination of the *Drosophila* Sex combs reduced embryonic phenotype. *Dev Biol* **180**, 131-42.
- Peterson, M. D., Rogers, B. T., Popadic, A. and Kaufman, T. C.** (1999). The embryonic expression pattern of labial, posterior homeotic complex genes and the teashirt homologue in an apterygote insect. *Dev Genes Evol* **209**, 77-90.
- Prpic, N. M., Wigand, B., Damen, W. G. and Klingler, M.** (2001). Expression of dachshund in wild-type and Distal-less mutant *Tribolium* corroborates serial homologies in insect appendages. *Dev Genes Evol* **211**, 467-77.
- Robertson, L. K., Bowling, D. B., Mahaffey, J. P., Imiolczyk, B. and Mahaffey, J. W.** (2004). An interactive network of zinc-finger proteins contributes to regionalization of the *Drosophila* embryo and establishes the domains of HOM-C protein function. *Development* **131**, 2781-9.
- Robertson, L. K., Dey, B. K., Campos, A. R. and Mahaffey, J. W.** (2002). Expression of the *Drosophila* gene disconnected using the UAS/GAL4 system. *Genesis* **34**, 103-6.
- Roder, L., Vola, C. and Kerridge, S.** (1992). The role of the teashirt gene in trunk segmental identity in *Drosophila*. *Development* **115**, 1017-33.
- Rogers, B. T. and Kaufman, T. C.** (1996). Structure of the insect head as revealed by the EN protein pattern in developing embryos. *Development* **122**, 3419-32.
- Ryder, E., Blows, F., Ashburner, M., Bautista-Llacer, R., Coulson, D., Drummond, J., Webster, J., Gubb, D., Gunton, N., Johnson, G. et al.** (2004). The DrosDel collection: a set of P-element insertions for generating custom chromosomal aberrations in *Drosophila melanogaster*. *Genetics* **167**, 797-813.
- Stahling-Hampton, K., Jackson, P. D., Clark, M. J., Brand, A. H. and Hoffmann, F. M.** (1994). Specificity of bone morphogenetic protein-related factors: cell fate and gene expression changes in *Drosophila* embryos induced by decapentaplegic but not 60A. *Cell Growth Differ* **5**, 585-93.

**Steller, H., Fischbach, K. F. and Rubin, G. M.** (1987). Disconnected: a locus required for neuronal pathway formation in the visual system of *Drosophila*. *Cell* **50**, 1139-53.

**Tautz, D. and 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**, 81-5.

**Tseng, H. and Green, H.** (1992). Basonuclin: a keratinocyte protein with multiple paired zinc fingers. *Proc Natl Acad Sci U S A* **89**, 10311-5.

**Weigmann, K. and Cohen, S. M.** (1999). Lineage-tracing cells born in different domains along the PD axis of the developing *Drosophila* leg. *Development* **126**, 3823-30.

**Williams, T. A. and Nagy, L. M.** (2001). Developmental modularity and the evolutionary diversification of arthropod limbs. *J Exp Zool* **291**, 241-57.

**Wu, J. and Cohen, S. M.** (2000). Proximal distal axis formation in the *Drosophila* leg: distinct functions of teashirt and homothorax in the proximal leg. *Mech Dev* **94**, 47-56.

**Xu, T. and Rubin, G. M.** (1993). Analysis of genetic mosaics in developing and adult *Drosophila* tissues. *Development* **117**, 1223-37.

## CHAPTER FOUR

### Functional analysis of the Disconnected protein

## Introduction

Previously we identified two redundant genes, *disconnected* (*disco*) and *disco-related* (*disco-r*), that are genetic cofactors with the HOM-C genes *Deformed* (*Dfd*) and *Sex combs reduced* (*Scr*) (Mahaffey et al., 2001; Robertson et al., 2004). *disco* and *disco-r* along with *Dfd* and *Scr* coordinately regulate target gene expression within the gnathal lobes. However, there is not much functional information of how *Disco* and *Disco-r* work to regulate target genes. To help our understanding it is useful to know how the *Disco* proteins work at the molecular level.

*disco* and *disco-r* encode conserved C2H2 zinc finger transcription factors. C2H2 zinc fingers are one of the most prevalent protein motifs in eukaryotic genomes; they are characterized by the ability to coordinate a zinc ion through the cysteine and histidine residues. In addition to the conserved cysteines and histidines, C2H2 zinc fingers contain conserved hydrophobic residues that help form the characteristic secondary structure consisting of an alpha-helix and a two-stranded antiparallel beta-sheet (Ryan and Darby, 1998). The consensus sequence for C2H2 zinc fingers is C-X<sub>2-5</sub>-C-X<sub>3</sub>-F/Y-X<sub>8</sub>-H-X<sub>3-5</sub>-H. Proteins that contain C2H2 zinc fingers are capable of regulating transcription by recognizing specific DNA sequences. The expansion of the C2H2 motif as organism complexity increases demonstrates the versatility of the motif and the ability to recognize multiple DNA sequences. Because of their ability to act as transcriptional regulators, proteins that contain zinc fingers also tend to be highly conserved, and *Disco* is no exception (Knight and Shimeld, 2001; Philipson and Suske, 1999). Knight and Shimeld (2001) demonstrated that *Disco* was evolutionarily conserved between *Drosophila*, *C.*

*elegans*, and *Homo sapiens*.

We present an alignment with more ecdysozoans, a lophotrochozoan *Aplysia californica*, and some deuterostomes to identify functionally important residues of the Disco zinc finger region. It has been well documented that the most important residues for DNA recognition by C2H2 zinc fingers are the residues at -1, 2, 3, and 6 positions relative to the start of the alpha-helix (Philipsen and Suske, 1999). If we observe lineage specific differences at these positions between deuterostomes compared to the ecdysozoa and lophotrochozoa, that residue may be functionally important. We demonstrate that the ecdysozoa and lophotrochozoa have a conserved cysteine, not found within deuterostomes, which appears to be functionally significant.

In addition to identifying functionally significant residues, we present a DNA recognition site for the zinc finger region. It has previously been shown that Disco is capable of binding DNA, though no binding site has been presented (Knight and Shimeld, 2001; Lee et al., 1999). Finally, we'll describe studies indicating that Disco can act as a potential repressor through its interaction with the corepressor CtBP.

## **Material and Methods**

### **Electromobility Shift Assays**

A fragment of the *reaper* enhancer region, *rpr4S3* was kindly provided by Ingrid Lohmann and William McGinnis. Approximately 2 ug of the *rpr4S3* fragment was digested with Pci I at 37°C and then precipitated and resuspended in dH<sub>2</sub>O. The digested product was end labeled using <sup>32</sup>P-CTP. The labeled fragment was purified by using the Qiagen PCR purification kit and resuspended in 30 ul dH<sub>2</sub>O. The resulting labeled DNA

was digested with the appropriate enzyme at 37°C and then run on a 5% nondenaturing polyacrylamide gel. The DNA was then gel extracted and eluted overnight in TE (10 mM Tris pH 7.5, 1 mM EDTA) 2 volumes of the gel slice at 37°C. The eluted labeled DNA was passed through a column of siliconized glass wool and then precipitated by adding 2 volumes of 100% EtOH. The DNA was resuspended in 200 ul of TE and precipitated again using 25 ul of 3M NaOAc pH 5.2 and 2 volumes of 100% EtOH. The pellet was washed with 70% EtOH and resuspended in an appropriate amount of sterile dH<sub>2</sub>O such that there was approximately 2000 cpm/ul of probe. The electromobility shift assays were carried out in 30 ul reactions containing 20 mM HEPES pH 7.5, 12% glycerol, 1 mM MgCl<sub>2</sub>, 15 mM DTT, 50 uM ZnCl<sub>2</sub>, 50 mM KCl, 1 ug dI/dC and 2,000 cpm of labeled probe. Various amounts of protein were added to the reactions and incubated at 4°C for 30 min. Sucrose loading dye was then added to each reaction and loaded onto a TBE pH 8.3 (89 mM Tris, 89 mM boric acid, 2 mM EDTA) 8% nondenaturing polyacrylamide gel which had been poured the day before and kept at 4°C. The gel was run in 0.5X TBE at 4°C. The gel was then placed on 3MM Whatman paper, dried, and placed onto X-ray film with an intensifier screen. The gel was exposed to the film overnight at -80°C and developed the next day.

### **DNase I footprinting**

For DNase protection assays end label DNA of ~1200 cpm was incubated with varying amounts of protein, 1ul of 1mg/ml of sonicated salmon sperm DNA, binding buffer (100 mM KCl, 12.5 mM MgCl<sub>2</sub>, 20% glycerol, 50 mM HEPES pH 7.5, 2 mM DTT), and sterile dH<sub>2</sub>O in a 25 ul reaction. The reaction was allowed to sit on ice for 1 hr at which

point 50 ul of 10 mM MgCl<sub>2</sub>, 5mM CaCl<sub>2</sub> solution was added and incubated at room temp. for 2 min. Three microliters of DNase I at 5:95 dilution (no protein) or 5:195 dilution (reactions with protein) of 1mg/mL DNase I was added to each reaction. The reactions were incubated for 15 min at 4°C at which point 75 ul of stop DNase (20 mM EDTA, 1% SDS, 0.2M NaCl, and 125 ug/ml yeast tRNA) was added. The DNA in the reaction was phenol: chloroform extracted, precipitated with 2.5 volumes 100% EtOH at -80°C, and washed with 70% EtOH, dried, and resuspended in 10 ul of sequencing dye. The resulting DNA fragments were run on a 8% 19:1 (acrylamide: bis-acrylamide) sequencing polyacrylamide gel. The gel was then placed on 3MM Whatman paper, dried, and placed onto X-ray film with an intensifier screen. The gel was exposed to the film for various lengths of times at -80°C and developed.

### **Protein production and purification**

Disco, Cys 102, and Cys 94-1 proteins were created using bacterial and lysate systems. For the bacterial system, the Eco RI-Not I fragment of each construct, which contains the zinc finger region and the N-terminal end, was cloned in the pET28a vector. BL21 cells that contained the pLysS vector were transformed with the various *disco* pET28a constructs. To produce protein, cultures of BL21 *disco*-pET28a cells were grown to an OD=0.6 at which point IPTG was added to a final concentration of 1 mM to begin induction of protein synthesis. The cells were induced for 3 hrs then spun down at 5,000 rpm for 5 min, and the pellets were frozen overnight. The cells were lysed by resuspending the pellet in 1/10 volume of the culture in 50 mM Tris-HCl pH 8.0, 2 mM EDTA that had a Complete Tablet (Roche) added to it and 1/10 of the current volume of

1% Triton X-100. The suspension was incubated at room temperature for 30 min. and then sonicated on ice using brief 10 sec pulses until the solution lost viscosity. The suspension was centrifuged for 15 min. at 12,000 rpm at 4°C. The supernatant was transferred to a new tube and a 20%-30% ammonium sulfate cut was done. The resulting pellet was resuspended in 50 mM Tris-HCl pH 8.0, 2 mM EDTA. The solution was then dialyzed against 20 mM HEPES pH 7.5, 12% glycerol, 1 mM MgCl<sub>2</sub>, 5 mM DTT, 50 μM ZnCl<sub>2</sub>, 50 mM KCl. For the lysate system the Eco RI-Not I fragment of *disco* and *Cys 94-1* were cloned in the pCITE4a vector. The TnT coupled reticulocyte lysate system protocol by Promega was used to produce protein. In both cases production of protein was verified by Western Blots.

#### **Site directed mutagenesis**

Site directed mutagenesis was carried out as described in Seyfang and Jin (2004). Briefly, 10 ul of 50 μM anchor primers were mixed with 3.4 ul of 50 μM mutagenic primer. Two microliters of this primer pool was phosphorylated in 25 ul of kinase buffer using T4 polynucleotide kinase. Afterwards a 20 ul reaction was set up that contained 80 ng of alkali denatured *disco* pBS, 5 ul of phosphorylated primer mix, 2 ul 10X annealing buffer and the rest dH<sub>2</sub>O. The primers were annealed by heating the mixture to 75°C for 5 min. followed by gradual cooling to 45°C within 30 min (-1°C/min.) and to 22°C within 12 min. (-2°C/min.) in a thermocycler. The subsequent synthesis reaction was carried out by adding 3 ul 10X synthesis buffer, 5 ul dH<sub>2</sub>O, 1 ul T4 DNA polymerase, and 1 ul T4 DNA ligase to the 20 ul anneal template-primer mixture, which was incubated at 37°C for 2 hours. The mutagenized strand was amplified by using 8 ul of

the template mutant strand in a standard 100 ul PCR reaction that contained 2 ul each of 50 uM PCR 5' anchor and PCR 3' anchor. Primers used for this experiment are as follows. For mutating the CtBP site in Disco: Disco CtBP 5' anchor: 5'-TATGTGCTCCATGAATTCCTCGAGTCTAGAATCGAT-3' Disco CtBP 3' anchor: 5'-GGATCCAGATCTAAGCTTTTAGACGGAAATCGGCGC-3' Disco CtBP mut: 5'-GCTTGTGAAGCCTGGCCTCGAGCTTCTGGTCAGCT-3' PCR 5' anchor: 5'-ATCGATTCTAGACTCGAG-3' PCR 3' anchor: 5'-GGATCCAGATCTAAGCTT-3'.

For mutating the Disco sites in the rpr4S3 fragment: 5' rpr anchor: 5'-TTTGTCTGGGACTGCCTCCTCGAGTCTAGAATCGAT-3' 3' rpr anchor: 5'-GGATCCCAGATCTAAGCTTGTATGTCCTTCGCGGTAA-3' disco mut primer 1: 5'-GCGGGTAATTATGGGAGTTTTATTC-3' disco mut primer 2: 5'-TAACTTGGGAGTTGTGG-3' disco mut primer 3: 5'-TTGAGTTGGGAGAATTGG-3' disco mut primer 4: 5'-CTTGTGATGGGAGTCTGTC-3' disco mut primer 5: 5'-GGGAAGTGGGAGAATCG-3' . For Cys 102 and Cys 94-1 constructs the site was mutated by amplifying a fragment of the *disco* ORF using the disco\_frame primer 5'-GGATCCGAATTCATGGAGCACATAATG-3' and either Cys 94-gly primer 5'-CTTAAGGGCGCCCTTGTCGCAGAAGGTCTTGAAGCAAATGGAGCCCTGTAC-3' or disco\_cys102gly\_reverse primer 5'-CTTAAGGGCGCCCTTGTCGCCGAAGGTCTT-3'. The amplified fragments were then cloned in the pGEM T easy vector. The Cys 102 pGEM construct was subsequently digested with Eco RI and Afl II and cloned into disco pBS that had also been digested with Eco RI and Afl II. The Cys 94 pGEM construct was also digested with Eco RI and

Afl II, but was cloned into the *disco*<sup>1</sup> pBS construct which contains the *disco*<sup>1</sup> mutation that had also been digested with Eco RI and Afl II.

### **Trangemics**

To create UAS lines, p-element mediated germ-line transformation was carried out essentially as described in Rubin and Spradling (1982) (Rubin and Spradling, 1982). The Cys 102 and Cys 94-1 pBS constructs were digested with Eco RI and Xba I and cloned in the pUAST. Disco CtBP<sup>+</sup> was digested with Eco RI and cloned in pUAST.

### **Yeast 2 Hybrid Assay**

For Disco and Disco CtBP constructs, *disco* pBS was put through the site directed mutagenesis protocol with or without the Disco CtBP mut primer. The resulting products were cloned in the pGEM-T easy vector. The products were cut with Eco RI and cloned in the yeast pB42AD vector. For CtBP, the CtBP pBTM116 clone was kindly provided by Susan Parkhurst, which was digested with Bam HI and cloned in the yeast pLexA vector. Sequential transformations were done according to the Matchmaker LexA Two Hybrid System manual (Clontech). First, EGY48[p8op-lacZ] cells were transformed with CtBP pLexA. Then cells that carried the CtBP pLexA vector were transformed with either Disco CtBP<sup>+</sup> pB42AD or Disco pB42AD. Interaction was tested by plating on SD-galactose –His Trp Ura Leu plates.

### **Results**

Disco contains two C2H2 zinc fingers that are separated by a six amino acid spacer (Heilig et al., 1991). Alignment of the Disco zinc finger region with the homologs of Disco in *Drosophila melanogaster*, *Tribolium castaneum*, *Anopheles gambiae*, *Bombyx*

*morii*, *Aplysia californica*, *Caenorhabditis elegans*, *Homo sapiens*, *Mus musculus*, *Danio rerio*, and *Ciona intestinalis* demonstrate that the Disco zinc finger region is highly conserved (Fig. 1). Among the ecdysozoa (members of the animal kingdom that produce and shed cuticles), there is 87% identity within the zinc finger region. Among vertebrates that percentage increases to 94%. There are two interesting differences between deuterostomes and the ecdysozoa/lophotrochozoa in the first zinc finger, which may be functionally significant. The first is the switch from an alanine in ecdysozoa and lophotrochozoa to a threonine in deuterostomes. The structurally nonconservative alanine to threonine switch is found at position 3 which interacts with DNA (Fig. 1). The other change is the presence of a cysteine at position 102 of the Disco protein that is found in examined ecdysozoa and in the mollusk *Aplysia* but not within deuterostomes (Fig. 1). Based on our alignment all the residues in the second zinc finger at the positions critical for DNA recognition are conserved.

### **Significance of Alanine to Threonine change**

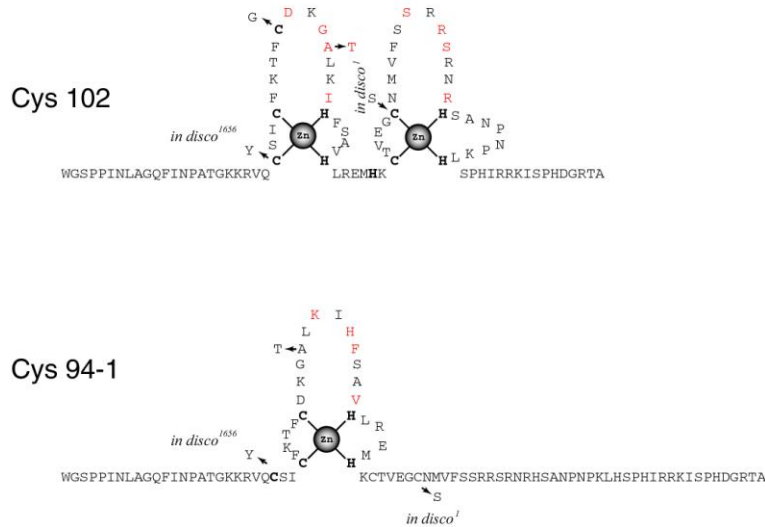
At position 3 in the first zinc finger there is an alanine in the ecdysozoa plus lophotrochozoa and a threonine in deuterostomes. In order to test if this change is significant we altered the alanine of Disco to threonine using a PCR based method, which was subsequently cloned in a UAS vector to create transgenic flies. The UAS-Gal4 system (Brand and Perrimon, 1993) was then used to express this variant to determine the functional significance of the change. Wild type UAS-Disco and UAS-Disco A->T were crossed to *prd*-Gal4, which drives expression in the posterior half of each alternating segment, and cuticle phenotypes were examined.

In wild type larval development as the germ band retracts the leading edge, which are cells that abut the amnioserosa, undergoes cell shape changes and elongations ultimately leading to an epidermal layer of cells enclosing the dorsal surface of the embryo. Once dorsal closure is complete the epidermis completely envelops the embryo and secretes cuticle (Scuderi and Letsou, 2005). In *prd*>>Disco embryos, the affected segments fail to undergo dorsal closure (Robertson et al., 2004). The resulting cuticle phenotype is one in which the affected segments fail to fuse at the dorsal midline leading to a disruption of the dorsal hairs of the cuticle (Fig. 2). The *prd*>>Disco A->T phenotype was identical to *prd*>>Disco (data not shown). Thus the change from A->T did not seem to affect Disco function in the *Drosophila* embryo. However, this change may be more significant for the vertebrate lineages where phosphorylation plays a larger role in regulatory mechanisms than in invertebrates (Dovat et al., 2002; Flint and Jones, 1991).

### **Potentially different isoforms of the Disco protein**

In addition to the alanine to threonine switch, we observed another significant difference between deuterostomes and the ecdysozoa plus lophotrochozoa. The ecdysozoa and lophotrochozoa contain a cysteine rather than a tyrosine at position 102 of the *Drosophila* Disco protein. Interestingly, this provides the potential to make different structural isoforms of the Disco zinc finger region (Fig. 1). With this conserved cysteine, the ecdysozoa and lophotrochozoa are capable of forming either a two zinc finger form or a single zinc finger form. Since the two-finger form is conserved, we expect this to be the main functional structure of the protein. However, to test this and to determine whether

KKRVQCSICFKTFCDKGALKIHFSAVHLREMHKCTVEGCNMFSSRRSRNRHSANPNPKLHSPHIRRKISPHDGRTAQ	D.m. disco
KKRVQCNVCLKTFCDKGALKIHFSAVHLREMHKCTVDGCSMMFSSRRSRNRHSANPNPKLHSPHLRRKISPHDGRSAQ	D.m. disco-r
KKRVQCNVCLKTFCDKGALKIHFSAVHLREMHKCTVEGCNMFSSRRSRNRHSANPNPKLHSPHLRRKISPHDGRSAQ	T.c.
KKRVQCNVCFKTFCDKGALKIHFSAVHLREMHKCTVEGCNMFSSRRSRNRHSANPNPKLHSPHLRRKISPHDGRSAQ	A.g.
KKRVQCNVCLKTFCDKGALKIHFSAVHLREMHKCTVEGCNMFSSRRSRNRHSANPNPKLHSPHLRRKISPHDGRSAQ	B.m.
KRRVACDICSKSFCDKGALKIHFSAVHLREMHKCTVTGCGKQFSSRRSRNRHSNNPKLHMPESLTL	C.e.
KKRVLCACNKTFCDKGALKIHYSAVHLKEMHKCTVEGCNMFSSRRSRNRHSANPNPKLHMPQKRKD	A.c.
KGRVFCNACGKTFYDKGTLKIHYNAVHLKIKHRCTIEGCNMFSSRLSRNRHSANPNPRLHMPMLRNN	H.s. Bnc2
KGRVFCNACGKTFYDKGTLKIHYNAVHLKIKHRCTIEGCNMFSSRLSRNRHSANPNPRLHMPMLRNN	M.m. Bnc2
KGRVFCNACGKTFYDKGTLKIHYNAVHLKIKHRCTIEGCNMFSSRLSRNRHSANPNPRLHMPMLRNN	H.s. Bnc
KGRVFCNACGKTFYDKGTLKIHYNAVHLKIKHRCTIEGCNMFSSRLSRNRHSANPNPRLHMPMLRNN	M.m. Bnc
KGRVFCNACGKTFYDKGTLKIHYNAVHLKIKHRCTIEGCNMFSSRLSRNRHSANPNPRLHMPMLRNN	D.r.
TGPVQCNICRMYSNKGLTRVHFKSVHLREMHQCTVPGCDMMFTSVRSRNRHSQNPN--LHRSIKI	C.i.



