ABSTRACT

GRUBBS, NATHANIEL PAUL. Expanding the Genetic Architecture Contributing to Appendage Development. (Under the direction of Dr. James W. Mahaffey).

Appendage development in the model organism *Drosophila melanogaster* is well studied. Many of the genes involved have been implicated in other processes, such as nervous system development, vertebrate appendage development and cancer. Though it has been well-studied, the model of Drosophila leg development is incomplete as evidenced by the recent discovery that the redundant genes *disconnected* (*disco*) and *disco-related* (*disco-r*) have a critical role in leg development in Drosophila. The single *disco* ortholog of another model organism, the beetle *Tribolium castaneum*, was similarly identified as being important for leg development in that species.

The goal of this work is to broaden our understanding of appendage development by specifically addressing three questions: What other functions of *disco* genes might be conserved between these species? How broadly is that function conserved outside of insects? Finally, what other genes may also be missing from models of leg development?

In Drosophila embryos, the *disco* genes interact with the body-wall gene *teashirt* (*tsh*) to create mutually exclusive domains of expression. The work presented here suggests that this interaction is not conserved in Tribolium. In addition, evidence is given for interactions between *disco* and the appendage gene *dachshund* (*dac*) in Tribolium leg development.

Given the conservation of *disco* gene function in insect leg development, it is important to determine how broadly this role has been conserved beyond insects. The expression patterns of potential *disco* orthologs were examined in the nematode species *Caenorhabditis elegans* and in the zebra fish, *Danio rerio*. While no appendage expression

was found in the fish, both species possessed expression in the nervous system, a feature found also in Drosophila.

Finally, this work describes the use of Genome Wide Association Studies (GWAS) to identify genes not previously associated with leg development. Natural variation in proportionality of leg segments was used to conduct association mapping with 117 inbred lines from the Drosophila Genetic Reference Panel. Femur, tibia, and tarsal segments of the first (T1) and second (T2) thoracic leg were measured from a total of 12 males and 12 females. Single Nucleotide Polymorphisms (SNPs) with the lowest P-values across male and female, T1 and T2 were chosen for further study. Genes near SNPs of interest were identified independently for each leg segment analyzed. After candidate genes were identified, we used in situ hybridization to determine gene expression in Drosophila imaginal discs, and genes with expression in the developing leg discs were tested for function using RNA interference (RNAi). RNAi lines were crossed with several GAL4 drivers to ascertain when and where gene products were needed during normal development. We identified genes that, when reduced, resulted in phenocopies with deleted leg segments, reduction of leg segment lengths and altered leg segment morphology, as well as some that altered antenna and wing morphology. Our findings include several newly discovered genes, implicating new genetic pathways with specific effects on leg development.

Expanding the Genetic Architecture Contributing to Appendage Development

by
Nathaniel Paul Grubbs

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

2012

APPROVED BY:

Dr. James W. Mahaffey Committee Chair	Dr. Patricia A. Estes
Dr. Jonathan M. Horowitz	Dr. Trudy F. Mackay

DEDICATION

I would like to dedicate this work to my fiancé, Justin Holt, whose love and support during the last two years has been invaluable to my success.

I would also like to dedicate this work to my parents, who were the first to support and foster my love for science. I can’t thank them enough for opening my mind to the world around me.

BIOGRAPHY

Nathaniel Grubbs was born in Raleigh, NC to the Reverend Paul Grubbs and his wife, Teresa (nee Caroon), on 9 February 1983. The family spent time in the Philippines and Bulgaria before returning to North Carolina. Nathaniel graduated from Northeastern High School in Elizabeth City, NC, in 2001. In 2005, he completed his degree in biology at Mount Olive College, Mount Olive, NC. That same year, he began graduate work in the Genetics Department of North Carolina State University, eventually joining the lab of Dr. Jim Mahaffey. During this time, Nathaniel went through a great deal of personal growth, a path that led him to find love with a wonderful man. The two have enjoyed many adventures, particularly exploring the outdoors and the culinary arts, as well as opening their home to two stray cats (in addition to the one they already had). Nathaniel completed work on his Doctor of Philosophy in May 2012.

ACKNOWLEDGMENTS

I am grateful to the Genetics Department for the opportunity to pursue advanced studies in Genetics and for all the help and support along the way. I would particularly like to thank my advisor, Jim Mahaffey, for his patience and advice, and the members of my advisory committee—Pat Estes, Jon Horowitz, and Trudy Mackay—who were always available to help, and provided assistance for many projects. Special thanks also go to Laura Mathies, who served for some time on my committee and who also helped with some of my projects.

I received a lot of help from many members of the Mahaffey lab, both past and present, through the course of my research. Laila Farzana was instrumental in guiding me through my early years in graduate school, as well teaching me how to work with Tribolium. Muk Patel and Jenny Hutcherson were also great teachers and guides in my early years in the lab. The GWAS studies could not have been completed without the hard work of Xin Su, Tiffany Petrisko, and Megan Leach. Juan Rosario has been an invaluable lab mate, providing advice on a variety of topics and acting as a sounding board for research ideas. Crystal Tabor and Callie Longo have done much to keep the lab running, even as they worked on their own projects. To all of them, I say, "Thank you."

Finally, I would like to acknowledge the source of funding, the National Science Foundation, whose grants to J.W. Mahaffey have made this research possible.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE

General Introduction

INTRODUCTION

Limbs are an extension of the body and nervous system that seem to have evolved from outgrowths present in the ancestor of bilateral animals (Panganiban, Irvine et al. 1997; Shubin, Tabin et al. 1997; Pueyo and Couso 2005). They serve a variety of purposes: as extensions for gathering sensory information, as a means of locomotion or as a means of gathering and moving objects. The prominence of limbs is such that they are useful as a defining feature in animals, as in how they are used for the classification of arthropod groups, helping to distinguish insects from chelicerates, etc. (Shubin, Tabin et al. 1997; Prpic, Janssen et al. 2003).

Amongst insects, limb development has been most studied in the fruit fly, *Drosophila melanogaster* (reviewed in (Lecuit and Cohen 1997; Panganiban 2000; Morata 2001; Kojima 2004; Pueyo and Couso 2005; Estella, Voutev et al. 2012)). However, limb development in fruit flies is highly derived. Unlike many insect species, Drosophila larvae lack limbs. Instead, their limbs develop from imaginal discs, specialized tissues one cell-layer thick, which are allocated in early embryogenesis and are patterned during the larval stages. Leg imaginal discs form where the segmentally repeated expression of the signaling factor Wingless (Wg) activates expression of *Distal-less* (*Dll*), which encodes a homeobox transcription factor. This initial activation of *Dll* is repressed by the morphogen Decapentapalegic (Dpp) dorsally, and by the epidermal growth factor receptor (EGFR) pathway ventrally (Fig 1; Similar signaling events establish the primordia of other appendages) (Panganiban 2000; Estella, Voutev et al. 2012).

FIGURE 1: Schematic of initiation of leg primordium formation during embryogenesis

These initial steps in embryogenesis establish the cells destined to form the adult legs, and set the stage for patterning, which begins in earnest during the larval stages (Fig 2). A current model of leg development proposes that the transcription factor, Engrailed (En), which is expressed in the posterior compartment of the developing discs, activates expression of *hedgehog* (*hh*), which encodes a signaling molecule (Morata 2001). Although *hh* is only expressed in posterior cells, its most critical role in the developing leg is seen in the anterior compartment. Hh is secreted from cells in the posterior of the disc, establishing a gradient in the anterior compartment. Just past the anterior-posterior boundary, Hh signaling activates expression of *wg* and *dpp*; *dpp* transcripts accumulate in a dorsal stripe and *wg* ventrally (Lecuit and Cohen 1997). The combined expression of both factors is strongest in the center of the disc, where they work together to activate *Dll*; the center of the disc will later form the distal portion of the leg (Lecuit and Cohen 1997; Panganiban 2000). Near the edge of the disc, where Wg and Dpp signal is low, the transcription factor Homothorax (HTH) is active and imports its cofactor, Extradenticle (EXD), to the nucleus to help pattern the proximal portion of the leg and its connection to the rest of the body wall (Kojima 2004; Estella, Voutev et al. 2012). Finally, *dachshund* (*dac*) is expressed in a region between *hth* and *Dll*, from which the medial portions of the leg develops (Kojima 2004). The limb tissue grows during the larval stages, then everts and expands during pupation (Kojima 2004).

While these genes are critical in proper patterning of the appendages, this model does not offer a complete picture. To better understand how the developing Drosophila leg achieves its final form, we need to know what else shapes these domains of expression, and how these genes and pathways interact to shape the proportions of the completed leg.

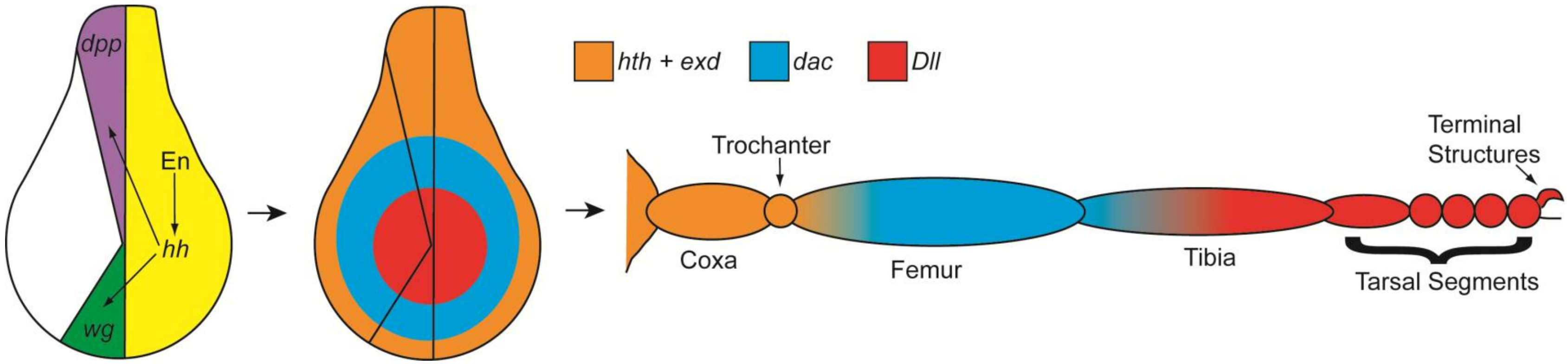


FIGURE 2: A model of larval imaginal disc patterning and adult leg development

Leg imaginal disc patterning begins (left) when En, present in the posterior (yellow) portion of the disc, activates *hh* expression. The Hh protein diffuses into the anterior portion of the disc. There it activates the expression of *dpp* (purple) dorsally and *wg* (green) ventrally. Dpp and Wg, in turn, diffuse throughout the dorsal and ventral portions, respectively, of the disc, creating a gradient of their mutual presence. This gradient is responsible for the pattern of expression of transcription factors (middle) that establish the proximal-distal axis of the leg (right). In the center of the disc, mutual Dpp and Wg is highest, activating the expression of *Dll*, which is responsible for patterning the most distal structures of the leg, including the terminal structures (claw and pulvillus). Reduced mutual Dpp and Wg results in the activation of *dac*, responsible for the patterning of the middle portions of the leg. Where there is almost no mutual Dpp or Wg, Hth and Exd are active, patterning the proximal leg portions and the junction with the rest of the body.

Fortunately, many of these pathways have been extensively studied, and many factors are known that shape their expression and field of function, some of which will be discussed below. This will be followed by an overview of some of what we know about the conservation of these genes in invertebrates and an introduction to my work in broadening our understanding of leg development.

SIGNALING PATHWAYS INVOLVED IN DROSOPHILA LEG DEVELOPMENT

The initiation and patterning of Drosophila legs rely on morphogen signals. Morphogens act in gradients, with cells closest to the origin receiving a higher signal than those farther away. Changes in gradient levels are often associated with changes in the activities of receiving cells (Estella and Mann 2008; Sopory, Kwon et al. 2010; Erickson 2011; Giorgianni and Mann 2011; Mulligan, Fuerer et al. 2012).

Hh pathway

One of the earliest morphogens involved in patterning the leg discs is Hh. Transcription of *hh* is initiated by the homeodomain transcription factor En, which is active in the posterior compartment of the imaginal disc (Morata 2001; Estella, Voutev et al. 2012). The Hh ligand is secreted into the anterior portion of the disc, where it is able to activate target genes (Morata 2001; Estella, Voutev et al. 2012).

Our understanding of the Hh pathway comes primarily from studies in embryogenesis and wing development (Jiang and Hui 2008). In the absence of Hh signal, the transmembrane protein Patched (Ptc) blocks activation of another transmembrane protein, Smoothened

(Smo), allowing for the cleavage of the transcription factor Cubitus interuptus (Ci). This results in a truncated form of Ci that acts as a repressor of some Hh targets (Lum, Yao et al. 2003; Jiang and Hui 2008; Camp, Currie et al. 2010; Wu, Lin et al. 2011). When the Hh ligand binds to Ptc, Ptc no longer deactivates Smo. Smo is then able to block the cleavage of Ci, allowing Ci to activate Hh target genes, like *wg* and *dpp* (Lum, Yao et al. 2003; Jiang and Hui 2008; Camp, Currie et al. 2010).

The cleavage of Ci is a critical step in regulating Hh signaling, and involves a number of protein interactions. In order for Smo to be activated, it must be phosphorylated, which is carried out by the kinases cAMP-dependent protein kinase 1 (Pka-C1) and Casein kinase Iα (CKIα). Otherwise, Smo is endocytosed and degraded, activities promoted by Ptc in the absence of Hh (Jiang and Hui 2008). In the absence of active Smo, Pka-C1 and CKIα, along with another kinase, Shaggy (Sgg), bind, with the aid of Costa (Cos) and Fused (Fu), then sequentially posphorylate Ci (Lum, Yao et al. 2003; Jiang and Hui 2008; Wu, Lin et al. 2011). This increases the affinity of Ci for the F-box protein Supernumerary limbs (Slmb), which is part of an ubiquitin ligase complex that marks Ci for cleavage by proteolysis (Lum, Yao et al. 2003; Jiang and Hui 2008; Wu, Lin et al. 2011). It has been shown that the Slmb-dependent ubiquitination of Ci is also negatively regulated by the Cop9 signalsome (Wu, Lin et al. 2011).

The Hh signal is also regulated at the level of Hh itself. The initial protein product of the *hh* gene first undergoes autocleavage, and is then palmitoylated by Rasp, before being released by the Ptc-like transmembrane protein Dispatched (Disp). However, these modifications restrict free dispersal of the protein (Jiang and Hui 2008). Several cell-surface proteins also affect the dispersal and signal strength of Hh. Amongst these are the heparin sulfate proteoglycans (HSPGs), Division abnormally delayed (Dally) and Dally-like (Dlp), which are necessary for proper Hh function. In this capacity, they appear to work to secure Hh to the cell surface, simultaneously strengthening the signal on that cell while limiting the dispersal of the ligand to other cells (Lum, Yao et al. 2003; Jiang and Hui 2008; Yan, Wu et al. 2010). The cell-surface proteins Interference hedgehog (Ihog) and Brother of Ihog (Boi) are also important for Hh signaling. These proteins act redundantly to increase binding of Ptc to the Hh ligand by forming a complex with Ptc and Hh. This activity also works to increase Hh signal while decreasing range (Yao, Lum et al. 2006; Camp, Currie et al. 2010; Yan, Wu et al. 2010; Zheng, Mann et al. 2010). Interestingly, Hh appears to regulate both sets of cell surface proteins, positively regulating *dally*/*dlp* while negatively regulating *ihog*/*boi* (Yan, Wu et al. 2010), meaning Hh regulates its own dispersal and signal strength.

Wg pathway

During embryogenesis, En and Hh work in concert with Wg to establish the polarity of embryonic segments. At this point, *wg* is expressed in a line of cells just anterior to the *en*-expressing cells, which are the posterior-most cells in the segment (Gonsalves and DasGupta 2008). It is this stripe of Wg that will produce the expression of Dll that marks the formation of the leg primordium (cells recruited to the primordium from the En domain continue to express *en* and form the posterior compartment of the disc) (Kubota, Goto et al. 2003; Kojima 2004; Estella, Voutev et al. 2012).

Wg acts as a morphogen to contribute to the proximodistal patterning of the leg (Lecuit and Cohen 1997; Morata 2001; Kojima 2004; Estella, Voutev et al. 2012). The receptor for the Wg ligand is the transmembrane protein, Frizzled (Fz), which works with the co-receptor Arrow (Arr) to stabilize the β-catenin homolog Armadillo (Arm). Stabilized Arm is able to enter the nucleus, where it binds with the transcription factor Pangolin (Pan) to help activate target genes (Theisen, Syed et al. 2007; Gonsalves and DasGupta 2008; Mulligan, Fuerer et al. 2012). In the absence of Wg, a complex containing Apc, Axin and the kinase Sgg promote the proteolysis of Arm (Gonsalves and DasGupta 2008). However, binding of the Wg ligand results in the phosphorylation of Arr and the recruitment of Dishevelled (Dsh) to the Fz-Arr complex. With the aid of Dsh, the Axin complex is transfered to the plasma membrane, preventing the degradation of Arm (Polakis 2007; Gao and Chen 2010).

Like Hh, the HSPG proteins Dally and Dlp play a role in Wg signaling, seemingly by stabilizing the ligand in extracellular space (Gonsalves and DasGupta 2008; Erickson 2011; Mulligan, Fuerer et al. 2012). Also like Hh, the Wg ligand undergoes the addition of palmitic/palmitoleic acid groups, which might limit its ability to diffuse through the extracellular matrix. However, the unmodified ligand does not diffuse well naturally. Rather, these modifications permit the high-affinity binding of an extracellular protein Secreted Wg-interacting molecule (Swim). Swim improves the long-range signaling activity of Wg, but can out-compete Fz in binding the ligand if Swim is in sufficiently high concentrations (Mulligan, Fuerer et al. 2012). Lipoprotein particles composed of Lipophorin may act in a way parallel to the activity of Swim to transport lipid-linked ligands, which include both Hh and Wg (Mulligan, Fuerer et al. 2012).

In addition to helping establish the proximodistal axis of the leg, Wg also defines the ventral side of the disc (Kojima 2004; Brook 2010; Estella, Voutev et al. 2012). This is done in part by the activation of two T-box transcription factors, Midline (Mid) and H15 (Brook 2010; Estella, Voutev et al. 2012).

Dpp pathway

Dpp signaling, activated by Hh, cooperates with the Wg pathway to promote the formation of the proximodistal axis of the leg (Lecuit and Cohen 1997; Morata 2001; Kojima 2004; Estella, Voutev et al. 2012). During leg development, this domain of Dpp signaling also defines the dorsal domain of the leg, a function that acts in opposition to the ventralizing function of Wg (Penton and Hoffmann 1996; Kojima 2004; Brook 2010; Estella, Voutev et al. 2012).

Dpp has two types of receptors which must work together to effect Dpp signaling; type I receptors Thickveins (Tkv) and Saxophone (Sax) and the type 2 receptor Punt (Penton and Hoffmann 1996; Affolter and Basler 2007; Schwank, Dalessi et al. 2011). Upon binding the ligand, Punt phosphorylates Tkv, which then phosphorylates Mothers against dpp (Mad) (Affolter and Basler 2007). This activates Mad, allowing it to bind the transcription factor Medea, at which point the two are transported into the nucleus to activate target genes (Penton and Hoffmann 1996; Affolter and Basler 2007; Schwank, Dalessi et al. 2011). One notable exception to the activator activities of this complex is their deployment to repress the transcription of *brinker* (*brk*). This repression also requires the activity of the zinc-finger transcription factor Schnurri (Shn) (Affolter and Basler 2007). *brinker* encodes a transcription factor that acts to repress Dpp target genes in the absence of Dpp signal. Since its transcription is negatively regulated by Dpp, Brk forms an inverse gradient to that of Dpp (Affolter and Basler 2007; Estella, Voutev et al. 2012). Amongst the cofactors that appear to work with Brk are Arm and Pan, which may work with Brk to directly repress *dpp* transcription (Theisen, Syed et al. 2007).

As with Wg and Hh, Dpp signal range seems to be controlled at the ligand level by the activity of Dally and Dlp (Affolter and Basler 2007; Sopory, Kwon et al. 2010; Erickson 2011; Schwank, Dalessi et al. 2011). Cleavage of the immature Dpp protein is also important for establishing a stable, long-range signal. There are two cleavage sites found in Dpp which are cleaved by proprotein convertase family members Furrin 1 and 2 (Sopory, Kwon et al. 2010). Although the immature Dpp is capable of signaling if not processed, it does not last long as a stable secreted molecule. Cleavage of the first site is sufficient for establishing a short-range signal, which may be used in a tissue-specific manner where long-range signal is not needed; the full range of Dpp signal cannot be achieved without proper cleavage of both sites (Sopory, Kwon et al. 2010).

Endocytosis seems to be important for signaling of Dpp, Wg, and probably other morphogens. It has been hypothesized that ligands may actually travel through cells in a process called transcytosis, rather than around cells in the extracellular matrix, to establish their gradients (Erickson 2011; Schwank, Dalessi et al. 2011). There is some evidence that endocytosis resulting in degradation of the ligand is necessary for proper signaling, both to establish a steep gradient and to strengthen the signal by maintaining the connection between ligand and receptor up until they are destroyed (Erickson 2011). However, research with Dpp indicates that transcytosis mediated by the receptors is an unlikely explanation for how the Dpp gradient is formed (Schwank, Dalessi et al. 2011).

EGFR pathway

One of the targets activated by mutual expression of Dpp and Wg in the center (distal) portion of the developing leg is the EGFR pathway (Campbell 2002; Galindo, Bishop et al. 2002). The EGFR pathway has five ligands. Four of these, Spitz (Spi), Keren (Krn), Gurken (Grk) and Vein (Vn) activate the pathway, while the other, Argos (Aos) is an antagonist of the pathway (Shilo 2003). The primary activating ligand is Spi (Shilo 2003; Shilo 2005). Activation of the pathway results in induction of pathway target genes primarily through the transcription factor Pointed (Pnt) (Shilo 2003; Shilo 2005). There are several ways in which this pathway is regulated. Expression of *aos* is induced by EGFR activation. Aos then binds Spi directly, limiting its signaling range (Shilo 2005). Kekkon is also activated by EGFR signaling, and it binds the receptor with a leucine rich repeat (Shilo 2005). However, the most critical method for regulating the EGFR pathway is in ligand expression and processing. Grk is primarily limited to the developing oocyte (Shilo 2003; Shilo 2005). As with Dpp and Hh, Spi, Grk and Krn must undergo cleavage in order to function. These EGF ligands are initially created as membrane-bound precursor molecules (Shilo 2003; Shilo 2005); the transmembrane domain is removed to create a functional ligand. First, the transmembrane protein Star shuttles the membrane-bound pre-ligands out of the endoplasmic reticulum and to the Golgi compartment. In this new compartment, the ligands can be cleaved by the transmembrane protein Rhomboid (Rho) (Shilo 2003; Shilo 2005). *Star* and *spi* are broadly expressed, but *rho* expression is highly restricted, making it the limiting factor in the pathway (Shilo 2003; Shilo 2005). Vn is also used in pathway modulation. Vn is produced as a secreted protein, so it does not require the processing necessary for the function of its fellow ligands. However, it provides a weaker signal and a broader diffusion than Spi, elements that could be employed to initiate or reinforce the primary signal of Spi. Expression of *vn* is also tightly controlled, sometimes subject to activation by the EGFR pathway (Shilo 2003; Shilo 2005).

EGF signaling is typically short-range. However, there are two notable exceptions (Shilo 2003). The first is the patterning of the embryonic ventral ectoderm by Spi, which is secreted from the ventral midline (Shilo 2003). It is this expression which also provides the ventral limits of *Dll* expression when it is first activated in the presumptive leg primordia (Kojima 2004; Estella, Voutev et al. 2012). The second is the use of Vn to help pattern the distal tip of the developing adult leg (Shilo 2003), which will be further described below.

Activation and targets of Dll

Expression of *Dll* has been used as one of the earliest markers of the formation of the leg primordium (Morata 2001; Kojima 2004; Estella, Voutev et al. 2012). Because of this, a number of studies have been done to analyze how expression of *Dll* is controlled. Further, *Dll* encodes a homeodomain transcription factor (Cohen, Bronner et al. 1989), making potential targets of great interest.

Expression of *Dll* has primarily been analyzed by piecing together the genetic domains, or enhancer elements, responsible for providing proper expression of *Dll* throughout leg development. The enhancer that provides the earliest known expression in the developing leg is called 304 (Bolinger and Boekhoff-Falk 2005; Estella, McKay et al. 2008; McKay, Estella et al. 2009; Galindo, Fernandez-Garza et al. 2011; Estella, Voutev et al. 2012). This element is activated by the Wg pathway, while being repressed by the Dpp pathway dorsally, by EGFR signaling ventrally, and by the Hox proteins Ultrabithorax (Ubx) and Abdominal-A (AbdA) in abdominal segments (Goto and Hayashi 1997; Estella, Voutev et al. 2012). Cells that activate 304 contribute to two other primordia in addition to the leg. A dorsal portion of the initial *Dll*-expressing cells eventually lose *Dll* expression and migrate dorsally to form the imaginal disc of the adult wing and haltere in the second and third thoracic segments, respectively (Morata 2001; Estella, Rieckhof et al. 2003; Bolinger and Boekhoff-Falk 2005; Estella, Voutev et al. 2012). Of the remaining cells, a certain portion lose the expression of *hth*, which at this point is ubiquitously expressed in the embryo; these cells give rise to the larval sensory organ called the Keillin’s Organ and are also marked by a switch from using the 304 enhancer to a tissue specific enhancer called DKO (Bolinger and Boekhoff-Falk 2005; McKay, Estella et al. 2009; Estella, Voutev et al. 2012). The cells that continue to express both *Dll* and *hth* are also marked by a switch of enhancers, this time to leg-specific enhancers called Leg Trigger (LT, expressed from this point through the final stages of disc patterning (Estella, McKay et al. 2008)) and Leg Primordium (LP, expressed only during embryogenesis (Galindo, Fernandez-Garza et al. 2011)). Unlike 304, LT and LP require positive inputs from both Dpp and Wg signaling (Estella, McKay et al. 2008; Galindo, Fernandez-Garza et al. 2011), and several binding sites for both Pan and Mad have been identified as having functional importance in LT (Estella, McKay et al. 2008). Dll is also required for expression from the LT and LP enhancers, suggesting that Dll helps restrict its own expression (Estella, McKay et al. 2008; Galindo, Fernandez-Garza et al. 2011). Although Dpp and Wg are not needed to maintain *Dll* expression throughout leg development, continuous input from both is necessary to maintain the expression of the LT enhancer on its own(Campbell 2002; Galindo, Bishop et al. 2002; Estella, McKay et al. 2008). In order for *Dll* expression to become Wg and Dpp independent, some other factor must act. Addition of another enhancer called Maintenance (M) was found to be necessary to maintain LT expression through the rest of leg development. This enhancer acts, in part, through positive regulation by Dll directly (Estella, McKay et al. 2008).