### Figure 1. Disco is an evolutionarily conserved protein.

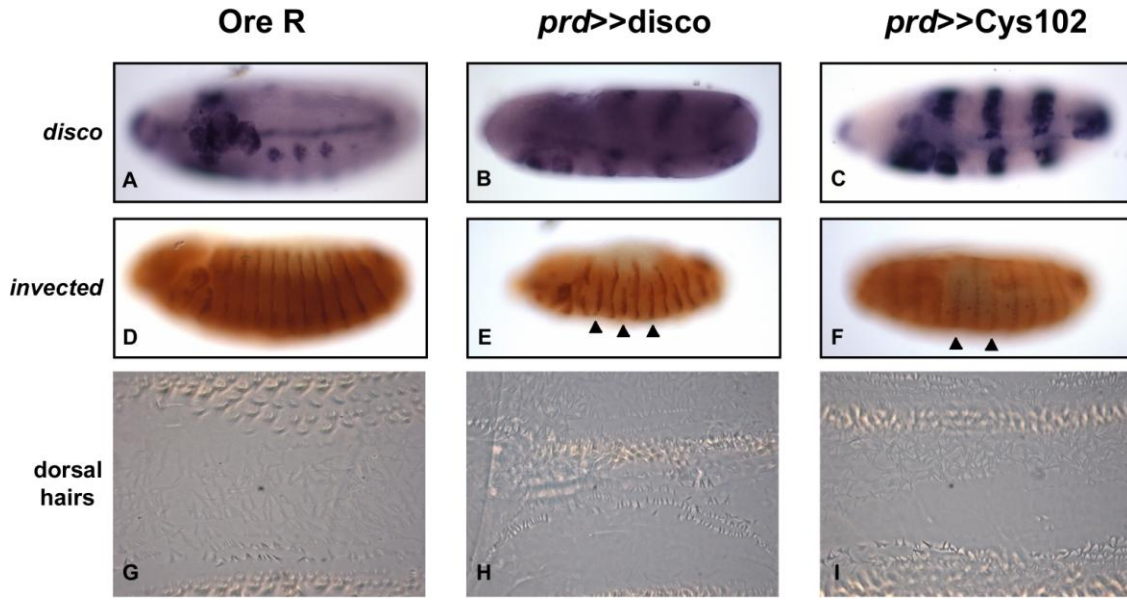
(Top) Alignment of the zinc finger region of Disco to *Drosophila melanogaster* (*D.m.*), *Tribolium castaneum* (*T.c.*), *Anopheles gambiae* (*A.g.*), *Bombyx moryii* (*B.m.*), *Caenorhabditis elegans* (*C.e.*), *Aplysia californica* (*A.c.*), *Homo sapien* (*H.s.*), *Mus musculus* (*M.m.*), *Danio rerio* (*D.r.*), and *Ciona intestinalis* (*C.i.*). The conserved cysteines and histidines are in red. -1, 2, 3, 6 indicate the important amino acids in DNA recognition. Light blue indicates deuterosomes and no color indicates the ecdysozoans and lophotrochozoans. Notice that the ecdysozoa and lophotrochozoa have a conserved cysteine that is capable of forming an alternate zinc finger as well as an alanine at position 3 of DNA contact residues that is not found with deuterosomes. These differences are highlighted in yellow. (Bottom) The potential zinc finger forms that can be created by Disco. The Cys at position 102 was altered to a glycine to create the Cys102 isoform which contains two zinc fingers and the Cys at position 94 was altered to a glycine and combined with the *disco*<sup>1</sup> mutation to create the Cys 94-1 isoform.

the single finger form is functional we changed certain cysteines so only one structure could be formed. In order to restrict the structure to the two zinc finger form, the cysteine at position 102 was changed to a glycine (Fig. 1). This is referred to as Cys102 below. To create the single zinc finger form, referred to as Cys 94-1, the cysteines at positions 94 and 127 were changed to a glycine and serine, respectively (Fig. 1), blocking formation of the conserved two zinc finger structures while maintaining the ability to form the alternative single finger form. A protein containing just a Cys 94 change still has the potential to form a double zinc finger that differs from Cys102. However, this structure is unlikely since only one amino acid would separate the first zinc finger from the second. To test the functional significance of these isoforms, each was cloned into pUAST and transgenic lines were created. The lines were crossed to the *prd-gal4* driver and compared to *prd>>Disco* whose phenotype was described above.

The *prd>>Cys 94-1* lines had no phenotype; the larval cuticular pattern appeared to be wild type and all larvae hatched indicating that the single zinc finger form has no function in this assay. The *prd>>Cys 102* lines produced a phenotype, but one significantly weaker than normal *prd>>Disco*. Normal *prd>>Disco* larva do not hatch, but by contrast most *prd>>Cys102* larva hatched. Some died during later larval stages, but many survived to pupation and even eclosed (Table 1). This was true for all four independent lines tested, though data for only two of these lines are presented in Table 1. The larva that did not hatch had very weak or no detectable dorsal closure defects (Fig. 2). This is supported by *in situ* detection of the engrailed/injected antibody to these embryos (Fig. 2). Engrailed is expressed in a dorsal-ventral stripe about two cells wide

Table 1. Hatch rate on *UAS-Cys102/Cyo X prd-GAL4/prd-GAL4*. Unlike *UAS-Disco X prd-GAL4*, where none of the ectopically expressing larva hatch, most of ectopically expressing Cys102 larva hatch. The hatch rate should be about 50% since the *Cyo/prd-GAL4* larva should hatch, pupate, and eclose. However, the overall hatch rate was 87.5% for line 1 and 79% for line 2. Based on the numbers some larva died during larval stages and some during pupation indicating that ectopic Cys102 can be lethal, but is much weaker than ectopic Disco.

	Hatched	Pupated	Eclosed	Straight Winged
Line 1	350/400	261	178	47
Line 2	316/400	219	158	24



**Figure 2. Ectopic expression of the different isoforms of Disco.**

Embryos are positioned such that anterior is to the left and dorsal is up. (A) Wild type *disco* expression in an Ore R embryo. (B) *prd*>>*disco* expression, note that *prd* only drives expression in alternating segments. (C) *prd*>>*Cys102* expression is also expressed in alternating segments similar to *prd*>>*disco*. (D) Wild type *Invected* pattern in an Ore R embryo. Expression in all segments extends dorsally during dorsal closure. (E) In *prd*>>*disco* the affected segments fail to undergo dorsal closure and *Invected* expression does not migrate dorsally (arrowheads). (F) Later embryo of *prd*>>*Cys102*, note that only some of the affected segment fails to undergo dorsal closure. (G) Wild type dorsal hair pattern of a 1<sup>st</sup> instar larva. (H) Dorsal hair pattern of *prd*>>*disco* larva. (I) The dorsal hair pattern of *prd*>>*Cys102* larva is less disrupted than *prd*>>*disco*. *prd*>>*Cys94-1* phenotype is wild type.

and marks the posterior compartment of each segment (DiNardo et al., 1985). In *prd>>Disco* embryos the engrailed stripes in the affected segments fail to extend dorsally (Robertson et al., 2004). This is not the case in the majority of *prd>>Cys102* embryos. In most *prd>>Cys102* embryos, Engrailed expression extended dorsally and fused at the dorsal midline. There were instances where engrailed extended but did not fuse at the dorsal midline (Fig. 2).

In *prd>>Disco* embryos, failure to undergo dorsal closure leads to a disruption of the dorsal hairs (Fig. 2). This disruption was very slight in *prd>>Cys102* embryos when it occurred, but often there was no disruption (Fig. 2). This suggests that the double zinc finger form is probably the functionally active form of Disco, but it also implies that the cysteine at position 102 is required for complete functionality, since the form lacking this cysteine was not able to recapitulate the wild type Disco function. Further evidence that the double zinc finger form is probably the functionally active form comes from DNA binding assays described below.

### **DNA binding by the zinc finger region of the Disco protein**

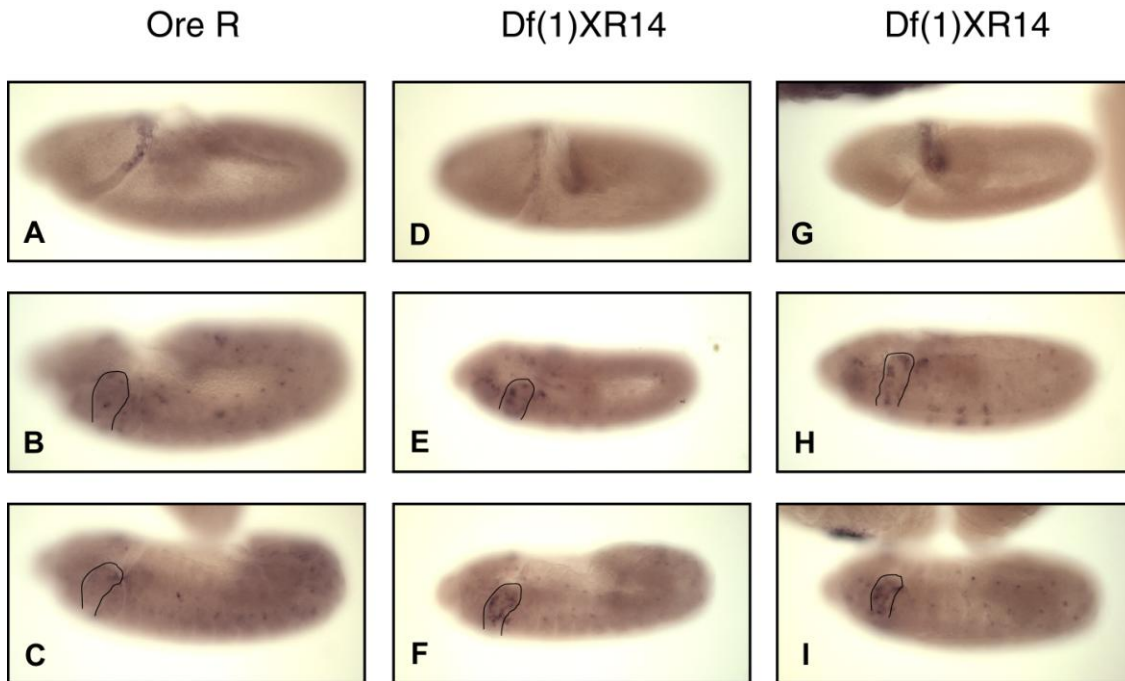
Since it is well known that many C2H2 zinc finger proteins regulate transcription by binding directly to DNA, the same might be expected for Disco. Indeed, Lee et al. (1999) demonstrated that Disco binds to two fragments at the *disco* locus, one that is ~2.5 kb upstream of the transcription start site and one within the intron. From their work we had a good candidate region to test for a DNA binding site. However, we wanted an enhancer region that could potentially respond to Disco and the Hox protein Dfd, and there was no indication that the Disco enhancer regions found by Lee et al. (1999) would

respond to Dfd. Our lab has also done several experiments to identify a potential direct interaction between Dfd and Disco, however, no such interaction has been detected.

Therefore, we set out to identify another good candidate enhancer region that we could possibly examine for a potential collaborative interaction between Disco and Dfd.

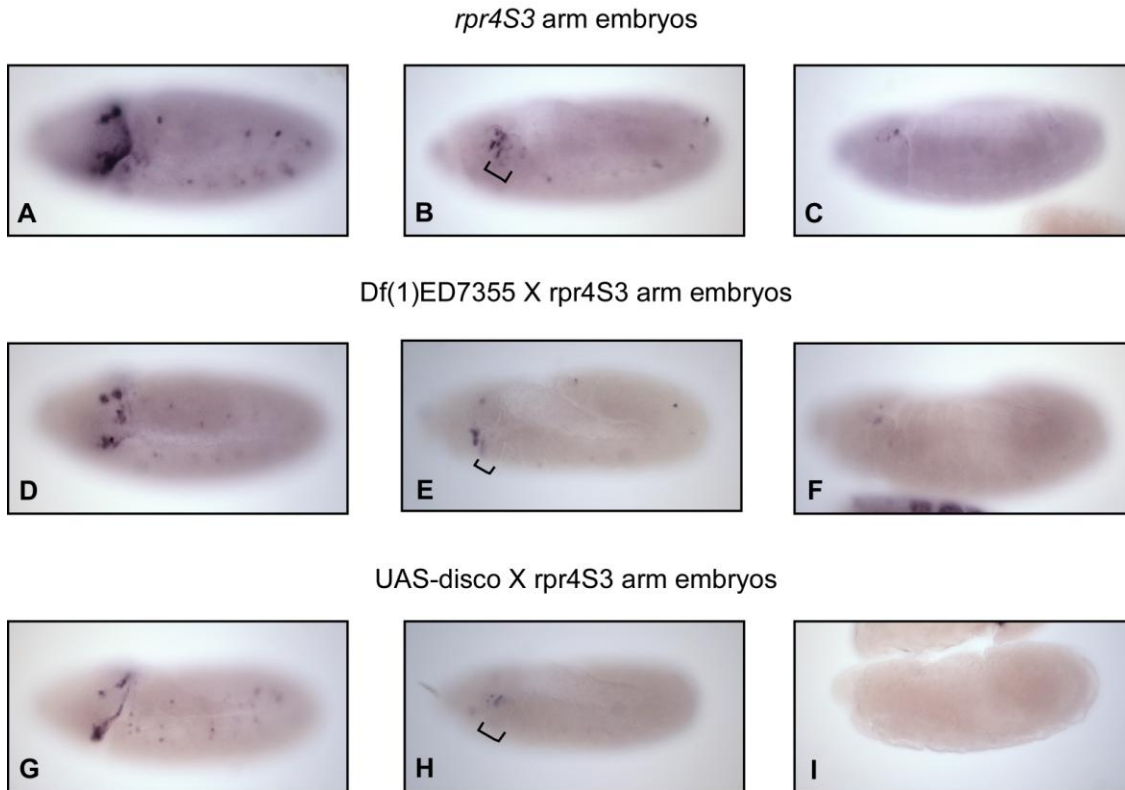
One such potential enhancer region is the *reaper* (*rpr*) enhancer. Lohmann et al. (2002) identified an 677 bp enhancer region of *rpr* that when placed in front of a *lacZ* reporter responded to Dfd within the mandibular/maxillary segment border, the *rpr4S3* reporter. *reaper* plays a role in the initiation of programmed cell death, and expression of *rpr* mimics the pattern of programmed cell death observed with acridine orange staining in wild type embryos (Abrams et al., 1993; White et al., 1994).

Endogenous *rpr* expression was affected in deficiencies that removed *disco* and *disco-r* (for example *Df(1)XR14*). In *Df(1)XR14* embryos loss of the *disco* genes lead to increased and more persistent *rpr* expression in the maxillary segment indicating that *disco* might be functioning as a repressor (Fig. 3). Interestingly, *rpr4S3* reporter expression was not influenced by loss of the *disco* genes (Fig. 4). *Df(1)ED7355* removes both *disco* and *disco-r* (as well as a few other neighboring genes with no developmental roles). In embryos that have both the deficiency and the *rpr4S3* reporter there was no difference in *lacZ* expression noted when compared to embryos with just the reporter. *rpr4S3* reporter expression in *Df(1)XR14* embryos was the same as in *Df(1)ED7355*. The lack of response to loss of the *disco* genes by the *rpr4S3* reporter may have been due to the fact that the cells expressing *lacZ* may not express *disco*, if this is the case then the *rpr4S3* fragment may still be affected by ectopic Disco expression.



**Figure 3. Disco represses *rpr* expression in the maxillary segment.**

(A,B,C) *rpr* RNA expression in early, mid, and late stages of embryogenesis, *rpr* expression mirrors the apoptotic pattern of cell death as witnessed by acridine orange staining, note in particular the staining within the maxillary segment which is bordered. (D,G) Early expression of *rpr* is similar in wild type embryos and Df(1)XR14 embryos which removes the *disco* genes. (E,H) More *rpr* expression can be observed in the maxillary segment of Df(1)XR14 embryos than in wild type during mid stages of development. (F,I) This increased expression is even more obvious at later stages of development where *rpr* expression is stronger and more persistent than in wild type embryos.



**Figure 4. *lacZ* expression in deficiency and ectopic Disco embryos.**

Embryos are positioned such that anterior is to the left and dorsal is up. (A,B,C) *lacZ* RNA expression of the *rpr4S3* reporter in Ore R embryos during early, mid, and late stages of embryogenesis. Early *lacZ* is detected along the cephalic furrow and spots within the thoracic segment, during mid stages of development expression is seen around the maxillary segment which is indicated by a bracket, and during late stages expression is between the maxillary segment and optic lobe. (D,E,F) *lacZ* expression of the *rpr4S3* reporter in the Df(1)ED7355 background which removes *disco* and *disco-r* as well as several other genes. Expression is mostly similar to Ore R embryos. (G,H,I) *lacZ* expression in ectopically expressing Disco embryos. (G) Early expression is similar to wild type. (H) However, at mid stages of development expression is reduced at the mandibular/maxillary border and (I) is also absent at later stages.

The *rpr4S3* reporter mimics only a very small portion of the *rpr* expression pattern. *rpr4S3* is expressed in the mandibular/maxillary border as well as a few maxillary and procephalic cells (Lohmann et al., 2002). When Disco was ectopically expressed using the *arm*-Gal4 driver in a background that contains the *rpr4S3* reporter, we observed reduced *lacZ* expression at the mandibular/maxillary boundary (Fig. 4). Since ectopic *disco* expression represses *rpr4S3*, and endogenous *rpr* expression is affected by loss of the *disco* genes, we suspected that the *rpr4S3* reporter might contain Disco binding sites. To test this hypothesis we conducted DNase protection and electromobility shift assays on the *rpr4S3* enhancer region from the reporter. DNase protection assays conducted by Barbara Imiolyzyk on the *rpr4S3* enhancer region showed that the zinc finger region of Disco can bind to two related sequences ATGACAA and TGACACA, both of which have TGACA at the core (Fig. 5). There are five sites within the *rpr4S3* enhancer, by chance only one site would be expected in this fragment. These sites are similar to what has been reported for the first pair of zinc fingers of Basonuclin, the vertebrate homolog of Disco, however slightly different binding sites have been reported for the full length Basonuclin protein which binds to sites with dyad symmetry (Iuchi and Green, 1999; Tian et al., 2001; Tseng et al., 1999).

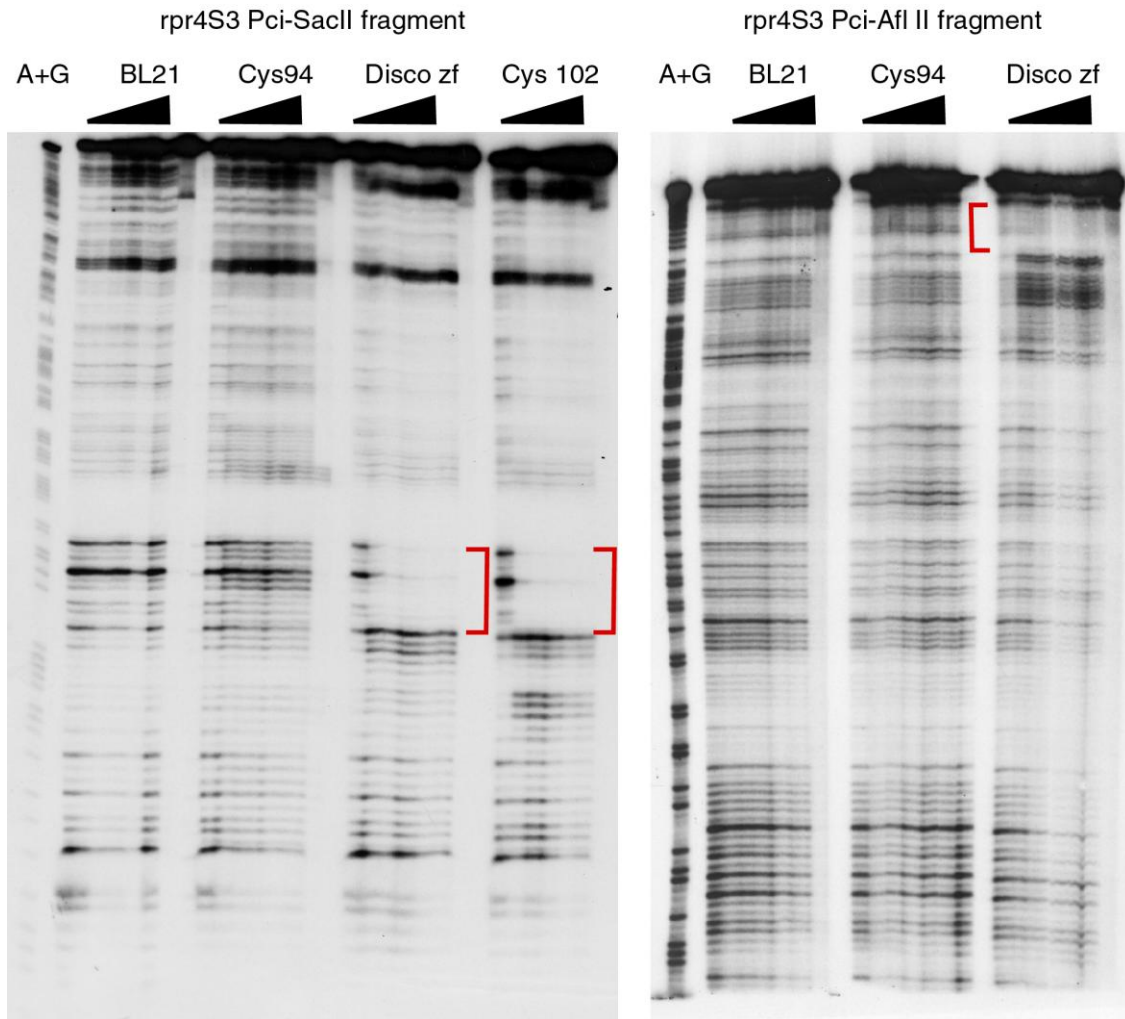
Gel shift experiments further verify that the zinc finger region of Disco binds to several sites in the *rpr4S3* fragment. Furthermore, these sequences are necessary for the zinc finger to bind the fragment. Gel shifts were conducted in which the probe contained normal Disco sites or sites where the GACA of the core sequence was mutated to CTCC. When the normal binding sites are present, Disco forms a complex with the DNA that

migrates slower than the probe alone (Fig. 6). When the binding sites were mutated we expected no DNA/protein complex to form and, as expected, there was no shift (Fig. 6). Gel shifts done with both transcription and translation kits and bacterially produced proteins yielded similar results. BL21 protein extract and luciferase were used as negative controls, both of which were incapable of shifting the DNA.