The cells which will contribute to the leg disc are marked by more than a switch from 304 to LT/LP. The zinc finger transcription factor Escargot (Esg) is also found in these cells, as well as those that will form the wing disc, giving this gene the reputation as marker of imaginal fate (Morata 2001; Bolinger and Boekhoff-Falk 2005; McKay, Estella et al. 2009; Estella, Voutev et al. 2012). Some future leg cells expressing *esg* eventually lose *Dll* expression, though they maintain *hth* expression. These cells eventually give rise to the proximal portions of the leg (Bolinger and Boekhoff-Falk 2005; McKay, Estella et al. 2009). The cells of the future distal leg continue to express *hth* until after embryogenesis, when *hth* expression is turned off by a Wg/Dpp dependent method, potentially involving Dll. This deactivation is necessary for the cells of the distal leg to proliferate (Bolinger and Boekhoff-Falk 2005; McKay, Estella et al. 2009).

The SP1 family of zinc-finger transcription factors, Sp1 and Buttonhead (Btd), are also needed for proper leg development and are sufficient to transform wings towards leg identity when ectopically expressed (Estella, Rieckhof et al. 2003). *Sp1* and *btd* are both expressed prior to *Dll* in the future leg primordium, and are necessary for its maintenance (Estella, Rieckhof et al. 2003). Interestingly, their embryonic expression is dependent on activation by Wg and is repressed by Dpp and Ubx, much like the early expression of *Dll* (Estella, Rieckhof et al. 2003; Estella, Voutev et al. 2012). Sp1 and Btd do not activate *Dll* expression, and are not needed for activation of 304 (Estella, Rieckhof et al. 2003; McKay, Estella et al. 2009; Estella, Voutev et al. 2012), but are necessary for activation of LT (McKay, Estella et al. 2009). Since *Dll* is initially expressed in the cells which will form the wing disc, and given the transformative potential of these genes when ectopically expressed, it has been suggested that the SP1 genes function to distinguish leg fate from that of the wing (Estella, Rieckhof et al. 2003; McKay, Estella et al. 2009; Estella, Voutev et al. 2012). Another set of zinc-finger transcription factors, Disconnected (Disco) and Disco-related (Disco-r) are also expressed in the early leg primordium and in the leg imaginal disc (Cohen 1990; Patel, Farzana et al. 2007). Further, ectopic expression of *disco* in the wing also results in the transformation of these tissues towards leg identity (Patel, Farzana et al. 2007), and these genes appear to be necessary for proper Dll function in the leg (Dey, Zhao et al. 2009), suggesting that the *disco* genes might also play an important role in specifying leg fate.

Recent studies have suggested that the expression of *Dll* that is necessary for specifying the distal portions of the leg imaginal disc, as well as the medial-leg expression of *dac*, does not so much rely on direct activation by the Dpp pathway. Rather, it seems that Brk actively represses these genes, and the intervention of the Dpp pathway is necessary to relieve this repression (Estella and Mann 2008; Estella, Voutev et al. 2012). This work also suggests that *dac* is less sensitive to this repression than is *Dll*, so *Dll* expression is found in the center of the disc where the Dpp signal is able to completely suppress *brk* while receiving positive input from the Wg pathway (Estella and Mann 2008; Estella, Voutev et al. 2012).

In addition to maintaining its own expression, Dll is also necessary to activate expression of *dac* (Giorgianni and Mann 2011). Analysis of *dac* enhancers suggests that *dac* is not activated by either Wg or Dpp activity, but is instead repressed by high levels of both morphogens. This repression is enacted directly through binding sites for Pan and Mad, just as is activation by Dll, which also has binding sites in the *dac* enhancer (Giorgianni and Mann 2011). While Dll is needed to activate *dac*, many *dac*-expressing cells lose *Dll* expression, and what maintains *dac* expression in these cells is unknown, though autoregulation is one possibility (Giorgianni and Mann 2011). Repression of *dac* by Wg and Dpp is also transient, active only when *dac* is first expressed. Afterwards, *dac* is repressed by the distal leg activities of the EGFR pathway (Campbell 2002; Galindo, Bishop et al. 2002; Giorgianni and Mann 2011), while Dac provides the proximal limit for the distal targets of EGFR (Galindo, Bishop et al. 2002).

Dll works with Dpp and Wg to activate the EGFR pathway in the center- (distal-) most portion of the leg disc (Campbell 2002; Galindo, Bishop et al. 2002; Kojima 2004; Estella, Voutev et al. 2012). EGFR signaling then works in a gradient-dependent manner to activate a series of genes that help pattern the future tarsus of the leg (Campbell 2002; Galindo, Bishop et al. 2002). The highest level of Vn activates expression of homoeodomain genes *aristaless* (*al*) and *Lim1*, which are necessary for proper formation of the distal tip of the leg, including the terminal claw. Adjacent to this domain, *Bar* is activated by slightly reduced EGFR signal (Campbell 2002; Galindo, Bishop et al. 2002; Kojima 2004). At the edge of the EGFR signal, the zinc-finger-encoding gene *rotund* (*rn*) is briefly activated, followed quickly by the expression of the *bric-a-brac* (*bab*) genes. These require *Dll* expression, and repress *dac*, which also reciprocally represses *bab* and *rn* expression. Although EGFR input is needed for their expression, high EGFR signal actually prevents *bab* and *rn* expression in the distal portions of the leg (Campbell 2002; Galindo, Bishop et al. 2002).

CONSERVED AND DIVERGENT FUNCTIONS OF LEG-PATTERNING GENES

Unlike the two-dimensional patterning used by Drosophila for its limb development, vertebrates and many other arthropods pattern their limbs in three-dimensional, cylindrical outgrowths during embryonic development. These appendages will then grow directly into their adult forms during pupation. Because of these spatial and temporal differences, Drosophila limb development is not expected to be representative of limb development in other taxa (Prpic, Wigand et al. 2001; Angelini and Kaufman 2005). Therefore, understanding of limb development is being sought in less traditional model organisms, through the use of such techniques as in situ hybridization and RNA interference (RNAi).

It is, perhaps, not surprising that current literature on limb development reveals two trends in other organisms, compared with Drosophila: (1) proximal-distal (P-D) patterning by *dll*, *dac*, and *hth*/*exd* appears to be highly conserved, while (2) elements upstream of this seem to have diverged (Pueyo and Couso 2005). This might be expected, since the regulatory changes necessary to form limbs from imaginal disc would probably have happened in events that take place before and after P-D patterning (Ober and Jockusch 2006). Indeed, *wg* and *dpp* expression are most conserved amongst animals with directly developing limbs than they are when these species are compared to the fly (Prpic, Janssen et al. 2003). Even examinations of other upstream factors, like various Hox genes, reveal some divergence. This pattern can be clearly seen in the literature with only a sampling of arthropod species. In particular, I will describe information from *Tribolium castaneum* (beetle), *Oncopeltus fasciatus* (hemipteran), *Glomeris marginata* (myriapod) and *Cupiennius salei* (spider).

Basic patterning

In the Tribolium embryo, *Tc-Dll* "is expressed in the most distal part of all… appendages except the mandibles," but this single area of expression is later altered into two domains referred to as a "'ring and sock' pattern" (Prpic, Wigand et al. 2001). Similarly, embryonic *Tc-dac* expression resembles that in the fly imaginal disc, occupying the middle portion of most limbs, except the mandibles, where *Tc-dac* expression extends to the distal tip. In the thoracic limbs and the maxilla, *Tc-dac* forms two distinct domains of expression. In the legs, the more distal domain forms within the early *Tc-Dll* expression domain, and eventually divides the *Tc-Dll* expression into the "sock" and the "ring" domains; *Tc-Dll* expression is eliminated from the area of *Tc-dac* expression, except for a small overlap at the distal end of the *Tc-dac* domain. Interestingly, without *Tc-Dll*, this more distal domain of *Tc-dac* appears not to form at all; instead, the appendage primordium does not seem to extend beyond the weaker proximal domain (Prpic, Wigand et al. 2001). This is consistent with a Drosophila leg development model that suggests *dac* is activated by *Dll* rather than directly by *wg* and *dpp* in the developing disc (Giorgianni and Mann 2011; Estella, Voutev et al. 2012), suggesting that this interaction is conserved. Reduction of the function of *Tc-Dll* and *Tc-dac* prior to pupation resultes in the loss of distal and medial (respectively) leg structures in adults, suggesting a conserved function for these genes in adult development (Angelini, Smith et al. 2012).

Tc-exd and *Tc-hth* expression also resemble the *Drosophila* counterparts. Throughout leg development, *Tc-exd* is expressed in the full length of the leg. Expression of *Tc-hth* is slightly more dynamic. At the earliest stages, *Tc-hth* is rather broadly expressed, except in the distal portions of the legs. Later, two domains form in the legs, both relatively proximal (Prpic and Tautz 2003). Interestingly, RNAi of *Tc-hth* prior to pupation does not result in a loss of proximal segments, but rather appears to transform these segments towards more distal like identities (Angelini, Smith et al. 2012).

The expression patterns of *Dll*, *dac* and *hth* orthologs in Oncopeltus are very similar to those in Tribolium, and have been functionally corroborated with RNAi experiments. *Of'Dll* is first expressed where the limb buds will eventually emerge. As the limbs elongate, *Of'Dll* becomes restricted to the distal end, forming a "ring and sock" pattern in the legs, as in Tribolium. Only the mandible lacks expression, while the maxilla continues to express *Of'Dll* in its entire length. *Of'Dll* RNAi resulted in defects in the distal portions of the legs, labium and antennae (Angelini and Kaufman 2004).

Of'dac is expressed in the middle portions of the legs, possibly lying between the two *Of'Dll* domains. RNAi results in defects in the leg's middle segments. *Of'hth* is expressed throughout the embryo, including most of the antennae, mandibles and maxilla, but the distal portions of the legs and labium remain free of expression. RNAi results in a number of defects, including homeotic transformations (ex: distal labium is transformed towards leg identity), proximal truncation and fusion of legs, and missing antennae (Angelini and Kaufman 2004). The homeotic transformations are perhaps not surprising, given the association of both *hth* and *exd* genes as cofactors for Hox genes.(Moens and Selleri 2006)

As in Oncopeltus, expression of the *Dll* ortholog in *Glomeris* is present in the presumptive limb primordia before the appendages are physically visible, possibly prefiguring the sight of limb development. As the limb buds take shape, expression is very dynamic in the mouthparts, but is restricted to the distal portion of the antennae and legs. In the legs, *Gm-Dll* appears to form two domains similar to those observed in Tribolium and Oncopeltus. *Gm-dac* is also activated prior to formation of the limb buds, but only in the presumptive appendage primorida of the head segments. When the remaining limb buds start to form, a ring of expression develops in the middle portion of each leg. There is also a proximal domain of *Gm-dac* expression that forms in the legs and antennae, probably where sensory organs form (Prpic and Tautz 2003).

Although the expression patterns of Glomeris orthologs of *Dll* and *dac* appear to be conserved between insects and the myriapod, expression of *Gm-hth* and *Gm-exd* are slightly different. Unlike the broadly expressed insect *exd* orthologs, *Gm-exd* is restricted from the distal portions of almost all limbs, expressing only in the most proximal portions of the legs. *Gm-hth*, however, is more broadly expressed in the all limbs, extending to the distal end in all but the leg primordia. But, as Prpic and Tautz (2003) point out, since *hth* and *exd* are cofactors, their combined expression domain in Glomeris is similar to the combined expression of *exd* and *hth* in insects. Therefore, Prpic and Tautz conclude that function of *hth*/*exd* is also conserved in this species.

Similarly, expression patterns of *hth* and *exd* in the spider, Cupiennius, are reversed from those in insects. *Cs-hth-1* is expressed throughout most of the limb primordia, except for the distal tip, while *Cs-exd-1* expression is restricted to the proximal portions of the limbs. As in Glomeris, the domain of mutual, and thus functional, expression suggests that the roles of these genes in proximal limb development are conserved between insects and these other arthropod species. Prpic et al. (2003) suggest that *exd* and *hth* were both restricted proximally in an arthropod ancestor, but since it is not necessary to restrict both genes, restriction of one was independently relaxed in the various arthropod lineages, resulting in the differences between insects and the other arthropod groups examined. Examination of more species in each lineage, and even in different lineages, could verify this hypothesis.

Expression patterns of *Dll* and *dac* orthologs appear to be mostly conserved in Cupiennius. *Cs-Dll* is expressed in a single distal domain. *Cs-dac* comes on later in an intermediate domain that overlaps with *Cs-Dll*; a more proximal spot develops later, as in Tribolium, while the overlap shifts distally. However, *Cs-dac* does not appear to split the *Cs-Dll* domain as it does in Oncopeltus and Tribolium (Prpic and Tautz 2003).

Prpic and Damen (2004) later examine *Dll*, *dac*, *exd* and *hth* expression the spider's other two appendage types, the leg-like pedipalps and the chelicerae. They find expression in the pedipalps resembles that in the legs. However, they find expression in the chelicerae to more closely resemble the expression patterns in fly antennae; *exd*/*hth* and *Dll* expression domains occupy most of the antenna and overlap extensively. In the legs of spiders and flies, *Dll*, *dac* and *hth*/*exd* form mutually exclusive domains of expression. Prpic and Damen suggest that alterations in this mutual exclusion are necessary to form non-leg appendages. Indeed, examination of the mouthparts in *Oncopeltus* (Angelini and Kaufman 2004) and the mouthparts and antennae in *Tribolium* (Prpic, Wigand et al. 2001; Prpic, Janssen et al. 2003) and *Glomeris* (Prpic and Tautz 2003) seem to agree with this conclusion.

Prpic and Damen (2003) also examine the expression patterns of *dpp* and *wg* orthologs in the spider, and observe some notable differences from their Drosophila counterparts. *Cs-dpp* is expressed in the distal tip of the outgrowing appendages. Later, *Cs-dpp* is expressed in rings that might be related to joint formation. *Cs-wg*, however, is expressed in an anterior-ventral stripe along the length of the appendages, in a region resembling Drosophila leg expression of *wg*, except for the termination of *Cs-wg* domain just before the distal tip. This group also observes similar patterns of expression of *wg* and *dpp* orthologs in Glomeris (Prpic and Tautz 2003).

The difference between expression of fly and spider *dpp*, while *wg* appears relatively conserved, prompts Prpic and Damen to propose a model that explains this difference via the topological differences in limb development between these species. In this model, the three-dimensional growth of the spider leg necessitates one of these two signaling molecules to be limited to the distal tip of the outgrowing limb. In the two-dimensional patterning of the Drosophila leg primordium, the dorsal and ventral expression of *dpp* and *wg* , respectively, results in a central domain receiving a high signal of both morphogens, which is necessary for activation of *Dll* (Fig 2). However, due to the three-dimensional patterning of the embryonic spider limbs, if both *Cs-wg* and *Cs-dpp* were expressed in ventral and dorsal stripes, respectively, in a way that would resemble fly expression, the genes that would be activated by their mutual presence (like *Dll*) would be expressed in the entire length of the spider's cylindrical limb, instead of just at the tip. Instead, *Cs-wg* is expressed in a ventral stripe and *Cs-dpp* in the distal tip. Therefore, the areas of mutual expression are strongest at the distal tip, fading toward the proximal portions, similar to the areas of mutual expression of the fly imaginal discs (Prpic, Janssen et al. 2003).

Expression patterns of *wg* and *dpp* orthologs in Oncopeltus and Tribolium were found to be very similar to those in *Cupiennius*, *wg* in a ventral stripe, *dpp* in the distal tip. However, functional analyses in these organisms have led the respective researchers to reject Prpic and Damen's topology model (Angelini and Kaufman 2005; Ober and Jockusch 2006).

Angelini and Kaufman (2005) state that the topology model makes two testable assumptions. First, *wg* and *dpp* "should be critical to the proper development of appendages" in other insects. Second, that the "genetic pathway of the Drosophila limb PD axis specification mechanism is conserved." Therefore, if the topology model is true, perturbations of *wg* or *dpp* orthologs should result both in failure of appendages to develop and in improper expression of P-D genes like the orthologs of *Dll*, *dac* and *exd*/*hth*. Unfortunately, *Of'dpp* RNAi results in early embryo termination because the germband failed to invaginate. *Of'wg* RNAi embryos develop, but do not display any limb defects. This result was sufficient for Angelini and Kaufman to reject the topology model as an explanation for the differences in *dpp* expression between Drosophila and Oncopeltus.

RNAi of the *wg* ortholog in *Tribolium* does result in leg defects. A high proportion (76%) of embryos had negligible limb outgrowth in gnathal and thoracic segments. *Tc-Dll* expression is also missing in these segments. However, *Tc-dpp* RNAi does not prevent leg formation, and only a very small portion (2%) of embryos exhibit possible defects in segmentation. More importantly, *Tc-Dll* expression was normal at all stages (Ober and Jockusch 2006). It seems, therefore, that *dpp* function in *Tribolium* does not fit the topology model, and it has been suggested that the appendage role of *wg* is recently acquired (Angelini and Kaufman 2005; Ober and Jockusch 2006; Grossmann, Scholten et al. 2009). Reduction of *Tc-dpp* mRNA prior to pupation results in loss or malformation of distal leg regions (Angelini, Smith et al. 2012), suggesting its function, at this stage, mirrors its expression pattern (which the authors note still possesses a distal bias).

Changes in selector gene functions

A handful of selector genes also show functional differences in Tribolium and Oncopeltus from their roles in Drosophila. Herke et al. (2005) define selector genes as "transcription factors that, in combinations, define the body plan by controlling the development of specific regions of the embryo." For example, in insects, these genes are believed to "modify the development of a default, or 'ground state', appendage into tagma-specific appendage such as a mouthpart, antenna or leg" (Herke, Serio et al. 2005). Specifically, *proboscipedia* (*pb*) and *Sex combs reduced* (*Scr*) are both necessary in *Drosophila* adults to specify proboscis development in the labial segment. In the absence of *pb*, the appendage is transformed towards first thoracic tarsus. Without *Scr,* the appendage is transformed towards maxillary identity. Clones mutant for both genes are transformed towards antennal identity (DeCamillis, Lewis et al. 2001). Similarly, *Ultrabithorax* (*Ubx*) and *abdominal-A* (*abd-A*) can each repress *Dll* expression in the abdomen of the fly, preventing appendage development, even though *Ubx* is also expressed in the third thoracic segment. Here, leg development is able to proceed because *Dll* is activated before *Ubx*, and is able to maintain its own expression in the presence of *Ubx* activity. If, however, *Ubx* is activated first, *Dll* does not get turned on (Lewis, DeCamillis et al. 2000). Each of these examples also shows notable functional changes in other insects.

Tribolium orthologs of *Scr* and *pb* (*Cephalothorax*, *Cx*, and *maxillopedia*, *mxp*, respectively) were examined for their affect on labial development. Loss of *mxp* results in transformation of the labial appendage into leg. Like in the fly, *Cx* is expressed in the first thoracic segment as well as the labial segment. Without *mxp* expression, *Cx* can only confer leg identity. However, the loss of *Cx*, as well as loss of both *Cx* and *mxp*, causes the labial appendages to develop like antennae. Examination of *mxp* expression in *Cx* null mutants revealed that *mxp* expression is lost. In essence, *Cx* mutants are nulls for both selectors. Similar expression results are seen in the labial segments of Oncopeltus and of Drosophila larvae, leading to the conclusion that the interactions between *Scr* and *pb* have been modified in probiscis development in the adult fly (DeCamillis, Lewis et al. 2001).

Differences have also been found in Tribolium for the mechanisms behind the prevention of abdominal appendage development. The *Ubx* ortholog, *Ultrathorax* (*Utx*), is expressed from the third thoracic segment posteriorly, and the expression of the *abd-A* ortholog, *Abdominal* (*A*), begins in the posterior half of the first abdominal segment (A1). Also in this segment, a pair of appendage-like pleuropodia develops in the embryo, at the only sites of *Tc-Dll* expression in the abdomen. The pleuropodia develop only in the anterior half of A1. Initially this development resembles that of the thoracic legs, but eventually the pleuropodia invaginate. Loss of *Utx* causes the pleuropodia to continue development towards leg identity, even developing subterminal tarsal claws. But this pair of abdominal “legs” does not develop in the posterior portion of A1, nor does *Tc-Dll* expression appear in any other abdominal segment. Loss of *A*, however, results in pairs of *Tc-Dll*-expressing sites in all abdominal segments. However, these regions develop like the pleuropodia of A1. Only with loss of both selectors are ectopic legs able to develop in the abdominal segments. These results suggest that that the ancestral role of *abd-A* was to block *Dll* expression in the abdomen, while the ancestral *Ubx* functioned more like a modifier, turning *Dll* expressing cells away from leg identity. Its role in blocking *Dll* expression is therefore thought to be new to Diptera and Lepidoptera (Lewis, DeCamillis et al. 2000).

BROADENING EVOLUTIONARY PERSPECTIVES IN LIMB DEVELOPMENT

The number of species in which the leg development genes have been studied is ever increasing. As with the species described above, the expression patterns of orthologs to the P-D patterning genes *exd*/*hth*, *dac* and *Dll* appear to be conserved in the developing limbs of the cockroach, *Periplaneta Americana* (Blagburn 2007), of the moth, *Manduca sexta* (Tanaka and Truman 2007), and of the crustacean species, *Triops longicaudatus* and *Thamnocephalus platyurus* (Williams, Nulsen et al. 2002; Sewell, Williams et al. 2008). So it is perhaps not too surprising that similar findings have been made in vertebrate species, including the zebra fish (Akimenko, Ekker et al. 1994; Popperl, Rikhof et al. 2000; Waskiewicz, Rikhof et al. 2001; Hammond, Hill et al. 2002), the chicken (Ferrari, Sumoy et al. 1995; Kida, Maeda et al. 2004; Capellini, Zappavigna et al. 2011; Sanchez-Guardado, Irimia et al. 2011) and the mouse (Hanson 2001; Pueyo and Couso 2005; Kraus and Lufkin 2006).

Summary

Many of the genes known to be involved in Drosophila leg development are either transcription factors or are members of the signaling pathways mentioned above, but their targets remain largely a mystery (Morata 2001). As has been seen, these signaling pathways and transcription factors have orthologs in a great variety of species. Further, vertebrate orthologs have been associated with cancer (Briscoe and Therond 2005; Polakis 2007; Morini, Astigiano et al. 2010; Popov, Wu et al. 2010). These genes and their orthologs also seem to be important in nervous system development (Dirksen, Mathers et al. 1993; Akimenko, Ekker et al. 1994; Panganiban, Irvine et al. 1997; Merlo, Zerega et al. 2000; Popperl, Rikhof et al. 2000; Waskiewicz, Rikhof et al. 2001; Hammond, Hill et al. 2002; Williams, Nulsen et al. 2002; Angelini and Kaufman 2004; Kraus and Lufkin 2006; Blagburn 2007). Identifying their targets in leg development would be helpful in expanding our understanding in these other areas (Panganiban 2000). One of the objectives of my research has been to expand our understanding of the genes involved in leg development.

We know that critical genes have been missed in these models for Drosophila leg development. The *disco* genes, *disconnected* and *disco-related*, encode C_2H_2 zinc finger proteins that have been shown genetically to be Hox co-factors (Mahaffey, Griswold et al. 2001; Robertson, Bowling et al. 2004). However, their importance in leg development has only been shown relatively recently, as has the conservation of this role in another insect species, *Tribolium castenum* (Patel, Farzana et al. 2007). Although there are no available mutations for *disco* in Tribolium, RNAi in this species is simplified by the ability to inject dsRNA into female pupae and examine the resulting phenocopy in her offspring, a phenomenon referred to as parental RNAi (Bucher, Scholten et al. 2002). There still remain questions of the degree to which the functions of *disco* genes are conserved between Tribolium and Drosophila. Chapter two explores in more depth what is conserved between these two species, particularly examining the interactions of Tribolium *disco* with another transcription factor, the Tribolium ortholog of *teashirt*.

Understanding of *disco* genes does not extend much beyond these two species. Vertebrates appear to have *disco* orthologs, the two *basonuclin* (*bnc*) genes, which have been studied to some degree in mice and zebra fish (Tseng 1998; Romano, Li et al. 2004; Vanhoutteghem and Djian 2004; Lang, Patterson et al. 2009; Vanhoutteghem, Bouche et al. 2011). Whether or not these genes are also involved in appendage development has not yet been clarified. Chapter three will examine the functions of *bnc1* and *bnc2* in the mouse and zebra fish, and what possible conserved roles they may share with their insect orthologs.

This research also seeks to return to the broader questions of leg development. Chapter four will demonstrate the successful use of genome wide association studies to identify genetic factors contributing to Drosophila leg development that had not previously been characterized in this role. This research took advantage of a set of lines called the Drosophila Genetic Reference Panel (DGRP (Mackay, Richards et al. 2012). These lines were created with isolated females caught from a wild population in Raleigh, NC. Their descendants were full-sib mated for 20 generations to create lines of mostly isogenic flies (Mackay, Richards et al. 2012). What makes these lines particularly useful is that although there is minimal genetic diversity within a line, the diversity amongst the lines is comparable to wild populations (Ayroles, Carbone et al. 2009; Mackay, Richards et al. 2012). The lack of genetic diversity within a line means that a complex trait can be measured multiple times on the same genetic background. This permits an examination of the role of environment upon the trait, since it is the primary variable within the line. Meanwhile, differences between lines for any particular trait can be attributed to genetic differences. The genome sequences of the lines can then be examined for any genetic markers (ex: single nucleotide polymorphisms, SNPs) associating with particular phenotypes. Many of these lines have been measured for a number of quantitative traits, including alcohol sensitivity and tolerance and transcript abundance, and show a great deal of between-line diversity (Ayroles, Carbone et al. 2009; Yamamoto, Anholt et al. 2009; Mackay, Richards et al. 2012). With respect to alcohol sensitivity and tolerance, these lines are described as showing a "range of variation…comparable to that achieved after 25 generations of selection."(Morozova, Ayroles et al. 2009).

Finally, the implications of this research in expanding our understanding of *disco* genes and leg development will be discussed in the closing chapter.

References:

Affolter, M. and K. Basler (2007). "The Decapentaplegic morphogen gradient: from pattern formation to growth regulation." Nat Rev Genet **8**(9): 663-74.

Akimenko, M. A., M. Ekker, et al. (1994). "Combinatorial expression of three zebrafish genes related to distal-less: part of a homeobox gene code for the head." J Neurosci **14**(6): 3475-86.

Angelini, D. R. and T. C. Kaufman (2004). "Functional analyses in the hemipteran Oncopeltus fasciatus reveal conserved and derived aspects of appendage patterning in insects." Dev Biol **271**(2): 306-21.

Angelini, D. R. and T. C. Kaufman (2005). "Functional analyses in the milkweed bug Oncopeltus fasciatus (Hemiptera) support a role for Wnt signaling in body segmentation but not appendage development." Dev Biol **283**(2): 409-23.

Angelini, D. R., F. W. Smith, et al. (2012). "Extent With Modification: Leg Patterning in the Beetle Tribolium castaneum and the Evolution of Serial Homologs." G3 (Bethesda) **2**(2): 235-48.

Ayroles, J. F., M. A. Carbone, et al. (2009). "Systems genetics of complex traits in Drosophila melanogaster." Nat Genet **41**(3): 299-307.

Blagburn, J. M. (2007). "Co-factors and co-repressors of Engrailed: expression in the central nervous system and cerci of the cockroach, Periplaneta americana." Cell Tissue Res **327**(1): 177-87.

Bolinger, R. A. and G. Boekhoff-Falk (2005). "Distal-less functions in subdividing the Drosophila thoracic limb primordium." Dev Dyn **232**(3): 801-16.

Briscoe, J. and P. Therond (2005). "Hedgehog signaling: from the Drosophila cuticle to anti-cancer drugs." Dev Cell **8**(2): 143-51.

Brook, W. J. (2010). "T-box genes organize the dorsal ventral leg axis in Drosophila melanogaster." Fly (Austin) **4**(2): 159-62.

Bucher, G., J. Scholten, et al. (2002). "Parental RNAi in Tribolium (Coleoptera)." Curr Biol **12**(3): R85-6.

Camp, D., K. Currie, et al. (2010). "Ihog and Boi are essential for Hedgehog signaling in Drosophila." Neural Dev **5**: 28.

Campbell, G. (2002). "Distalization of the Drosophila leg by graded EGF-receptor activity." Nature **418**(6899): 781-5.

Capellini, T. D., V. Zappavigna, et al. (2011). "Pbx homeodomain proteins: TALEnted regulators of limb patterning and outgrowth." Dev Dyn **240**(5): 1063-86.

Cohen, S. M. (1990). "Specification of limb development in the Drosophila embryo by positional cues from segmentation genes." Nature **343**(6254): 173-7.

Cohen, S. M., G. Bronner, et al. (1989). "Distal-less encodes a homoeodomain protein required for limb development in Drosophila." Nature **338**(6214): 432-4.

DeCamillis, M. A., D. L. Lewis, et al. (2001). "Interactions of the Tribolium Sex combs reduced and proboscipedia orthologs in embryonic labial development." Genetics **159**(4): 1643-8.

Dey, B. K., X. L. Zhao, et al. (2009). "Mutual regulation of the Drosophila disconnected (disco) and Distal-less (Dll) genes contributes to proximal-distal patterning of antenna and leg." Cell Tissue Res **338**(2): 227-40.

Dirksen, M. L., P. Mathers, et al. (1993). "Expression of a Xenopus Distal-less homeobox gene involved in forebrain and cranio-facial development." Mech Dev **41**(2-3): 121-8.

Erickson, J. L. (2011). "Formation and maintenance of morphogen gradients: an essential role for the endomembrane system in Drosophila melanogaster wing development." Fly (Austin) **5**(3): 266-71.

Estella, C. and R. S. Mann (2008). "Logic of Wg and Dpp induction of distal and medial fates in the Drosophila leg." Development **135**(4): 627-36.

Estella, C., D. J. McKay, et al. (2008). "Molecular integration of wingless, decapentaplegic, and autoregulatory inputs into Distalless during Drosophila leg development." Dev Cell **14**(1): 86-96.

Estella, C., G. Rieckhof, et al. (2003). "The role of buttonhead and Sp1 in the development of the ventral imaginal discs of Drosophila." Development **130**(24): 5929-41.

Estella, C., R. Voutev, et al. (2012). "A dynamic network of morphogens and transcription factors patterns the fly leg." Curr Top Dev Biol **98**: 173-98.

Ferrari, D., L. Sumoy, et al. (1995). "The expression pattern of the Distal-less homeobox-containing gene Dlx-5 in the developing chick limb bud suggests its involvement in apical ectodermal ridge activity, pattern formation, and cartilage differentiation." Mech Dev **52**(2-3): 257-64.

Galindo, M. I., S. A. Bishop, et al. (2002). "Leg patterning driven by proximal-distal interactions and EGFR signaling." Science **297**(5579): 256-9.

Galindo, M. I., D. Fernandez-Garza, et al. (2011). "Control of Distal-less expression in the Drosophila appendages by functional 3' enhancers." Dev Biol **353**(2): 396-410.

Gao, C. and Y. G. Chen (2010). "Dishevelled: The hub of Wnt signaling." Cell Signal **22**(5): 717-27.

Giorgianni, M. W. and R. S. Mann (2011). "Establishment of medial fates along the proximodistal axis of the Drosophila leg through direct activation of dachshund by Distalless." Dev Cell **20**(4): 455-68.

Gonsalves, F. C. and R. DasGupta (2008). "Function of the wingless signaling pathway in Drosophila." Methods Mol Biol **469**: 115-25.

Goto, S. and S. Hayashi (1997). "Specification of the embryonic limb primordium by graded activity of Decapentaplegic." Development **124**(1): 125-32.

Grossmann, D., J. Scholten, et al. (2009). "Separable functions of wingless in distal and ventral patterning of the Tribolium leg." Dev Genes Evol **219**(9-10): 469-79.

Hammond, K. L., R. E. Hill, et al. (2002). "Isolation of three zebrafish dachshund homologues and their expression in sensory organs, the central nervous system and pectoral fin buds." Mech Dev **112**(1-2): 183-9.

Hanson, I. M. (2001). "Mammalian homologues of the Drosophila eye specification genes." Semin Cell Dev Biol **12**(6): 475-84.

Herke, S. W., N. V. Serio, et al. (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**(1): 27-34.

Jiang, J. and C. C. Hui (2008). "Hedgehog signaling in development and cancer." Dev Cell **15**(6): 801-12.

Kida, Y., Y. Maeda, et al. (2004). "Chick Dach1 interacts with the Smad complex and Sin3a to control AER formation and limb development along the proximodistal axis." Development **131**(17): 4179-87.

Kojima, T. (2004). "The mechanism of Drosophila leg development along the proximodistal axis." Dev Growth Differ **46**(2): 115-29.

Kraus, P. and T. Lufkin (2006). "Dlx homeobox gene control of mammalian limb and craniofacial development." Am J Med Genet A **140**(13): 1366-74.

Kubota, K., S. Goto, et al. (2003). "The role of Wg signaling in the patterning of embryonic leg primordium in Drosophila." Dev Biol **257**(1): 117-26.

Lang, M. R., L. B. Patterson, et al. (2009). "Basonuclin-2 requirements for zebrafish adult pigment pattern development and female fertility." PLoS Genet **5**(11): e1000744.

Lecuit, T. and S. M. Cohen (1997). "Proximal-distal axis formation in the Drosophila leg." Nature **388**(6638): 139-45.

Lewis, D. L., M. DeCamillis, et al. (2000). "Distinct roles of the homeotic genes Ubx and abd-A in beetle embryonic abdominal appendage development." Proc Natl Acad Sci U S A **97**(9): 4504-9.

Lum, L., S. Yao, et al. (2003). "Identification of Hedgehog pathway components by RNAi in Drosophila cultured cells." Science **299**(5615): 2039-45.

Mackay, T. F., S. Richards, et al. (2012). "The Drosophila melanogaster Genetic Reference Panel." Nature **482**(7384): 173-8.

Mahaffey, J. W., C. M. Griswold, et al. (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**(1): 225-36.

McKay, D. J., C. Estella, et al. (2009). "The origins of the Drosophila leg revealed by the cis-regulatory architecture of the Distalless gene." Development **136**(1): 61-71.

Merlo, G. R., B. Zerega, et al. (2000). "Multiple functions of Dlx genes." Int J Dev Biol **44**(6): 619-26.

Moens, C. B. and L. Selleri (2006). "Hox cofactors in vertebrate development." Dev Biol **291**(2): 193-206.

Morata, G. (2001). "How Drosophila appendages develop." Nat Rev Mol Cell Biol **2**(2): 89-97.

Morini, M., S. Astigiano, et al. (2010). "Mutually exclusive expression of DLX2 and DLX5/6 is associated with the metastatic potential of the human breast cancer cell line MDA-MB-231." BMC Cancer **10**: 649.

Morozova, T. V., J. F. Ayroles, et al. (2009). "Alcohol sensitivity in Drosophila: translational potential of systems genetics." Genetics **183**(2): 733-45, 1SI-12SI.

Mulligan, K. A., C. Fuerer, et al. (2012). "Secreted Wingless-interacting molecule (Swim) promotes long-range signaling by maintaining Wingless solubility." Proc Natl Acad Sci U S A **109**(2): 370-7.

Ober, K. A. and E. L. Jockusch (2006). "The roles of wingless and decapentaplegic in axis and appendage development in the red flour beetle, Tribolium castaneum." Dev Biol **294**(2): 391-405.

Panganiban, G. (2000). "Distal-less function during Drosophila appendage and sense organ development." Dev Dyn **218**(4): 554-62.

Panganiban, G., S. M. Irvine, et al. (1997). "The origin and evolution of animal appendages." Proc Natl Acad Sci U S A **94**(10): 5162-6.

Patel, M., L. Farzana, et al. (2007). "The appendage role of insect disco genes and possible implications on the evolution of the maggot larval form." Dev Biol **309**(1): 56-69.

Penton, A. and F. M. Hoffmann (1996). "Decapentaplegic restricts the domain of wingless during Drosophila limb patterning." Nature **382**(6587): 162-4.

Polakis, P. (2007). "The many ways of Wnt in cancer." Curr Opin Genet Dev **17**(1): 45-51.

Popov, V. M., K. Wu, et al. (2010). "The Dachshund gene in development and hormone-responsive tumorigenesis." Trends Endocrinol Metab **21**(1): 41-9.

Popperl, H., H. Rikhof, et al. (2000). "lazarus is a novel pbx gene that globally mediates hox gene function in zebrafish." Mol Cell **6**(2): 255-67.

Prpic, N. M., R. Janssen, et al. (2003). "Gene expression in spider appendages reveals reversal of exd/hth spatial specificity, altered leg gap gene dynamics, and suggests divergent distal morphogen signaling." Dev Biol **264**(1): 119-40.

Prpic, N. M. and D. Tautz (2003). "The expression of the proximodistal axis patterning genes Distal-less and dachshund in the appendages of Glomeris marginata (Myriapoda: Diplopoda) suggests a special role of these genes in patterning the head appendages." Dev Biol **260**(1): 97-112.

Prpic, N. M., B. Wigand, et al. (2001). "Expression of dachshund in wild-type and Distal-less mutant Tribolium corroborates serial homologies in insect appendages." Dev Genes Evol **211**(10): 467-77.

Pueyo, J. I. and J. P. Couso (2005). "Parallels between the proximal-distal development of vertebrate and arthropod appendages: homology without an ancestor?" Curr Opin Genet Dev **15**(4): 439-46.

Robertson, L. K., D. B. Bowling, et al. (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**(12): 2781-9.

Romano, R. A., H. Li, et al. (2004). "Identification of Basonuclin2, a DNA-binding zinc-finger protein expressed in germ tissues and skin keratinocytes." Genomics **83**(5): 821-33.

Sanchez-Guardado, L. O., M. Irimia, et al. (2011). "Distinct and redundant expression and transcriptional diversity of MEIS gene paralogs during chicken development." Dev Dyn **240**(6): 1475-92.

Schwank, G., S. Dalessi, et al. (2011). "Formation of the long range Dpp morphogen gradient." PLoS Biol **9**(7): e1001111.

Sewell, W., T. Williams, et al. (2008). "Evidence for a novel role for dachshund in patterning the proximal arthropod leg." Dev Genes Evol **218**(6): 293-305.

Shilo, B. Z. (2003). "Signaling by the Drosophila epidermal growth factor receptor pathway during development." Exp Cell Res **284**(1): 140-9.

Shilo, B. Z. (2005). "Regulating the dynamics of EGF receptor signaling in space and time." Development **132**(18): 4017-27.

Shubin, N., C. Tabin, et al. (1997). "Fossils, genes and the evolution of animal limbs." Nature **388**(6643): 639-48.

Sopory, S., S. Kwon, et al. (2010). "Regulation of Dpp activity by tissue-specific cleavage of an upstream site within the prodomain." Dev Biol **346**(1): 102-12.

Tanaka, K. and J. W. Truman (2007). "Molecular patterning mechanism underlying metamorphosis of the thoracic leg in Manduca sexta." Dev Biol **305**(2): 539-50.

Theisen, H., A. Syed, et al. (2007). "Wingless directly represses DPP morphogen expression via an armadillo/TCF/Brinker complex." PLoS One **2**(1): e142.

Tseng, H. (1998). "Basonuclin, a zinc finger protein associated with epithelial expansion and proliferation." Front Biosci **3**: D985-8.

Vanhoutteghem, A., C. Bouche, et al. (2011). "Basonuclins and disco: Orthologous zinc finger proteins essential for development in vertebrates and arthropods." Biochimie **93**(2): 127-33.

Vanhoutteghem, A. and P. Djian (2004). "Basonuclin 2: an extremely conserved homolog of the zinc finger protein basonuclin." Proc Natl Acad Sci U S A **101**(10): 3468-73.

Waskiewicz, A. J., H. A. Rikhof, et al. (2001). "Zebrafish Meis functions to stabilize Pbx proteins and regulate hindbrain patterning." Development **128**(21): 4139-51.

Williams, T. A., C. Nulsen, et al. (2002). "A complex role for distal-less in crustacean appendage development." Dev Biol **241**(2): 302-12.

Wu, J. T., W. H. Lin, et al. (2011). "CSN-mediated deneddylation differentially modulates Ci(155) proteolysis to promote Hedgehog signalling responses." Nat Commun **2**: 182.

Yamamoto, A., R. R. Anholt, et al. (2009). "Epistatic interactions attenuate mutations affecting startle behaviour in Drosophila melanogaster." Genet Res **91**(6): 373-82.

Yan, D., Y. Wu, et al. (2010). "The cell-surface proteins Dally-like and Ihog differentially regulate Hedgehog signaling strength and range during development." Development **137**(12): 2033-44.

Yao, S., L. Lum, et al. (2006). "The ihog cell-surface proteins bind Hedgehog and mediate pathway activation." Cell **125**(2): 343-57.

Zheng, X., R. K. Mann, et al. (2010). "Genetic and biochemical definition of the Hedgehog receptor." Genes Dev **24**(1): 57-71.

CHAPTER TWO

Molecular Analysis of the *disconnected* Ortholog in the Beetle, *Tribolium castaneum*

INTRODUCTION

In 2001, a study was published revealing a pair of genes to be involved in proper head development of *Drosophila melanogaster* larvae (Mahaffey, Griswold et al. 2001); these genes were identified as *disconnected* (*disco*) and its redundant paralog, *disco-related* (*disco-r*). Disco had been identified previously as a C_2H_2 zinc finger transcription factor (Heilig, Freeman et al. 1991; Lee, Mukhopadhyay et al. 1999) important for proper migration of the Bolwig nerve during embryonic development in Drosophila (Steller, Fischbach et al. 1987). However, the role in embryonic head development requires the removal of both genes.

In specifying the larval head, the *disco* genes seem to act in a dichotomous role opposite another zinc-finger transcription factor, *teashirt* (*tsh*), where *tsh* is responsible for proper trunk development (Fasano, Roder et al. 1991; Robertson, Bowling et al. 2004). Ectopic expression of *tsh* in the developing head causes those segments to take on characteristics of trunk segments. Conversely, over expression of *disco* genes in trunk segments seem to cause head-like structures to develop. To determine how conserved these functions and interactions are, there first needed to be an examination of the role of *disco* and *tsh* orthologs in other insect species.

In her doctoral dissertation, Lisa K. Robertson (2005) described in detail the embryonic expression pattern and RNA-knockdown phenocopy of the single *disco* ortholog in the red flour beetle, *Tribolium castaneum* (*Tc-disco*). She showed *Tc-disco* to be expressed in all the developing appendages of the Tribolium embryo, quickly becoming restricted to the medial portion of most appendages. To confirm the role of *Tc*-disco in Tribolium leg development, she used RNA interference (RNAi) to reduce *Tc-disco* expression in both embryos and larvae (the effects of which are seen in adults). In both cases, she discovered that the appendages that developed were truncated and were missing medial portions. This work supplemented other work in the lab, which indicated a role for the *disco* genes in Drosophila appendage development. Together, they suggested that the conserved role for *disco* genes was in insect appendage development.

In preparing the Tribolium work for publication (Patel, Farzana et al. 2007), we sought a molecular marker that would complement the morphological features described by Robertson in her RNAi phenocopies. In addition to finalizing the publication of these data, we also wanted to further explore what other functions might be conserved between the *disco* genes of Drosophila and Tribolium. Our primary question concerned the conservation of the dichotomy between *disco* and *tsh* genes (*Tc-tiotsh* in Tribolium (Shippy, Tomoyasu et al. 2008)). In the Drosophila embryo, Tsh represses *disco* expression, and reduction of *tsh* expression results in expansion of *disco* (Robertson, Bowling et al. 2004). Would the same be true for Tribolium?

In the following sections, I describe my use of the Tribolium *dachshund* (*Tc-dac*) gene as a molecular marker of medial appendage fate in *Tc-disco* RNAi embryos. I also discuss the results of my RNAi knock-down of *Tc-tiotsh* and the effect on *Tc-disco* expression.

MATERIALS AND METHODS

Tribolium Stocks and Culture

Tribolium (GA1 wild-type) were reared at 30 °C on whole-wheat flour supplemented with 5% yeast powder (Berghammer, Bucher et al. 1999).

RNAi Injections

Injections for *Tc-disco* were performed as described for parental RNAi (Bucher, Scholten et al. 2002); for *Tc-tiotsh* adult injections were as in Shippy et al. (2008). Approximately 0.2 µl of dsRNA (0.5 µg/µl to 3 µg/µl) 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. Controls were performed previously with respect to *Tc-disco*, or by Shippy et al. (2008). Embryos were fixed for *in situ* hybridization; digoxigenin-labeled antisense RNAs were prepared and in situ hybridizations were done essentially as in (Tautz and Pfeifle 1989). Embryos for cuticle preps were allowed to age for four days after collection. Larvae and any remaining unhatched embryos were soaked in a mixture of lactic acid and acetic acid. I confirmed that the RNAi was giving the expected phenocopy by examining the cuticle preps in three-dimensions in watch-glass dishes.

RESULTS

Medial leg fate marker *Tc-dac* is reduced in *Tc-disco* RNAi.

The expression of *Tc-dac* has been previously described (Prpic, Wigand et al. 2001). I focused on the developing legs, where expression begins at the base of the forming appendage. As development progresses, two domains become visible, a proximal and a distal domain. The distal domain is in the middle of the developing leg and grows stronger as development progresses.

These features were used to compare with *Tc-dac* expression in *Tc-disco* RNAi (Fig 1). Surprisingly, *Tc-dac* expression appeared normal in the early stages of development. Both domains of expression were present in younger embryos (Fig 1A-B). However, the distal domain in mid-stage RNAi embryos did not appear as strong as in wild type (Fig 1C-D), and seemed very faint in RNAi embryos at the end of germband retraction (Fig 1E-F). A proximal domain remained, and was presumably the proximal domain that formed early in development. The apparent loss of the distal domain of *Tc-dac* supports the finding, using morphological features, that *Tc-disco* is important for medial leg development in Tribolium.

Tc-tiotsh specifies body wall.

Since the appendage role of *disco* genes seemed conserved between Tribolium and Drosophila, I sought to address whether other aspects of *disco* function were also conserved. Specifically, I wanted to address the conservation of its interaction with *tsh*. To do this I sought to replicate the *Tc-tsh* RNAi experiments initiated by our collaborators (Shippy, Tomoyasu et al. 2008). While examining the cuticle preps to ensure the efficacy and

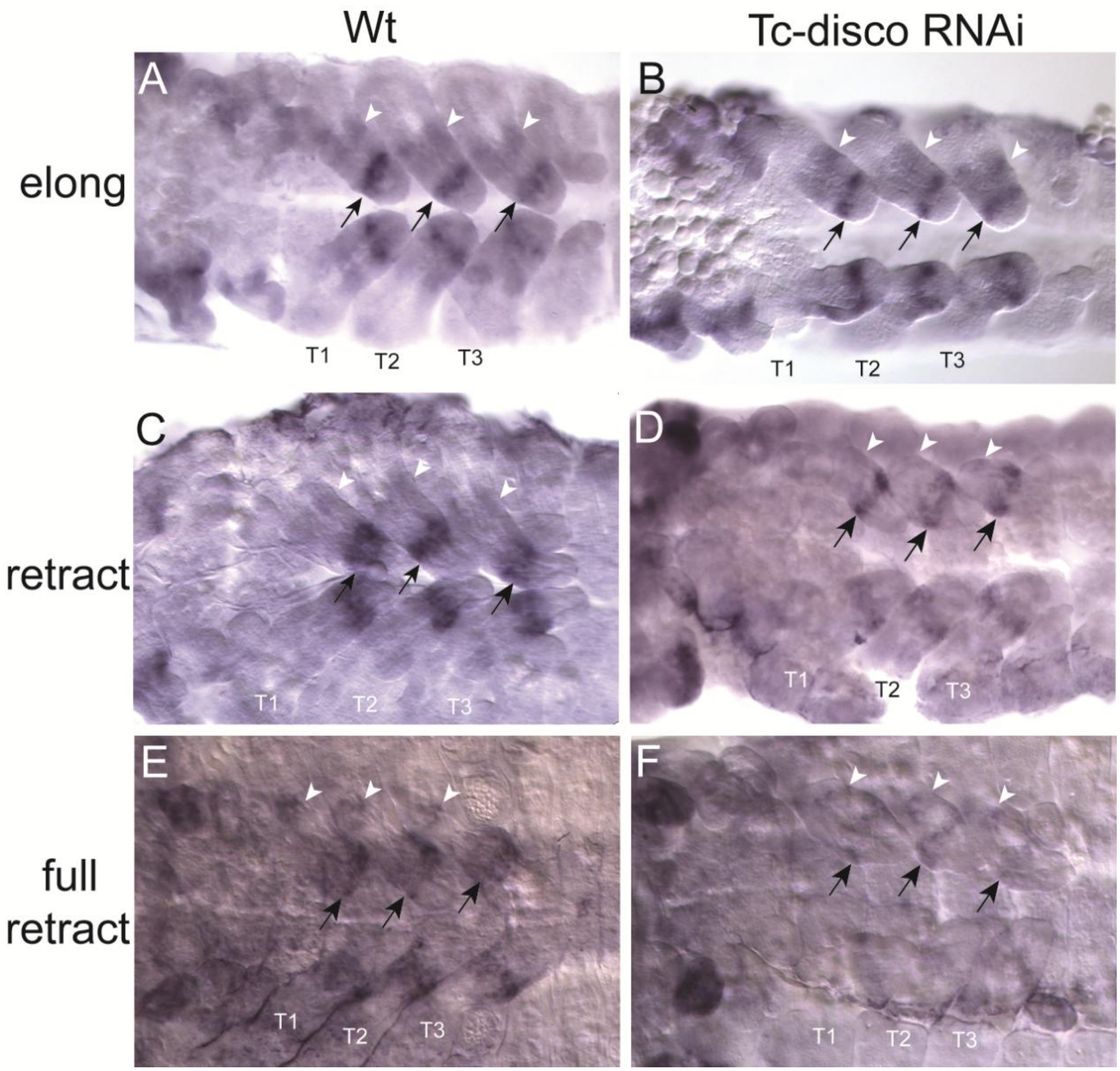


FIGURE 1: Comparison of *Tc-dac* mRNA accumulation in developing legs of wild type and *Tc-disco* RNAi embryos

Wild type embryos (A, C, E) compared to *Tc-disco* RNAi embryos (B, D, F) stained against *Tc-dac*. Arrows indicate the distal domain of *dac* expression and arrowheads point to the weaker, proximal domain. During germband elongation (A-B), expression in RNAi embryos is nearly indistinguishable from that of wild type. By early retraction stages (C-D), the distal domain is not as strong in RNAi embryos, though the proximal domain appears comparable. By the time the embryos have completed retraction (E-F), the distal domain is faint or absent. T#, thoracic segments.

accuracy of the RNAi, I noted a number of unhatched embryos with cuticle tears. These tears occurred most often along the ventral midline (Figure 2), but were also found in dorsal portions of the cuticle. Table 1 provides an overview of numbers and percentages observed. This trait was never observed in either untreated or *Tc-disco* RNAi cuticle preps.

These embryos appeared ready to hatch; their development seemed otherwise complete and their posture indicated preparation for hatching. It seems possible that these embryos tore themselves trying to exert the force required to tear open the egg case. If true, this would provide additional evidence for a role of *Tc-tiotsh* in forming body wall. Such a conclusion would be in agreement with the reports of adult cuticle defects and of limb truncation by reduction/fusion of proximal elements (Shippy, Tomoyasu et al. 2008). In Tribolium, then, *Tc-disco* acts to specify medial appendage while *Tc-tiotsh* acts to specify proximal limb and its connection to the body wall.

Tc-disco expression does not expand in the absence of _Tc-tiotsh_.

Given the conservation of roles of *disco* and *tsh* between Tribolium and Drosophila, it seemed possible that their mutual exclusion might also be conserved. To determine if this is true, I examined the expression of *Tc-disco* in *Tc-tiotsh* RNAi embryos (Fig 3). *Tc-disco* expression did not appear to expand beyond the field of the developing appendages in these embryos. It is possible that *Tc-disco* expression in the appendages is closer to the body of RNAi embryos, but this may be a result of the loss of proximal tissues due to loss of *Tc-tiotsh*.