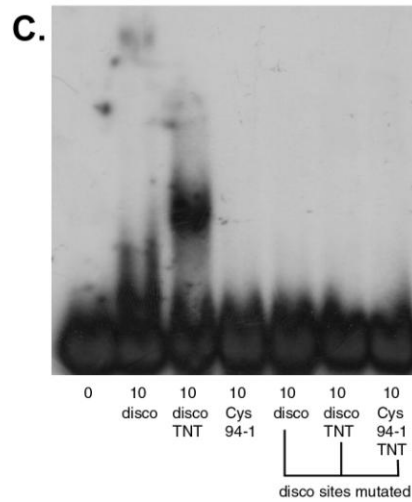
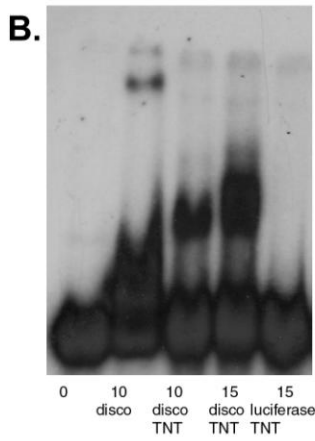
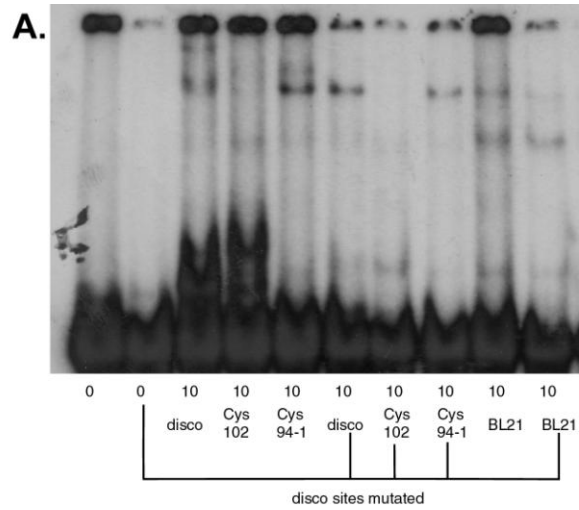
Evidence from the ectopic expression studies of the different zinc finger isoforms suggested that the two finger form was the active form of Disco and that the single zinc finger was inactive. This hypothesis was further supported by DNase protection and electromobility shift assays carried out with the *rpr4S3* enhancer using the Cys 102 and Cys 94-1 protein isoforms described above. The Cys 102 isoform protected the same region as the wild type Disco protein; whereas the Cys 94-1 isoform did not protect any region indicating that this form was not capable of binding DNA in our assays (Fig. 5). Likewise, in gel shift studies the Cys 102 isoform formed complexes with DNA when Disco binding sites were present (Fig. 6), but the Cys 94-1 isoform was unable to bind and shift the *rpr4S3* fragment.

### ***In vivo* significance of Disco binding site**

Identifying the *in vitro* binding site for Disco is an important step to understanding its function, but it is more important to understand the significance of the binding site *in vivo*. To test the *in vivo* significance of the Disco binding sites we attempted to mutate the sites and reintroduce the reporter back into *Drosophila*. Up until recently, we have been unable to produce any transgenics. We are able to easily produce transgenics using the original reporter and have produced many other transgenics in the



**Figure 5. Disco binds to the site TGACA within the *rpr4S3* enhancer region.** DNase protection assay on *rpr4S3* Pci I-Sac II (left) and Pci-Afl II (right) fragments. The protected areas are bracketed in red. (Left) From left to right A+G ladder, BL21 bacterial protein extract, Cys 94-1 the single zinc finger isoform, Disco (R1-Not), which encodes the 5' end and the zinc finger region of the protein, and Cys102 the double zinc finger isoform. In each section protein concentration increases from left to right as indicated by the triangles and the last lane is undigested DNA. The main region of protection is 5'-TTGACACACTTCC-3' within the Pci-Sac II fragment. The Cys 94-1 isoform is not capable of binding DNA and the Cys 102 isoform binds the same regions as the wild type Disco R1-Not I zinc finger region. (Right) Disco protects a slightly different region in the *rpr4S3* Pci-Afl II fragment, the region that is protected corresponds to 5'-ATGACAATAATTA-3'. Both regions that are protected have 5'-TGACA-3' at the core.



**Figure 6. The Disco binding site is necessary for Disco to bind DNA.**

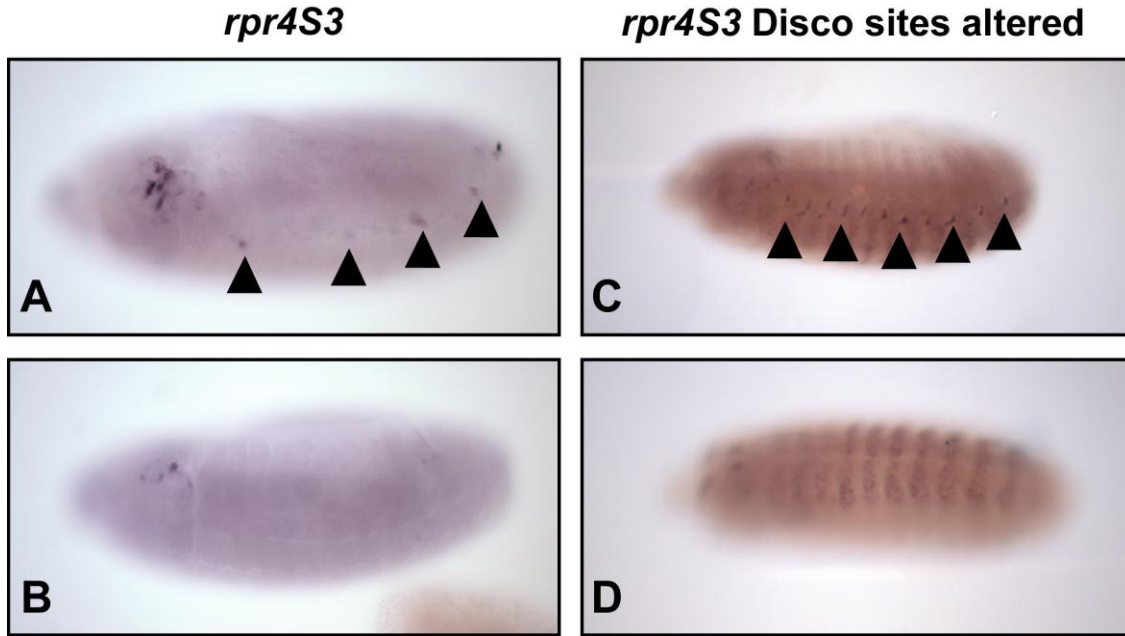
(A) Gel shifts using bacterially expressed proteins of Disco, Cys 102, Cys 94-1. The numbers indicate uL of protein extract added to each lane. Only Disco and Cys 102 are capable of binding DNA when the Disco binding sites are present in the rpr4S3 Pci I-Afl II fragment. When the GACA of the binding site is mutated to CTCC neither Disco nor Cys 102 can bind DNA. BL21 extract is used as a negative control and should not bind the DNA. (B) Comparison between bacterially expressed protein and TNT reticulolysate protein. Luciferase is used as a negative control. (C) The same results as in (A) are observed if TNT expressed proteins are used.

lab. Therefore, it is not clear why, after several attempts by more than one person, we have been unable to gain more than one insert of this particular fragment.

The one transgenic line that was obtained came from approximately 150 injected adults. However, our normal rate is about 2-3 out of 10 adults. Why this is so low for the mutated *rpr4S3* construct is not clear. Preliminary results with this line indicate *lacZ* is expressed more persistently within the abdominal segments compared to wild type and in addition we observe expression ventrally which is not observed at in the the wild type *rpr4S3* reporter (Fig. 7).

### **Disco interacts with the corepressor CtBP**

Some of the above data indicate that Disco might function as a repressor. Indeed, the Disco protein has a PLDFS peptide motif which is close to the PLDLS sequence known as the interaction site for the corepressor CtBP. Similar sites are conserved in all ecdysozoans that have a single Disco-like protein, but in ecdysozoans that have two Disco-like proteins only one protein has the CtBP interaction site. Thus Disco may have the potential to act as a repressor by forming a protein complex with CtBP. To verify that CtBP does interact with Disco through the PLDFS domain we conducted a yeast-2-hybrid assay. In yeast-2-hybrid assays the DNA binding domain of the yeast LexA protein is fused to a protein referred to as the bait. In this manner, the bait protein can bind to LexA binding sites within yeast gene promoters. The other protein to be tested for an interaction is fused to a transcriptional activation domain, and this construct is referred to as the prey. The LexA DNA binding domain and the activation domain are not capable of interacting on their own. If, however, the two proteins linked to these fragments can

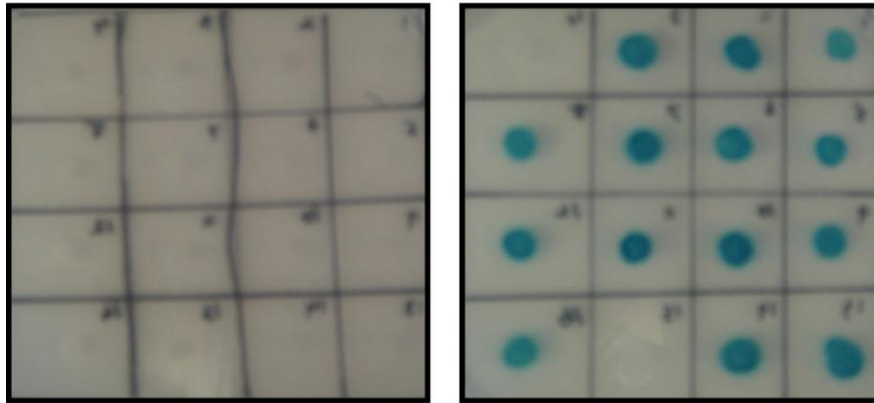


**Figure 7. Expression of *lacZ* from the *rpr4S3* reporter with Disco sites mutated.** Anterior is to the left, dorsal is up except in (D) where ventral is up. (A, B) Expression of *lacZ* from the normal *rpr4S3* reporter during midembryogenesis and late embryogenesis, respectively. Note the expression of *lacZ* in the abdominal segments (arrowheads) is not observed during late embryogenesis. (C) Expression within the abdominal segments is observed during late embryogenesis in embryos carrying the *rpr4S3* reporter with the Disco sites mutated. (D) In addition, *lacZ* expression is observed ventrally which is not observed with the normal *rpr4S3* reporter.

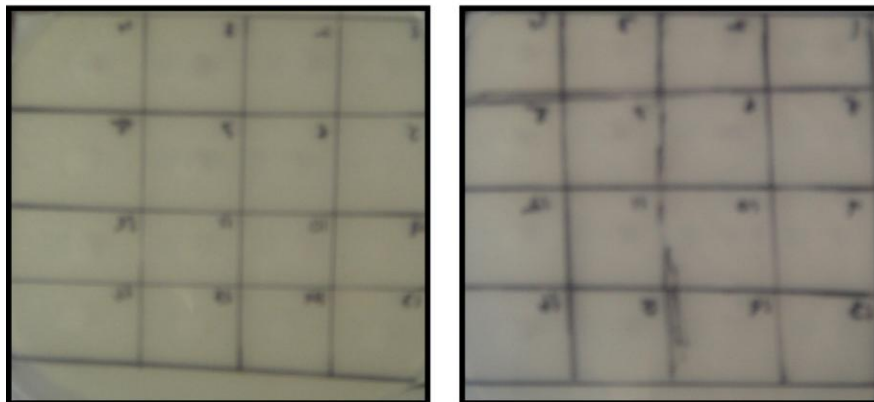
interact, the activation domain will be brought into close proximity with the DNA binding domain at target genes where the complex will, in turn, activate the expression of reporter constructs. In our case, interactions were assessed on the activation of genes in the leucine biosynthesis pathway and of a  $\beta$ -galactosidase reporter.

CtBP was fused to the LexA DNA binding domain from a cDNA construct provided by Dr. Sue Parkhurst and this was used in a cotransformation with either 1) full length Disco fused to the activation domain or 2) full length Disco with the PLDFS motif altered fused to the activation domain. This exact change, PLDFS to ALAFS, has been shown to completely prevent interaction between ZNF366 and CtBP (Lopez-Garcia et al., 2006). Individual colonies were picked and grown on SD-glucose –His Trp Ura Leu and SD-galactose –His Trp Ura Leu plates that contained X-gal. Expression of the prey construct is induced only on galactose plates, thus there should be no colony growth on glucose plates. Further, since the test plates lack leucine, and leucine can be produced only by activating the leucine synthesis pathway, colonies should grow only on galactose plates if interaction occurs yielding a functional transcription unit. In the case of CtBP with Disco full length, we did not observe colony growth on SD-glucose plates but there was strong blue growth on SD-galactose plates, indicating an interaction between CtBP and Disco and activation of both the leucine and  $\beta$ -gal reporters (Fig. 8). When the PLDFS site in Disco was mutated there was no colony growth on either SD-glucose or SD-galactose plates (Fig. 8). Therefore, the PLDFS domain is necessary for the interaction between Disco and CtBP. The next question was whether this interaction might have an *in vivo* function.

CtBP pLexA + Disco pB42AD  
 SD-glucose -His Trp Ura Leu    SD-galactose -His Trp Ura Leu



CtBP pLexA + Disco CtBP<sup>-</sup> pB42AD  
 SD-glucose -His Trp Ura Leu    SD-galactose -His Trp Ura Leu



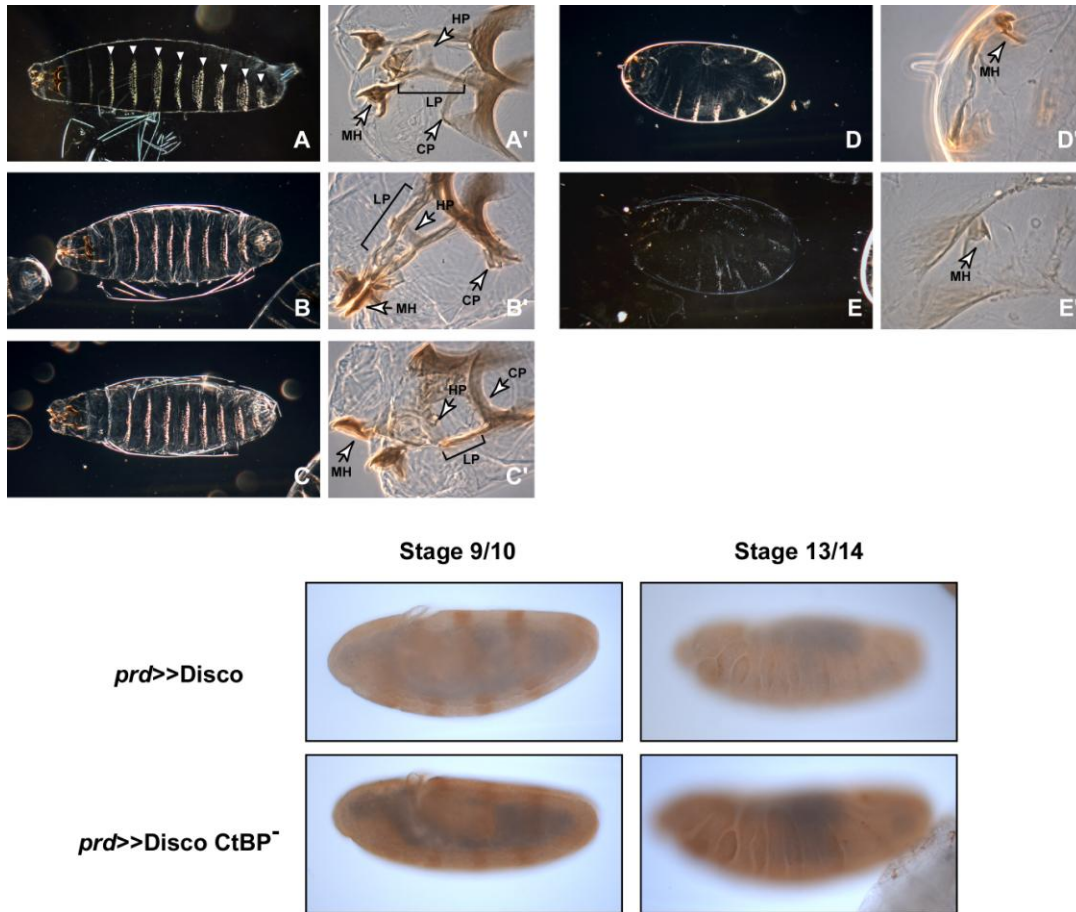
**Figure 8. CtBP is capable of interacting with Disco only when the PLDFS site is present.**

Yeast 2 Hybrid assay testing the interaction between CtBP and Disco. (Top,Left) No interaction should be detected on SD-glucose plates, however, there is slight colony growth on these plates possibly due to the activation potential of Disco. (Top, Right) When Disco is induced on SD-galactose plates it is capable of interacting with CtBP and strongly activating both the Leucine and B-gal reporters. (Bottom, Left) Once again there should be now colony growth on SD-glucose plates. (Bottom, Right) When the PLDFS site is mutated to ALAFS in Disco, Disco is no longer capable interacting with CtBP and activating the Leucine and B-gal reporters on SD-galactose plates.

To test the *in vivo* requirement for this interaction, we created transgenic UAS-*disco* constructs with the altered PLDFS domain, referred to as UAS-*disco CtBP*<sup>-</sup>. These lines were crossed to *arm-Gal4*, which drives expression throughout all cells of the embryo. Wild type UAS-*disco* crossed to *arm-Gal4* results in embryonic lethality, and these embryos have defects in germ band retraction, dorsal closure, and head involution. In addition, the cuticles have a decreased number of denticles whose organization is disrupted, often having gaps along the ventral midline. Most mouth parts do not develop except for mouth hooks and cirri (Robertson et al., 2002). When UAS-Disco CtBP<sup>-</sup> lines were crossed to *arm-Gal4*, a range of phenotypes with differing severity was obtained depending on the line tested.

In all, five separate lines were examined. Two of these lines when crossed to *arm-Gal4* exhibited no phenotype; they were viable and produced fertile adults. The other three lines when crossed to *arm-Gal4* produced three separate phenotypes. One line produced a very mild phenotype in which the lateral process had a brittle appearance, i.e. the lateral process did not appear to be as solid as in wild type embryos (Fig. 9). In another line most of the lateral process was absent or disrupted. In both of these lines the germ band did retract but segments failed to complete dorsal closure (Fig. 9). The final line exhibited a phenotype that was almost identical to *arm>>Disco*, however, reduction of the denticle belts and gaps in the denticle pattern along the ventral midline were never observed in these embryos (Fig. 9).

We suspected that the range in phenotype may have been due to differences in the amount of protein that was expressed. To address this idea protein expression was



**Figure 9. Ectopic expression of UAS-Disco CtBP<sup>-</sup> and protein expression levels.** (Top) (A) Cuticle of wild type 1<sup>st</sup> instar larva and (A') 100x magnification of the mouth. MH-mouth hook, LP-lateral process, HP-H-piece, and CP-cephalopharyngeal plates (B,C,D) Ectopic expression of Disco CtBP<sup>-</sup> using the *arm*-Gal4 driver in three separate lines, severity of phenotype increases from B-D. This is particularly evident in the mouth. (B') In the weakest phenotype all mouth structures are present, only the LP is affected having a brittle appearance compared to wild type. (C') The moderate phenotype is missing the lateral process anterior to the HP which is out of focus. (D') The most severe phenotype is similar to ectopic expression of Disco (E') where only the MH are present. (D) In addition to the mouth defects the germ band fails to retract similar to ectopic Disco expression (E). In the weaker phenotypes the germ band retracts but not all the segments complete dorsal closure leading to a disruption of the dorsal hairs between the denticle belts which are indicated by white arrowheads (B,C). (Bottom) Comparison of protein levels in a *prd*-Gal4 driven Disco line and the strongest Disco CtBP<sup>-</sup> line at two different stages. Expression is observed as stripes in alternating segments at stage 9/10 and at stage 13/14 as expected. Taking background into consideration the two lines appear to express protein at the same levels.

examined in a strong Disco line and three out of five of the Disco CtBP<sup>-</sup> lines crossed to *prd-Gal4* using a rabbit anti-Disco antibody (Lee et al., 1991) (Fig. 9). The *prd-Gal4* driver was used because the stripes of expression produced by using the *prd* driver are easily recognizable. The Disco and the Disco CtBP<sup>-</sup> proteins were expressed at comparable levels in the three lines examined, but only the expression in the strongest line is shown (Fig. 9). These data indicates the variability in phenotypes can not be attributed to differences in protein expression levels.