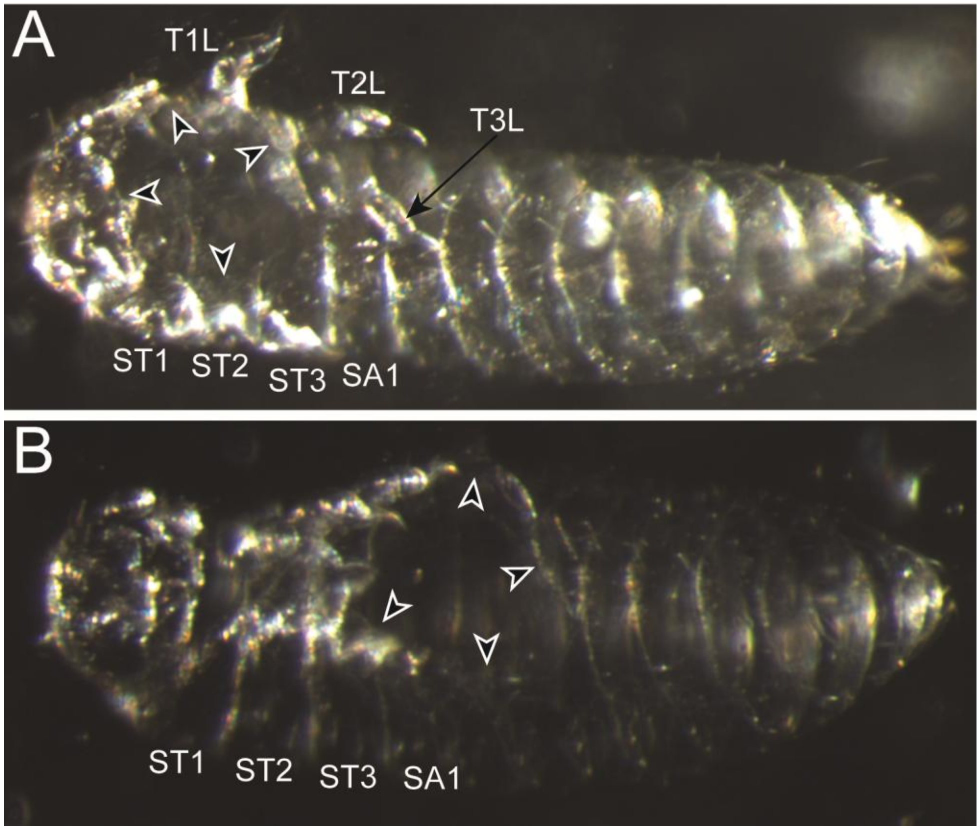


FIGURE 2: Cuticle preperations of *Tc-tiotsh* RNAi embryos showing cuticle tears

A) A tear occurring between the appendages of the throracic and head segments. Legs of one side are visible in the focal plane and are marked by T#L. B) A tear starting in the third thoracic segment and spreading into the abdomen. Arrowheads indicate edges of the tears. The three thoracic segments (ST) as well as the first abdominal segment (SA1) are labeled.

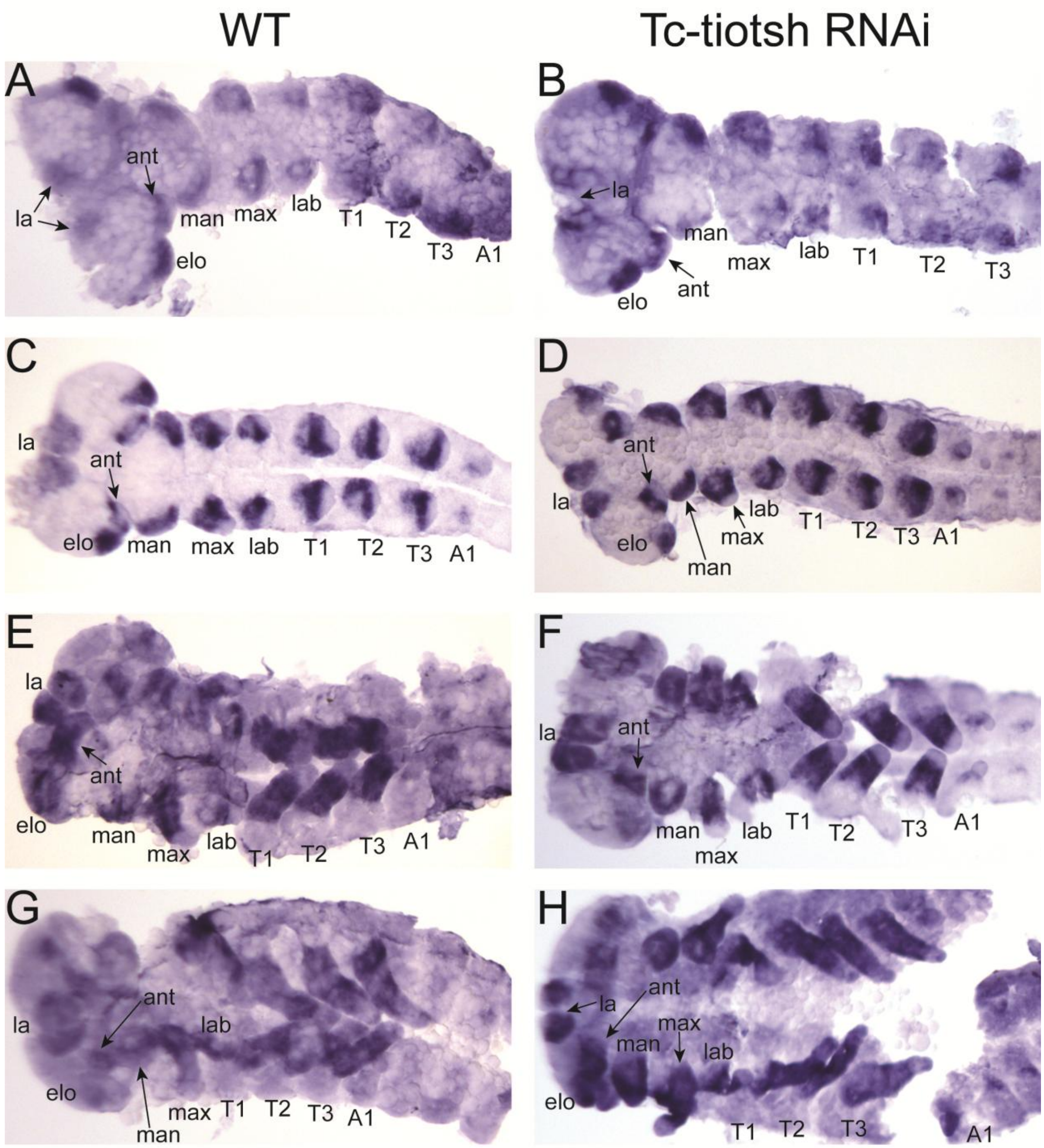
TABLE 1: Number and percentages of embryos with torn cuticles

Numbers represent 13 consecutive days of collection, starting on the second day after injected adults started laying. % Unhatched is calculated as Torn/Unhatched*100. % Total is calculated as Torn/Total*100.

Collection	Torn Cuticle	Unhatched	% Unhatched	Total	%Total
1	4	33	12.12	50	8.00
2	9	29	31.03	46	19.57
3	2	26	7.69	43	4.65
4	6	26	23.08	46	13.04
5	15	52	28.85	95	15.79
6	9	38	23.68	84	10.71
7	12	37	32.43	66	18.18
8	23	92	25.00	168	13.69
9	15	52	28.85	100	15.00
10	2	13	15.38	29	6.90
11	5	8	62.50	26	19.23
12	4	9	44.44	28	14.29
13	5	7	71.43	24	20.83
Total	111	422	26.30	805	13.79

FIGURE 3: Comparison of *Tc-disco* mRNA accumulation in developing legs of wild type and *Tc-tiotsh* RNAi embryos

Wild type embryos (A, C, E, G) compared to *Tc-tiotsh* RNAi embryos (B, D, F, H) stained against *Tc-disco*. Comparable stages are paired horizontally (A and B, C and D, E and F, G and H), with age increasing from top to bottom (A and B being the youngest, G and H the oldest). Expression of *disco* mRNA does not appear to expand beyond the developing appendages. Abbreviations: elo, eyelobe; la, labrum; ant, antenna; mn, mandible; mx, maxillary; lab, labial; T1–3, thoracic; A1 first abdominal.



WT
Tc-tiotsh RNAi
A
la
ant
man
max
lab
T1
T2
T3
A1
elo
B
la
man
max
lab
T1
T2
T3
ant
elo
C
la
ant
elo
man
max
lab
T1
T2
T3
A1
D
ant
la
lab
max
man
T1
T2
T3
A1
elo
E
la
ant
elo
man
max
lab
T1
T2
T3
A1
F
la
ant
man
max
lab
T1
T2
T3
A1
G
la
ant
lab
elo
max
man
T1
T2
T3
A1
H
la
ant
man
max
lab
elo
T1
T2
T3
A1

DISCUSSION

This work provides further evidence for the conserved roles of *disco* and *tsh* genes as appendage and body wall factors, respectively. Using *Tc-dac* as a marker for medial appendage, I was able to confirm that reduction of *Tc-disco* in the developing embryo results in the loss of the medial portions of the appendage. My examination of *Tc-tiotsh* knock-down larvae revealed a high incidence of torn cuticles. At the same time, I note that there does not appear to be conservation of the negative regulation of *disco* by *ts*h.

In Tribolium, *Tc-disco* appears to be necessary for the maintenance of *Tc-dac* expression in the developing leg. In *Tc-disco* RNAi embryos, the distal domain of *Tc-dac* appears normal in early developmental stages, but is reduced in later stages of development. The resulting legs are truncated, missing medial portions. This mimics the results of clonal analyses performed in Drosophila, where clones lacking disco are absent in the medial and distal portions of legs (Patel, Farzana et al. 2007). What is not clear is what is happening to the cells that initially express *Tc-dac*. It is possible that *Tc-disco* is required for the maintenance of *Tc-dac* expression, but not for its initiation. In this case, cells eventually stop expressing *Tc-dac*. However, this alone does not account for the truncation of appendages resulting from *Tc-disco* RNAi. It seems likely that *Tc-disco* is also necessary for the proliferation or survival of those cells. These are not mutually exclusive possibilities, and there is evidence in other organisms supporting both proliferation (Tseng and Green 1992; Tseng 1998; Vanhoutteghem, Maciejewski-Duval et al. 2009) and cell survival (Lang, Patterson et al. 2009). It would be intriguing to distinguish between these possibilities, particularly with respect to the potential role of disco genes in mammalian cancers (Cui, Elsam et al. 2004; Shames, Girard et al. 2006; Akagi, Ito et al. 2009; Dunwell, Hesson et al. 2009; Song, Ramus et al. 2009). Teasing apart the role of disco genes in cell survival and proliferation could offer information useful in pursuing reliable diagnostics and therapeutics. One option, then, could be to monitor cell death in Tribolium, which might provide clues about what is happening to *Tc-dac* expressing cells. In conjunction with clonal analyses in Drosophila, such a study might offer further hints as to the broader conserved function of genes in the *disco* family.

I have also reported additional evidence supporting a role of *Tc-tiotsh* in proper development of the body wall. Specifically, since I examined unhatched embryos in a watchglass rather than flattened on a slide, I was able to identify cuticle tears in many of the affected progeny. I believe these tears indicate in important role for *Tc-tiotsh* in larval cuticle development. This finding is complimentary to the results of *Tc-tiotsh* RNAi in developing larvae and adults reported by Shippy and colleagues (2008). In larva, they found truncation of the appendages by reductions and fusions of proximal portions. In adults, they observed improper development of the cuticle in addition to defects of the proximal leg. Altogether, these experiments demonstrate a role for the Tribolium *tsh* gene in body wall formation and support the conclusion that *tsh* genes have a conserved role in the development of body wall in insects. However, if the respective conserved roles of *disco* and *tsh* genes are in appendage and body wall development, we must explain why Drosophila larvae appear to have evolved their respective head and trunk roles for these genes.

Patel et al. (2007) suggest that the apparent anterior/posterior roles of *tsh* and *disco* in the developing Drosophila embryo are artifacts of the shift made by dipterans from larvae with limbs towards the vermiform maggot larva. Because Drosophila larvae have evolved to move through rotting fruit, appendages were unnecessary, and perhaps even harmful. So, evolution favored those with internalized mouth-parts and reduced trunk appendages. *disco* mainly seems to specify head, but is, in truth, necessary for the development of the limb-derived larval mouth parts. This derivation required the reduction of body wall for head involution, the process by which the larval mouth parts are internalized. At the same time, appendage development was restricted in the remaining body by limiting *disco* expression in this region, and body wall is promoted via *tsh*. In adult development, these appendage roles are more obvious and more closely resemble the orthologous roles in Tribolium.

In spite of this conservation, some changes are expected. One possible change is the role of *tsh* repressing *disco* expression in the developing larval trunk of Drosophila. *Tc-disco* expression did not seem to expand beyond the limb buds, in spite of the reduction of *Tc-tiotsh* expression. Indeed, it was difficult to distinguish *Tc-tiotsh* RNAi embryos from non-treated embryos based on *Tc-disco* expression. Since this study was conducted using RNAi, it is possible that there was sufficient *Tc-tiotsh* expression to block expansion of *Tc-disco* expression. However, given the variety of RNAi phenotypes, it would seem likely that some embryos would possess near-complete knockdown of the target. In such a case, I would expect *Tc-disco* expression to expand if indeed it were controlled by *Tc-tiotsh*. It is also possible that expansion occurred in the developing appendages, but did not stray into the body proper. Since *Tc-tiotsh* RNAi results in loss or reduction of the proximal appendage regions, it would be difficult to tell if such expansion occurred.

It is intriguing to consider that repression of *disco* by Tsh is derived in Drosophila, especially given the conservation of their limb development roles. This repression might be the result of the shift towards the maggot larval form. In the evolutionary change, it may have become necessary to acquire additional restrictions to *disco* expression, and *tsh* was recruited for this role. In order to confirm that this change in *disco/tsh* interactions is real, and to determine its significance, we would need to address a few questions. It would be important to know when, evolutionarily, this change took place. We would also need to determine whether this change was unique to Drosophila or if the change occurred early in Dipteran evolution. Given that this comparison has only been made in two species, it is also possible that this repression role was lost in the Tribolium lineage rather than gained in Drosophila. Other insect species would need to be examined and compared before conclusions can be drawn.

Conclusion

The preceding data support a conserved role for *disco* and *tsh* genes in specifying appendage and body wall identity, respectively. However, the conservation of their roles and interactions does not seem to be complete. One change appears have occurred by the addition of a role for Tsh repressing *disco* expression in the developing larval trunk of Drosophila. Another alteration that seems to have occurred is the expansion of the *disco* genes' expression to encompass much of the leg imaginal discs of Drosophila, as well as a corresponding shift in their role towards specifying both distal and medial portions of adult legs (Fig 4). As we have seen in Tribolium, the expression and RNAi effect of *Tc-disco* was limited to the medial portion of appendages (Fig 4). However, these minor evolutionary changes should not overshadow the importance of the conserved roles that have been observed. To further confirm and dissect these roles, it would be helpful to expand our understanding of these genes and their orthologs beyond these two species, especially moving beyond insects. Such studies will help illumine the functions of *disco* genes that have remained conserved throughout history while helping address the reasons behind changes in roles seen between species.

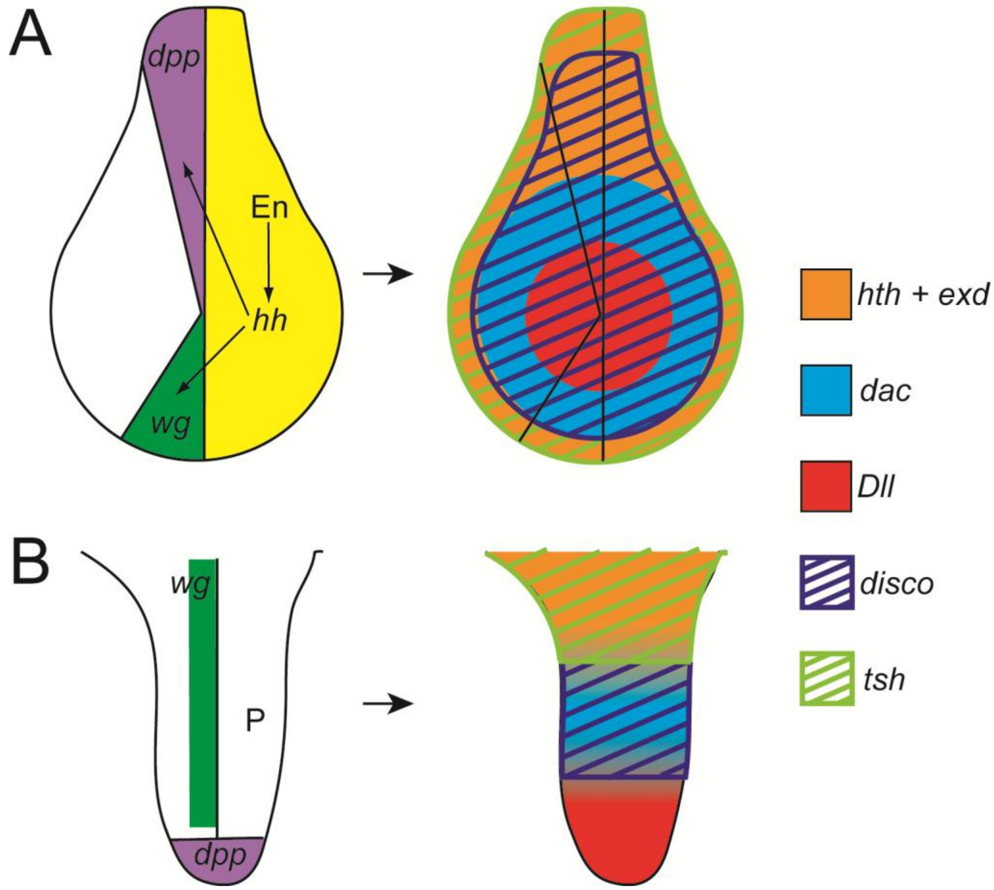


FIGURE 4: Comparison of models of leg patterning in Drosophila and Tribolium

A) (Left) In Drosophila, leg discs are patterned when En expression in the posterior portion of the disc activates Hh signaling (yellow), which in turn activates *dpp* (purple) dorsally and *wg* (green) ventrally, in the anterior of the disc. (Right) These morphogens then pattern the concentric rings of *Dll* (red), *dac* (blue), and *hth/exd* (orange). Expression of *disco* (dark blue stripes) is also present in much of the disc, while *tsh* (lime stripes) is limited to the outer edge, which will form the proximal appendage. B) (Left) Ventral view of Tribolium limb bud. Expression of *dpp* (purple) is at the leg tip, while *wg* (green) is expressed in an anterior, ventral stripe. P = posterior. Proximal at top. (Right) The expression patterns of *Dll* (red), *dac* (blue), and *hth/exd* (orange) resemble their Drosophila orthologs. *tsh* (lime stripes), as in Drosophila, is limited to the proximal portion, but *disco* (dark blue stripes) is only expressed in the middle leg domain.

References:

Akagi, T., T. Ito, et al. (2009). "Chromosomal abnormalities and novel disease-related regions in progression from Barrett's esophagus to esophageal adenocarcinoma." Int J Cancer **125**(10): 2349-59.

Berghammer, A., G. Bucher, et al. (1999). "A system to efficiently maintain embryonic lethal mutations in the flour beetle Tribolium castaneum." Dev Genes Evol **209**(6): 382-9.

Bucher, G., J. Scholten, et al. (2002). "Parental RNAi in Tribolium (Coleoptera)." Curr Biol **12**(3): R85-6.

Cui, C., T. Elsam, et al. (2004). "Gli proteins up-regulate the expression of basonuclin in Basal cell carcinoma." Cancer Res **64**(16): 5651-8.

Dunwell, T. L., L. B. Hesson, et al. (2009). "Epigenetic analysis of childhood acute lymphoblastic leukemia." Epigenetics **4**(3): 185-93.

Fasano, L., L. Roder, et al. (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**(1): 63-79.

Heilig, J. S., M. Freeman, et al. (1991). "Isolation and characterization of the disconnected gene of Drosophila melanogaster." EMBO J **10**(4): 809-15.

Lang, M. R., L. B. Patterson, et al. (2009). "Basonuclin-2 requirements for zebrafish adult pigment pattern development and female fertility." PLoS Genet **5**(11): e1000744.

Lee, K. J., M. Mukhopadhyay, et al. (1999). "Autoregulation of the Drosophila disconnected gene in the developing visual system." Dev Biol **214**(2): 385-98.

Mahaffey, J. W., C. M. Griswold, et al. (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**(1): 225-36.

Patel, M., L. Farzana, et al. (2007). "The appendage role of insect disco genes and possible implications on the evolution of the maggot larval form." Dev Biol **309**(1): 56-69.

Prpic, N. M., B. Wigand, et al. (2001). "Expression of dachshund in wild-type and Distal-less mutant Tribolium corroborates serial homologies in insect appendages." Dev Genes Evol **211**(10): 467-77.

Robertson, L. K., D. B. Bowling, et al. (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**(12): 2781-9.

Shames, D. S., L. Girard, et al. (2006). "A genome-wide screen for promoter methylation in lung cancer identifies novel methylation markers for multiple malignancies." PLoS Med **3**(12): e486.

Shippy, T. D., Y. Tomoyasu, et al. (2008). "Do teashirt family genes specify trunk identity? Insights from the single tiptop/teashirt homolog of Tribolium castaneum." Dev Genes Evol **218**(3-4): 141-52.

Song, H., S. J. Ramus, et al. (2009). "A genome-wide association study identifies a new ovarian cancer susceptibility locus on 9p22.2." Nat Genet **41**(9): 996-1000.

Steller, H., K. F. Fischbach, et al. (1987). "Disconnected: a locus required for neuronal pathway formation in the visual system of Drosophila." Cell **50**(7): 1139-53.

Tautz, D. and C. Pfeifle (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**(2): 81-5.

Tseng, H. (1998). "Basonuclin, a zinc finger protein associated with epithelial expansion and proliferation." Front Biosci **3**: D985-8.

Tseng, H. and H. Green (1992). "Basonuclin: a keratinocyte protein with multiple paired zinc fingers." Proc Natl Acad Sci U S A **89**(21): 10311-5.

Vanhoutteghem, A., A. Maciejewski-Duval, et al. (2009). "Basonuclin 2 has a function in the multiplication of embryonic craniofacial mesenchymal cells and is orthologous to disco proteins." Proc Natl Acad Sci U S A **106**(34): 14432-7.

CHAPTER THREE

Expression Studies of Potential *disconnected* Orthologs in Multiple Species

INTRODUCTION

Some of the most important questions in genetic evolution center on questions of conservation. Conservation of genetic sequence is used to support proximity of inter-species relationships, or to demonstrate the functional importance of a particular gene product. Similarly, conservation in protein sequence is used to exemplify the importance of functional domains, while conservation of protein functions and interactions serves as evidence for the homology of tissues or organs between species (Shubin, Tabin et al. 1997; Shubin, Tabin et al. 2009). However, this information can be misleading. The information used to reach the conclusion of conservation might be inadequate, such as too small a sample or too short a sequence. An improper interpretation of conservation can lead to inaccurate interpretations of the relationships above (see, for example, the use of beta-thymosin in the debate over the placement of *C. elegans* in ecdysozoa (Manuel, Kruse et al. 2000; Telford 2004)). Therefore, to ensure proper understanding of evolution, by way of a proper understanding of conservation, information should be expanded by exploring sequences and gene networks in new species, as well as improving knowledge of well studied species. Until recently, one particular family of transcription factors has been explored very little outside of the genetic model organism *Drosophila melanogaster*. These transcription factors are C_2H_2 zinc-finger proteins related to the Drosophila protein Disconnected (Disco), and the vertebrate orthologs, the Basonuclin (BNC) proteins (Tseng and Green 1992).

Disco, a C_2H_2 zinc finger transcription factor (Heilig, Freeman et al. 1991; Lee, Mukhopadhyay et al. 1999), was first discovered as being important for proper migration of the Bolwig nerve during embryonic development in Drosophila (Steller, Fischbach et al.

1987). Later, along with its paralog, Disco-related (Disco-r), it was implicated in proper head development of the Drosophila embryo (Mahaffey, Griswold et al. 2001; Robertson, Bowling et al. 2004). This established the *disco* genes in a dichotomous role opposite another zinc-finger transcription factor, *teashirt* (*tsh*), where *tsh* was responsible for proper trunk development (Fasano, Roder et al. 1991; Robertson, Bowling et al. 2004; Patel, Farzana et al. 2007; Shippy, Tomoyasu et al. 2008). Most recently, the *disco* genes were identified as important to adult limb development, resolving into an exclusive domain with *tsh*, where *tsh* is expressed in the proximal portions of leg imaginal discs while the *disco* genes were expressed in the medial and distal portions of these discs (Wu and Cohen 2000; Patel, Farzana et al. 2007). This work created new questions about the role of the *disco* genes in development. Examination of the embryonic development of the flour beetle, *Tribolium castaneum*, shed some light on the subject: reduction of *Tc-disco* mRNA by RNAi revealed that it is important for (medial) limb development of all ventral appendages, but not in head development specifically. This led to the conclusion that the head role of the *disco* genes is only an artifact of the highly derived development of the Drosophila larva, and that the conserved role of the *disco* genes is in limb development (Patel, Farzana et al. 2007).

Indeed, this conclusion is not surprising when regarding studies of the *Teashirt* gene family. *Teashirt* related genes have been associated with limb development in a wide range of species, including Tribolium (Shippy, Tomoyasu et al. 2008), the milkweed bug, *Oncopeltus fasciatus* (Herke, Serio et al. 2005), the zebrafish, *Danio rerio* (Wang, Lee et al. 2007), and the mouse, *Mus musculus* (Long, Park et al. 2001; Caubit, Tiveron et al. 2005; Pueyo and Couso 2005).

BNC1 and BNC2, on the other hand, have been found to be important in epithelial development in humans and mice. Epithelial expression has been found in a variety of tissues from the digestive tract to the corneal surface (Tseng and Green 1992; Teumer, Tseng et al. 1997; Tseng 1998; Vanhoutteghem and Djian 2004). Specifically, it has been suggested that BNC1 may play a role in the proliferation of karatinocytes (Tseng and Green 1992; Tseng 1998). However, there is evidence of BNC expression in a great variety of tissues not necessarily related to epithelium or proliferation, like the heart and brain (Tseng 1998). Exploration of these areas seems limited.

The *disco* genes are useful for evolutionary study because they are present in a great variety of species and they interact with the networks of the developmentally and evolutionarily important Hox genes (Mahaffey, Griswold et al. 2001; Robertson, Dey et al. 2002; Patel, Farzana et al. 2007). However, a complete understanding of the role these genes have played in the evolution of development is important to piecing together the functions of these genes in these species. Unfortunately, information from vertebrates and other species outside of hexapodia is limited, so it is difficult to make accurate conclusions about conserved function. Specifically, we would like to know if the role in limb development is conserved outside of insects. If not, we would want to discover a conserved role that can account for the activities of *disco* genes throughout the evolution of development. To expand our limited knowledge, we attempted to explore the function of *disco* genes in two vertebrate species, zebra fish and mice. We also sought to understand the role of a sequence in the limbless nematode species, *Caenorhabditis elegans*, which appears to code for a Disco-like protein.

MATERIALS AND METHODS

C. elegans strains and RNAi

A translational reporter GFP-construct was previously assembled by JM Mahaffey, containing the majority of the *Ce-disco* coding sequence and 2419 bp upstream from the predicted ATG. This construct was injected into N2 worms along with the co-injection marker *rol-6(su1006)* (Boulin, Etchberger et al. 2006). Three deficiency strains were obtained from the Japanese National BioResource Project, which are missing some portion of the presumptive Ce-disco gene. Deficiency tm3030 removes 52 bp upstream of translation start site through 225 bp into the coding sequence, deleting the first exon and most of the second; deficiency tm3053 removes bases 110 through 318 of the coding sequence, deleting most of the second exon; deficiency tm3087 removes bases 23 through 349 of the coding sequence, deleting most of the first exon and all of the second.

RNAi experiments took advantage of the RNAi-hypersensitive strain VH624 (Schmitz, Kinge et al. 2007). Constructs for making dsRNA used an approximately 500bp-long sequence from cDNA template of *Ce-disco*. Worms were either fed dsRNA-expressing bacteria or injected with dsRNA, essentially as in (Ahringer 2006). Worms to be fed were placed on plates with the bacteria and allowed to lay progeny, which were maintained on those plates. The original parent and their progeny were examined for phenocopy. In addition, we injected greater than 20 worms with dsRNA at concentrations greater than 500ng/µL. Progeny were collected for 24 hrs following a 24hr recovery period. Distilled water served as a negative control while *unc-119* served as a positive control for the efficacy of the hypersensitive strain.

Cloning zebra fish *bnc* genes and *in situ* hybridization process

Zebra fish embryos were obtained from the lab of Dr. Jon Horowitz at NCSU. Coding sequence from each zebra fish *bnc* gene was cloned using the following primer pairs: *bnc1* GTCCAGGACGTGGCTGGTAAG/AAACGCAGCATTGCAGCCGTC and *bnc2* GGCAAGGTTCTGGACAGATGG/GAAAGCAGCATTGCATCCAGA. The *in situ* hybridization was performed essentially as in (Thisse, Thisse et al. 1993).

Cloning mouse *bnc* genes and *in situ* hybridization process

Mouse embryos were obtained with the help of Dr. David Threadgill's lab at NCSU. Coding sequence from each mouse *bnc* gene was cloned previously by Mukund Patel using the following primer pairs: TTCGCTTTGGAGAGACCAA/TGTGTGTTCTTCTCATGCGT for *bnc1* and TGTCTCGGGAAGAAGAAATCA/TGCAGACGTGCATCTCTTTTA for *bnc2*. The *in situ* hybridization was performed essentially as in (Wilkinson 1992).

RESULTS

Analysis of expression and function of a possible *disco* gene in *Caenorhabditis elegans*

BLAST against the completed *C. elegans* genome revealed a potential *disco* ortholog in *C. elegans*, WormBase predicted gene F55C5.11, which I will refer to as *Ce-disco*. A GFP expression construct showed GFP to be expressed in the nucleus of cells that appeared to be part of the *C. elegans* neural network (Fig 1B). Comparison with constructs that express in motor neurons suggested that these cells could have been motor neurons (Fig 1A).

Since expression suggested a neural role, I sought to determine if there were any behavioral or movement phenotypes in worms deficient for *Ce-disco*. I was able to obtain and examine three deficiency strains that remove some portion of the first two exons (Fig 2). These worms behaved normally and did not exhibit abnormalities in locomotion. However, these deficiencies did not affect the third exon, which would encode for the zinc-fingers of the protein, so it is possible that these deficiencies were not true nulls. Therefore I attempted to use RNA interference against this third exon to ensure a more effective reduction of expression. Since neuronal RNAi is limited in *C. elegans*, we used the RNAi-hypersensitive strain VH624 (Schmitz, Kinge et al. 2007). I did not observe any effect from the RNAi using either feeding or injection techniques to knock down *Ce-disco*.

RNAi and mutant analyses seem to indicate that there is no function for the putative *Ce-disco* under our laboratory conditions. However, since RNAi even with this worm strain is not always completely penetrant, it is also possible that I was unsuccessful in actually stopping or reducing gene function. I cannot at this time distinguish between these two possibilities.

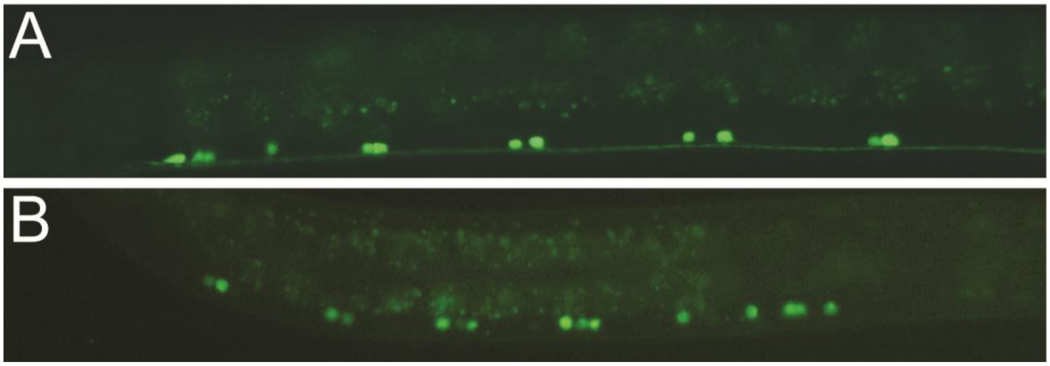


FIGURE 1: GFP expression driven by *Ce-disco*

Ce-disco expression appears to be similar to expression of known neuronal gene. A: expression of 31E strain from Miller lab, showing pattern of expression in some motor neurons of L4 larva. B: *Ce-disco*-GFP expression in adult. Dorsal at top.

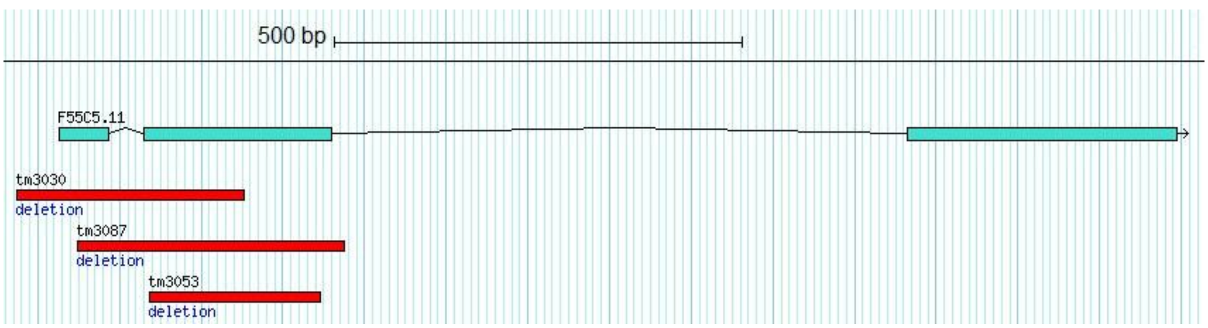


FIGURE 2: Location of deletions relative to C. elegans *disco*-like gene F55C5.11

Exons are in bright blue, deletions are in red. Image modified from wormbase.org.

BNC expression in the zebra fish (*Danio rerio*)

If the *disco* genes are true appendage development genes, then I would expect to find expression of the *disco* orthologs, the *basonuclin* genes, in vertebrate appendages. The zebra fish pectoral fin seemed to be a readily accessible and analogous model in which to test this hypothesis. I focused my attention on stages Prim 5 to Long Pec. The earliest signs of the pectoral fin formation can be seen in Prim 5, while the bud takes shape in the immediately-following Prim 15 stage. The bud continues to elongate and develop in the two consecutive stages, High Pec and Long Pec. As can be seen in Figures 3 and 4, both *bnc* genes are expressed in several tissues. Both are found in various parts of the eye and brain, as well as possible expression in the branchial arches.

In spite of these extensive and dynamic patterns, I was unable to identify any expression of either gene in the developing pectoral fins. However, the relationship between zebra fish fins and land animal limbs makes zebra fish hard to compare to terrestrial vertebrates when studying limb development. Specifically, the anatomical origins of land animal hands and feet (autopodia) appear to have been lost in zebra fish, not gained in autopods (Dahn, Davis et al. 2007; Davis, Dahn et al. 2007). Given the possibility that zebra fish fins represent a reduction at the cost of distal features, it seems likely that the zebra fish is not a good model for examination of appendage development, and it is perhaps not so surprising that I found no *bnc* expression in the developing fins. Another group has since identified *bnc2* expression in zebra fish fins, but since this expression was not identified until the larva-to-adult transformation, it is questionable that this expression has any bearing on fin development beyond a role in pigmentation (Lang, Patterson et al. 2009).

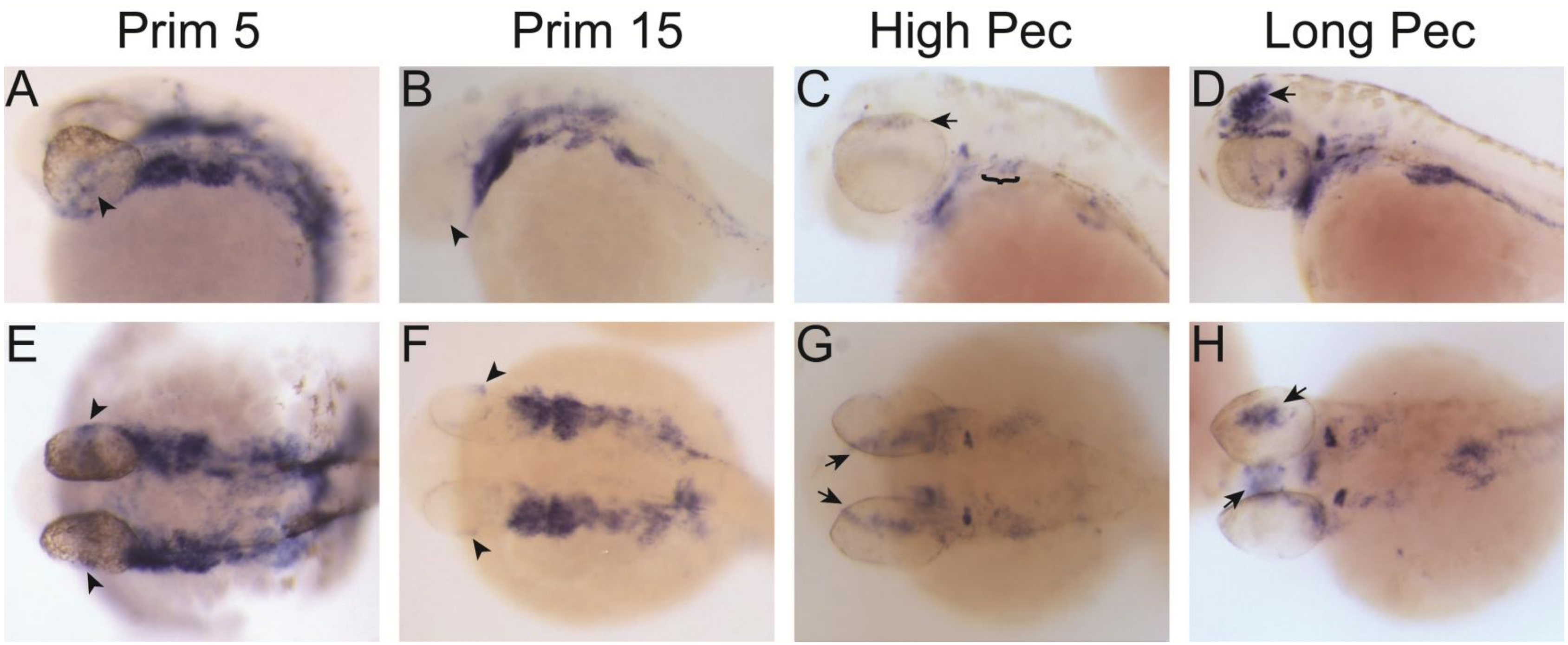


FIGURE 3: Expression of *bnc1* in the zebra fish embryo

bnc1 expression can be found in the eye (arrowheads), brain (arrows) and potential branchial arches (brackets) of developing embryos. A-D show lateral views and E-H show dorsal views of the same embryos, respectively. Stage is indicated at the top.

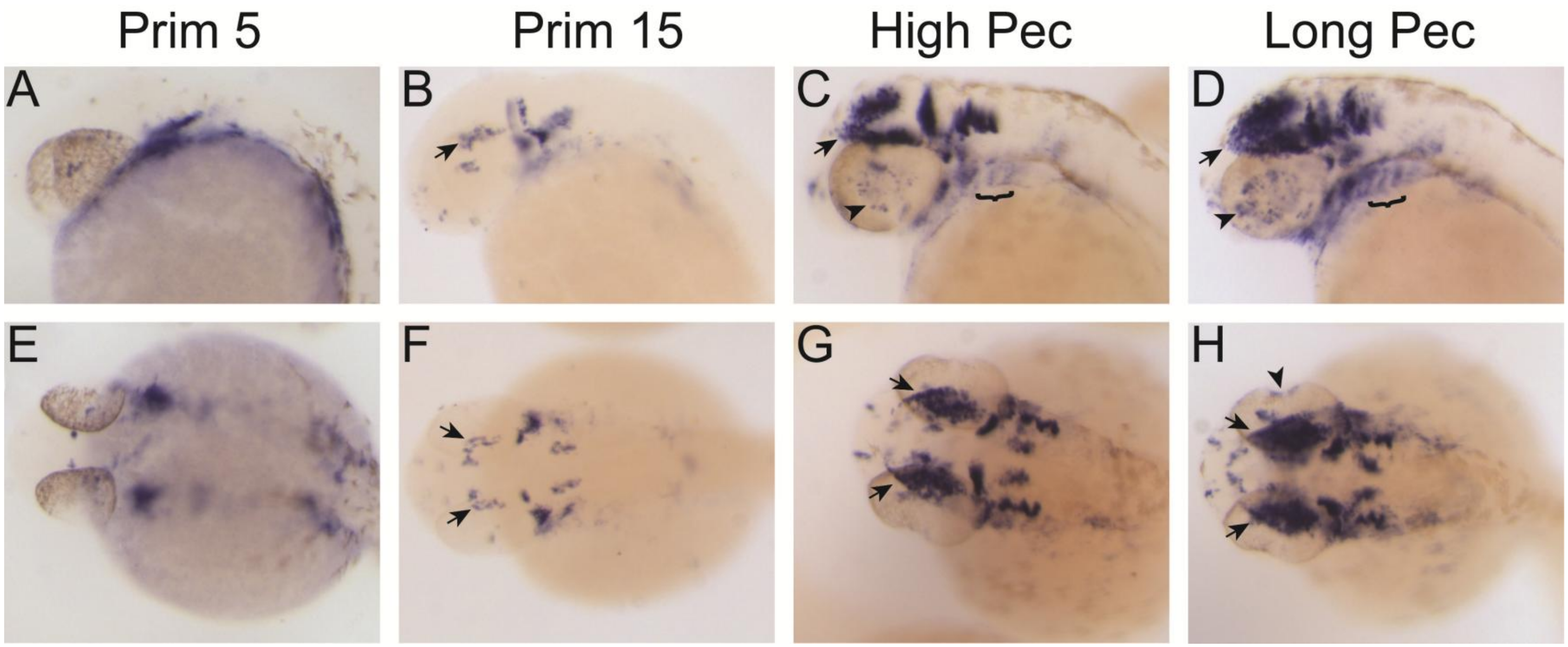


FIGURE 4: Expression of *bnc2* in the zebra fish embryo

bnc2 expression can be found in the eye (arrowheads), brain (arrows) and potential branchial arches (brackets) of developing embryos. A-D show lateral views and E-H show dorsal views of the same embryos, respectively. Stage is indicated at the top.

BNC expression in the mouse (*Mus musculus*)

Since the nature of the pectoral fin as an appendage model is in question, I turned my attention to another vertebrate model, the mouse. No group had previously published data showing whole-mount embryonic expression of the murine basonuclin genes. In order to determine if there is a conserved role for *disco* gene family members in vertebrate limb development, I sought to examine expression of the BNC genes in mice. We collected embryos ranging 9.5 (when the limb bud begins forming) to 13.5 days post conception, and stained them with probes against mRNA of BNC1 or BNC2. However, I was unable to produce reliable staining in these embryos. Because other groups have shown BNC expression in several tissues (see below), I suspect that the technique was not properly performed.

DISCUSSION

The preceding work offers few clues to the conserved function of *disco* genes. However, recent studies in mice and zebra fish may help shed some light on the evolutionary history of these genes.

Mouse *bnc* genes are involved in proliferation

In 1992, the first vertebrate *basonuclin* gene was identified as a gene expressed in cultured human epidermal keratinocytes (Tseng and Green 1992), and has been linked to epidermal proliferation in mice (Tseng and Green 1992; Tseng 1998; Zhang and Tseng 2007). Most notably, a mouse line with a null mutation for *basonuclin1* was constructed (Zhang and

Tseng 2007). Though infertile, these mutants were viable. These mice also possessed thinner corneal epithelium and delayed healing of that tissue. These effects were the result of reduced and delayed cell proliferation.

The second *basonuclin* gene was identified independently by two separate groups (Romano, Li et al. 2004; Vanhoutteghem and Djian 2004). BNC2 was shown not only to be closely related to BNC1, but to also possess extreme conservation across species, including insect *disco* genes (Romano, Li et al. 2004; Vanhoutteghem and Djian 2004; Vanhoutteghem and Djian 2006; Vanhoutteghem, Bouche et al. 2011). Function of BNC2 in mouse development was mostly speculative, based on similarities with BNC1, until a null mutation was constructed (Vanhoutteghem, Maciejewski-Duval et al. 2009). Homozygous mutants died within 24 hours of birth, and showed extensive cranio-facial abnormalities, including cleft palates and over-all smaller heads, resulting from improper proliferation of cells contributing to these features. The authors interpreted these results as a head-specifying function, and offered them as proof of the homology between the *bnc* genes and the 'head-specifying' *disco* genes (Vanhoutteghem, Maciejewski-Duval et al. 2009; Vanhoutteghem, Bouche et al. 2011). However, this interpretation fails to fully account for more recent research in Drosophila and Tribolium that suggests the *disco* genes of these species function in appendage development rather than development of the head proper (Patel, Farzana et al. 2007).

Earlier research had indicated that the *disco* genes were important in the development of the Drosophila larval head (Mahaffey, Griswold et al. 2001; Robertson, Bowling et al. 2004). However, Patel et al (2007) showed that the single Tribolium *disco* gene primarily plays a role in limb development, not of the head proper. They also demonstrated that the Drosophila *disco* genes were important for adult ventral limb development (antennae, mouthparts and legs). So they suggest that the presumptive head-development role of *disco* was an artifact of the derived form of the Drosophila larva.

The Drosophila larva has a number of features that distinguish it from many insect larval forms. Most insect larvae develop with several head and thoracic appendages (antennae, mouthparts and legs). However, Drosophila larvae develop without overt appendages. Further, the embryonic head segments internalize before development is completed. These two features in particular allow the larvae of the fruit fly to live in their preferred environment of aging and rotten fruit. The absence of external appendages, achieved through reduction of thoracic appendages and internalization of mouthparts, facilitates easier movement through this medium. Patel and company suggest that the function of *disco* in the developing head is to specify the formation of the appendage-derived mouthparts that will eventually be internalized. So they conclude that even in the developing Drosophila larva, the role of the *disco* genes is appendage development.

In identifying conserved roles between *disco* and *bnc* genes, Vanhoutteghem and company (2009, 2011) also ignore a number of other results. One is the *bnc2*-driven expression of lacZ in the joints of the developing mouse limb, suggesting some limb development role in mice. Although their *bnc2* mutant did not show limb defects, it is possible there is some redundancy in function with *bnc1*. Further, the craniofacial features that are affected in their mutants seem to be branchial arch derivatives. Branchial arches, though not appendages, have been shown to employ aspects of the limb development program (Panganiban, Irvine et al. 1997; Shubin, Tabin et al. 1997; Shubin, Tabin et al. 2009), and such co-option of this program for non-limb outgrowths is not unheard of (Moczek 2006). My own observations of the expression of bnc genes in zebra fish suggests these genes may be expressed in the branchial arches of this species (Fig 3, 4). Further, the *distalless* gene family has been shown to be important in limb development in both Drosophila and mice (in which the genes are referred to as *Dlx*). However, various combinations of *Dlx* mutations result not only in appendage malformations, but also skull and craniofacial malformations (reviewed in (Kraus and Lufkin 2006)) similar to those described for *bnc2*. So, it is possible that the head development role claimed by Vanhoutteghem is really more reflective of appendage development than they credit. Double-knockouts of both *bnc* genes might be necessary to reveal an appendage development role.

Zebra fish studies suggest a conserved neural role for *bnc* genes

In 2009, another group published results of *bnc2* mutants, this time in the zebra fish (Lang, Patterson et al. 2009). While screening for adult pigmentation mutants, they identified a mutation they called *bonaparte*. This mutation was allelic to another pigment mutant, *poppy*, and both alleles cause female infertility. The *bonaparte* allele was mapped to a mutation in *bnc2*. Interestingly, the color pattern of larval fish was not affected. Instead, chromatophores which were meant to contribute to adult coloration were lost during pigment pattern metamorphosis. It should be noted that these pigment cells, while not directly decended from neural crest cells, are nevertheless, descendents of that lineage.

Lang and company also examined the expression of the two zebra fish *bnc* genes. Both genes were expressed in a variety of tissues, including the eye, the dorsal neural tube and the brain (Lang, Patterson et al. 2009). Similar sites of expression have been claimed for the mouse (Vanhoutteghem, Bouche et al. 2011). Indeed, a number of appendage genes have also been observed in the devoping nervous system of insects (Angelini and Kaufman 2004; Blagburn 2007), including *disco* itself (Lee, Freeman et al. 1991). This trend is not limited to insects, either. Similar discoveres have been made in other invertebrates (Panganiban, Irvine et al. 1997; Williams, Nulsen et al. 2002). Vertebrate orthologs to known appendage development genes have also been shown to be expressed in the developing nervous systems of zebra fish (Akimenko, Ekker et al. 1994; Popperl, Rikhof et al. 2000; Waskiewicz, Rikhof et al. 2001; Hammond, Hill et al. 2002), mice (Merlo, Zerega et al. 2000; Kraus and Lufkin 2006) and frogs (*Xenopus laevis*) (Dirksen, Mathers et al. 1993). Finally, the putative *Ce-disco* also appears to be expressed in the worm nervous system (Fig 1B).

It has been proposed that limbs originated from nervous system outgrowths (Panganiban, Irvine et al. 1997; Pueyo and Couso 2005), and the expression of appendage factors in the nervous system certainly lends credence to this model. So it should not be surprising to find an appendage development gene being expressed in the nervous system, as has been discovered in the mouse and zebrafish for the *disco* orthologs, the *bnc* genes. However, the *bnc2* mutant fish were not reported to show neuronal phenotypes (Lang, Patterson et al. 2009). This may be due to redundancies with *bnc1*, since the authors report partial overlap in the domains of expression in the brain. Mutants with knock-outs of both genes will certainly shed light on this.

However, we may be able gather clues from the chromophore phenotype of *bnc2* mutants. Neural crest cells derive from the developing neural tube. The affected chromophores are neural crest descendents, and these cells died in the absence of *bnc2*. In addition to having neuronal origin, neural crest cells contribute to parts of the brain and peripheral nervous system which might suffer a similar fate in the absence of both *bnc* genes.

Conclusion

While my work was unable to shed any light into the conserved role of *disco* genes using non-insect models, research from other groups have provided interesting clues. Most importantly, *bnc* genes have extensive expression in the nervous systems of zebra fish and mice (Lang, Patterson et al. 2009; Vanhoutteghem, Bouche et al. 2011). Although the *disco* genes have been shown to have important roles in the development of appendages in Drosophila and Tribolium (Patel, Farzana et al. 2007), the initial discovery of *disco* was as a player in neural development of the Drosophila larva and adult (Steller, Fischbach et al. 1987) and expression of *disco* has been observed in the nervous system of the developing embryo (Lee, Freeman et al. 1991). This neural role seems to be conserved, and may shed more light on the origins of appendages in bilateral animals.

ACKNOWLEDGEMENTS

Special thanks to Dr. David Miller of Vanderbilt for helping us try to identify the cells expressing *Ce-disco*-driven GFP and for providing us with a strain expressing GFP in motor neurons (as seen in Figure 1).

References:

Ahringer, J., Ed. (2006). Reverse genetics. *WormBook*, ed., The *C. elegans* Research Community, WormBook.

Akimenko, M. A., M. Ekker, et al. (1994). "Combinatorial expression of three zebrafish genes related to distal-less: part of a homeobox gene code for the head." J Neurosci **14**(6): 3475-86.

Angelini, D. R. and T. C. Kaufman (2004). "Functional analyses in the hemipteran Oncopeltus fasciatus reveal conserved and derived aspects of appendage patterning in insects." Dev Biol **271**(2): 306-21.

Blagburn, J. M. (2007). "Co-factors and co-repressors of Engrailed: expression in the central nervous system and cerci of the cockroach, Periplaneta americana." Cell Tissue Res **327**(1): 177-87.

Boulin, T., J. F. Etchberger, et al. (2006). "Reporter gene fusions." WormBook: 1-23.

Caubit, X., M. C. Tiveron, et al. (2005). "Expression patterns of the three Teashirt-related genes define specific boundaries in the developing and postnatal mouse forebrain." J Comp Neurol **486**(1): 76-88.

Dahn, R. D., M. C. Davis, et al. (2007). "Sonic hedgehog function in chondrichthyan fins and the evolution of appendage patterning." Nature **445**(7125): 311-4.

Davis, M. C., R. D. Dahn, et al. (2007). "An autopodial-like pattern of Hox expression in the fins of a basal actinopterygian fish." Nature **447**(7143): 473-6.

Dirksen, M. L., P. Mathers, et al. (1993). "Expression of a Xenopus Distal-less homeobox gene involved in forebrain and cranio-facial development." Mech Dev **41**(2-3): 121-8.

Fasano, L., L. Roder, et al. (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**(1): 63-79.

Hammond, K. L., R. E. Hill, et al. (2002). "Isolation of three zebrafish dachshund homologues and their expression in sensory organs, the central nervous system and pectoral fin buds." Mech Dev **112**(1-2): 183-9.

Heilig, J. S., M. Freeman, et al. (1991). "Isolation and characterization of the disconnected gene of Drosophila melanogaster." EMBO J **10**(4): 809-15.

Herke, S. W., N. V. Serio, et al. (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**(1): 27-34.

Kraus, P. and T. Lufkin (2006). "Dlx homeobox gene control of mammalian limb and craniofacial development." Am J Med Genet A **140**(13): 1366-74.

Lang, M. R., L. B. Patterson, et al. (2009). "Basonuclin-2 requirements for zebrafish adult pigment pattern development and female fertility." PLoS Genet **5**(11): e1000744.

Lee, K. J., M. Freeman, et al. (1991). "Expression of the disconnected gene during development of Drosophila melanogaster." EMBO J **10**(4): 817-26.

Lee, K. J., M. Mukhopadhyay, et al. (1999). "Autoregulation of the Drosophila disconnected gene in the developing visual system." Dev Biol **214**(2): 385-98.