## **Discussion**

Disco is a C2H2 zinc finger protein that contains a pair of zinc fingers that is evolutionarily conserved among the ecdysozoa, lophotrochozoa, and some deuterostomes (Disco is not found in sea urchins) (Knight and Shimeld, 2001; Tseng and Green, 1992). Even though Disco is highly conserved, this conservation is limited mostly to the C2H2 zinc finger region. Further, there are also some important differences. First, all known vertebrate lineages contain three pairs of C2H2 zinc fingers whereas in the ecdysozoa and lophotrochozoa lineages the number of pairs of zinc fingers ranges from one to three, and sometimes only a single finger remains of the third pair. Second, deuterostomes do not have any obvious sequences similar to the CtBP interaction site while ecdysozoans with a single Disco-like protein do and when there are two Disco-like proteins one of the two has a site similar to PLDLS. Lastly, the ecdysozoan plus the lophotrochozoan lineages contain a conserved cysteine not found in deuterostome lineages. This change gives the ecdysozoans and lophotrochozoans the potential to form multiple zinc finger isoforms. Our studies indicate that without this cysteine Disco cannot maintain its wild type

function. When this cysteine is replaced, Disco can still form the pair of zinc fingers and bind DNA *in vitro*; however, this protein does not fully function *in vivo*. Ectopic expression of this form produces a much weaker phenotype than wildtype Disco, indicating that the cysteine is somehow necessary for complete function. The ability to form an alternate single zinc finger isoform did not seem to have any significance, at least as far as the studies we have conducted, since ectopic expression of this isoform does not produce any phenotype and is also incapable of binding DNA. This could indicate that the single-finger version alone is not functional but that it can be when the two-finger form is also present.

One of the more important questions we wanted to address in regards to Disco function was how Disco regulates gene expression. Disco and Disco-r are redundant proteins that play important roles in the development of the *Drosophila* larva and adult. In order for the HOM-C proteins, DFD and SCR, to properly regulate downstream targets either Disco or Disco-r must be present. Also, Disco can regulate the transcription of several genes as evidenced by ectopic expression studies in the larva and adult, but it is not known whether any of these are direct (Patel et al., 2007; Robertson et al., 2004). However, not much is known as to how Disco functions to regulate gene expression and which genes are directly regulated by Disco. To this end we demonstrate that Disco can bind DNA and has the potential to act as a repressor.

Our data show that Disco binds to the sequences ATGACAA and TGACACA both of which have TGACA at the core based on the strand of DNA we labeled, it is possible that Disco is actually binding to the opposite strand. Identifying binding sites

for Disco is a key step towards understanding Disco function because it allows for the identification of potential direct targets of Disco. The only potential direct target of Disco that has been identified thus far is *disco* itself, where it is required to maintain *disco* mRNA and protein expression within the optic lobe primordia (Lee et al., 1999). Lee et al. (1999), demonstrated that Disco binds to the enhancer region of *disco* ~2.5 kb upstream of the transcription start site and within the intron. If we examine 5kb upstream of *disco*, there are three sites that have TGACA at the core, one of which lies within the 120bp fragment bound by Disco in the experiments done by Lee et al. (1999). There are also four sites within the intron, one of which lies within the 280 bp fragment bound by Disco. Other potential target genes are known. Robertson et al. (2004), showed that *pannier* is repressed when Disco is ectopically expressed. There is only ~1.8kb between *pnr* and the next upstream gene CG17631. Examining this upstream region of *pannier* reveals that there are three TGACAA sites. This site would be expected to occur randomly throughout the genome roughly once every 1kb, so the occurrence of this site within the upstream region is higher than expected by chance, suggesting that these sites may be true Disco binding sites and that repression of *pannier* is directly mediated by Disco.

Our studies demonstrate that ectopic expression of Disco leads to repression of the *rpr4S3-LacZ* reporter, however, nothing is known as to how this repression is mediated. One potential mechanism might be through the use of the transcriptional corepressor CtBP. CtBP was first identified as a 48kDa cellular phosphoprotein that binds to the C-terminal region of human adenovirus E1A protein (Chinnadurai, 2002). CtBP

interacts with the motif PLDLS, which is conserved in the E1A proteins of human and primate adenoviruses (Chinnadurai, 2002). It is believed that CtBP acts through local interaction with histone deacetylases or through a mechanism called quenching (Courey and Jia, 2001; Nibu and Levine, 2001; Nibu et al., 2003; Nibu et al., 1998). In quenching, activators and repressors co-occupy closely linked sites and then the repressor inhibits the adjacent activator. Indeed Disco is capable of interacting with CtBP through a similar site, PLDFS, as indicated by the yeast-2-hybrid assay.

However, the significance of Disco interaction with CtBP remains to be determined since ectopic expression of Disco CtBP<sup>-</sup> can produce a phenotype similar to ectopic expression of Disco. There are a few possibilities as to why Disco CtBP<sup>-</sup> can behave similarly to Disco. The first is that Disco has another potential corepressor site. The site FSMDQLL in Disco is similar to the Engrailed homology 1 motif (eh1) which is recognized by the corepressor Groucho (Courey and Jia, 2001). Unlike CtBP, Groucho is a long range repressor capable of silencing transcription of linked promoters in a distance and orientation independent manner. Whether Disco actually interacts with Groucho remains to be tested. Evidence that a protein can interact with both corepressors is seen with several proteins, including Hairless, Hairy, Brinker, and TCF-4 (Hasson et al., 2001; Koelzer and Klein, 2006; Nagel et al., 2005; Phippen et al., 2000; Valenta et al., 2003). It could be that Disco interacts with Groucho and that CtBP is required to attenuate the activity of Groucho, which is seen with the protein Hairy. Another possibility is that Disco may have a repressive activity separate from CtBP and Groucho, which is evidenced with Brinker. In addition, Disco may still be able to bind CtBP through

another motif besides the PLDFS as is the case with ZN217; however, this is unlikely since it should have been detected in the yeast-2-hybrid assay. Lastly, it may be possible that Disco CtBP<sup>-</sup> can still autoactivate the endogenous *disco* gene leading to endogenous Disco protein which causes the phenotype or that association with CtBP is important for protein stability somehow.

With evidence that Disco is capable of acting as a repressor, how do we reconcile the fact that *disco* is required with Dfd and Scr to properly activate the *Dll*, *serrate*, *1.28*, and *pb* genes? In other words, how can Disco act as a repressor and an activator? One likely model is that the ability to act as a repressor or an activator is context dependent. Evidence for this can be seen in the Wnt signaling pathway and also in regulation of *spalt* expression within the wing and haltere. In the Wnt pathway, CtBP is recruited to Wnt-regulated enhancers in the absence of Wingless signaling by an unknown protein, which in conjunction with TCF/Groucho complex, acts to repress gene transcription as seen with the *naked* gene (Fang et al., 2006). However, when Wingless signaling is active, CtBP is brought by Armadillo and other factors to the TCF binding site where it acts to activate transcription, which is the case seen with *CG6234*, suggesting that CtBP does not always act as a repressor (Fang et al., 2006).

Context dependent activation or repression is also observed in the differential regulation of *spalt* (*sal*) expression within the wing and haltere. *sal* expression is activated within the wing and repressed within the haltere, and this expression is dependent on two Smad proteins, Mad and Medea, and the Hox protein, Ubx (Walsh and Carroll, 2007). In the wing, the two Smad proteins are genetically required for the

expression of *sal*, however, in the haltere these two proteins act with Ubx to repress *sal* expression. This difference in regulation is due to the spacing between the Mad/Med binding site and the Ubx binding site. Thus, the topology of the binding sites determined whether *sal* was either repressed or activated. Could Disco be acting in a similar manner to the situations described above? There are still many questions that remain to be answered in order to fully understand how Disco functions, but initial steps have been taken to help our understanding.

## REFERENCES

- Abrams, J. M., White, K., Fessler, L. I. and Steller, H.** (1993). Programmed cell death during *Drosophila* embryogenesis. *Development* **117**, 29-43.
- Brand, A. H. and Perrimon, N.** (1993). Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* **118**, 401-15.
- Chinnadurai, G.** (2002). CtBP, an unconventional transcriptional corepressor in development and oncogenesis. *Mol Cell* **9**, 213-24.
- Courey, A. J. and Jia, S.** (2001). Transcriptional repression: the long and the short of it. *Genes Dev* **15**, 2786-96.
- DiNardo, S., Kuner, J. M., Theis, J. and O'Farrell, P. H.** (1985). Development of embryonic pattern in *D. melanogaster* as revealed by accumulation of the nuclear engrailed protein. *Cell* **43**, 59-69.
- Dovat, S., Ronni, T., Russell, D., Ferrini, R., Cobb, B. S. and Smale, S. T.** (2002). A common mechanism for mitotic inactivation of C2H2 zinc finger DNA-binding domains. *Genes Dev* **16**, 2985-90.
- Fang, M., Li, J., Blauwkamp, T., Bhambhani, C., Campbell, N. and Cadigan, K. M.** (2006). C-terminal-binding protein directly activates and represses Wnt transcriptional targets in *Drosophila*. *EMBO J* **25**, 2735-45.
- Flint, K. J. and Jones, N. C.** (1991). Differential regulation of three members of the ATF/CREB family of DNA-binding proteins. *Oncogene* **6**, 2019-26.
- Hasson, P., Muller, B., Basler, K. and Paroush, Z.** (2001). Brinker requires two corepressors for maximal and versatile repression in Dpp signalling. *EMBO J* **20**, 5725-36.
- Heilig, J. S., Freeman, M., Lavery, T., Lee, K. J., Campos, A. R., Rubin, G. M. and Steller, H.** (1991). Isolation and characterization of the disconnected gene of *Drosophila melanogaster*. *EMBO J* **10**, 809-15.
- Iuchi, S. and Green, H.** (1999). Basonuclin, a zinc finger protein of keratinocytes and reproductive germ cells, binds to the rRNA gene promoter. *Proc Natl Acad Sci U S A* **96**, 9628-32.
- Knight, R. D. and Shimeld, S. M.** (2001). Identification of conserved C2H2 zinc-finger gene families in the Bilateria. *Genome Biol* **2**, RESEARCH0016.

- Koelzer, S. and 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.
- Lee, K. J., Freeman, M. and Steller, H.** (1991). Expression of the disconnected gene during development of *Drosophila melanogaster*. *EMBO J* **10**, 817-26.
- Lee, K. J., Mukhopadhyay, M., Pelka, P., Campos, A. R. and Steller, H.** (1999). Autoregulation of the *Drosophila* disconnected gene in the developing visual system. *Dev Biol* **214**, 385-98.
- Lohmann, I., McGinnis, N., Bodmer, M. and McGinnis, W.** (2002). The *Drosophila* Hox gene deformed sculpts head morphology via direct regulation of the apoptosis activator reaper. *Cell* **110**, 457-66.
- Lopez-Garcia, J., Periyasamy, M., Thomas, R. S., Christian, M., Leao, M., Jat, P., Kindle, K. B., Heery, D. M., Parker, M. G., Buluwela, L. et al.** (2006). ZNF366 is an estrogen receptor corepressor that acts through CtBP and histone deacetylases. *Nucleic Acids Res* **34**, 6126-36.
- Mahaffey, J. W., Griswold, C. M. and Cao, Q. M.** (2001). The *Drosophila* genes disconnected and disco-related are redundant with respect to larval head development and accumulation of mRNAs from deformed target genes. *Genetics* **157**, 225-36.
- Nagel, A. C., Krejci, A., Tenin, G., Bravo-Patino, A., Bray, S., Maier, D. and Preiss, A.** (2005). Hairless-mediated repression of notch target genes requires the combined activity of Groucho and CtBP corepressors. *Mol Cell Biol* **25**, 10433-41.
- Nibu, Y. and Levine, M. S.** (2001). CtBP-dependent activities of the short-range Giant repressor in the *Drosophila* embryo. *Proc Natl Acad Sci U S A* **98**, 6204-8.
- Nibu, Y., Senger, K. and Levine, M.** (2003). CtBP-independent repression in the *Drosophila* embryo. *Mol Cell Biol* **23**, 3990-9.
- Nibu, Y., Zhang, H., Bajor, E., Barolo, S., Small, S. and Levine, M.** (1998). dCtBP mediates transcriptional repression by Knirps, Kruppel and Snail in the *Drosophila* embryo. *EMBO J* **17**, 7009-20.
- Patel, M., Farzana, L., Robertson, L. K., Hutchinson, J., Grubbs, N., Shepherd, M. N. and Mahaffey, J. W.** (2007). The appendage role of insect disco genes and possible implications on the evolution of the maggot larval form. *Dev Biol* **309**, 56-69.
- Philipsen, S. and Suske, G.** (1999). A tale of three fingers: the family of mammalian Sp/XKLF transcription factors. *Nucleic Acids Res* **27**, 2991-3000.

- Phippen, T. M., Sweigart, A. L., Moniwa, M., Krumm, A., Davie, J. R. and Parkhurst, S. M.** (2000). Drosophila C-terminal binding protein functions as a context-dependent transcriptional co-factor and interferes with both mad and groucho transcriptional repression. *J Biol Chem* **275**, 37628-37.
- Robertson, L. K., Bowling, D. B., Mahaffey, J. P., Imiolczyk, B. and Mahaffey, J. W.** (2004). An interactive network of zinc-finger proteins contributes to regionalization of the Drosophila embryo and establishes the domains of HOM-C protein function. *Development* **131**, 2781-9.
- Robertson, L. K., Dey, B. K., Campos, A. R. and Mahaffey, J. W.** (2002). Expression of the drosophila gene disconnected using the UAS/GAL4 system. *Genesis* **34**, 103-6.
- Rubin, G. M. and Spradling, A. C.** (1982). Genetic transformation of Drosophila with transposable element vectors. *Science* **218**, 348-53.
- Ryan, R. F. and Darby, M. K.** (1998). The role of zinc finger linkers in p43 and TFIIIA binding to 5S rRNA and DNA. *Nucleic Acids Res* **26**, 703-9.
- Scuderi, A. and Letsou, A.** (2005). Amnioserosa is required for dorsal closure in Drosophila. *Dev Dyn* **232**, 791-800.
- Tian, Q., Kopf, G. S., Brown, R. S. and Tseng, H.** (2001). Function of basonuclin in increasing transcription of the ribosomal RNA genes during mouse oogenesis. *Development* **128**, 407-16.
- Tseng, H., Biegel, J. A. and Brown, R. S.** (1999). Basonuclin is associated with the ribosomal RNA genes on human keratinocyte mitotic chromosomes. *J Cell Sci* **112 Pt 18**, 3039-47.
- Tseng, H. and Green, H.** (1992). Basonuclin: a keratinocyte protein with multiple paired zinc fingers. *Proc Natl Acad Sci U S A* **89**, 10311-5.
- Valenta, T., Lukas, J. and Korinek, V.** (2003). HMG box transcription factor TCF-4's interaction with CtBP1 controls the expression of the Wnt target Axin2/Conductin in human embryonic kidney cells. *Nucleic Acids Res* **31**, 2369-80.
- Walsh, C. M. and Carroll, S. B.** (2007). Collaboration between Smads and a Hox protein in target gene repression. *Development* **134**, 3585-92.
- White, K., Grether, M. E., Abrams, J. M., Young, L., Farrell, K. and Steller, H.** (1994). Genetic control of programmed cell death in Drosophila. *Science* **264**, 677-83.

## CHAPTER FIVE

### Identification of novel C2H2 zinc finger genes in epidermal pattern formation

## Introduction

In 1980 Christiane Nusslein-Volhard and Eric Wieschaus reported on an embryonic lethal screen to identify novel genes involved in segment pattern formation of the *Drosophila* larva (Nusslein-Volhard and Wieschaus, 1980). They had hoped to achieve saturation with this screen, meaning they wanted to identify all genes involved in segment pattern formation. From this screen, most of the alleles affecting segmentation on the second chromosome were isolated; they believed they had identified only 50% of the segmentation loci on the third and first chromosomes (Nusslein-Volhard and Wieschaus, 1980). Thus they believed it was possible to identify all genetic components involved in embryonic segment pattern formation.

Nusslein-Volhard and Wieschaus did not achieve saturation in this particular screen, and other scientists have since identified more genes involved in segment pattern formation. Even today, it is hard to say when saturation will be reached due to the restrictions of classical genetic approaches. In fact, genes with redundant function would be overlooked by classical approaches. Two genes are considered to be redundant when one gene is capable of carrying out the function of another gene. In such cases a phenotype would manifest only if both genes are mutated, and in some instances more than two genes can have redundant function (Dolezal et al., 2003; Schmid et al., 2002). However, with the advent of technologies such as RNAi and full genome sequencing, problems posed by redundancy might be resolved and saturation could be achieved.

Two genes with redundant function are *disconnected* (*disco*) and *disco-related* (*disco-r*). *disco* encodes a C2H2 zinc finger transcription factor and was first identified

in a screen for X chromosome structural brain mutants (Steller et al., 1987). In loss-of-function *disco* mutants the larval optic nerves, also known as Bolwig nerves, occasionally fail to make stable connections with target cells in the brain while remaining semi-viable (Heilig et al., 1991; Steller et al., 1987). Still, *disco* did not appear to have any role in patterning of the *Drosophila* embryonic head until *disco-r* was discovered, though based on its expression in the gnathal segments one might have expected it (Lee et al., 1991). *disco-r* was discovered in a deficiency screen for potential cofactors for the HOM-C gene *Dfd* (Mahaffey et al., 2001). Similar to *disco*, *disco-r* encodes a C2H2 zinc finger protein, but *disco-r* has two pairs of zinc fingers instead of one pair. *disco-r* expression is mostly identical to *disco* expression with some differences (Mahaffey et al., 2001).

Using RNAi to knockdown *disco-r* function in a *disco*<sup>1</sup> mutant background leads to head segmentation defects which are not observed in individual mutants alone (Mahaffey et al., 2001). This indicates the presence of *disco* or *disco-r* is necessary for proper head development and the two genes are functionally redundant. Without RNAi, mutants in *disco-r* would have to be generated and then combined with the *disco*<sup>1</sup> mutation to observe the embryonic head phenotype, this would have taken considerably more time with classical genetic approaches. Although it is possible to use deficiencies which are large deletions of the chromosome, they often remove many more genes than the genes of interest complicating the analysis of phenotypes. With RNAi it is possible to specifically knockdown or knockout function of more than one gene at a time.

RNAi is a mechanism by which genes can be silenced. The RNAi mechanism has been observed in many organisms including *Caenorhabditis elegans* and *Drosophila*

*melanogaster* and is related to posttranscriptional gene silencing in plants (for review see (Hannon, 2002)). Double stranded RNA is introduced into the *Drosophila* embryo through injection where it is broken down by the endonuclease Dicer into 21-22 bp fragments called small interfering RNAs (siRNAs) (Elbashir et al., 2001; Ketting et al., 2001; Kim et al., 2006). The siRNAs then associate with the RNA-induced silencing complex (RISC) leading to complementary base pairing between the siRNA and the endogenous mRNA and subsequent degradation of the endogenous mRNA (Hammond et al., 2000). RNAi is such a powerful tool that many labs are using it as means of screening genes that are important to the biological process they are studying. For example, RNAi screens have been used to identify genes important for nervous system development, chromatin regulators, and even genes necessary for RNAi (Dorner et al., 2006; Koizumi et al., 2007; Lu et al., 2007). There has even been a *Drosophila* RNAi screening center established to help with the annotation of many uncharacterized genes (Flockhart et al., 2006). In addition, a genomic set of RNAi p-element insertions, which covers approximately 88% of the protein coding genes in *Drosophila*, has been created allowing possible examination of gene function at different stages and tissues (Dietzl et al., 2007).

With the full genome sequence of *Drosophila melanogaster* available, there are now many predicted genes with very little known about their function. Some of these genes encode proteins with high sequence similarity, indicating that they may be functionally redundant. Some might play a role in embryonic pattern formation. Once again, even individual RNAi screens would overlook potentially redundant genes. We

proposed to identify previously uncharacterized genes involved in epidermal embryonic pattern formation that might have been overlooked due to redundancy.

Redundancy often results from the process of gene duplication. When a gene is duplicated it can take several paths. Either copy may be maintained as a functional backup (this is what is recognized as redundancy). Sometimes the paralogs, homologous genes found within the same genome, may diverge and acquire new functions (neofunctionalization), or they may subdivide the original gene's function (subfunctionalization). Alternatively, one of the genes may become a pseudogene (Cooke et al., 1997; Guan et al., 2007; Nowak et al., 1997). In most cases the duplicated gene will become a pseudogene and eventually be lost.

Since redundancy can result from duplication, we wanted to identify potentially duplicated C2H2 zinc finger proteins within the *Drosophila* genome. Specifically, we wanted to examine genes that encoded C2H2 zinc finger proteins that contain five or fewer zinc fingers. C2H2 zinc finger proteins make up one of the largest classes of transcription factors and, in addition to *disco* and *disco-r*, several other C2H2 zinc finger transcription factors have been identified with roles in patterning that are partially redundant. These include the gene pairs *teashirt* and *tiptop*, *buttonhead* and *D-Sp1*, *knirps* and *knirps-related*, *spalt major* and *spalt related*, and the *Odd-skipped* family (Dong et al., 2003; Gonzalez-Gaitan et al., 1994; Hart et al., 1996; Laugier et al., 2005; Reuter et al., 1996; Schock et al., 1999).

To identify potentially redundant genes, genes encoding related proteins were identified using the BLAST algorithm and Clustal X (Altschul et al., 1990; Thompson et

al., 1997). Once duplicated genes were identified, *in situ* hybridizations were done to *Drosophila* embryos and the subsequent expression patterns were examined. If aspects of the expression patterns were similar, RNAi was then used to examine individual mutant phenotypes and in combination.

Many of the gene pairs that we tested did not have related expression patterns. Some were ubiquitously expressed in the embryo or had expression patterns in the CNS/PNS. These genes might be interesting, but we were aimed at identifying uncharacterized epidermal patterning genes. Since our initial approach did not identify any potential epidermal patterning genes, we decided to BLAST the original zinc finger protein list from *Drosophila melanogaster* to the *Anopheles gambiae* (the mosquito) genome. Since *Drosophila* and *Anopheles* are approximately 180 million years diverged (Whiteley and Kassis, 1997), we assumed that genes that are important for epidermal pattern formation would be highly conserved between the two species. Using this approach we were able to identify several C2H2 zinc finger genes that were highly conserved including two particular genes, CG11798 and CG5249, which might play a role in epidermal pattern formation.

## **Materials and Methods**

### **Identification of potentially redundant C2H2 zinc finger genes**

Genes with C2H2 zinc finger motifs were identified using the Ensembl website (<http://www.ebi.ac.uk/ensembl/>). A search of the *Drosophila* genome using C2H2 zinc finger as search criteria produces links to all genes that contain C2H2 zinc finger domains. Protein sequences for all proteins that had 5 or fewer zinc fingers were

obtained and pasted into a notepad. These sequences were further edited by eliminating sequences outside of the zinc fingers since such sequences would skew alignments due to most identity being within the zinc finger regions. The sequences were then aligned and a neighbor joining tree was created using Clustal X (Thompson et al., 1997). Based on the alignment and phylogeny, potentially redundant genes were identified. The list of C2H2 zinc finger genes was also used to BLAST against the *Anopheles gambiae* genome using a stand alone BLAST program (<http://www.ncbi.nlm.nih.gov/blast/download.shtml>)(Altschul et al., 1990). Genes with a very high BLAST cutoff score  $E > 1e^{-50}$  were examined more closely.