Long, Q., B. K. Park, et al. (2001). "Expression and regulation of mouse Mtsh1 during limb and branchial arch development." Dev Dyn **222**(2): 308-12.

Mahaffey, J. W., C. M. Griswold, et al. (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**(1): 225-36.

Manuel, M., M. Kruse, et al. (2000). "The comparison of beta-thymosin homologues among metazoa supports an arthropod-nematode clade." J Mol Evol **51**(4): 378-81.

Merlo, G. R., B. Zerega, et al. (2000). "Multiple functions of Dlx genes." Int J Dev Biol **44**(6): 619-26.

Moczek, A. P. (2006). "Integrating micro- and macroevolution of development through the study of horned beetles." Heredity **97**(3): 168-78.

Panganiban, G., S. M. Irvine, et al. (1997). "The origin and evolution of animal appendages." Proc Natl Acad Sci U S A **94**(10): 5162-6.

Patel, M., L. Farzana, et al. (2007). "The appendage role of insect disco genes and possible implications on the evolution of the maggot larval form." Dev Biol **309**(1): 56-69.

Popperl, H., H. Rikhof, et al. (2000). "lazarus is a novel pbx gene that globally mediates hox gene function in zebrafish." Mol Cell **6**(2): 255-67.

Pueyo, J. I. and J. P. Couso (2005). "Parallels between the proximal-distal development of vertebrate and arthropod appendages: homology without an ancestor?" Curr Opin Genet Dev **15**(4): 439-46.

Robertson, L. K., D. B. Bowling, et al. (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**(12): 2781-9.

Robertson, L. K., B. K. Dey, et al. (2002). "Expression of the drosophila gene disconnected using the UAS/GAL4 system." Genesis **34**(1-2): 103-6.

Romano, R. A., H. Li, et al. (2004). "Identification of Basonuclin2, a DNA-binding zinc-finger protein expressed in germ tissues and skin keratinocytes." Genomics **83**(5): 821-33.

Schmitz, C., P. Kinge, et al. (2007). "Axon guidance genes identified in a large-scale RNAi screen using the RNAi-hypersensitive Caenorhabditis elegans strain nre-1(hd20) lin-15b(hd126)." Proc Natl Acad Sci U S A **104**(3): 834-9.

Shippy, T. D., Y. Tomoyasu, et al. (2008). "Do teashirt family genes specify trunk identity? Insights from the single tiptop/teashirt homolog of Tribolium castaneum." Dev Genes Evol **218**(3-4): 141-52.

Shubin, N., C. Tabin, et al. (1997). "Fossils, genes and the evolution of animal limbs." Nature **388**(6643): 639-48.

Shubin, N., C. Tabin, et al. (2009). "Deep homology and the origins of evolutionary novelty." Nature **457**(7231): 818-23.

Steller, H., K. F. Fischbach, et al. (1987). "Disconnected: a locus required for neuronal pathway formation in the visual system of Drosophila." Cell **50**(7): 1139-53.

Telford, M. J. (2004). "The multimeric beta-thymosin found in nematodes and arthropods is not a synapomorphy of the Ecdysozoa." Evol Dev **6**(2): 90-4.

Teumer, J., H. Tseng, et al. (1997). "The human basonuclin gene." Gene **188**(1): 1-7.

Thisse, C., B. Thisse, et al. (1993). "Structure of the zebrafish snail1 gene and its expression in wild-type, spadetail and no tail mutant embryos." Development **119**(4): 1203-15.

Tseng, H. (1998). "Basonuclin, a zinc finger protein associated with epithelial expansion and proliferation." Front Biosci **3**: D985-8.

Tseng, H. and H. Green (1992). "Basonuclin: a keratinocyte protein with multiple paired zinc fingers." Proc Natl Acad Sci U S A **89**(21): 10311-5.

Vanhoutteghem, A., C. Bouche, et al. (2011). "Basonuclins and disco: Orthologous zinc finger proteins essential for development in vertebrates and arthropods." Biochimie **93**(2): 127-33.

Vanhoutteghem, A. and P. Djian (2004). "Basonuclin 2: an extremely conserved homolog of the zinc finger protein basonuclin." Proc Natl Acad Sci U S A **101**(10): 3468-73.

Vanhoutteghem, A. and P. Djian (2006). "Basonuclins 1 and 2, whose genes share a common origin, are proteins with widely different properties and functions." Proc Natl Acad Sci U S A **103**(33): 12423-8.

Vanhoutteghem, A., A. Maciejewski-Duval, et al. (2009). "Basonuclin 2 has a function in the multiplication of embryonic craniofacial mesenchymal cells and is orthologous to disco proteins." Proc Natl Acad Sci U S A **106**(34): 14432-7.

Wang, H., E. M. Lee, et al. (2007). "Isolation and expression of zebrafish zinc-finger transcription factor gene tsh1." Gene Expr Patterns **7**(3): 318-22.

Waskiewicz, A. J., H. A. Rikhof, et al. (2001). "Zebrafish Meis functions to stabilize Pbx proteins and regulate hindbrain patterning." Development **128**(21): 4139-51.

Wilkinson, D. G., Ed. (1992). In Situ Hybridization: A Practical Approach. Practical Approach Series. Oxford, Oxford University Press.

Williams, T. A., C. Nulsen, et al. (2002). "A complex role for distal-less in crustacean appendage development." Dev Biol **241**(2): 302-12.

Wu, J. and S. M. Cohen (2000). "Proximal distal axis formation in the Drosophila leg: distinct functions of teashirt and homothorax in the proximal leg." Mech Dev **94**(1-2): 47-56.

Zhang, X. and H. Tseng (2007). "Basonuclin-null mutation impairs homeostasis and wound repair in mouse corneal epithelium." PLoS One **2**(10): e1087.

CHAPTER FOUR

Using Association Mapping to Identify

Unknown Leg Development Genes

Several people contributed to this work and will be included as authors when it is published. Xin Su and Tiffany Petrisko prepared the legs and made the measurements used in the analysis for this study. Dr. Megan Leach helped with candidate identification and analysis.

INTRODUCTION

All multicellular life starts as a single cell capable of growing into a whole complex organism. In most cases, after a few cell divisions, expression of certain genes establishes the poles and orientations of the embryo. The process continues giving rise to new domains, new organs and new tissues until a completed organism is created. The genetic control of this process has been characterized best in studying the development of organs from certain organisms (ex: heart development in Drosophila and Mammals (Medioni, Senatore et al. 2009; Munoz-Chapuli and Perez-Pomares 2010)). In cases like this, genes with large effect (ex: defining a particular domain or organ) have easily been found. Genes that are important but have a less significant role are harder to identify and their developmental roles are more difficult to understand. The fruit fly *Drosophila melanogaster* has been studied extensively, and much is known about how the whole embryo is patterned, as well as about how specific tissues are established and patterned. The development of adult legs is one such system.

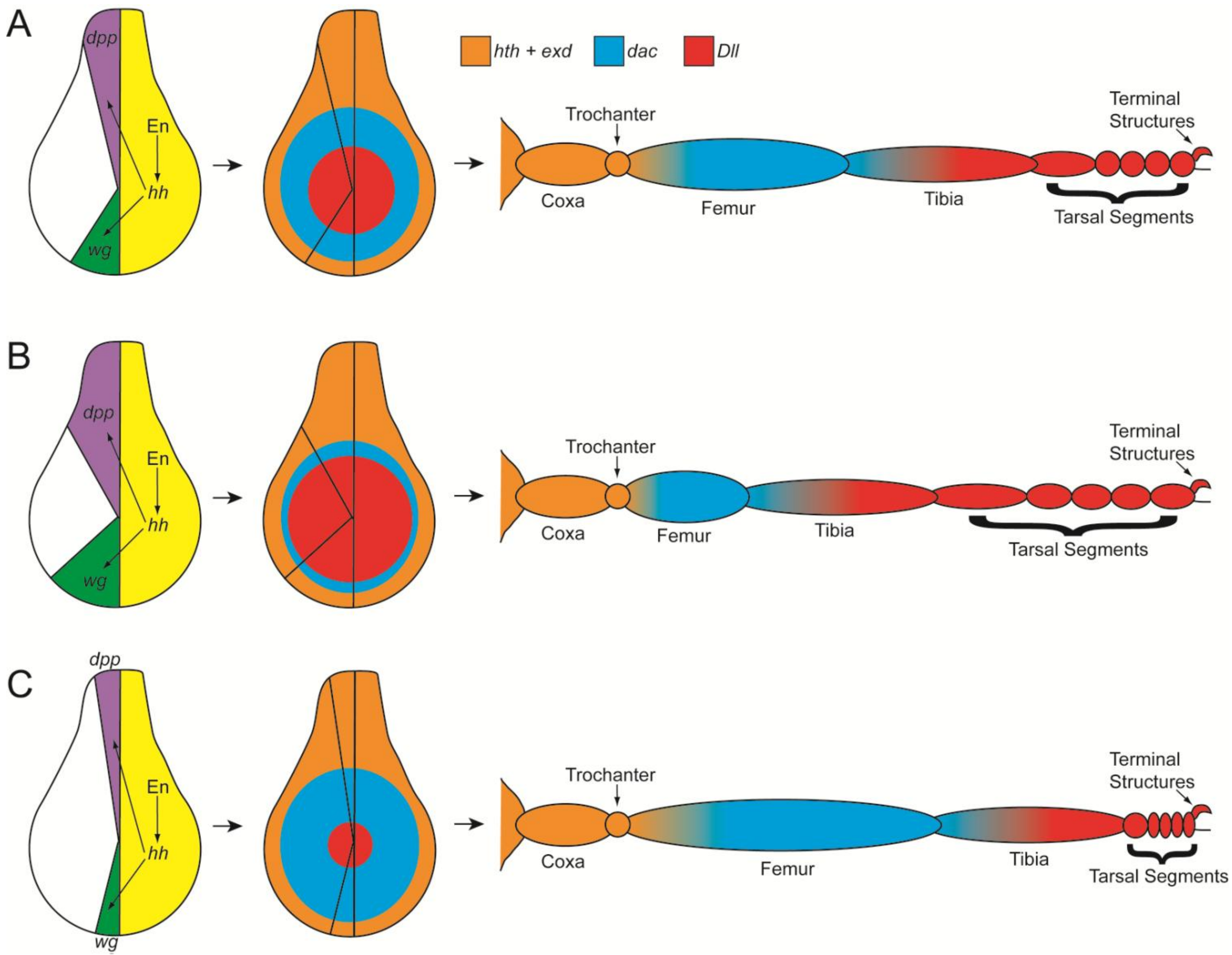
In the fruit fly, patterning of the adult limbs begins in small, flat areas of tissue called imaginal discs, which are allocated in early embryogenesis and are patterned during the larval stages (Reviewed in (Panganiban 2000; Morata 2001; Kojima 2004; Pueyo and Couso 2005; Estella, Voutev et al. 2012)). The primordia of the leg imaginal discs arise in the thoracic segments at positions where segmentally repeated expression of the signaling factor Wingless (Wg) activates the expression of *Distal-less* (*Dll*), which encodes a transcription factor. Abdominal Hox genes block *Dll* expression posterior to the thoracic segments, while the Decapentaplegic (Dpp) and the Epidermal Growth Factor Receptor (EGFR) pathways limit *Dll* dorsally and ventrally, respectively (Panganiban 2000; Estella, Voutev et al. 2012).

After the disc primordia are established during embryogenesis, they begin receiving patterning instructions via a series of signaling molecules, morphogens and transcription factors expressed during the larval stages (Fig 1A). This patterning begins when the transcription factor Engrailed (En), defining cells of the posterior compartment of the discs, activates the expression of *hedgehog* (*hh*), which encodes a morphogen that travels into the anterior compartment (Morata 2001). There, Hh signal activates the expression of *wg* and *dpp*, *dpp* transcripts forming a dorsal stripe and *wg* transcripts a ventral one, with the two meeting in the center of the disc (Lecuit and Cohen 1997). In addition to defining the dorsal-ventral (D-V) axis, these morphogens also work together to define the proximal-distal (P-D) axis of the leg by activating concentric expression patterns of transcription factors. Where cells experience high levels of both Wg and Dpp signal, in the center of the disc, *Dll* is activated to pattern the distal portion of the leg (Lecuit and Cohen 1997; Panganiban 2000). As the mutual presence of Wg and Dpp decreases, a threshold is reached that permits the activation of *dachshund* (*dac*), which is responsible for patterning the middle leg segments (Kojima 2004; Estella, Voutev et al. 2012). Near the edge of the disc, where the mutual expression of *wg* and *dpp* is weakest, the transcription factor, *homothorax* (*hth*), is expressed. There, the Hth protein imports Extradenticle (EXD) into the nucleus, specifying the proximal portion of the leg and its connection with the body wall (Kojima 2004; Estella, Voutev et al. 2012).

Many of these genes and pathways have been shown to be important in the development and growth of appendages and appendage-like structures in other metazoans (Panganiban, Irvine et al. 1997; Morata 2001). Misexpression of many of these genes also

FIGURE 1: A model of larval imaginal disc patterning and adult leg development, and a possible mode of change in proportion variation

A) Leg imaginal disc patterning begins (left) when En, present in the posterior (yellow) portion of the disc, activates *hh* expression. The Hh ligand diffuses into the anterior portion of the disc. There it activates the expression of *dpp* (purple) dorsally and *wg* (green) ventrally. Dpp and Wg, in turn, diffuse throughout the dorsal and ventral portions, respectively, of the disc, creating a gradient of mutual signal. This gradient is responsible for the pattern of expression of transcription factors (middle) that establish the proximal-distal axis of the leg (right). In the center of the disc, mutual Dpp and Wg is highest, activating the expression of *Dll* (red) which is responsible for patterning the most distal structures of the leg, including the terminal structures (claw and pulvillus). Reduced mutual Dpp and Wg results in the activation of *dac* (blue) responsible for the patterning of the middle portions of the leg. Where there is almost no mutual Dpp or Wg, Hth and Exd are active (orange), patterning the proximal leg portions and the junction with the rest of the body. B) Expansion of the expression or distribution of Dpp and Wg (left) could result in an expansion of *Dll* expression, at the expense of *dac* (middle). This could, in turn, cause the tarsal segments to make up a larger portion of the leg without changing leg length (right). C) Similarly, the proportion of the femur could be expanded (right) if *dac* expression were expanded (middle). This might result from a decrease in expression or distribution of Dpp and Wg (left).



A
dpp
En
hh
wg
hth + exd
dac
Dll
Trochanter
Terminal
Structures
Coxa
Femur
Tibia
Tarsal Segments
B
dpp
En
hh
wg
Trochanter
Terminal
Structures
Coxa
Femur
Tibia
Tarsal Segments
C
dpp
En
hh
wg
Trochanter
Terminal
Structures
Coxa
Femur
Tibia
Tarsal Segments

contributes to certain diseases, like cancer (Morata 2001; Briscoe and Therond 2005; Polakis 2007; Morini, Astigiano et al. 2010; Popov, Wu et al. 2010).

In spite of our extensive knowledge of Drosophila leg development, gaps still remain. Many of the genes known to be involved encode transcription factors, but their targets remain largely a mystery (Morata 2001). Even major players in leg development have been missed; the *disco* genes, *disconnected* (*disco*) and *disco-related* (*disco-r*), were only relatively recently discovered to be critical for the maintenance of cells contributing to the developing legs and other ventral appendages (Patel, Farzana et al. 2007). It is plausible that other important genes or pathways may have been missed as well. Since many leg development genes play an important role in metazoan development, developmental disorders, and cancer, it is critical to fill these gaps to fully understand appendage development.

However, there are limitations to techniques currently used to study development, even in Drosophila. For example, the redundancy of the *disco* genes with respect to leg development masked their importance in this role in traditional mutant analyses. The limitations of traditional methods requires employment of new techniques to explore these otherwise well-studied systems.

Therefore, in order to identify new genes contributing to leg development, we have employed a genome-wide association study (GWAS) to identify single-nucleotide polymorphisms (SNPs) associated with proportionality variation of the femur, tibia and tarsal segments relative to total leg length. Candidate genes were identified based on proximity to SNPs and were further tested using *in situ* hybridization and RNA interference (RNAi) to determine a role in leg development. In this study we identified some genes from known pathways that were not previously associated with leg development. We also identified previously uncharacterized genes and demonstrated their role in this process.

MATERIALS AND METHODS

Growth of Drosophila

All flies used in these experiments were reared at room temperature, about 23ºC, on standard cornmeal-agar-molasses medium.

Statistical analyses

Factorial, mixed model ANOVAs of form $Y = \mu + S + L + V(L) + S{\times}L + S{\times}V(L) + \varepsilon$ were used to partition variation in leg segment proportion between sexes (fixed), DGRP lines (random), vial from which the measured fly came (random), the S×L interaction (random), the S×V interaction (random), and the error variance (ε).

Association between variability of leg segment proportionality and SNP markers was tested using SAS software. Markers were tested for association using a simple ANOVA model: phenotype=SNP, assayed independently by position, by sex, and by leg type (i.e. T1 or T2). Sites with a minor allele frequency of less than four were removed from analysis. We used data from Freeze One of the Drosophila Genetic Reference Panel (DGRP) (Mackay, Richards et al. 2012) for our analysis.

Identifying candidate genes

From association mapping, SNP positions that had the lowest combined P-values for male, female, T1 and T2 were identified independently for the femur, tibia, and tarsus, using Fisher's method for combining P-values. This ensured we were selecting general leg factors, rather than sex specific or body-segment specific genes. The result was three tables, one for each leg segment, similar to what is seen in Table 4. We focused our attentions on SNPs at the top of each table (i.e. the lowest P-values in all four categories), with two exceptions. Candidates with leg imaginal expression (see below) identified during early rounds of experiments were not abandoned if they had lower significance during later rounds. Candidates that failed to clone (see below) were set aside for later analysis, due to the large number of candidates being analyzed.

The GBrowse program on FlyBase (www.flybase.org, (McQuilton, St Pierre et al. 2012)) was used to identify candidate genes by proximity to SNP positions. If a SNP was found between two neighboring genes, both genes were analyzed as candidates. Genes near SNPs of interest with a known function in leg development were not examined further. SNP positions too far from any gene to successfully identify a candidate (≥8Kbp) were discarded.

Cloning and *in situ* hybridization

Genes were cloned using primers designed from exon sequences on FlyBase (McQuilton, St Pierre et al. 2012). Vector NTI (Life Technologies Corporation) was used to help design primers for a coding region of each candidate gene. These regions were then amplified using PCR of Drosophila genomic DNA prepared in the lab from an Ore-R fly. PCR products were then cloned into a pGEM vector using the pGEM-T Easy Vector System (Promega).

in situ hybridization was used to determine expression patterns of candidate genes in imaginal discs and embryos. Drosophila imaginal discs from third instar larvae were dissected in PBS and placed into standard in situ fixative. Digoxigenin-labeled antisense

RNAs were prepared and *in situ* hybridization was done essentially as in (Tautz and Pfeifle 1989).

RNAi analysis

RNAi lines were obtained from the Vienna Drosophila RNAi Center (VDRC) (Dietzl, Chen et al. 2007) directly, and from the Transgenic RNAi Project (TRiP) from Harvard Medical School through the Bloomington Stock Center. Males from RNAi lines were crossed to virgin females from five Gal4 lines: *rotund* (*rn*) (St Pierre, Galindo et al. 2002), *decapentaplegic* (*dpp*) (Staehling-Hampton, Jackson et al. 1994), *armadillo* (*arm*)(Sanson, White et al. 1996), *Distal-less* (*Dll*) and *teashirt* (*tsh*) (Calleja, Moreno et al. 1996). Progeny from crosses were analyzed to determine the effect of silencing candidate genes in Gal4 driver domains. Efficacy of RNAi was confirmed by examining expression of the target gene in larvae from these crosses using *in situ* hybridization.

Potential paralogs of candidate genes were identified using BLAST (Altschul, Gish et al. 1990; Altschul, Madden et al. 1997) to search the encoded protein sequence against a database of translated *Drosophila melanogaster* DNA sequences on FlyBase.

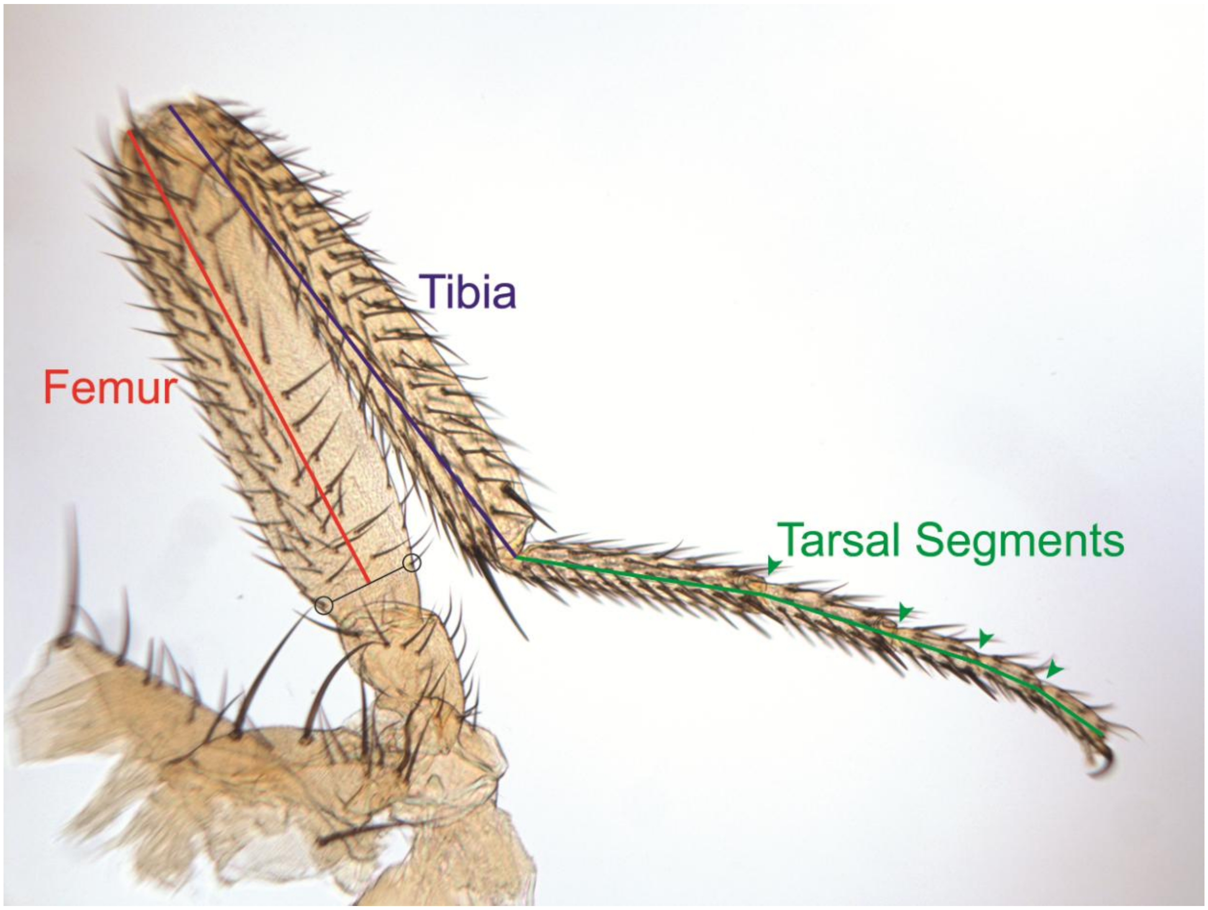


FIGURE 2: Measuring legs

An example DGRP female, T2 leg marked to show the measuring method. In the femur, a black circle marks each of the two most proximal bristles. Measurement begins in the middle of the line (black) between these two bristles, and ends at the end of the segment (red line). The tibia is measured from its beginning, in the joint with the femur, and ends where it forms a joint with the first tarsal segment (blue line). Each tarsal segment was measured separately (green lines), beginning with the joint to the more proximal segment and ending with the joint to the more distal segment (green arrow heads). In the case of the first tarsal segment, measurement began at the joint with the tibia. Measuring the final tarsal segment ended at the tip of the leg, not including terminal structures.

RESULTS

Quantitative genetic variation in leg proportions of the DGRP

First (T1) and second (T2) thoracic legs were measured from adult flies from 117 lines of the Drosophila Genetic Reference Panel (DGRP) (Ayroles, Carbone et al. 2009; Mackay, Richards et al. 2012). Three segments, femur, tibia and tarsus were measured separately (Fig 2) to attain both the total leg length and the proportion each segment contributed to that length (X. Su and J.W. Mahaffey, personal communication). We found significant variation in leg segment proportion amongst the DGRP lines (Fig 3, Table 1), so variation in proportionality of the femur, tibia, and tarsus was used to conduct association mapping. We examined proportionality of leg segments instead of total leg length to avoid confounding results with non-genetic sources of variability in length. These environmental influences on body size resulted in a low repeatability of length measurements. However, leg proportions were highly repeatable (X.S. and J.W.M., personal communication).

Studying proportionality variation also brought the focus of our study on identifying leg patterning genes. Such patterning genes could potentially play a role in shaping the domains of expression and activity of the various morphogens, signaling molecules, and transcription factors that pattern the leg in the model of Drosophila leg development presented above (Fig 1). A change in the signal strength of Dpp or Wg, for example, could alter the proportion of Drosophila imaginal discs expressing the leg determinant genes *hth*, *dac*, or *Dll*, which could, in turn, alter the proportions of leg segments without necessarily affecting over-all length (Fig 1). Therefore, variation in proportion seemed the more informative trait to study.

FIGURE 3: Variation in proportions of leg segments

Mean leg segment proportions for each measured DGRP line are graphed on the Y-axis, organized by proportion in females (red) for each trait, in each leg type. Males are shown in blue. The P-value for line effect is also shown, and was very significant in all cases.

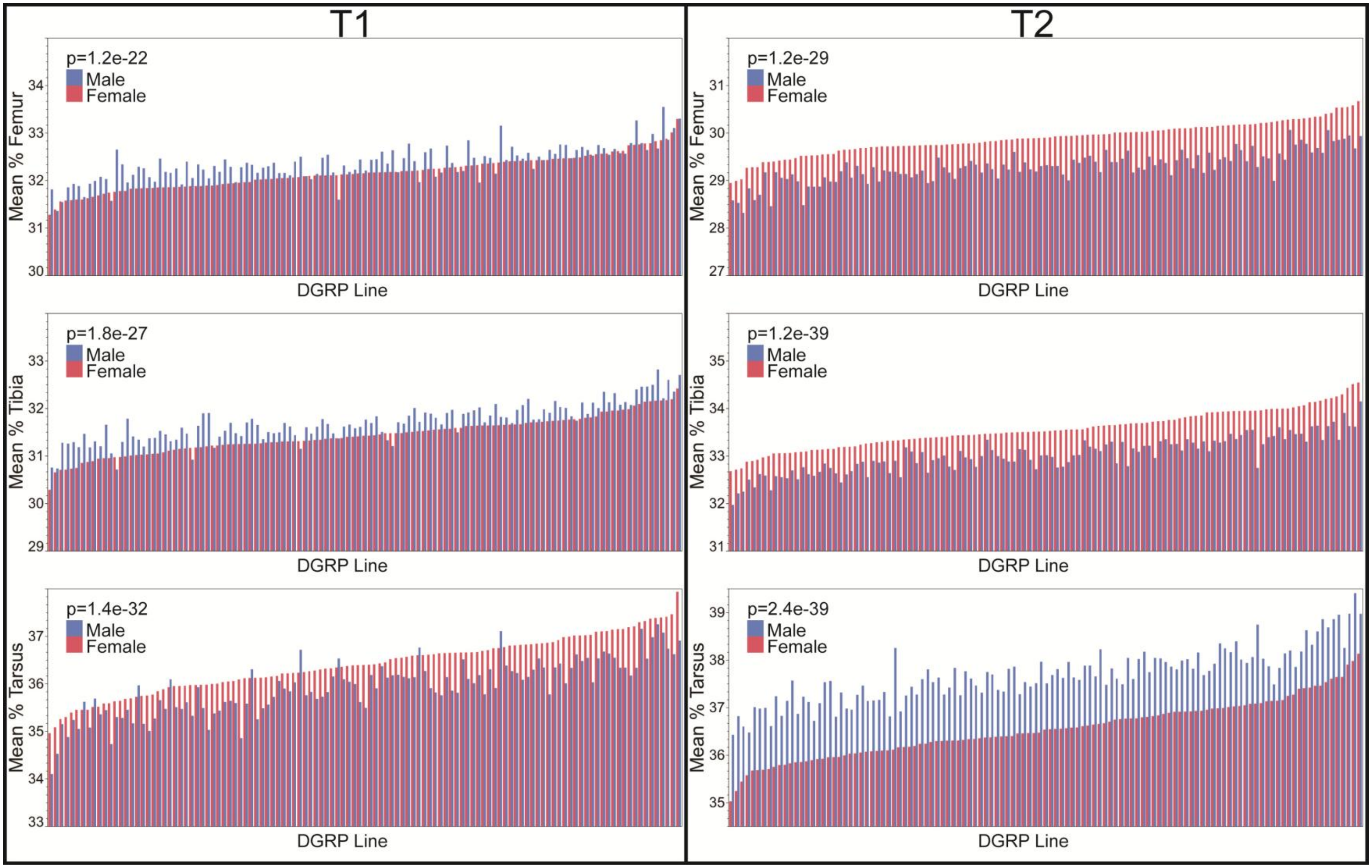

T1
T2
p=1.2e-22
Male
Female
Mean % Femur
DGRP Line
p=1.2e-29
p=1.8e-27
Mean % Tibia
p=1.2e-39
p=1.4e-32
Mean % Tarsus
p=2.4e-39

TABLE 1: Analysis of variance of leg segment proportions

DF: degrees of freedom; MS: Type III Mean Squares; σ^2: variance component.

Trait	Leg Type	Source of Variation	DF	MS	*F*	p-value	σ^2
% Femur	T1	Sex	1	26.687	82.262	<0.001	fixed
		Line	116	2.761	8.234	<0.001	0.100
		Sex * Line	116	0.324	1.566	0.001	0.010
		Vial(Line)	351	0.218	1.053	0.315	0.002
		Sex * Vial(Line)	351	0.207	1.035	0.331	0.002
		Error	1872	0.200	-	-	0.200
	T2	Sex	1	254.533	1272.846	<0.001	fixed
		Line	116	2.438	11.372	<0.001	0.092
		Sex * Line	116	0.200	1.568	0.001	0.006
		Vial(Line)	351	0.142	1.113	0.158	0.002
		Sex * Vial(Line)	351	0.128	1.180	0.019	0.006
		Error	1872	0.108	-	-	0.108
% Tibia	T1	Sex	1	51.852	222.765	<0.001	fixed
		Line	116	3.452	15.750	<0.001	0.133
		Sex * Line	116	0.233	1.248	0.065	0.004
		Vial(Line)	351	0.173	0.927	0.761	0
		Sex * Vial(Line)	351	0.186	1.314	<0.001	0.013
		Error	1872	0.142	-	-	0.142
	T2	Sex	1	199.525	1169.220	<0.001	fixed
		Line	116	3.385	19.279	<0.001	0.133
		Sex * Line	116	0.171	1.813	<0.001	0.006
		Vial(Line)	351	0.099	1.052	0.317	0.001
		Sex * Vial(Line)	351	0.094	1.099	0.118	0.003
		Error	1872	0.086	-	-	0.086
% Tarsus	T1	Sex	1	152.936	253.203	<0.001	fixed
		Line	116	7.450	11.401	<0.001	0.281
		Sex * Line	116	0.604	2.109	<0.001	0.026
		Vial(Line)	351	0.336	1.173	0.068	0.008
		Sex * Vial(Line)	351	0.286	1.128	0.066	0.011
		Error	1872	0.254	-	-	0.254
	T2	Sex	1	904.771	2037.206	<0.001	fixed
		Line	116	7.648	16.354	<0.001	0.297
		Sex * Line	116	0.444	2.210	<0.001	0.020
		Vial(Line)	351	0.225	1.117	0.150	0.004
		Sex * Vial(Line)	351	0.201	1.331	<0.001	0.017
		Error	1872	0.151	-	-	0.151

Correlations

We found significant correlations between traits (i.e. proportions of each segment; Fig 4), between leg types (i.e. T1 and T2; Fig 5), and between the sexes (Fig 6). In all cases, correlations were significantly different from unity (Table 2). General leg patterning genes would be expected to be in common in both sexes and in both leg types, so such strong correlations are not surprising.

The correlation between proportion of tibia and proportion of femur was relatively low, and in one case (between tibia and femur of male T1), it was not significant (Table 2). This suggests there is a weak relationship between the patterning of these two leg segments. Correlations between the other trait pairs (femur and tarsus, tibia and tarsus) were negative. This relationship between the proportions of femur and tarsus might be expected. As shown in Figure 1, genetic factors expanding the proportion of one segment might simultaneously cause an equivalent reduction in the other. It is particularly interesting to consider the relationship of the tibia with the other two segments. As with the femur, the tibia has a negative correlation with the tarsus, suggesting the tibia is more strongly influenced by the same patterning elements that shape the femur. The weak, positive correlation between femur and tibia suggests that this patterning may establish a broad field distinct from the tarsus from which the femur and tibia will form later.

GWAS in the DGRP

We ran twelve separate GWAS to identify SNPs associating with variation in leg segment proportions. For each trait, we separated the sexes and leg types, analyzing 2,285,856 SNPs

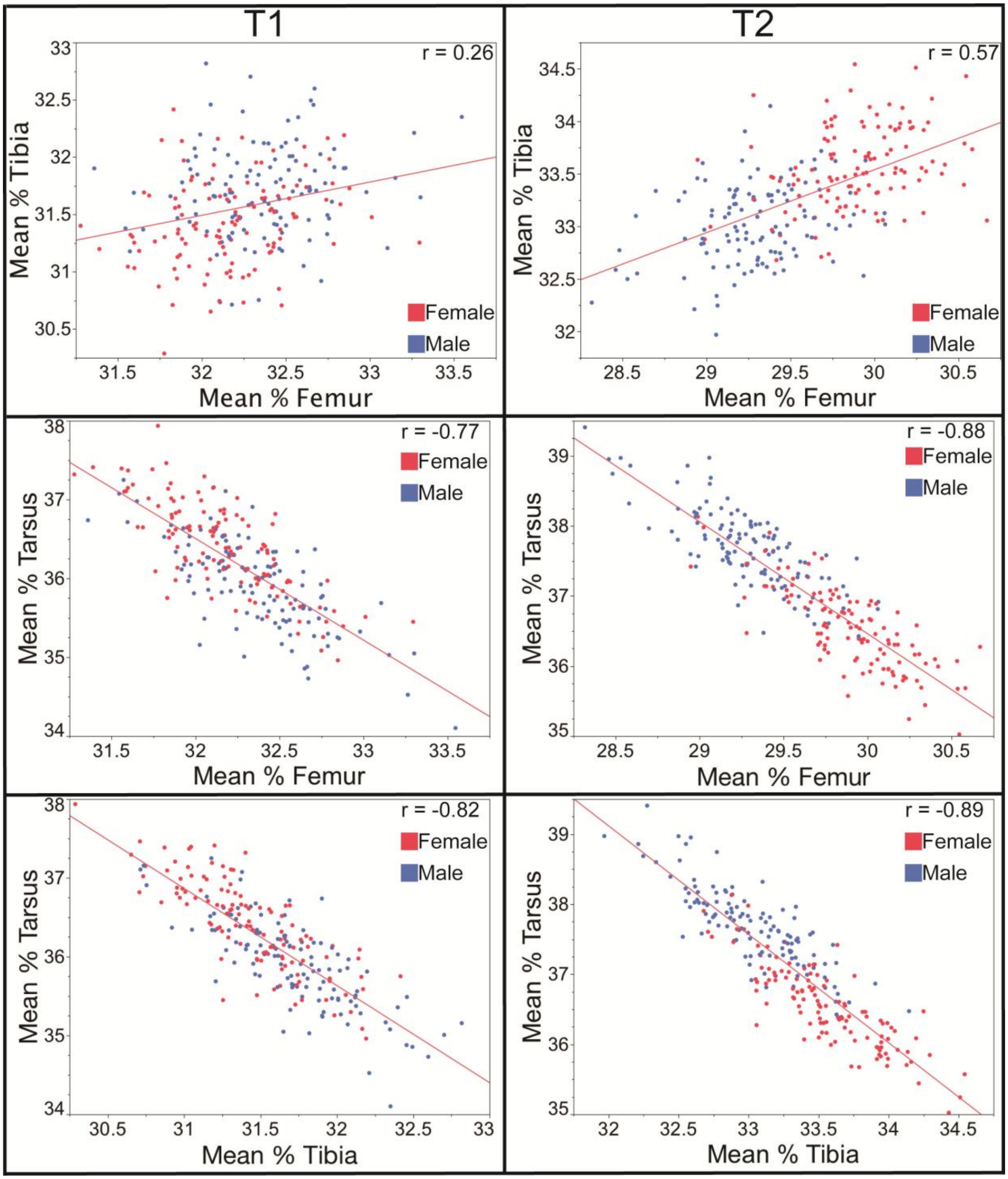


FIGURE 4: Correlations between traits, separated by leg type

Values of r presented in graphs represents the combined r of both sexes. Individual correlations for each sex can be found in Table 2.

TABLE 2: Correlations between traits, between sexes, and between leg types

Correlations for each relationship were calculated independently with respect to trait, sex and leg type, and these are recorded in column ‘r’. Degree of significance from the lowest possible value (0) and the highest possible value (1) were also calculated (recorded in columns ‘t’ and ‘P-value’). Abbreviations: Fe, femur; Ta, tarsus; Ti, tibia; T1, first thoracic leg; T2, second thoracic leg; F, female; M, Male; A, both sexes.

Correlation between	Separated By	In	r-squared	r	n	Interrelatedness (relationship to 0)		Independence (relationship to 1)	
						t	P-value	t	P-value
TixFe	T1	M	0.021	0.144	117	1.560	0.061	10.612	<0.01
		F	0.057	0.239	117	2.639	<0.01	10.413	<0.01
		A	0.068	0.260	234	4.102	<0.01	14.708	<0.01
	T2	M	0.087	0.294	117	3.303	<0.01	10.249	<0.01
		F	0.106	0.325	117	3.686	<0.01	10.142	<0.01
		A	0.328	0.573	234	10.643	<0.01	12.485	<0.01
TaxFe	T1	M	0.530	0.728	117	11.392	<0.01	7.350	<0.01
		F	0.589	0.767	117	12.836	<0.01	6.875	<0.01
		A	0.589	0.767	234	18.232	<0.01	9.765	<0.01
	T2	M	0.598	0.774	117	13.089	<0.01	6.796	<0.01
		F	0.607	0.779	117	13.339	<0.01	6.719	<0.01
		A	0.777	0.881	234	28.428	<0.01	7.193	<0.01
TaxTi	T1	M	0.613	0.783	117	13.504	<0.01	6.669	<0.01
		F	0.650	0.806	117	14.600	<0.01	6.348	<0.01
		A	0.670	0.819	234	21.710	<0.01	8.748	<0.01
	T2	M	0.695	0.833	117	16.169	<0.01	5.927	<0.01
		F	0.716	0.846	117	17.006	<0.01	5.720	<0.01
		A	0.796	0.892	234	30.058	<0.01	6.885	<0.01
T1xT2	Fe	M	0.521	0.722	117	11.187	<0.01	7.421	<0.01
		F	0.475	0.690	117	10.209	<0.01	7.767	<0.01
	Ti	M	0.562	0.749	117	12.139	<0.01	7.100	<0.01
		F	0.528	0.727	117	11.352	<0.01	7.364	<0.01
	Ta	M	0.624	0.790	117	13.818	<0.01	6.575	<0.01
		F	0.658	0.811	117	14.876	<0.01	6.271	<0.01
MxF	Fe	T1	0.624	0.790	117	13.809	<0.01	6.577	<0.01
		T2	0.720	0.848	117	17.189	<0.01	5.676	<0.01
	Ti	T1	0.764	0.874	117	19.302	<0.01	5.208	<0.01
		T2	0.818	0.904	117	22.734	<0.01	4.575	<0.01
	Ta	T1	0.723	0.850	117	17.311	<0.01	5.647	<0.01
		T2	0.793	0.891	117	21.009	<0.01	4.875	<0.01

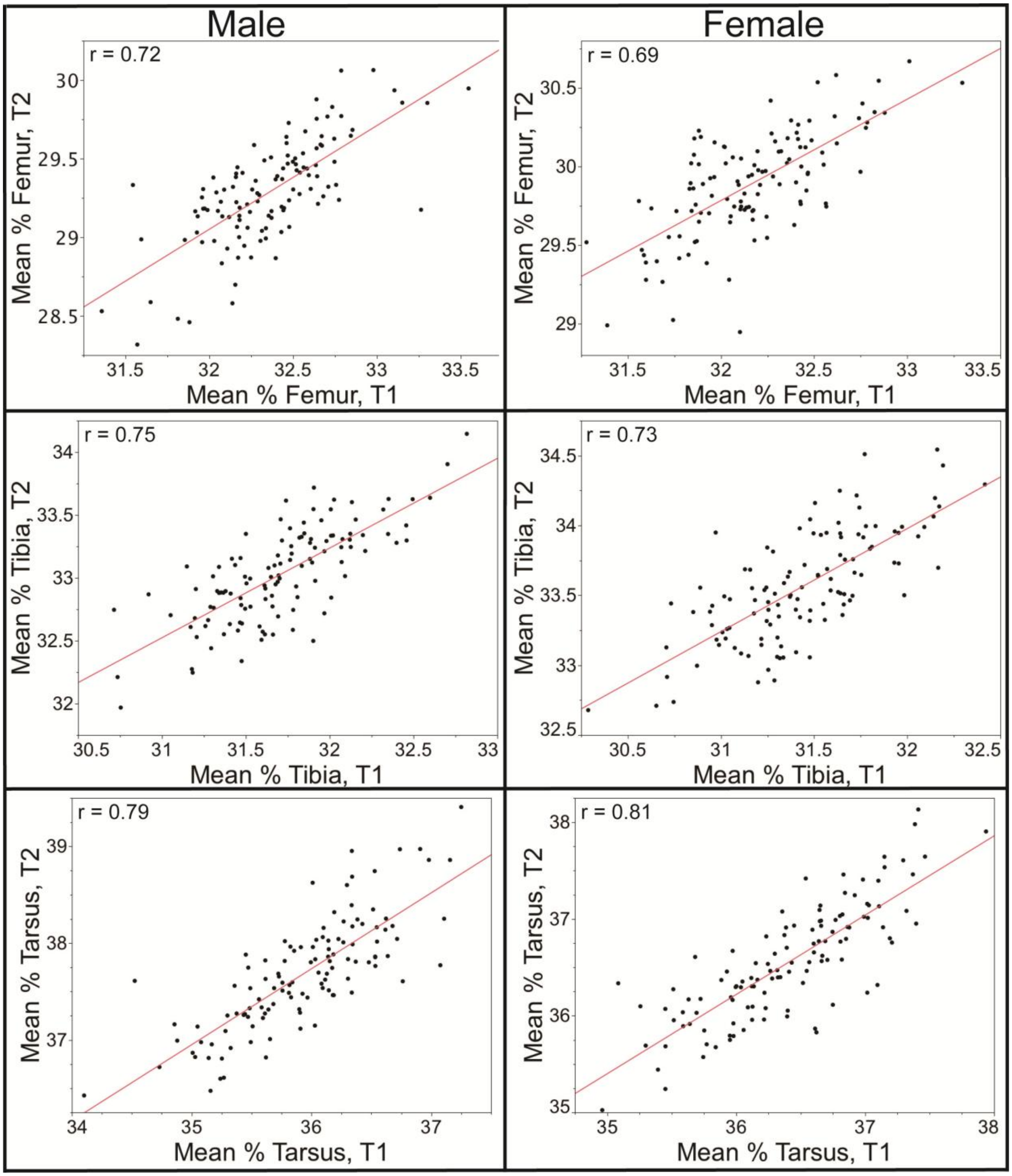


FIGURE 5: Correlations for traits between leg types, separated by sex

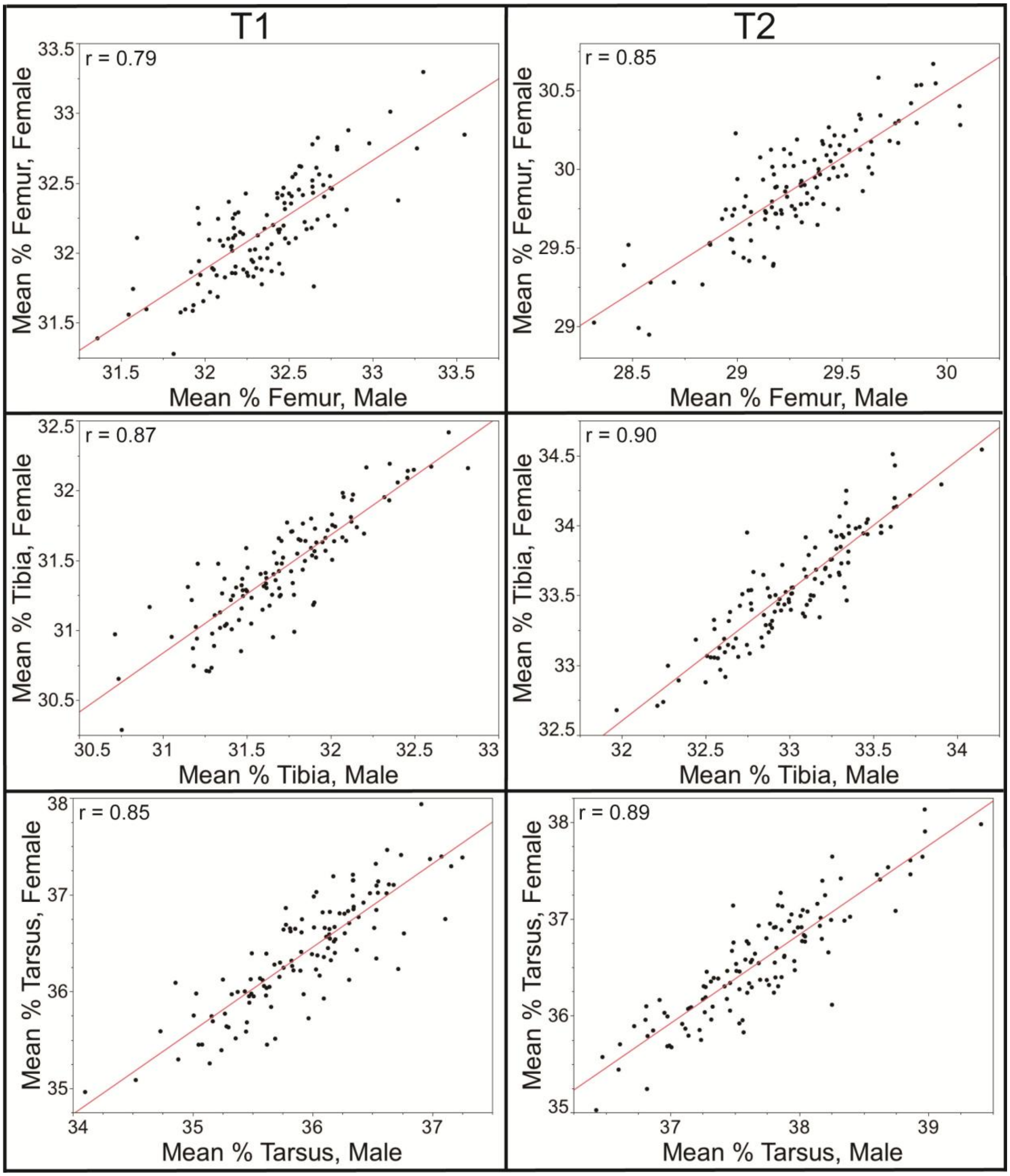


FIGURE 6: Correlations for traits between sexes, separated by leg type

with a minor allele frequency of four or greater. The number of SNPs with P-values less than 10^{-5} can be found in Table 3. In all cases we expected approximately 23 SNPs below this cutoff. In most cases, we found fewer SNPs than expected, resulting in high false discovery rates (FDR; Table 3). Since the expected number is an estimate that assumes all tests are independent, it is expected to be an overestimate. Specifically, linkage disequilibrium between neighboring SNPs would indicate that certain SNPs were not independent, suggesting that the actual number of analyzed SNPs is lower than that used for the estimation. However, the goal of this research was to provide candidate genes for functional study, and as is described below, the FDR was not indicative of success finding candidates with roles in leg development.

In order to improve our ability to find functional candidates, we sought to identify genes in common by trait across sexes and leg types. Such genes would be expected to be general leg development factors, and should have a role in patterning all leg discs. Therefore, we picked genes close to SNPs that had the lowest P-values for both males and females, and both T1 and T2 for each of the traits.

Functional tests

Table 4 shows the top 20 SNP positions associated with proportion variation for each of the three measured segments. Candidate genes were identified based on their proximity to a SNP. Although some SNPs were not near any known or predicted genes, many SNP positions were found directly in some gene sequences including introns, exons, and protein coding regions, while other SNPs were found upstream or downstream of genes.

TABLE 3: False discovery rates for GWAS analyses

The P-value cutoff used for this analysis was 10^{-5}. In all tests, we expected ~23 SNPs below the cutoff.

Leg Segment	Leg Type	Sex	#SNPs below Actual	#SNPs below Expected	FDR
Femur	T1	M	46	22.859	0.497
		F	34	22.859	0.672
	T2	M	35	22.859	0.653
		F	27	22.859	0.847
Tibia	T1	M	19	22.859	1.203
		F	14	22.859	1.633
	T2	M	14	22.859	1.633
		F	18	22.859	1.270
Tarsus	T1	M	15	22.859	1.524
		F	15	22.859	1.524
	T2	M	29	22.859	0.788
		F	21	22.859	1.089

TABLE 4: Top 20 candidate genes for each leg segment

Each SNP site (represented in columns ‘Chr’ and ‘Position’) was analyzed independently for contribution to the proportion of each of the three measured leg segments, femur, tibia and tarsus (column ‘Associated segment’). Effects of T1 v. T2, and sex (male v. female) were separated out, so have been given separate P-values (columns ‘T1F’, ‘T1M’, ‘T2F’, and ‘T2M’). Candidates were sorted using Fisher’s method for combining P-values to identify SNPs with the lowest P-values in all four columns. Candidates (column ‘Gene’) were identified by proximity to SNPs (column ‘Relative location’). Some candidates were near multiple SNPs, but only the most significant SNP is shown.