### ***In Situ* Hybridizations**

Digoxigenin labeled antisense RNAs were prepared and in situ hybridizations were done as in Tautz and Pfeifle, 1989 (Tautz and Pfeifle, 1989). If a cDNA was available it was obtained from the Berkley Drosophila Genome Project ([www.fruitfly.org](http://www.fruitfly.org)). All other probes were generated by PCR using Drosophila genomic DNA as a template. For a list of primers used see Table 1. For some genes, expression data was obtained from the Berkley Drosophila Genome Project.

### **RNAi protocol**

Production of dsRNA and injections were done essentially as described in Brown et al. 1999. Buffer alone was used as a control for defects observed due to the injection procedure.

### **Fly Crosses**

Df(3L)ZN47, ry<sup>506</sup>/TM3, Sb<sup>1</sup> (64C;65C) (FBst0003096) and y<sup>1</sup>w<sup>\*</sup>, Dp(1;Y) y<sup>+</sup>;

**Table 1: Primers used for PCR**

Gene	Forward Primer 5'-3'	Reverse Primer 5'-3'
<i>CG10440</i>	TTAAGAACATCCTGGCAGC	CGGAAACAGAAGCCACAGAA
<i>CG3625</i>	AATCTCGTTTTGGGGCTCTG	GTTTATTCCAAGCTTCCCACC
<i>CG11275</i>	TATCGCTTCTATTTCGCCCG	TTGAACTCTGAGGCAGCGTA
<i>CG30283</i>	TGGCTCATAGCCTCATTGAG	CTGGGAGTGAAATTGTGCACT
<i>CG30287</i>	GGCACTTGAGTTCTCAGTCAG	CAAGAGAGACGTCAACAGCA
<i>CG4707</i>	TCTTCATGCACGCCCTGT	TGCTTCTGTTCCGACTGTGGG
<i>CG4282</i>	CAGCTTGCACCTGCTATTGTG	CAATTATTTGGATCGGAGC
<i>CG11695</i>	GGCATCTCTGTGCTCGAAGT	AGTCATTTCATGTCGCCCCG
<i>CG11696</i>	GGAGGTCTCTTTTCGCGGT	CACAAAGTCGCAGGGACAGA
<i>CG8145</i>	CGCTGTGAATGGACATTAGCC	CTTCCTGACACAAAAGGTCG
<i>CG11762</i>	TGCCTTTCATCAGTCGAG	GCTTGAGCTATTTAATCCTTC
<i>CG8159</i>	CTTCGTGGCAAATGGCAA	GGCTTTAAAGGTTGAGATGG
<i>CG9793</i>	GGAAGGAAAAGTGTGTCGC	GGTAATGCAACCGTTTTTC
<i>CG9797</i>	CGGGGCGAATGGCAAATA	GTCGGTCTAGCATTGCCTGT
<i>CG6813</i>	GCTCGAAACGTAAACAGT	CAGAGTTGGCAGAAATCG
<i>CG14711</i>	GGACGTGCAAGATAATGCC	TTCGACTAATAGCGCCCC
<i>CG6808</i>	GGGACTTCCTGCGGAAAT	TGTCCATTCTACCGTTTCG
<i>CG14710</i>	CCTTAATTTCCAGCCCC	GCTCTCTAACATTTCCGGAG
<i>CG17568</i>	GGGTCTTATTGGGTGTTGTG	TTTGCTACTTGGCAAGCA
<i>CG9895</i>	AAGTGGAGTGGCAAGCCTG	TGAGCTACTTAGGCGACATGG
<i>CG17326</i>	CGGGCAGCGTCCCTAAATT	GTGGTGATGCAGGTGCGTT
<i>CG11243</i>	TTACAGATGCCTCGCTAGC	TCGTATGGAGAACCCTTC
<i>CG18081</i>	GCAGCGAAACAAAACCCG	TGCTAGTTGCCTTTCAGGG
<i>CG6930</i>	GCATCAGATTAAGTGCTGCAA	CAGTCAGTCGTATCCGTATAC
<i>CG7368</i>	CTAGTGTGTGTGAGTGAGTGA	ATTCCATCGCGAGCTTAGGAT
<i>CG10714</i>	ACCCATCGAAGATCCGTTTCG	GGAATTGTCGAGGACACGCA
<i>CG3850</i>	ATTACCCGAACCTCTGCAGGC	AAGAGCTGGGAAACAGCGCA
<i>CG5249</i>	AAATCTCCCCATTTCCCCC	TCTGGTCTGCAGCTGCTCCAT
<i>CG9650</i>	GTTCCGCCACCATGGTTTCA	GGCACGGGAAAACCACTGAA
<i>CG12769</i>	TGTGGGTTTCAGTTGGTTA	CGTAGCCCCATTTGATAA
<i>CG17181</i>	CATCACGTGCGCATTTAGCAG	TCCGGTGTCTTTTGCGTGGT
<i>CG11798</i>	CAGTTTGCCACCTGTTTTGT	GTCAAATGCCTAATGCATTCC

Df(2R)XTE-58/Cyo (51D1-2;51E5) (FBst0006504) were obtained from the Drosophila Stock Center and raised on standard cornmeal-molasses-agar media. Each stock was outcrossed to Ore-R to remove balancer chromosomes. The resulting progeny Df(3L)ZN47, ry<sup>506</sup>/+ and Df(2R)XTE-58/+ were then crossed to themselves. Cuticle preparations were then done on homozygous deficiency embryos as described in Pederson et al. (1996).

## **Results**

### **Identification and expression patterns of potentially duplicated genes**

C2H2 zinc fingers are one of the most prevalent protein motifs in eukaryotic genomes; they are characterized by the ability to coordinate a zinc ion through cysteine and histidine residues. In addition to the conserved cysteines and histidines, C2H2 zinc fingers contain conserved hydrophobic residues four amino acids C-terminal to the second cysteine that help form the characteristic secondary structure consisting of an alpha-helix and a two stranded antiparallel beta-sheet (Ryan and Darby, 1998). The consensus sequence for C2H2 zinc fingers is C-X<sub>2-5</sub>-C-X<sub>3</sub>-F/Y-X<sub>8</sub>-H-X<sub>3-5</sub>-H. Proteins that contain C2H2 zinc fingers are capable of activating or repressing transcription by recognizing specific DNA sequences. In addition, some zinc finger domains are also capable of protein-protein interactions (Kelley et al., 1998; McCarty et al., 2003; Morgan et al., 1997). The expansion of the C2H2 zinc finger motif as organism complexity increases demonstrates the versatility of the motif. The Drosophila genome, which contains 234-357 zinc finger genes, is no different where zinc finger proteins comprise one of the largest classes of proteins within the proteome (Schmidt and Durrett, 2004).

Many of the C2H2 zinc finger proteins in *Drosophila* have unknown function; due to sequence similarity, many are likely to be redundant. Considering the number of redundant genes that have pattern formation roles, we wondered whether there was a correlation between pattern formation and redundancy; therefore, we developed a method to identify potentially redundant, uncharacterized C2H2 zinc finger genes involved in epidermal pattern formation. To do this a list of C2H2 zinc finger genes with five or fewer C2H2 zinc fingers was compiled, and an alignment and a neighbor-joining phylogenetic tree were created using Clustal X (Thompson et al., 1997). The resulting phylogeny was examined to identify pairs of potentially redundant genes that were examined more carefully using BLAST. From this data we compiled a list of 26 gene pairs or groups that had highly conserved zinc finger regions. Next, expression data were obtained for most of these genes either by *in situ* hybridization with RNA probes to *Drosophila* embryos or from expression data from the Berkley *Drosophila* Genome Project ([www.fruitfly.org](http://www.fruitfly.org)). A summary of these data can be seen in Table 2.

In addition to the pairs of zinc finger genes, we also identified five clustered regions that contained 4-8 related zinc finger genes each. Genes within these clusters were oriented in both directions, and though the clusters appeared to be related, the orientations were not conserved. These genes were all on chromosome arm 3R. Expression of genes within two of these clusters was examined using *in situ* hybridization with RNA probes (Table 2). Out of the nine genes examined from the two clusters only one, CG9797, gave any sort of expression pattern. Expression of the other three clusters was not examined due to our initial observations with the first two clusters. Since nearly

**Table 2-Expression of duplicated genes.**

Expression of potentially redundant pairs or groups of genes based on RNA *in situ* hybridization data to Drosophila embryos or data from the Berkley Drosophila Genome Project.

Gene	<i>In Situ</i> Pattern	Gene	<i>In Situ</i> Pattern	Gene	<i>In situ</i> Pattern
<i>CG17803</i>	no data	<i>CG6689</i>	ubiquitous, too weak		
<i>CG15703</i>	maternal, ubiquitous (mistake?)	<i>CG8388</i>	anterior anlage, cellular blastoderm, faint stripes in segments		
<i>CG10959</i>	maternal, no staining	<i>CG2129</i>	mesoderm at germband extending?, very little staining		
<i>CG4936</i>	maternal, ubiquitous	<i>CG4413</i>	maternal, no staining		
<i>CG5245</i>	no data	<i>CG3485</i>	no data		
<i>CG18379</i>	this is <i>Lola</i> gene	<i>CG4411</i>	maternal, weak CNS, polar buds		
<i>CG5204</i>	no pic, array data looks horrible	<i>CG9932</i>	array data says absent at all stages		
<i>CG18764</i>	maternal, no staining	<i>CG14712</i>	maternal, weak nonspecific	<i>CG1792</i>	no pic array data only
<i>CG10669</i>	maternal, ubiquitous, embryonic hindgut, midgut, and larval muscle system	<i>CG17822</i>	no staining (Production problem)	<i>CG11902</i>	CNS/ubiquitous
<i>CG10568</i>	no data	<i>CG10348</i>	no data	<i>CG6813</i>	no data
<i>CG1603</i>	maternal, no staining	<i>CG1605</i>	no staining	<i>CG1602</i>	no pic, array data
<i>CG17802</i>	maternal, no staining	<i>CG7357</i>	maternal, no staining, no pic	<i>CG4820</i>	no data
<i>CG10040</i>	CNS/PNS	<i>l(3)02102</i>	PNS		
<i>CG1130</i>	CNS/PNS	<i>CG12605</i>	CNS/PNS		
<i>CG7368</i>	no staining	<i>CG6930</i>	CNS/PNS		
<i>CG15715</i>	no staining/ubiquitous	<i>CG18081</i>	no staining/ubiquitous		
<i>CG4707</i>	no staining	<i>CG4282</i>	no stain/ubiquitous		
<i>CG11696</i>	no staining/ubiquitous	<i>CG11695</i>	no staining/ubiquitous		
<i>l(2)35Ea</i>	no staining/ubiquitous	<i>CG17568</i>	no staining/ubiquitous		

Table 2 (continued)

<i>CG4936</i>	no staining/ubiquitous	<i>CG4413</i>	not done		
<i>CG10016</i> ( <i>sob</i> )	Segment polarity, ring gland, gut	<i>CG31216</i> ( <i>drm</i> )	Segment polarity, ring gland, gut		
<i>CG9895</i>	no staining/trachea?	<i>CG12029</i>	not done		
<i>CG3065</i>	no staining/ubiquitous				
<i>CG5669</i>	no staining/ubiquitous				
<i>CG11243</i>	PNS	<i>CG10714</i> ( <i>Ly</i> )	not done	<i>CG17326</i>	CNS/PNS
Clustered genes					
Group 1					
<i>CG8136</i>	not done	<i>CG8145</i>	no staining	<i>CG11762</i>	no staining
<i>CG8159</i>	no staining	<i>CG9797</i>	CNS/PNS	<i>CG9793</i>	no staining
Group 2					
<i>CG14710</i>	no staining	<i>CG6808</i>	no staining	<i>CG6813</i>	no staining
<i>CG14711</i>	no staining/ubiquitous				

all of the genes that we examined either had no expression, were ubiquitously expressed, or were expressed in the CNS/PNS, which was out of the realm of this study, we did not conduct RNAi on any of these genes.

### **Highly conserved genes between *Anopheles* and *Drosophila***

Due to our initial screen not providing any candidates for novel epidermal patterning genes, we decided to take another approach. Using the BLAST algorithm we identified highly conserved C2H2 zinc finger genes between *Drosophila* and *Anopheles* assuming that genes important for pattern formation would be highly conserved. Using BLAST with a cutoff score of  $E > 1e^{-50}$  to identify highly conserved C2H2 zinc finger genes between *Drosophila* and *Anopheles* produced a list of 20 genes that met the criteria. Eliminating genes with known function and genes that are potentially duplicated in *Anopheles* narrowed the list down to 12 genes. Eight of these genes were expressed in the CNS/PNS, two were not expressed during embryogenesis, and two were expressed in the mesoderm. A description of the expression patterns of these genes are shown in Table 3. Two of these genes, *CG5249* and *CG11798*, gave expression patterns suggesting involvement in epidermal pattern formation. These two genes were further examined in more detail.

### **Expression pattern of *CG5249* (*Blimp-1*) and RNAi phenotypes**

*CG5249* has a very dynamic expression pattern that was first observed as two faint stripes in the anterior end of the embryo during early blastoderm stages (Fig. 1A). At later blastoderm stages a cap of expression was seen at the anterior end of the embryo as well as three broad stripes similar to gap genes (Fig. 1B). As gastrulation began,

**Table 3-Expression patterns of highly conserved C2H2 zinc finger genes between *Drosophila melanogaster* and *Anopheles gambiae*.**

Gene	<i>In situ</i> pattern	Gene	<i>In situ</i> pattern
<i>CG12769</i>	No stain/ubiquitous	<i>CG5249</i>	Gap pattern, PNS, CNS, segment polarity?
<i>CG9650</i>	CNS/PNS	<i>CG3850</i>	Mesoderm, posterior PNS
<i>CG12605</i>	CNS/PNS	<i>CG1130</i>	CNS/PNS
<i>CG17181</i>	Mesoderm	<i>CG11798</i>	CNS/PNS, segmental
<i>CG10040</i>	CNS/PNS	<i>l(3)02102</i>	PNS
<i>CG7368</i>	No staining	<i>CG6930</i>	CNS/PNS

strong expression was observed at the posterior pole of the embryo and faint expression near the cephalic furrow. It also appeared that as development proceeded the broad stripes each divided into two fainter stripes (Fig. 1C). As the germ band started to extend, spots of expression appeared in the head in the clypeolabrum, mandibular, maxillary, and labial segments (Fig. 1D). During germ band extension the spots of expression expanded to all the abdominal segments and faintly in the thoracic segments (Fig. 1E). These spots of expression were later shown to be in sensory organ promordia of the PNS (Ng et al., 2006). Expression then changed from spots to stripes about 2-3 cells wide in each segment as the germ band retracted (Fig. 1F). Once the germ band fully retracted expression was along the boundaries of each segment and after head involution there was expression in what appeared to be the brain (Fig. 1G).

Using RNAi to knockout the function of *CG5249* was embryonic lethal, though no segmental defects occurred (data not shown). This phenotype was supported by a deficiency obtained from the Drosophila Stock Center, *Df(3L)ZN47, ry<sup>506</sup>/TM3, Sb<sup>1</sup>* (64C;65C) (referred to as *Df(3L)ZN47*), which removes *CG5249* as well as many other genes. *Df(3L)ZN47* homozygous embryos had no obvious segmental defects, but were embryonic lethal. This result was somewhat surprising given the expression pattern. Work done by Ng et al. (2006) supported our initial findings as they also demonstrated that loss of *CG5249* leads to embryonic lethality with no obvious segmentation defects (see discussion for more information) (Ng et al., 2006).

### **Expression pattern of *CG11798* and RNAi phenotypes**

Expression of *CG11798* was first detected during early blastoderm stages almost

throughout the entire embryo with the exception of the anterior and posterior ends (Fig. 2A). The expression pattern appeared to refine to faint pair-rule stripes that were still present during gastrulation (Fig. 2B). By germ band extended stage, *CG11798* was expressed throughout the mandibular, maxillary, and labial segments and in a striped pattern in all the thoracic and abdominal segments (Fig. 2C). Finally, after head involution occurred, *CG11798* was expressed throughout the CNS and PNS (Fig. 2F).

Unlike *CG5249*, *CG11798* did have an RNAi phenotype. The *Drosophila* mouth has several major cuticular structures, which are derived from the mandibular, maxillary, and labial segments. These include the cirri, mouth hooks, median tooth, H-piece, lateral processes, cephalopharyngeal plates, and several sensory organs including the antennal, maxillary, and labial sense organs (Campos-Ortega, 1997; Merrill et al., 1987; Regulski et al., 1987). In *CG11798* RNAi embryos, the larval mouth was disrupted. In particular the lateral processes were missing or shortened, the median tooth was absent, while the mouth hooks were present but slightly disrupted. In most cases the cirri and all the sense organs were absent, but in some instances the antennal and maxillary sense organs were present, likely the result of the efficacy of the RNAi. In addition to the mouth defects, we observed holes in the cuticle near the mouth (Fig. 3). In some cases the denticles were disrupted or fused, however, these were likely due to the injection process. The RNAi results were once again supported by deficiency  $y^1w^*$ ,  $Dp(1;Y) y^+$ ;  $Df(2R)XTE-58/Cyo$  (51D1-2;51E5) (referred to as *Df(2R)XTE-58*) available from the *Drosophila* Stock center. In homozygous *Df(2R)XTE-58* embryos the same structures were absent and disrupted, although *Df(2R)XTE-58* embryos appeared to be a little more severe since the

```

Sequences producing significant alignments:
Score E
(bits) Value
gb|AAAB01008960.1| Anopheles gambiae str. PEST, whole genome sho... 358 1e-
098
>gb|AAAB01008960.1| Anopheles gambiae str. PEST, whole genome shotgun sequence
Length = 23099915

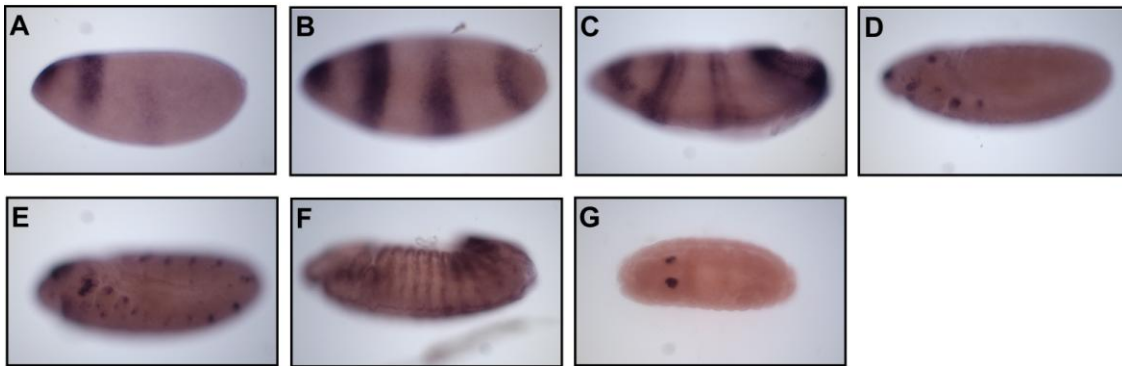
Score = 358 bits (918), Expect = 1e-098
Identities = 162/169 (95%), Positives = 165/169 (96%)
Frame = -1

Query: 1 S P L S P N S L A S R G Y R S L P Y P L R K K D G K M H Y E C N V C C T F G Q L S N L K V H L R T H S G E R P F F K C N 60
SP+SP S SRGYRSLPYPL+K+DGKMHYECNVC KTFQLSNLKVLHRLTHSGERPFKCN
Sbjct: 20486099 S P M S P G S P N S R G Y R S L P Y P L R R D G K M H Y E C N V C S K T F G Q L S N L K V H L R T H S G E R P F F K C N
20485920

Query: 61 V C T K S F T Q L A H L Q K H H L V H T G E K P H Q C D I C K R F S S T S N L K T H L R L H S G Q K P Y A C D L C P Q 120
V C T K S F T Q L A H L Q K H H L V H T G E K P H Q C D I C K R F S S T S N L K T H L R L H S G Q K P Y A C D L C P Q
Sbjct: 20485919 V C T K S F T Q L A H L Q K H H L V H T G E K P H Q C D I C K R F S S T S N L K T H L R L H S G Q K P Y A C D L C P Q
20485740

Query: 121 K P T Q F V H L K L H R L R L H T N D R P Y V C G C D K K Y I S A S G L R T H W K T T S C K P N N 169
K P T Q F V H L K L H R L R L H T N D R P Y V C G C D K K Y I S A S G L R T H W K T T S C K P N N
Sbjct: 20485739 K P T Q F V H L K L H R L R L H T N D R P Y V C G C D K K Y I S A S G L R T H W K T T S C K P N N 20485593

```



**Figure 1. Alignment of *CG5249* to the *Anopheles* genome and *Drosophila* embryonic expression.**

(Top) *CG5249* shares 95% identity to *Anopheles* within the zinc finger region and has an E value of  $1e^{-098}$ . (Bottom) Expression of *CG5249* RNA in *Drosophila* embryos. Anterior is to the left and dorsal is up. (A) Expression is first observed during blastoderm stages in the anterior end of the embryo. (B) Expression then becomes three broad stripes in the embryo with a cap of expression at the anterior. (C) During gastrulation the three stripes appear to divide into two fainter stripes a piece. (D) Once the germ band starts to extend expression is observed as spots within the sensory organ PNS. (E) By germ band extended stage expression has spread to the abdominal and thoracic segments (F) During germband retraction expression borders the segmental furrows. (G) After head involution is complete expression is seen possibly within the lobes of the brain.

```

Sequences producing significant alignments:
Score E
(bits) Value
gb|AAAB01008966.1| Anopheles gambiae str. FEST, whole genome sho... 238 3e-
062
>gb|AAAB01008966.1| Anopheles gambiae str. FEST, whole genome shotgun sequence
Length = 3863510
Score = 238 bits (608), Expect = 3e-062
Identities = 131/328 (39%), Positives = 164/328 (49%), Gaps = 61/328 (18%)
Frame = +2
Query: 27 KRYACTHCPYSTDRRDLVTRHENIHKDEKPFQCVACLKQFNADHVKKHFLRMHRELYD 86
KRY+CT CPY+TDRRDL+TRHENIHKDEKPFQCVACLK FNADHVKKHFLRMHREL+YD
Sbjct: 727676 KRYSCTSCPVTDRRDLVTRHENIHKDEKPFQCVACLKQFNADHVKKHFLRMHRELYD
727855

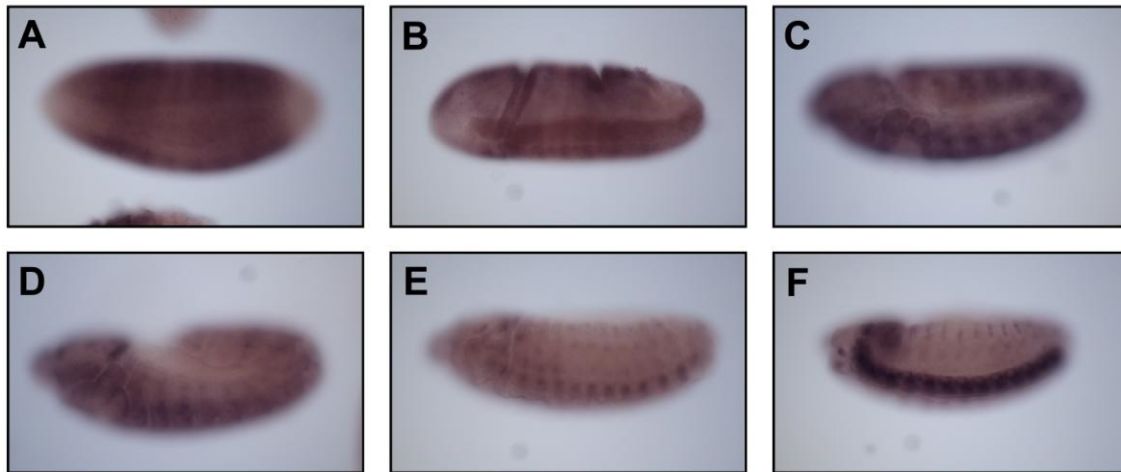
Query: 87 INKTRRH 93
I KTRRH
Sbjct: 727856 IAKTRRH 728063

Query: 189 -----XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX 227
K K EK+P C YCFW
Sbjct: 728216 QOTHANLHPAQLAIKQERKSSGSLNANSAGGSPDQYAKKI EGKRFCTCYCPWAGKR
728395

Query: 228 -----SGADWGLKSHLNTHTKPFVCLLDYKAARSERLATHV 265
+GAD WGLSHLNTHTKPFVCLLDYKAARSERLATHV
Sbjct: 728396 VWEVDFKTAREPHTNCNRSPTGADWGLKSHLNTHTKPFVCLLDYKAARSERLATHV
728575

Query: 266 LKVNKRCACSCSYLADTQREYQAIMSD 293
LKVNK+ACSKC++ A+ Q + +AR+ +
Sbjct: 728576 LKVNKRCACSKCNFFADQQAQLAHLQE 728659

```



**Figure 2. Alignment of *CG11798* to the *Anopheles* genome and *Drosophila* embryonic expression.**

(Top) *CG11798* is highly conserved within the zinc finger regions, but between zinc fingers there are considerable gaps, even with the gaps the E values is  $3e^{-062}$ . (Bottom) Expression of *CG11798* in *Drosophila* embryos. Anterior is to the left and dorsal is up. (A) *CG11798* is expressed almost entirely throughout the embryo during blastoderm stages, expression is not observed in the anterior and posterior ends. Faint stripes can also be observed within the expression pattern (B) During gastrulation the faint stripes are still visible. (C) By germ band extended stage expression is throughout the gnathal lobes and as broad stripes throughout all abdominal and thoracic segments. (D) As the germ band retracts the expression in the abdominal and thoracic segments is refined to narrower stripes. (E,F) As the embryo undergoes dorsal closure expression is limited to the CNS and PNS.

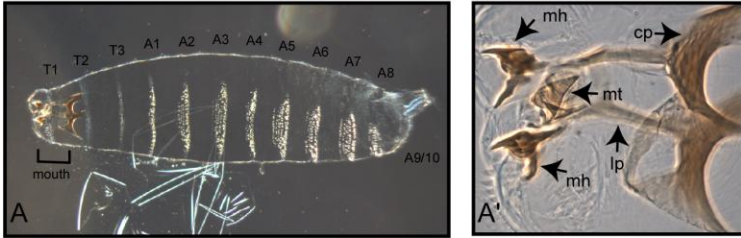
cephalopharyngeal plates also seemed to be disrupted (Fig. 3). Since our initial studies with this gene, a more thorough study has been done by Escudero et al. (2005)

## **Discussion**

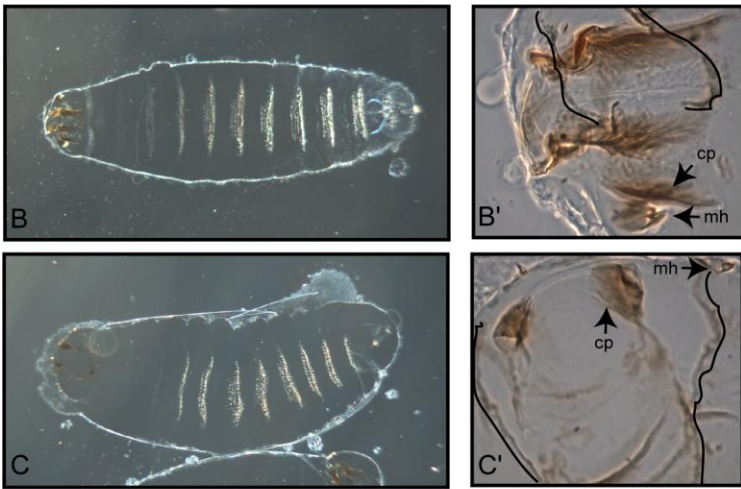
The goal of our research was to identify uncharacterized epidermal patterning genes that may have been overlooked due to redundancy. Our initial approach did not produce any likely candidates for epidermal pattern formation genes. Very few of the genes we examined were expressed in the embryo and of the few that were, most were expressed in the CNS or PNS. So how do we account for the results that we obtained? One of the problems may have arisen from the initial way we conducted our analysis to identify potentially duplicated genes. In order to get a better idea if a gene was truly duplicated it would have been more informative to make comparisons with other insect species. For example if we suspect a gene may have arisen due to duplication in *Drosophila melanogaster*, we might expect that earlier insect species would have only one copy of the gene indicating that the *Drosophila melanogaster* gene is a duplicate. In addition, comparison with other species would allow insight into when the duplication event took place. Also, this method would reduce the chances of us identifying potential duplicates due to domain shuffling. However, at the time that this study was conducted, the genome sequences of many other insect species were not available, so comparisons could only be made between *Drosophila melanogaster* and *Anopheles gambiae*. Another possibility is that these genes simply are not expressed during embryogenesis and are expressed later in development. Finally, we were dealing with computed genes; which may not actually be transcribed and translated as described *in vivo*.

**Figure 3. *Df(2R)XTE-58* and RNAi phenotypes of *CG11798* compared to wild type.** (A, A') Wild type *Drosophila* larval cuticle. (A) T1-T3 indicate the thoracic denticle belts, A1-A9/10 indicate the abdominal denticle belts. (A') 100x magnification of the the mouth. The abbreviations are as follows mh-mouth hooks, lp-lateral processes, mt-median tooth, cp-cephalopharyngeal plates. (B,B', C, C') Two different *Df(2R)XTE-58* homozygous embryos. The median tooth as well as the lateral processes are missing. The mouth hooks are occasionally there but they are abnormal as are the cephalopharyngeal plates. In addition there appear to be holes in the cuticle which are outlined. (D, D', E, E') Two larval RNAi phenotypes of *CG11798*. The mouth and cuticle are disrupted as in the deficiency. However, the denticle belts are also disrupted, this is most likely due to the injection procedure.

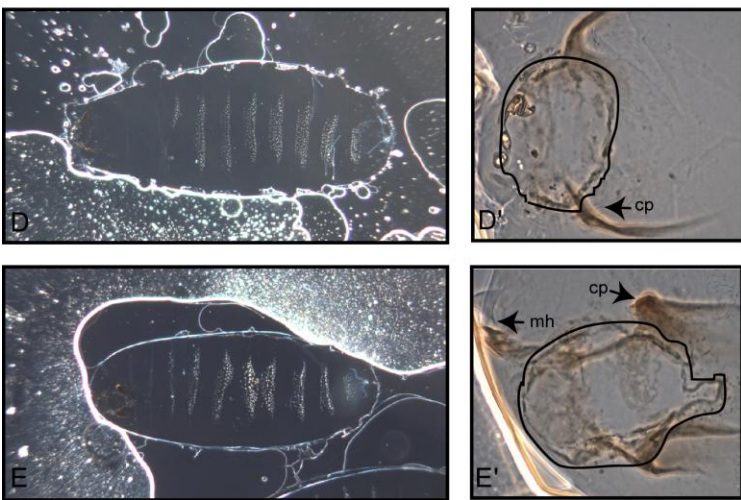
## Wild Type



## Deficiency Phenotypes



## RNAi Phenotypes



Our second approach was more fruitful than the first. By using BLAST to compare between *Anopheles gambiae* and *Drosophila melanogaster*, we were able to identify 12 genes that were highly conserved between the two species. Interestingly, all but two of the genes identified through this method gave expression patterns, and most were expressed in the CNS/PNS. Perhaps it is not too surprising that most of the genes gave CNS/PNS expression patterns, because like any other developmental pathway that is necessary for survival, essential genes within those pathways would be expected to be highly conserved across species. Genes important for *Drosophila* development such as *disco*, *Odd-skipped*, and *D-Sp1* are all conserved in *C. elegans* and humans (Knight and Shimeld, 2001). We were particularly interested in *CG5249* and *CG11798* because their expression pattern indicated they might be involved in segmental pattern formation.

RNAi with *CG5249* produced embryonic lethality, however, there were no obvious segmentation defects. It was later demonstrated that the lethality resulted from tracheal abnormalities in *CG5249* RNAi embryos (Ng et al., 2006). Ng et al. named *CG5249*, *blimp-1*, which is homologous to the *blimp-1* gene found in other invertebrates and vertebrates. In vertebrates, this gene has a variety of functions ranging from specification of slow twitch muscles and neural crest progenitors in zebrafish to conferring forebrain or head identity to certain tissues in *Xenopus* embryos (Ng et al., 2006; Roy and Ng, 2004). How or if these functions correlate to *Drosophila* remains to be seen. Though neither Ng et al. (2006) nor we have been able to show a role for *blimp-1* in segmental patterning, there might still be one. Redundancy with other genes may account for no obvious segmentation defects. Two possible candidates are *CG15269* and

*CG17368* to which it shares 43-48% identity within the zinc finger regions, though we have not examined their expression patterns.

*CG11798* has been examined more thoroughly by Escudero et al. (2005), since our initial studies. Though identified in our screen, *CG11798* was initially identified in a P-element screen for peripheral nervous system development mutants and was referred to as the *charlatan* locus (Kania et al., 1995). The *CG11798* transcript is an alternatively spliced form of the *charlatan* gene (Escudero et al., 2005). Escudero et al. (2005) examined the expression of *charlatan* and mutant phenotypes during embryonic and adult development. The embryonic expression pattern from our studies corresponds exactly with those from Escudero et al. (2005). Escudero et al. (2005) were also able to generate a null allele of *charlatan* by imprecise excision of the p-element insertion *Pl(2)42/18* (Escudero et al., 2005). The imprecise excision produced a potential null allele of *charlatan* since part of the promoter was removed and *in situ* hybridization to mRNA in homozygous embryos does not produce a signal. Embryos that are homozygous null for *charlatan* are missing many neurons and the chordotonal lateral neurons are abnormal. However, there is no mention of disrupted head segmentation as we observed with our RNAi phenotypes or with *Df(2R)XTE-58*. Since Escudero et al. (2005) were focused on the nervous system aspect of *charlatan* they may have overlooked the head segmentation defects, especially if they never observed cuticle phenotypes of *charlatan* mutants. A head segmentation role might also be expected, because there is strong expression within the mandibular, maxillary, and labial segments and these are the exact segments that will produce many of the mouth cuticular structures and sensory organs that are disrupted by

RNAi and the *Df(2R)XTE-58* (Panzer et al., 1992; Regulski et al., 1987). It would be interesting to obtain this allele of *charlatan* and examine the cuticle phenotype.

Since we were able to identify only very few genes that might be involved in segmental pattern formation, does this mean that we are close to achieving saturation since Nusslein-Volhard and Wieschaus conducted their initial screens? Not necessarily. One obvious reason is that in our screen we have examined only a subset of genes, specifically those that encode for C2H2 zinc finger transcription factors. Studies similar to the one we conducted would have to be carried out for all proteins in the *Drosophila* genome, and perhaps they should be. There have been many studies examining duplication, but usually the studies focus on a few sets of genes. Of the studies that do focus on duplication on a genome wide scale, very few examine the functional significance of the duplication event (Thomas, 2006; Tischler et al., 2006). Conducting studies similar to the one conducted here with other classes of proteins would help in our understanding of pattern formation.

## REFERENCES

- Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J.** (1990). Basic local alignment search tool. *J Mol Biol* **215**, 403-10.
- Brown, S. J., Mahaffey, J. P., Lorenzen, M. D., Denell, R. E. and Mahaffey, J. W.** (1999). Using RNAi to investigate orthologous homeotic gene function during development of distantly related insects. *Evol Dev* **1**, 11-5.
- Campos-Ortega, J. A. a. H., V.** (1997). The embryonic development of *Drosophila melanogaster*. Berlin; New York: Springer.
- Cooke, J., Nowak, M. A., Boerlijst, M. and Maynard-Smith, J.** (1997). Evolutionary origins and maintenance of redundant gene expression during metazoan development. *Trends Genet* **13**, 360-4.
- Dietzl, G., Chen, D., Schnorrer, F., Su, K. C., Barinova, Y., Fellner, M., Gasser, B., Kinsey, K., Oppel, S., Scheiblaue, S. et al.** (2007). A genome-wide transgenic RNAi library for conditional gene inactivation in *Drosophila*. *Nature* **448**, 151-6.
- Dolezal, T., Gazi, M., Zurovec, M. and Bryant, P. J.** (2003). Genetic analysis of the ADGF multigene family by homologous recombination and gene conversion in *Drosophila*. *Genetics* **165**, 653-66.
- Dong, P. D., Todi, S. V., Eberl, D. F. and Boekhoff-Falk, G.** (2003). *Drosophila* spalt/spalt-related mutants exhibit Townes-Brocks' syndrome phenotypes. *Proc Natl Acad Sci U S A* **100**, 10293-8.
- Dorner, S., Lum, L., Kim, M., Paro, R., Beachy, P. A. and Green, R.** (2006). A genomewide screen for components of the RNAi pathway in *Drosophila* cultured cells. *Proc Natl Acad Sci U S A* **103**, 11880-5.
- Elbashir, S. M., Lendeckel, W. and Tuschl, T.** (2001). RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes Dev* **15**, 188-200.
- Escudero, L. M., Caminero, E., Schulze, K. L., Bellen, H. J. and Modolell, J.** (2005). Charlatan, a Zn-finger transcription factor, establishes a novel level of regulation of the proneural achaete/scute genes of *Drosophila*. *Development* **132**, 1211-22.
- Flockhart, I., Booker, M., Kiger, A., Boutros, M., Armknecht, S., Ramadan, N., Richardson, K., Xu, A., Perrimon, N. and Mathey-Prevot, B.** (2006). FlyRNAi: the *Drosophila* RNAi screening center database. *Nucleic Acids Res* **34**, D489-94.

- Gonzalez-Gaitan, M., Rothe, M., Wimmer, E. A., Taubert, H. and Jackle, H.** (1994). Redundant functions of the genes knirps and knirps-related for the establishment of anterior Drosophila head structures. *Proc Natl Acad Sci U S A* **91**, 8567-71.
- Guan, Y., Dunham, M. J. and Troyanskaya, O. G.** (2007). Functional analysis of gene duplications in *Saccharomyces cerevisiae*. *Genetics* **175**, 933-43.
- Hammond, S. M., Bernstein, E., Beach, D. and Hannon, G. J.** (2000). An RNA-directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells. *Nature* **404**, 293-6.
- Hannon, G. J.** (2002). RNA interference. *Nature* **418**, 244-51.
- Hart, M. C., Wang, L. and Coulter, D. E.** (1996). Comparison of the structure and expression of odd-skipped and two related genes that encode a new family of zinc finger proteins in *Drosophila*. *Genetics* **144**, 171-82.
- Heilig, J. S., Freeman, M., Lavery, T., Lee, K. J., Campos, A. R., Rubin, G. M. and Steller, H.** (1991). Isolation and characterization of the disconnected gene of *Drosophila melanogaster*. *EMBO J* **10**, 809-15.
- Kania, A., Salzberg, A., Bhat, M., D'Evelyn, D., He, Y., Kiss, I. and Bellen, H. J.** (1995). P-element mutations affecting embryonic peripheral nervous system development in *Drosophila melanogaster*. *Genetics* **139**, 1663-78.
- Kelley, C. M., Ikeda, T., Koipally, J., Avitahl, N., Wu, L., Georgopoulos, K. and Morgan, B. A.** (1998). Helios, a novel dimerization partner of Ikaros expressed in the earliest hematopoietic progenitors. *Curr Biol* **8**, 508-15.
- Ketting, R. F., Fischer, S. E., Bernstein, E., Sijen, T., Hannon, G. J. and Plasterk, R. H.** (2001). Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in *C. elegans*. *Genes Dev* **15**, 2654-9.
- Kim, K., Lee, Y. S., Harris, D., Nakahara, K. and Carthew, R. W.** (2006). The RNAi pathway initiated by Dicer-2 in *Drosophila*. *Cold Spring Harb Symp Quant Biol* **71**, 39-44.
- Knight, R. D. and Shimeld, S. M.** (2001). Identification of conserved C2H2 zinc-finger gene families in the Bilateria. *Genome Biol* **2**, RESEARCH0016.
- Koizumi, K., Higashida, H., Yoo, S., Islam, M. S., Ivanov, A. I., Guo, V., Pozzi, P., Yu, S. H., Rovescalli, A. C., Tang, D. et al.** (2007). RNA interference screen to identify genes required for *Drosophila* embryonic nervous system development. *Proc Natl Acad Sci U S A* **104**, 5626-31.

**Laugier, E., Yang, Z., Fasano, L., Kerridge, S. and Vola, C.** (2005). A critical role of teashirt for patterning the ventral epidermis is masked by ectopic expression of tiptop, a paralog of teashirt in *Drosophila*. *Dev Biol* **283**, 446-58.

**Lee, K. J., Freeman, M. and Steller, H.** (1991). Expression of the disconnected gene during development of *Drosophila melanogaster*. *EMBO J* **10**, 817-26.

**Lu, J., Ruhf, M. L., Perrimon, N. and Leder, P.** (2007). A genome-wide RNA interference screen identifies putative chromatin regulators essential for E2F repression. *Proc Natl Acad Sci U S A* **104**, 9381-6.

**Mahaffey, J. W., Griswold, C. M. and Cao, Q. M.** (2001). The *Drosophila* genes disconnected and disco-related are redundant with respect to larval head development and accumulation of mRNAs from deformed target genes. *Genetics* **157**, 225-36.

**McCarty, A. S., Kleiger, G., Eisenberg, D. and Smale, S. T.** (2003). Selective dimerization of a C2H2 zinc finger subfamily. *Mol Cell* **11**, 459-70.

**Merrill, V. K., Turner, F. R. and Kaufman, T. C.** (1987). A genetic and developmental analysis of mutations in the Deformed locus in *Drosophila melanogaster*. *Dev Biol* **122**, 379-95.

**Morgan, B., Sun, L., Avitahl, N., Andrikopoulos, K., Ikeda, T., Gonzales, E., Wu, P., Neben, S. and Georgopoulos, K.** (1997). Aiolos, a lymphoid restricted transcription factor that interacts with Ikaros to regulate lymphocyte differentiation. *EMBO J* **16**, 2004-13.

**Ng, T., Yu, F. and Roy, S.** (2006). A homologue of the vertebrate SET domain and zinc finger protein Blimp-1 regulates terminal differentiation of the tracheal system in the *Drosophila* embryo. *Dev Genes Evol* **216**, 243-52.

**Nowak, M. A., Boerlijst, M. C., Cooke, J. and Smith, J. M.** (1997). Evolution of genetic redundancy. *Nature* **388**, 167-71.

**Nusslein-Volhard, C. and Wieschaus, E.** (1980). Mutations affecting segment number and polarity in *Drosophila*. *Nature* **287**, 795-801.

**Panzer, S., Weigel, D. and Beckendorf, S. K.** (1992). Organogenesis in *Drosophila melanogaster*: embryonic salivary gland determination is controlled by homeotic and dorsoventral patterning genes. *Development* **114**, 49-57.

**Pederson, J. D., Kiehart, D. P. and Mahaffey, J. W.** (1996). The role of HOM-C genes in segmental transformations: reexamination of the *Drosophila* Sex combs reduced embryonic phenotype. *Dev Biol* **180**, 131-42.

- Regulski, M., McGinnis, N., Chadwick, R. and McGinnis, W.** (1987). Developmental and molecular analysis of Deformed; a homeotic gene controlling Drosophila head development. *EMBO J* **6**, 767-777.
- Reuter, D., Kuhnlein, R. P., Frommer, G., Barrio, R., Kafatos, F. C., Jackle, H. and Schuh, R.** (1996). Regulation, function and potential origin of the Drosophila gene spalt adjacent, which encodes a secreted protein expressed in the early embryo. *Chromosoma* **104**, 445-54.
- Roy, S. and Ng, T.** (2004). Blimp-1 specifies neural crest and sensory neuron progenitors in the zebrafish embryo. *Curr Biol* **14**, 1772-7.
- Ryan, R. F. and Darby, M. K.** (1998). The role of zinc finger linkers in p43 and TFIIIA binding to 5S rRNA and DNA. *Nucleic Acids Res* **26**, 703-9.
- Schmid, A., Schindelholz, B. and Zinn, K.** (2002). Combinatorial RNAi: a method for evaluating the functions of gene families in Drosophila. *Trends Neurosci* **25**, 71-4.
- Schmidt, D. and Durrett, R.** (2004). Adaptive evolution drives the diversification of zinc-finger binding domains. *Mol Biol Evol* **21**, 2326-39.
- Schock, F., Purnell, B. A., Wimmer, E. A. and Jackle, H.** (1999). Common and diverged functions of the Drosophila gene pair D-Sp1 and buttonhead. *Mech Dev* **89**, 125-32.
- Steller, H., Fischbach, K. F. and Rubin, G. M.** (1987). Disconnected: a locus required for neuronal pathway formation in the visual system of Drosophila. *Cell* **50**, 1139-53.
- Tautz, D. and 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**, 81-5.
- Thomas, J. H.** (2006). Analysis of homologous gene clusters in Caenorhabditis elegans reveals striking regional cluster domains. *Genetics* **172**, 127-43.
- Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. and Higgins, D. G.** (1997). The CLUSTAL\_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* **25**, 4876-82.
- Tischler, J., Lehner, B., Chen, N. and Fraser, A. G.** (2006). Combinatorial RNA interference in Caenorhabditis elegans reveals that redundancy between gene duplicates can be maintained for more than 80 million years of evolution. *Genome Biol* **7**, R69.

**Whitely, M., and Kassis, J. A.** (1997). Rescue of *Drosophila engrailed* mutants with a highly divergent mosquito *engrailed* cDNA using a homing, enhancer-trapping transposon. *Development*. 124: 1531-1541.

## **CHAPTER SIX**

### **Summary and Future Work**

## Summary

Development of the *Drosophila* embryo occurs through the coordination of very complex processes. The initial processes are the establishment of the anterior-posterior and dorsal-ventral axes. After determining axes, the segments must be specified and given identity. The Hox proteins, which are homeodomain-containing transcription factors, play a critical role in specification of segment identity. Hox proteins, initially identified in *Drosophila*, have been extensively studied and found to be conserved across many phyla (Lewis, 1978; McGinnis and Krumlauf, 1992). Expression of a Hox protein outside its normal domain often leads to a transformation in segmental identity due to differential expression of Hox target genes. Each Hox protein is thought to regulate a battery of target genes to produce a distinct segmental identity.

Though extensive work has been done to determine how Hox proteins control segment identity, numerous questions still remain as to how they function. In particular, questions about how Hox proteins regulate target gene expression and the identity of Hox target genes remain. *In vitro* Hox proteins bind to similar sites which contain TAAT at the core; this lack of binding specificity cannot account for their high developmental specificity (Biggin and McGinnis, 1997; Ekker et al., 1994). Due to this paradox much effort has been put into identifying cofactors that might increase binding specificity, target gene selection, and/or recruitment of other transcription factors that might affect Hox protein function. To date, however, very few cofactors have been identified, and only Extradenticle (Exd) has been well characterized (Chan et al., 1994; Chan et al., 1996; Rauskolb et al., 1993; Rauskolb and Wieschaus, 1994; Ryoo and Mann, 1999;

Ryoo et al., 1999).

The *Drosophila* gene *disconnected* (*disco*), which encodes a C2H2 zinc finger transcription factor (Heilig et al., 1991), behaves as a genetic cofactor for the *Hox* genes *Deformed* (*Dfd*) and *Sex combs reduced* (*Scr*) (Mahaffey et al., 2001). In the absence of *disco* and its redundant partner *disco-related* (*disco-r*), *Dfd* and *Scr* cannot properly regulate several target genes such as *Dll*, *1.28*, and *Serrate* within the gnathal segments.

The research presented here is 1) a molecular characterization of Disco as a potential transcription factor and 2) a description of Disco's role in adult *Drosophila* development. The data indicate that Disco can bind *in vitro* to sites on DNA that contain 5'-TGACA-3' and has the potential to act as a repressor, based on its interaction with the corepressor C-terminal Binding Protein (CtBP). The interaction between CtBP and Disco does appear to affect denticle morphology. Ectopic Disco expression leads to fewer often disrupted denticles, this is not the case observed with ectopic expression of Disco CtBP, where the denticles appear to be normal most of the time. The conservation of the PLDLS site in ecdysozoa and lophotrochozoa suggests that it is functionally significant. In addition, the ecdysozoan and lophotrochozoan Disco proteins have a conserved cysteine (at position 102 in *Drosophila*) of the protein that is somehow important for function. During adult *Drosophila* development, ectopic Disco transforms dorsal appendages toward a ventral fate. This transformation is most obvious in the wing where ectopic Disco often transforms the wing into a well-formed leg. Furthermore, cells that lack Disco and *Disco-r* do not contribute to the ventral appendages. These data combined with work done in *Tribolium castaneum* by Lisa Robertson and Nathaniel Grubbs led us

to rework our previous model in which *disco*, along with *tsh*, act as anterior-posterior patterning factors to a model in which they act as proximal-distal factors. In our current model, we propose that genes such as *disco* and *tsh* are acting as appendage factors in the epidermis of the *Drosophila* embryo as well as in the adult, and that this appendage role is masked by the highly derived maggot-style larval body of flies.

All organisms are patterned along three axes: the anterior-posterior (A-P), dorsal-ventral (D-V), and proximal-distal (P-D). A-P and D-V axes in *Drosophila* larvae are easily identified; however, patterning along the P-D axis is not as obvious since the *Drosophila* larva lacks overt appendages, unlike most insects. The Keilin's organ, a sensory structure located in the thoracic segments, is thought to be a remnant of larval legs (Angelini and Kaufman, 2005; Bolinger and Boekhoff-Falk, 2005; Cohen and Jurgens, 1989). In embryos lacking the *disco* genes, development of the Keilin's organ is incomplete, suggesting that the appendage role observed in *Tribolium* is conserved in *Drosophila* (Patel et al., 2007). We extend this into the head segments where we propose that *disco* is still acting as an appendage factor within the gnathal segments and that this role is masked by the highly derived maggot-style larval body. This work contributes to our overall knowledge of *disco*'s biological function and understanding of Disco as a transcription factor. However, it also opens many new avenues of possible research.

#### **Future work for Disco's role in gene regulation**

Since Disco binding sites have been identified, genes that are potentially directly regulated by Disco can be identified. Several possible candidates include *pannier* (*pnr*), *Dll*, and *dachshund* (*dac*). *Dll* and *dac* are ectopically activated when *disco* is ectopically

expressed in the imaginal wing disc, while *pnr* is repressed by ectopic *disco* expression in the embryo. Whether Disco directly regulates any of these genes remains to be determined: however, these few genes are probably not the only ones regulated by Disco. It would be useful to take several of the binding site prediction programs to identify regions of the *Drosophila* genome that contain clusters of Disco binding sites. These data can be subsequently used to identify potential targets of Disco. One such program that might be used is Cis-analyst (<http://rana.lbl.gov/cis-analyst/>), which allows input of user specified DNA binding sites and allows up to four variations that it then uses to search the genomes of *Drosophila melanogaster* and *Drosophila pseudoobscura* for conserved binding sites. This sort of analysis could also be done using all 12 *Drosophila* species, which would provide more accurate predictions (Stark et al., 2007a; 2007b).

While the binding sites that we identified are sites bound *in vitro*, it is also important to identify sites that will be bound *in vivo*. A potential method to identify *in vivo* binding sites is chromatin immunoprecipitation. However, this method requires a good Disco antibody, which the lab has been unable to produce. One way to get around this problem might be to add an epitope tag to Disco and use an antibody to the tag in chromatin immunoprecipitation studies. This procedure would not only allow *in vivo* identification of Disco binding sites but the sequences derived from these studies could be used in the prediction programs to identify potential targets of Disco. Another possibility might be to use Disco-r since it is highly conserved and there are more *disco-r* like genes in other species.

### ***Distal-less* as a potential direct target of Disco**

In our studies we identified *disco* as being able to repress expression of the *rpr4S3* reporter and endogenous *rpr* expression. However, *disco* is also required for the activation of genes such as *Distal-less (Dll)* and *1.28* within the maxillary segment, though it is not known whether this activation is direct (Mahaffey et al., 2001). So how do we reconcile the two apparent contrasting roles of *disco*? The likely answer is that *disco*'s function is context dependent. The *Dll* and *1.28* genes are also regulated by *Dfd* suggesting that Disco plus the Hox protein *Dfd* are required for the proper activation of these two genes (O'Hara et al., 1993; Pederson et al., 2000). To date, no direct interaction between *Dfd* and Disco, which one would expect for a cofactor, has been demonstrated (Imolczyk B, Patel MK, and Mahaffey JW unpublished observation). Recently, however, it has been suggested that too much emphasis has been placed on cooperative binding between Hox proteins and cofactors to DNA and rather than on what has been termed collaborative interactions that are just as important (Walsh and Carroll, 2007).

In collaborative interactions, the identity of the collaborative protein and/or topology of cis-regulatory elements play a critical role in determining whether the target gene is activated or repressed (Walsh and Carroll, 2007). For example, homologs of Mothers against Dpp (Smad proteins) can activate or repress target genes, and the ability to do one or the other depends on the topology of the binding sites. The topology of the sites allows recruitment of the corepressor Schnurri (*Shn*) and if this topology is disrupted the target gene is activated (Gao et al., 2005; Pyrowolakis et al., 2004; Walsh

and Carroll, 2007). Could Disco be acting in a similar manner? The ability of ectopic Disco to transform the wing into a T2 leg, indicates that Disco might be able to act as a potential cofactor with several Hox proteins although no direct interaction with any Hox protein has been demonstrated, suggesting that collaborative interactions may play an important role in Disco function.

The *rpr4S3* reporter is a good candidate to address this question because it contains both Dfd and Disco binding sites. Interestingly, there is a Disco binding site adjacent to a Dfd binding site in this reporter that is conserved in twelve *Drosophila* species. However, the significance of this relationship cannot be tested because we have been unable to create transgenic flies carrying the mutated Disco binding sites. Therefore, perhaps attention should be directed at finding other potential target genes. Another candidate enhancer region to address this question might be the *Dll* enhancer region, ETD6, which controls expression of *Dll* in the ventral lateral regions of the maxillary segment. These regions are partially responsible for producing the cirri (O'Hara et al., 1993). Several lines of evidence suggest that this might be a good candidate region. First *Dll* expression in the ventral lateral region appears to be affected by loss of the *disco* genes and Dfd, respectively (Mahaffey et al., 2001; O'Hara et al., 1993). Second, the expression of *Dll* in the ventral lateral region also overlaps the expression of *disco* and *Dfd*. Third, loss of the *disco* genes or *Dfd* leads to loss of cirri (Mahaffey et al., 2001; Merrill et al., 1987; Regulski et al., 1987). Lastly, ectopic Dfd produces cirri only in regions where Disco is expressed (Robertson et al., 2004).

A search of the ETD6 enhancer reveals that there are six potential Disco binding

sites within this region, however, ETD6 is roughly 6 kb long, and one would expect to observe six binding sites similar to Disco binding sites by chance. The significance of these binding sites remains to be tested. Reporter expression driven by the ETD6 regulatory region is dependent on Dfd protein expression and the enhancer contains Dfd binding sites. Dependence of ETD6 reporter expression on Disco along with whether or not Disco can bind to this regulatory region is unknown. If reporter expression of ETD6 is affected by loss or gain of function of Disco, and Disco can bind the enhancer region, ETD6 could be used to examine transcriptional regulation by Dfd and Disco while also addressing questions about the importance of topology of Disco and Dfd binding. Insight into Disco functioning as a collaborative factor also might be gained. Unfortunately, fly stocks containing this enhancer construct have been lost and must be reconstructed.

### **Further evidence to support that Disco acts as an appendage factor during embryogenesis**

In our model Disco is acting as an appendage factor during embryogenesis and, thus far, the evidence provided for this is minimal. The argument can be strengthened by demonstrating that *Tribolium* Disco can rescue *disco/disco-r* mutants. If Disco's role in *Drosophila* embryogenesis is a derived function specific to *Drosophila*, it might be expected that the *Tribolium disco* gene could not rescue *disco/disco-r* mutant embryos. However, if Disco is acting as an appendage factor during embryogenesis then *Tribolium* Disco should be able to rescue *disco/disco-r* mutants. In order to test this hypothesis *Tribolium* Disco would have to be cloned into a UAS vector and crossed to a Gal4 driver that drives expression similar to endogenous *disco*. At the moment there is no such Gal4

construct, however, Dll-Gal4 drives expression in a subset of cells that also express Disco, specifically within the cells of the thoracic limb primordia and some regions of the maxillary lobe. Rescue of the cirri and Keilin's organ might be expected if *Dll*-Gal4 driven *Tribolium* Disco can rescue *disco/disco-r* mutants.

### **Other potential protein interactions with Disco**

One of the keys to understanding Disco function is to understand which protein-protein interactions are important. Information on two other Disco interacting proteins have been published; Disco interacting protein 1 (DIP1) and Disco interacting protein 2 (DIP2) (DeSousa et al., 2003; Mukhopadhyay et al., 2002). *dip2* encodes a protein with no known functional domains that is evolutionarily conserved (Mukhopadhyay et al., 2002). It is expressed in the nervous system of both *Drosophila* and mice. *dip1* encodes a protein that has two double stranded RNA binding domains, which can bind dsRNA *in vitro*, as well as three other potential biologically significant motifs (Bondos et al., 2004; DeSousa et al., 2003). In addition to interacting with Disco, DIP1 can also interact with Ubx, this interaction prevents transcriptional activation by Ubx. Ectopic expression of DIP1 in the eye-antennal disc produces ectopic bristles on the head capsule, ectopic formation of an apparent antenna, duplication of antennal palps and deletion of arista. Within the wing and haltere, ectopic expression of DIP1 produces shriveled wings with patches of bristles and halteres that are misshapen (Bondos et al., 2004; DeSousa et al., 2003). Interestingly, even though DIP1 and DIP2 were identified as interacting with Disco, no data have been presented on how these two proteins might affect Disco function. Also no data have been presented about which regions within Disco are

required for the interaction.

Although the CtBP interaction has been identified, Disco also has the potential to interact with another corepressor, Groucho. Disco contains a site similar to the engrailed homology 1 motif which interacts with Groucho (Flores-Saaib et al., 2001). Several other proteins also contain CtBP and Groucho interaction motifs such as Hairless, Hairy, TCF-4, and Brinker (Hasson et al., 2001; Koelzer and Klein, 2006; Nagel et al., 2005; Phippen et al., 2000; Valenta et al., 2003). It has been suggested that in situations where a protein can interact with both CtBP and Groucho, CtBP may be acting as an attenuator of Groucho activity, which is ubiquitously expressed (Phippen et al., 2000).

Another potential interaction that should be examined is that of Disco with Scr. Disco is a potential cofactor for Scr, but direct interaction tests have yet to be done. A direct interaction of Disco with Scr would be particularly interesting because it has already been demonstrated that Tsh can directly interact with Scr (Taghli-Lamalle et al., 2007). Does differential interaction between Scr and Disco lead to differential regulation than Scr and Tsh? Genetic studies indicate, yes. Does this account for differential gene regulation between the labial and prothoracic segments? It has also been demonstrated that Disco and Scr can activate the transcription of *pb* better than Scr alone (Robertson et al., 2004). However, nothing is known as to how this regulation occurs and, perhaps even more importantly, how this regulation differs from Disco with Dfd? The *pb* enhancer region has been characterized to some extent and could provide a good candidate region to start examining these questions (Kapoun and Kaufman, 1995).

## **Involvement of appendage genes in cardiac development**

In addition to the roles of *Disco* in embryonic head development and adult appendage development, *Disco* might also have roles in the formation of the *Drosophila* embryonic heart, the dorsal vessel, the peripheral nervous system, and the visceral mesoderm. Both *disco* genes are expressed in these regions. The *disco* genes appear to be expressed in the cardiac primordia of the dorsal vessel, however, the exact cells within which this expression occurs remain to be determined. Co-localization studies with genes that are known to be expressed within specific cells of the dorsal vessel would help clarify the exact expression domain of the *disco* genes. Potential candidate genes for these studies are *tinman*, *pnr*, *wingless*, *hedgehog* and the *Hox* genes *Antp*, *Ubx*, *abd-A*, and *Abd-B*, which mark either cardiac precursors or specific regions of the dorsal vessel (Bodmer, 1993; Gajewski et al., 1999; Lo et al., 2002; Park et al., 1996). These genes would also be useful in identifying structures that might be disrupted in embryos lacking the *disco* genes.

Interestingly, other appendage genes are also involved in the cardiac system. The human homolog of *Dachshund*, a protein required for medial appendage development in *Drosophila*, is expressed in the human heart, but its functional role within the vertebrate cardiac system is unknown (Kozmik et al., 1999). No expression for *Dac* within the *Drosophila* embryonic heart has been reported. Perhaps more interesting considering its relationship with *Disco*, is that *Teashirt* is expressed in the dorsal vessel, perhaps in the pericardial cells and alary muscles of the dorsal vessel (Patel, MK and Mahaffey, JW unpublished observation). To verify that *Teashirt* is expressed in these regions,

colocalization with *tailup* should be examined. Lastly, in the *Drosophila* embryo, *buttonhead*, another appendage factor, promotes hematopoietic differentiation through the activation of *serpent* (Yin et al., 1997). The role of these appendage genes within the cardiac system may be due to cooption events, but in order to make this determination they must be more carefully studied, placing *disco* into the proper framework of heart development

### **The role and function of Disco-r**

Much of the focus of this research has been on *disco*, however, many questions still remain for its redundant partner *disco-r*, particularly the reason why both are highly conserved in all *Drosophila* species examined. Are they truly redundant? And if so, why has one not been lost. Only recently has our lab created transgenic lines with *disco-r* for ectopic studies. Initial studies show that *disco-r* may have functions that are distinct from *disco*. Ectopically expressed *disco-r* in the wing imaginal discs does not produce as well formed legs as does ectopic *disco*. More surprisingly, ectopic *disco-r* induces ectopic maxillary palps on the head cuticle, something that we have not observed with ectopic *disco* expression. These data raise several questions. Can ectopic Disco-r activate genes such as *dac* and *Dll* as efficiently as Disco? What changes in gene regulation occur to produce ectopic maxillary palps? Does ectopic *disco* and *disco-r* expression in the embryo produce different phenotypes?

Disco contains a pair of C2H2 zinc finger motifs; however, Disco-r has two pairs. The first pair of Disco-r is highly conserved with the pair from Disco, and the residues that contact DNA are completely conserved. This suggest that Disco-r should be capable

of binding to the same sites as Disco, but the second pair of zinc fingers also suggests that Disco-r might be able to bind to sites that are not bound by Disco. These are also highly conserved in other insects with a second pair of fingers, and in homologs like Basonuclin, which has three pairs (includes deuterostomes and more basal insects). Additional pairs of zinc fingers may also be involved in protein-protein interactions (McCarty et al., 2003). Clearly, many questions need to be addressed in regards to Disco-r function. It would be interesting to ask whether the single *Tribolium disco* gene (actually more similar to *basonuclin* than *disco* and *disco-r*) is capable of performing both the functions of Disco and Disco-r. For example, would ectopic *Tribolium disco* expression in a *Drosophila* adult be able to produce ectopic maxillary palps and transform the wing into a leg?

### **Combinatorial expression of Disco with other Hox proteins**

Ectopic expression of Disco within the wing produces a well-formed second thoracic leg. This is important to note, because the wing and second thoracic legs are specified by the Hox protein Antennapedia (Gibson, 1988; Struhl, 1981). Thus ectopic Disco expression in the presence of Antp can specify second thoracic leg identity. Would ectopic expression of Disco plus other Hox proteins within the wing lead to different transformations? For example, would ectopic expression of Disco with Scr or Ubx produce prothoracic or metathoracic leg identities respectively? Would the same results be observed with Disco-r? The answers to these questions might provide additional evidence that Disco is acting as a collaborative factor. Perhaps in the adult, Disco is required for ventral appendage development but the specific identity of that ventral

appendage is determined by which Hox protein is functionally active within that region. Does this imply parallel pathways, or could this still imply integration of the proximal/distal axis and Hox functions?

In conclusion, the work from this dissertation presents a novel role for Disco in appendage development and provides data that will be useful in further studies examining Disco target gene identification and function. In addition, these studies and future studies prompted by this work will contribute to the overall knowledge and understanding of highly conserved developmental mechanisms.

## REFERENCES

- Angelini, D. R. and Kaufman, T. C.** (2005). Insect appendages and comparative ontogenetics. *Dev Biol.*
- Biggin, M. D. and McGinnis, W.** (1997). Regulation of segmentation and segmental identity by Drosophila homeoproteins: the role of DNA binding in functional activity and specificity. *Development* **124**, 4425-33.
- Bodmer, R.** (1993). The gene tinman is required for specification of the heart and visceral muscles in Drosophila. *Development* **118**, 719-29.
- Bolinger, R. A. and Boekhoff-Falk, G.** (2005). Distal-less functions in subdividing the Drosophila thoracic limb primordium. *Dev Dyn* **232**, 801-16.
- Bondos, S. E., Catanese, D. J., Jr., Tan, X. X., Bicknell, A., Li, L. and Matthews, K. S.** (2004). Hox transcription factor ultrabithorax Ib physically and genetically interacts with disconnected interacting protein 1, a double-stranded RNA-binding protein. *J Biol Chem* **279**, 26433-44.
- Chan, S. K., Jaffe, L., Capovilla, M., Botas, J. and Mann, R. S.** (1994). The DNA binding specificity of Ultrabithorax is modulated by cooperative interactions with extradenticle, another homeoprotein. *Cell* **78**, 603-15.
- Chan, S. K., Popperl, H., Krumlauf, R. and Mann, R. S.** (1996). An extradenticle-induced conformational change in a HOX protein overcomes an inhibitory function of the conserved hexapeptide motif. *EMBO J* **15**, 2476-87.
- Cohen, S. M. and Jurgens, G.** (1989). Proximal-distal pattern formation in Drosophila: cell autonomous requirement for Distal-less gene activity in limb development. *EMBO J* **8**, 2045-2055.
- DeSousa, D., Mukhopadhyay, M., Pelka, P., Zhao, X., Dey, B. K., Robert, V., Pelisson, A., Bucheton, A. and Campos, A. R.** (2003). A novel double-stranded RNA-binding protein, disco interacting protein 1 (DIP1), contributes to cell fate decisions during Drosophila development. *J Biol Chem* **278**, 38040-50.
- Ekker, S. C., Jackson, D. G., von Kessler, D. P., Sun, B. I., Young, K. E. and Beachy, P. A.** (1994). The degree of variation in DNA sequence recognition among four Drosophila homeotic proteins. *EMBO J* **13**, 3551-60.
- Flores-Saaib, R. D., Jia, S. and Courey, A. J.** (2001). Activation and repression by the C-terminal domain of Dorsal. *Development* **128**, 1869-79.

**Gajewski, K., Fossett, N., Molkentin, J. D. and Schulz, R. A.** (1999). The zinc finger proteins Pannier and GATA4 function as cardiogenic factors in *Drosophila*. *Development* **126**, 5679-88.

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

**Gibson, G., and Gehring, W.J.** (1988). Head and thoracic transformations caused by ectopic expression of *Antennapedia* during *Drosophila* development. *Development* **102**, 657-675.

**Hasson, P., Muller, B., Basler, K. and Paroush, Z.** (2001). Brinker requires two corepressors for maximal and versatile repression in Dpp signalling. *EMBO J* **20**, 5725-36.

**Heilig, J. S., Freeman, M., Lavery, T., Lee, K. J., Campos, A. R., Rubin, G. M. and Steller, H.** (1991). Isolation and characterization of the disconnected gene of *Drosophila melanogaster*. *EMBO J* **10**, 809-15.

**Kapoun, A. M. and Kaufman, T. C.** (1995). A functional analysis of 5', intronic and promoter regions of the homeotic gene proboscipedia in *Drosophila melanogaster*. *Development* **121**, 2127-41.

**Koelzer, S. and 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.

**Kozmik, Z., Pfeffer, P., Kralova, J., Paces, J., Paces, V., Kalousova, A. and Cvekl, A.** (1999). Molecular cloning and expression of the human and mouse homologues of the *Drosophila dachshund* gene. *Dev Genes Evol* **209**, 537-45.

**Lewis, E. B.** (1978). A gene complex controlling segmentation in *Drosophila*. *Nature* **276**, 565-570.

**Lo, P. C., Skeath, J. B., Gajewski, K., Schulz, R. A. and Frasch, M.** (2002). Homeotic genes autonomously specify the anteroposterior subdivision of the *Drosophila* dorsal vessel into aorta and heart. *Dev Biol* **251**, 307-19.

**Mahaffey, J. W., Griswold, C. M. and Cao, Q. M.** (2001). The *Drosophila* genes disconnected and disco-related are redundant with respect to larval head development and accumulation of mRNAs from deformed target genes. *Genetics* **157**, 225-36.

**McCarty, A. S., Kleiger, G., Eisenberg, D. and Smale, S. T.** (2003). Selective dimerization of a C2H2 zinc finger subfamily. *Mol Cell* **11**, 459-70.

- McGinnis, W. and Krumlauf, R.** (1992). Homeobox genes and axial patterning. *Cell* **68**, 283-302.
- Merrill, V. K., Turner, F. R. and Kaufman, T. C.** (1987). A genetic and developmental analysis of mutations in the Deformed locus in *Drosophila melanogaster*. *Dev Biol* **122**, 379-95.
- Mukhopadhyay, M., Pelka, P., DeSousa, D., Kablar, B., Schindler, A., Rudnicki, M. A. and Campos, A. R.** (2002). Cloning, genomic organization and expression pattern of a novel *Drosophila* gene, the disco-interacting protein 2 (dip2), and its murine homolog. *Gene* **293**, 59-65.
- Nagel, A. C., Krejci, A., Tenin, G., Bravo-Patino, A., Bray, S., Maier, D. and Preiss, A.** (2005). Hairless-mediated repression of notch target genes requires the combined activity of Groucho and CtBP corepressors. *Mol Cell Biol* **25**, 10433-41.
- O'Hara, E., Cohen, B., Cohen, S. M. and McGinnis, W.** (1993). Distal-less is a downstream gene of Deformed required for ventral maxillary identity. *Development* **117**, 847-56.
- Park, M., Wu, X., Golden, K., Axelrod, J. D. and Bodmer, R.** (1996). The wingless signaling pathway is directly involved in *Drosophila* heart development. *Dev Biol* **177**, 104-16.
- Patel, M., Farzana, L., Robertson, L. K., Hutchinson, J., Grubbs, N., Shepherd, M. N. and Mahaffey, J. W.** (2007). The appendage role of insect disco genes and possible implications on the evolution of the maggot larval form. *Dev Biol* **309**, 56-69.
- Pederson, J. A., LaFollette, J. W., Gross, C., Veraksa, A., McGinnis, W. and Mahaffey, J. W.** (2000). Regulation by homeoproteins: a comparison of deformed-responsive elements. *Genetics* **156**, 677-86.
- Phippen, T. M., Sweigart, A. L., Moniwa, M., Krumm, A., Davie, J. R. and Parkhurst, S. M.** (2000). *Drosophila* C-terminal binding protein functions as a context-dependent transcriptional co-factor and interferes with both mad and groucho transcriptional repression. *J Biol Chem* **275**, 37628-37.
- Pyrowolakis, G., Hartmann, B., Muller, B., Basler, K. and Affolter, M.** (2004). A simple molecular complex mediates widespread BMP-induced repression during *Drosophila* development. *Dev Cell* **7**, 229-40.
- Rauskolb, C., Peifer, M. and Wieschaus, E.** (1993). extradenticle, a regulator of homeotic gene activity, is a homolog of the homeobox-containing human proto-oncogene pbx1. *Cell* **74**, 1101-12.

**Rauskolb, C. and Wieschaus, E.** (1994). Coordinate regulation of downstream genes by extradenticle and the homeotic selector proteins. *EMBO J* **13**, 3561-9.

**Regulski, M., McGinnis, N., Chadwick, R. and McGinnis, W.** (1987). Developmental and molecular analysis of Deformed; a homeotic gene controlling Drosophila head development. *EMBO J* **6**, 767-777.

**Robertson, L. K., Bowling, D. B., Mahaffey, J. P., Imiolczyk, B. and Mahaffey, J. W.** (2004). An interactive network of zinc-finger proteins contributes to regionalization of the Drosophila embryo and establishes the domains of HOM-C protein function. *Development* **131**, 2781-9.

**Ryoo, H. D. and Mann, R. S.** (1999). The control of trunk Hox specificity and activity by Extradenticle. *Genes Dev* **13**, 1704-16.

**Ryoo, H. D., Marty, T., Casares, F., Affolter, M. and Mann, R. S.** (1999). Regulation of Hox target genes by a DNA bound Homothorax/Hox/Extradenticle complex. *Development* **126**, 5137-48.

**Stark, A., Kheradpour, P., Roy, S. and Kellis, M.** (2007a). Reliable prediction of regulator targets using 12 Drosophila genomes. *Genome Res.*

**Stark, A., Lin, M. F., Kheradpour, P., Pedersen, J. S., Parts, L., Carlson, J. W., Crosby, M. A., Rasmussen, M. D., Roy, S., Deoras, A. N. et al.** (2007b). Discovery of functional elements in 12 Drosophila genomes using evolutionary signatures. *Nature* **450**, 219-32.

**Struhl, G.** (1981). A homoeotic mutation transforming leg to antenna in Drosophila. *Nature* **292**, 635-8.

**Taghli-Lamalle, O., Gallet, A., Leroy, F., Malapert, P., Vola, C., Kerridge, S. and Fasano, L.** (2007). Direct interaction between Teashirt and Sex combs reduced proteins, via Tsh's acidic domain, is essential for specifying the identity of the prothorax in Drosophila. *Dev Biol* **307**, 142-51.

**Valenta, T., Lukas, J. and Korinek, V.** (2003). HMG box transcription factor TCF-4's interaction with CtBP1 controls the expression of the Wnt target Axin2/Conductin in human embryonic kidney cells. *Nucleic Acids Res* **31**, 2369-80.

**Walsh, C. M. and Carroll, S. B.** (2007). Collaboration between Smads and a Hox protein in target gene repression. *Development* **134**, 3585-92.

**Yin, Z., Xu, X. L. and Frasch, M.** (1997). Regulation of the twist target gene tinman by modular cis-regulatory elements during early mesoderm development. *Development* **124**, 4971-82.

## Appendix

Expression profiling is commonly used to detect transcriptional changes that are relevant to the process being studied. Using this method scientists have studied a range of topics from the changes in gene regulation during embryogenesis to identifying Hox target genes (Hueber et al., 2007; Tomancak et al., 2007). We used microarrays to examine the transcriptional changes that occur when Disco is ectopically activated throughout the embryo using the *armadillo*-Gal4 driver. The results from this study are presented below.

## **Materials and Methods**

### **Embryo preparation**

Embryos for the experiment were generated by crossing UAS-*disco*<sup>C2</sup>/UAS-*disco*<sup>C2</sup> flies to either *arm*-Gal4/*arm*-Gal4 (experimental embryos) or *w/w* (control embryos) flies. Embryos from these crosses were collected for either 2 hours and then aged 6 hours (6-8hr embryos) or collected for 6 hours and then aged 7 hours (7-13 hr embryos). Embryos were washed with NaCl/Triton several times, then dechorinated using a commercial bleach, NaCl/Triton solution (0.7% NaCl, 0.04% Triton X-100) that was 50% bleach and 50% NaCl/Triton for 5 minutes. Dechorinated embryos were washed several times with NaCl/Triton and with dH<sub>2</sub>O three times.

### **cDNA synthesis and dye coupling**

Total RNA was extracted from these embryos using the Qiagen RNeasy kit. Several collections were done and the total RNA was pooled. cDNA synthesis and dye coupling were done using the Superscript<sup>TM</sup> Plus indirect cDNA Labeling System from Invitrogen and their protocol. Pool Cy3 and Cy5 labeled cDNAs. Enough reactions were

done such that dyes could be swapped between the experimental and control cDNAs and the experiment repeated once. In all, four slides were hybridized for 6-8hr embryos or 7-13hr embryos.

### **Slide Hybridization**

DGRC-1 slides were obtained from the Drosophila Genomics Resource Center. The protocol used for hybridizations is the Amino Alkyl Labeling and Hybridization Protocol by Justen Andrews, Kevin Bogart, James Costello, and Brian Eads from the Drosophila Genomics Research Center with slight modifications ([http://cgb.indiana.edu/publications/alias/amino\\_alkyl/version/current](http://cgb.indiana.edu/publications/alias/amino_alkyl/version/current)). The Sodium Dodecyl Sulfate solution used had a pH of 7.2 and no formamide was added to the prehybridization solution.

### **Scanning slides and data analysis**

Slides were scanned using the GSI Lumonic ScanArray 4000 chip scanner and the Perkin Elmer ScanArray Express 3.0 software. Spot intensities were obtained using the UCSF Spot program (Jain et al., 2002). The data were log<sub>2</sub> transformed and analyzed using a SAS program kindly provided by Dr. Ian Dworkin.

### **Results**

The changes in gene expression are listed below in Table 1 and Table 2. There was no statistically significant change in expression obtained from this particular study. However, an approximately four fold change in *disco* was detected in 6-8 hour embryos and approximately a two fold change in 7-13 hour embryos. *pnr* expression is known to be affected by ectopic Disco expression, but this was not detected in our study

(Robertson et al., 2004). The inability to conduct statistical analysis on these hybridizations was due to the irreproducibility between slides. This problem has been noted for this generation of slides by other researchers.

**Table 1. Transcriptional Changes in 6-8 hr Ectopically Expressing Disco embryos.**

List of transcriptional changes that occurred in 6-8 hour embryos. Estimate is the log<sub>2</sub> fold change in expression. If there has been a molecular function that has been associated with the gene it is listed under the GO Molecular Function column.

Estimate	Gene Name	GO Molecular Function
-1.37343	<i>Trfp</i>	transcription ; GO:0006350, transcription initiation from Pol II promoter ; GO:0006367
-1.20503	<i>CG8506</i>	
-1.18279		
-1.13255	<i>CG4004</i>	
-1.06902	<i>Obp56c</i>	transport ; GO:0006810
-1.00966	<i>CG9065</i>	copper ion transport ; GO:0006825, cytochrome c oxidase biogenesis ; GO:0008535
-0.994	<i>T:EcoNacZ</i>	null
-0.99349	<i>CG13226</i>	
-0.97077	<i>CG1441</i>	
-0.96935	<i>CG14369</i>	
-0.96233	<i>amon</i>	hatching behavior ; GO:0035187, peptide hormone processing ; GO:0016486, proteolysis and peptidolysis ; GO:0006508
-0.96119	<i>CG4090</i>	chitin metabolism ; GO:0006030
-0.95424	No Amplicon Gene Matches	
-0.92449	<i>CG13488</i>	
-0.92356		
-0.91889	<i>CG11192</i>	proteolysis and peptidolysis ; GO:0006508
-0.91226	<i>CG30179</i>	cell proliferation ; GO:0008283, development ; GO:0007275, mesoderm development ; GO:0007498, muscle development ; GO:0007517
0.901816	<i>T:ScerIGAL4</i>	null
0.903376	<i>ImpE2</i>	imaginal disc eversion ; GO:0007561
0.907604	No Amplicon Gene Matches	
0.913347	<i>CG8756</i>	carbohydrate metabolism ; GO:0005975, chitin metabolism ; GO:0006030
0.924602	<i>m4</i>	Notch signaling pathway ; GO:0007219, cell fate specification ; GO:0001708, sensory organ development ; GO:0007423
0.926623	<i>T:ScerIGAL4</i>	null
0.944142	<i>T:ScerIGAL4</i>	null
0.95729	<i>Obp99c</i>	transport ; GO:0006810
0.972951	<i>Cirl</i>	G-protein coupled receptor protein signaling pathway ; GO:0007186, neuropeptide signaling pathway ; GO:0007218, neurotransmitter secretion ; GO:0007269, vesicle-mediated transport ; GO:0016192
0.973924	<i>Gr59c</i>	perception of taste ; GO:0050909, protein amino acid prenylation ; GO:0018346
0.976188	<i>dro5</i>	defense response ; GO:0006952, pathogenesis ; GO:0009405
1.020759	-	
1.030932	<i>CG4455</i>	
1.04178	<i>Obp99a</i>	transport ; GO:0006810

Table 1 (continued)

1.056388	<i>CG6218</i>	carbohydrate metabolism ; GO:0005975, monosaccharide metabolism ; GO:0005996
1.097566	<i>CG6631</i>	
1.114871	<i>T:Scer\GAL4</i>	null
1.12944	<i>T:Scer\GAL4</i>	null
1.144799	<i>CG17048</i>	protein ubiquitination ; GO:0016567
1.176038	<i>Brd</i>	Notch signaling pathway ; GO:0007219, cell fate specification ; GO:0001708, sensory organ development ; GO:0007423, sensory organ precursor cell fate determination ; GO:0016360
1.177579	<i>RfaBp</i>	lipid transport ; GO:0006869, transport ; GO:0006810
1.184306	<i>CAP</i>	MAPKKK cascade ; GO:0000165, cytoskeleton organization and biogenesis ; GO:0007010
1.197705	<i>T:Scer\GAL4</i>	null
1.694294	<i>bnb</i>	development ; GO:0007275, gliogenesis ; GO:0042063
2.031263	<i>disco</i>	brain development ; GO:0007420, circadian rhythm ; GO:0007623, eclosion rhythm ; GO:0008062, locomotor rhythm ; GO:0045475, photoreceptor maintenance ; GO:0045494

**Table 2. Transcriptional Changes in 7-13 hr Ectopically Expressing Disco embryos.**

List of transcriptional changes that occurred in 7-13 hour embryos. Estimate is the log<sub>2</sub> fold change in expression. If there has been a molecular function that has been associated with the gene it is listed under the GO Molecular Function column.

Estimate	Gene Name	Go Molecular Function
-1.81445	<i>CG14756</i>	
-1.56026	<i>CG13159</i>	
-1.35579	<i>CG14110</i>	
-1.20587	<i>CG5171</i>	disaccharide metabolism ; GO:0005984, trehalose biosynthesis ; GO:0005992
-1.15811	<i>CG10591</i>	
-1.10419	<i>smg</i>	negative regulation of translation ; GO:0016478
-1.08599	<i>CG1291</i>	polysaccharide metabolism ; GO:0005976, protein amino acid glycosylation ; GO:0006486
-1.05514	<i>CG1943</i>	
-1.03653	<i>CG6693</i>	
-1.03295	<i>CG6503</i>	
-1.00596	<i>dUTPase</i>	dUTP metabolism ; GO:0046080, nucleobase, nucleoside, nucleotide and nucleic acid metabolism ; GO:0006139
-0.99428	<i>CycB</i>	G2/M transition of mitotic cell cycle ; GO:0000086, attachment of spindle microtubules to kinetochore ; GO:0008608, cellular physiological process ; GO:0050875, cytokinesis ; GO:0000910, cytokinesis after mitosis ; GO:0000281, embryonic development (sensu Insecta) ; GO:0001700, mitotic anaphase B ; GO:0000092, mitotic chromosome movement ; GO:0007079, regulation of cell cycle ; GO:0000074
-0.9717	<i>CG30337</i>	DNA metabolism ; GO:0006259, cell cycle ; GO:0007049, intracellular protein transport ; GO:0006886, mitosis ; GO:0007067, nucleobase, nucleoside, nucleotide and nucleic acid metabolism ; GO:0006139, protein targeting ; GO:0006605
-0.96093	<i>polo</i>	cell cycle ; GO:0007049, cytokinesis ; GO:0000910, embryonic development ; GO:0009790, female meiosis II ; GO:0007147, female meiosis II spindle assembly (sensu Metazoa) ; GO:0007058, intracellular signaling cascade ; GO:0007242, male meiosis ; GO:0007140, male meiosis chromosome segregation ; GO:0007060, mitosis ; GO:0007067, pronuclear fusion ; GO:0007344, pronuclear migration ; GO:0035046, protein amino acid phosphorylation ; GO:0006468, protein localization ; GO:0008104, sperm aster formation ; GO:0035044
-0.95254	<i>smg</i>	negative regulation of translation ; GO:0016478
-0.94667	<i>CG33012</i>	
-0.92218	<i>CG5735</i>	RNA localization ; GO:0006403, embryonic development ; GO:0009790, mRNA polyadenylation ; GO:0006378
0.901191	<i>CG14956</i>	
0.905294	<i>CG11126</i>	nucleobase, nucleoside, nucleotide and nucleic acid metabolism ; GO:0006139, nucleotide catabolism ; GO:0009166
0.907056	<i>CG33519</i>	

Table 2 (continued)

0.908298	<i>CG6036</i>	protein amino acid dephosphorylation ; GO:0006470
0.909062	<i>Obp56a</i>	
0.920967	<i>Pros54</i>	protein metabolism ; GO:0019538, proteolysis and peptidolysis ; GO:0006508, ubiquitin-dependent protein catabolism ; GO:0006511
0.930139	<i>dro5</i>	defense response ; GO:0006952, pathogenesis ; GO:0009405
0.962076	<i>CG7149</i>	phospholipid biosynthesis ; GO:0008654, phospholipid metabolism ; GO:0006644
0.979688	<i>CG9592</i>	
0.980084	<i>CG11419</i>	mitotic anaphase ; GO:0000090, regulation of cell cycle ; GO:0000074, regulation of mitotic metaphase/anaphase transition ; GO:0030071
1.023976	<i>CG8038</i>	tRNA processing ; GO:0008033
1.023995	<i>CG6410</i>	intracellular signaling cascade ; GO:0007242, protein targeting ; GO:0006605
1.032231	<i>Obp99c</i>	transport ; GO:0006810
1.056716	<i>Peritrophin-15b</i>	
1.065306	<i>CG31660</i>	metabotropic glutamate receptor signaling pathway ; GO:0007216
1.079977	<i>capu</i>	cell organization and biogenesis ; GO:0016043, eggshell formation (sensu Insecta) ; GO:0007304, intracellular signaling cascade ; GO:0007242, oogenesis (sensu Insecta) ; GO:0009993, pole plasm RNA localization ; GO:0007316, pole plasm assembly (sensu Insecta) ; GO:0048113, pole plasm assembly ; GO:0007315, pole plasm oskar mRNA localization ; GO:0045451
1.097988	<i>Hil</i>	
1.106737	<i>CG30419</i>	
1.117279	<i>GlcAT-S</i>	glycoprotein biosynthesis ; GO:0009101, glycosphingolipid biosynthesis ; GO:0006688, proteoglycan biosynthesis ; GO:0030166
1.127242	<i>CG15870</i>	
1.129298	<i>Amplicon Hits Multiple Gene Regions</i>	
1.132163	<i>CG2256</i>	intracellular signaling cascade ; GO:0007242
1.150576	<i>disco</i>	brain development ; GO:0007420, circadian rhythm ; GO:0007623, eclosion rhythm ; GO:0008062, locomotor rhythm ; GO:0045475, photoreceptor maintenance ; GO:0045494
1.154001	<i>CG15208</i>	
1.158283	<i>CG32554</i>	biological_process unknown ; GO:0000004
1.171807	<i>CG5398</i>	cell homeostasis ; GO:0019725, transport ; GO:0006810
1.178419	<i>CG3635</i>	lipid metabolism ; GO:0006629
1.254357	<i>CG11409</i>	
1.547893	<i>CG3898</i>	
1.553368	<i>CG14660</i>	
1.633405	<i>Nep4</i>	proteolysis and peptidolysis ; GO:0006508, signal transduction ; GO:0007165

2.360331	CG10761	
----------	---------	--

## REFERENCES

- Hueber, S. D., Bezdán, D., Henz, S. R., Blank, M., Wu, H. and Lohmann, I.** (2007). Comparative analysis of Hox downstream genes in *Drosophila*. *Development* **134**, 381-92.
- Jain, A. N., Tokuyasu, T. A., Snijders, A. M., Segraves, R., Albertson, D. G. and Pinkel, D.** (2002). Fully automatic quantification of microarray image data. *Genome Res* **12**, 325-32.
- Robertson, L. K., Bowling, D. B., Mahaffey, J. P., Imiolczyk, B. and Mahaffey, J. W.** (2004). An interactive network of zinc-finger proteins contributes to regionalization of the *Drosophila* embryo and establishes the domains of HOM-C protein function. *Development* **131**, 2781-9.
- Tomancak, P., Berman, B. P., Beaton, A., Weiszmann, R., Kwan, E., Hartenstein, V., Celniker, S. E. and Rubin, G. M.** (2007). Global analysis of patterns of gene expression during *Drosophila* embryogenesis. *Genome Biol* **8**, R145.