TABLE 4: continued

Chr	Position	Gene	Relative location	T1 F	T1 M	T2 F	T2 M	Associated segment
3L	18875409	CG3961	intron	3.6327E-07	1.1494E-06	1.0852E-05	0.000266	Femur
1	2362060	boi	intron	1.7787E-06	0.00069416	8.4422E-07	3.6315E-05	Femur
3L	2800633	CG2083	~500bp US	7.89E-06	3.0983E-06	1.5868E-05	0.00010452	Femur
2R	19724480	CG5549	exon	0.00011193	0.00044607	2.0068E-06	2.8281E-06	Femur
1	8773421	Moe	intron	0.00075804	1.6311E-05	5.6257E-06	4.2685E-06	Femur
3L	6503528	sfl	intron	1.194E-05	0.00570599	2.0227E-06	6.163E-06	Femur
2R	19736539	ytr; eIF6	intron; 200bpDS	0.00014809	0.00111063	8.6398E-06	3.0685E-06	Femur
2R	10582312	mspo	intron	0.00832881	0.00148859	1.5156E-07	2.504E-06	Femur
1	14348501	dpr8	intron	0.0001148	0.00070997	9.9562E-06	7.94E-06	Femur
2R	1780461	nearest genes	>30kb away	1.782E-05	0.00018921	0.00021714	1.1643E-05	Femur
2L	1717804	CG18317	exon	1.6782E-05	1.3391E-05	0.00045409	9.8201E-05	Femur
2R	17964922	GlcT-1	exon	1.5283E-05	2.7251E-05	0.00034151	7.5782E-05	Femur
3L	10876047	CG7949	20bp US	4.0506E-07	9.9785E-05	0.00029779	0.00103257	Femur
1	2224434	CG33218	intron	0.00027346	0.0031249	8.0155E-06	1.8304E-06	Femur
3L	12174296	CAH2	intron	0.00052098	9.6949E-05	0.00010393	2.4369E-06	Femur
3L	5751724	Ubp64E	intron	0.00015224	3.7656E-05	3.0113E-05	9.8367E-05	Femur
3L	10836603	tna	intron	7.9965E-06	1.5841E-05	0.0001429	0.00099556	Femur
2R	13435137	RpL18A; MESR4	140bpUS; 200bpUS	3.8438E-06	0.00031841	0.00016845	0.00017155	Femur
2R	19727517	snoRNA		0.00011094	0.00062378	3.6411E-05	1.5082E-05	Femur
2L	2419744	CG34447	exon	0.00013538	8.4427E-06	0.00089305	3.7825E-05	Femur
3L	1251041	CG9129	CDS	2.1704E-05	0.00025437	6.1282E-06	7.623E-07	Tibia
3R	14444206	CG14304; CG7695	intron; 1kbUS	2.0877E-06	7.1018E-06	0.00022058	6.5055E-05	Tibia
1	15999736	CG42354/3	1kbUS	1.6534E-05	5.5085E-05	2.415E-05	1.0806E-05	Tibia
3R	26494838	nearest genes	>15kb away	2.124E-05	0.00013078	6.923E-05	1.5006E-06	Tibia
2L	12445302	Oatp33Ea	intron	6.2303E-05	8.83E-06	4.8204E-05	1.232E-05	Tibia
2L	3857356	CG3964	5kbUS	2.0054E-05	0.00011088	2.7001E-05	9.9566E-06	Tibia
3L	1277747	CG32333	intron	7.2559E-05	0.00038284	2.45E-05	1.4615E-06	Tibia
2R	4155864	CG30371	intron	0.00014176	7.7269E-06	2.6735E-05	4.1389E-05	Tibia
3R	26752160	Ptx1	intron	3.5133E-07	6.6009E-05	0.00076645	8.6764E-05	Tibia
3L	1666520	snoRNA		9.7198E-05	0.00072016	5.1695E-06	4.3465E-06	Tibia

TABLE 4: continued

Chr	Position	Gene	Relative location	T1 F	T1 M	T2 F	T2 M	Associated segment
2L	8169673	CG8486	intron	0.00020348	2.9755E-05	4.1527E-05	1.0246E-05	Tibia
2R	18998636	GR59c/d	6kbUS	3.7256E-05	0.00014232	6.8021E-05	7.237E-06	Tibia
3L	1263330	CG9134	intron	2.7362E-05	2.4967E-05	5.6536E-05	7.1087E-05	Tibia
3R	15651069	bnl	intron	2.8227E-05	2.5929E-05	9.4102E-06	0.00048908	Tibia
2R	14945788	CG15116; CG15109	2kbUS; 1kbUS	1.3777E-05	2.0375E-05	0.0002924	4.2161E-05	Tibia
3L	9034005	Doc1	470bpUS	9.1999E-05	9.3694E-06	0.00023759	2.6386E-05	Tibia
3L	12503800	sti	CDS	0.00058375	0.00150594	1.6744E-06	3.8769E-06	Tibia
2R	4153744	CG30375	CDS	0.00036704	3.7961E-05	1.0131E-05	4.2791E-05	Tibia
2R	7954509	CG8888	intron	7.1541E-05	1.248E-06	0.00104275	8.4542E-05	Tibia
2L	14776072	CG4650	1kbUS	0.00014735	0.00043289	4.4269E-06	3.2224E-05	Tibia
2R	10863972	ckn	intron	1.1454E-07	1.0871E-06	1.4321E-06	4.9078E-06	Tarsus
3L	17934562	CG43173/4	intron	2.0164E-06	1.0248E-05	1.8813E-06	2.0016E-05	Tarsus
1	16033169	disco-r	intron	9.6851E-06	0.00041928	1.1066E-05	1.1579E-06	Tarsus
2L	1644332	chinmo	7kbUS	2.9778E-05	2.5557E-06	0.00038901	1.4339E-05	Tarsus
2L	8037802	CG7367; Cka	intron; CDS	1.9688E-05	8.0011E-05	1.593E-05	1.9973E-05	Tarsus
3L	4839460	CG13707	intron	1.7343E-05	8.755E-05	1.2102E-05	3.5277E-05	Tarsus
3L	4291273	CG15012; Hexo1	3'UTR; 60bpDS	1.5209E-05	1.0519E-05	0.00014965	5.6546E-05	Tarsus
2R	18561440	RYBP	700bpDS	5.4674E-05	0.00011744	2.384E-05	1.114E-05	Tarsus
3L	18897328	CSN1b	CDS	0.0003216	0.00128683	1.6292E-06	5.3113E-06	Tarsus
1	7962846	mys	intron	1.8522E-05	9.5452E-06	0.00074614	2.8078E-05	Tarsus
3L	18898485	CG6841	5'UTR	0.00039964	0.00188576	1.5055E-06	4.7842E-06	Tarsus
3L	10628517	E(z)	CDS	0.00018675	5.2229E-05	4.6703E-05	1.4759E-05	Tarsus
3L	1079042	bab1	intron	6.2107E-05	5.2294E-05	0.00016025	1.4412E-05	Tarsus
2R	16131320	CG11208	CDS	0.00051985	7.9403E-05	1.3399E-05	1.5017E-05	Tarsus
3L	18901900	IR75d; CG14077	intron; intron	0.00051194	0.00160653	4.1264E-06	2.7945E-06	Tarsus
3L	2995056	CG2113	100bpUS	5.0753E-05	1.7827E-05	0.00018043	6.024E-05	Tarsus
2R	20188513	CG3394	3kbDS	7.9437E-06	0.00026184	0.0001221	3.9494E-05	Tarsus
2R	16130908	CG11208	CDS	0.00024044	0.00012171	2.339E-05	1.7541E-05	Tarsus
1	15999736	CG42353/4	800bpUS	4.8046E-05	0.00020418	3.3407E-05	4.998E-05	Tarsus
3L	13400560	CG10222	4kbUS	9.3488E-05	3.935E-05	1.0957E-05	0.00040675	Tarsus

Some candidate genes we identified were already known to be involved in leg development, including *Moesin* (*Moe*) (Speck, Hughes et al. 2003), *disco-r* (Patel, Farzana et al. 2007), *Enhancer of zeste* (*E(z)*) (Shearn, Hersperger et al. 1978; Phillips and Shearn 1990), and *bric à brac 1* (*bab1*) (Godt, Couderc et al. 1993). Identification of these genes helped to validate our methods, but they were not examined further in this study, with the exception of *disco-r*, for which there is no established mutant phenotype. Other genes were identified with known functions but their role in leg development was unclear or had not been examined. These were also selected as candidate genes to determine their role in leg development. In addition, several genes were identified that had not been previously characterized.

Expression of candidates during Drosophila development

In order to begin to determine if our candidate genes played a role in leg development, we first examined distribution of mRNA encoded by these genes using *in situ* hybridization, looking for expression in the developing imaginal discs of larvae or the forming disc primordia in embryos. Results of our *in situ* study are summarized in Table 5. Of the 38 examined genes, only eight had no staining in the leg imaginal discs (*CG5549*, *mspo*, *dpr8*, *CG33218*, *CG7949*, *CAH2*, *CG42353*, and *CG42354*). These genes also had no expression in the wing, haltere or eye-antennal imaginal discs. Of these, only *mspo* and *CG7949* had embryonic expression. *CG794*9 was expressed in the early mesoderm while *mspo* was primarily in the developing nervous system, but no expression of either was seen in the disc

TABLE 5: Observed expression of analyzed candidates

Genes are organized by segment of association and significance of candidate.

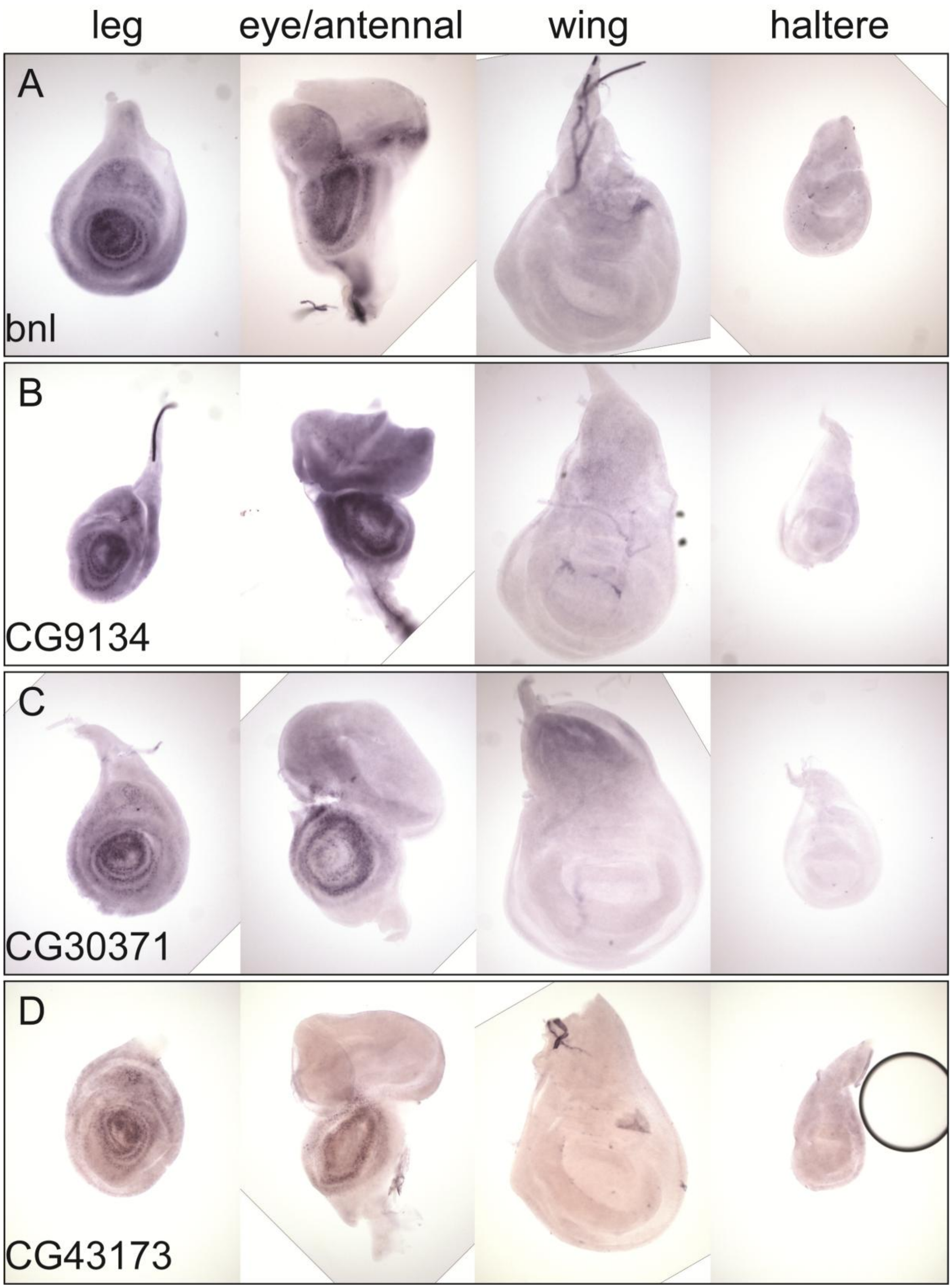
Gene	Expression in embryos	Expression in discs	Associated segment
CG3961	Patterned	Ubiquitous in eye/antenna, leg, wing and haltere discs	Femur
boi	No expression	Ubiquitous in eye/antenna, leg, and haltere discs; Patterned in wing discs	Femur
CG43444	Patterned	Ubiquitous in antenna, leg, wing and haltere discs; Patterned in eye disc	Femur
CG5549	No expression	No expression	Femur
ytr	Patterned	Ubiquitous in eye/antenna, leg, wing and haltere discs	Femur
eIF6	Patterned	Ubiquitous in eye/antenna, leg, wing and haltere discs; enhanced in ventral discs	Femur
mspo	Patterned	No expression	Femur
dpr8	No expression	No expression	Femur
CG18317	Ubiquitous	Ubiquitous in eye/antenna, leg, wing and haltere discs	Femur
GlcT-1	Patterned	Ubiquitous in eye/antenna, leg, wing and haltere discs	Femur
CG7949	Patterned	No expression	Femur
CG33218	No expression	No expression	Femur
CAH2	No expression	No expression	Femur
CG10947	No expression	Ubiquitous in eye/antenna, leg, wing and haltere discs	Femur
sick	Ubiquitous	Ubiquitous in eye/antenna, leg, wing and haltere discs	Femur
PI31	No expression	Patterned in antenna and tarsus segment of leg discs	Femur
CG9129	No expression	Ubiquitous in eye/antenna, leg, wing and haltere discs	Tibia
CG42353	No expression	No expression	Tibia
CG42354	No expression	No expression	Tibia
CG7695	No expression	Ubiquitous in eye/antenna, leg, wing and haltere discs	Tibia
CG32333	No expression	Ubiquitous in eye/antenna, leg, wing and haltere discs; enhanced in ventral discs	Tibia
CG30371	No expression	Patterned in antenna and tarsus segment of leg discs	Tibia
CG8486	Ubiquitous	Ubiquitous in eye/antenna, leg, wing and haltere discs; enhanced in ventral discs	Tibia
CG9134	Patterned	Patterned in eye/antenna and tarsus segment of leg discs	Tibia
bnl	Published	Patterned in antenna, leg, and wing discs	Tibia
ckn	Ubiquitous	Ubiquitous in eye/antenna, leg, wing and haltere discs	Tarsus
CG43173	No expression	Patterned in tarsus segment of leg discs and antenna	Tarsus
CG43174	No expression	Ubiquitous in eye/antenna, leg, wing and haltere discs	Tarsus
disco-r	Patterned	Patterned in tarsus segment of leg discs and antenna	Tarsus
chinmo	No expression	Ubiquitous in eye/antenna, leg, wing and haltere discs	Tarsus
Cka	No expression	Ubiquitous in eye/antenna, leg, wing and haltere discs	Tarsus
CG13707	No expression	Ubiquitous in eye/antenna, leg, wing and haltere discs	Tarsus
CG15012	Ubiquitous	Ubiquitous in eye/antenna, leg, wing and haltere discs	Tarsus
RYBP	Ubiquitous	Ubiquitous in eye/antenna, leg, wing and haltere discs	Tarsus
CSN1b	Ubiquitous	Ubiquitous in eye/antenna, leg, wing and haltere discs	Tarsus
mys	Patterned	Ubiquitous in eye/antenna, leg, wing and haltere discs	Tarsus
Hexo1	Ubiquitous	Ubiquitous in eye/antenna, leg, wing and haltere discs	Tarsus
CG6841	Ubiquitous	Ubiquitous in eye/antenna, leg, wing and haltere discs	Tarsus

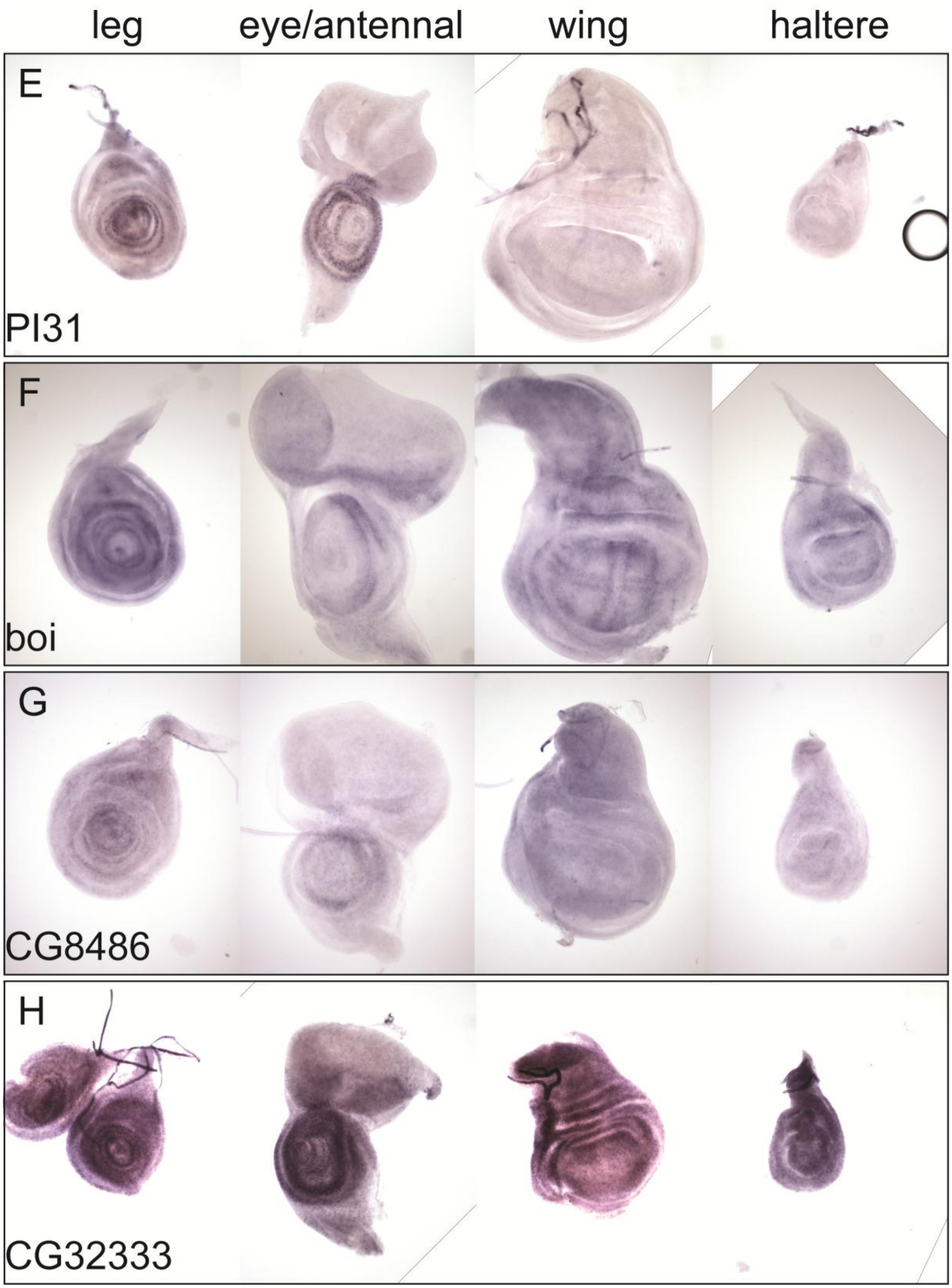
primordia. Since none of these eight were expressed at any observed point in the developing leg, they were excluded from further analysis.

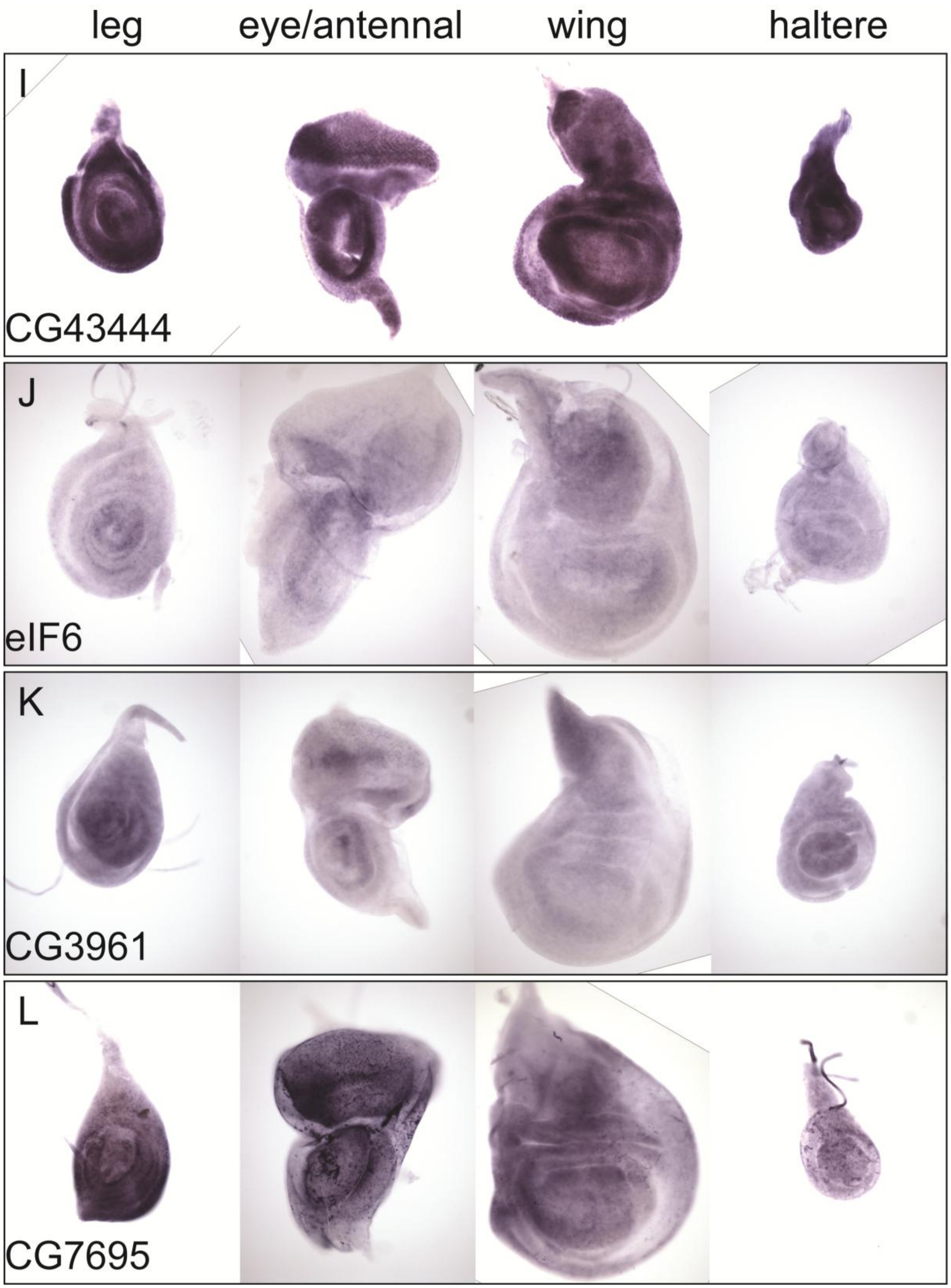
Of the genes that did have expression in the imaginal discs, the staining patterns could be grouped into three categories (Fig 7). Some genes were expressed only or mostly in the antennal and leg discs, with little or no staining in the eye, wing and haltere discs; leg expression was predominantly limited to the presumptive tarsus (Fig 7A-E). These genes include *CG9134, CG30371*, *CG43173, bnl*, and *PI31*. Other genes were expressed in all examined discs, with enhancements in one or more disc types (Fig 7F-J). Of these genes, *CG8486*, *CG32333* and *eIF6* had enhanced expression in the antenna and the presumptive tarsus of the leg. This category also contains *CG43444,* which had enhanced expression in the developing omatidia of the eye, and *boi*, which had enhanced expression in a cross pattern in the future wing blade (which appeared to outline the expression domains of *wg* and *dpp* in this region). The remaining genes had relatively even expression in all discs examined (Fig 7K-S). Of the 30 genes which had expression in the imaginal discs, a little over half (17) were observed to have embryonic expression (examples in Fig 8). Eight of these genes had patterned embryonic expression. The embryonic expression of *bnl* has been previously published (Sutherland, Samakovlis et al. 1996). Expression of *CG9134* was first observed in the pharynx floor, and expression in the hindgut developed later. *CG4344* and *ytr* had patterned expression in the developing nervous system. Expression of *CG3961* was limited to germ cells in early embryos, but this expression broadened to include most of the mesoderm in later stages, when enhanced expression can be seen in the presumptive optic lobe. Initial *GlcT-1* expression was observed in the early embryonic head, which was replaced by broad

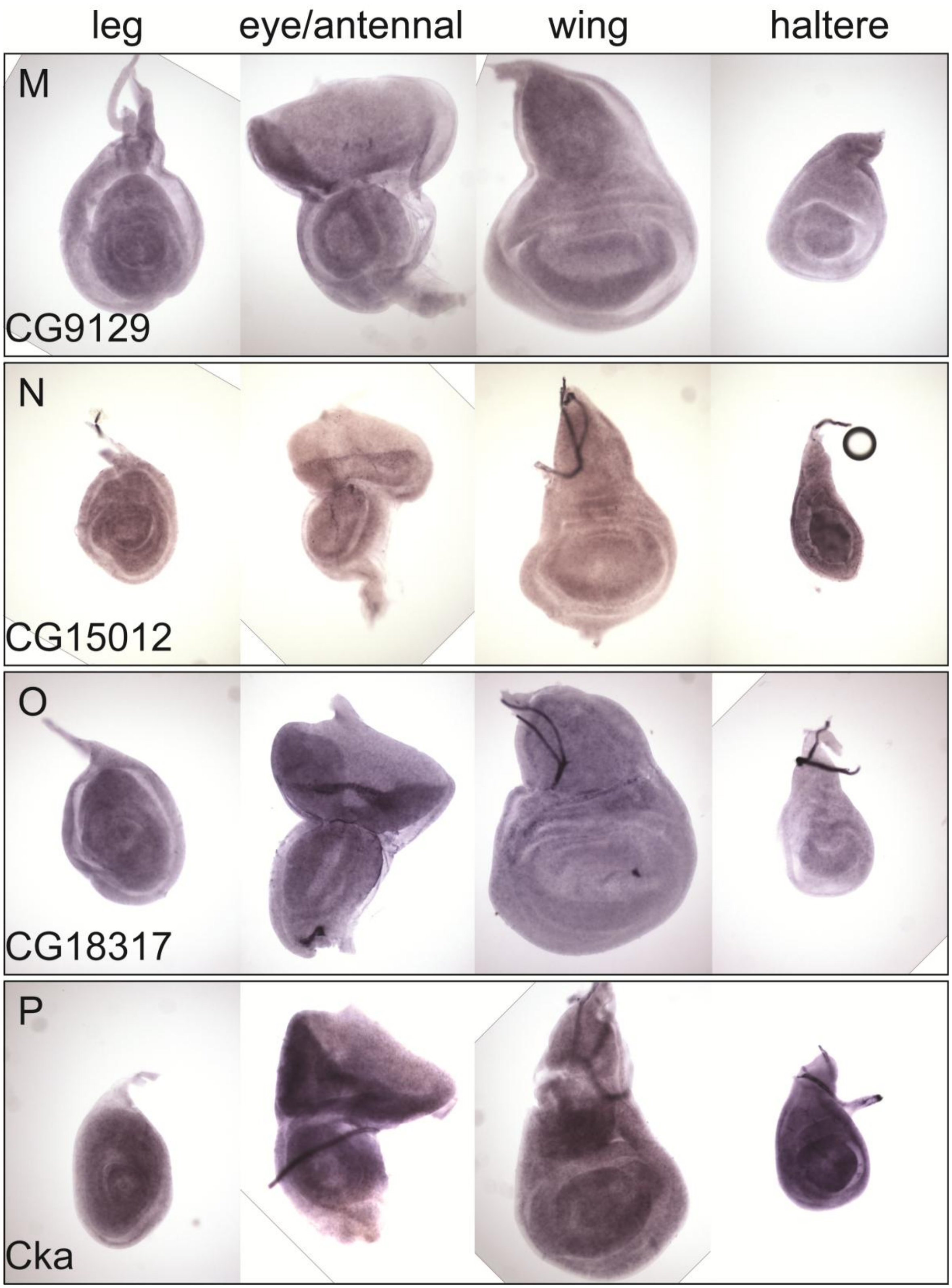
FIGURE 7: Expression of candidate genes in imaginal discs

Wild-type third instar imaginal discs stained for expression of the labeled candidate gene. Expression patterns can be grouped into three large categories. A-E are genes with little-to-no expression in the wing and haltere discs, but strong expression in the antennae and center leg discs. F-J are genes with some degree of expression in all discs, but with enhancements in one or more disc; all but *boi* (F – enhancement in ‘cross’ pattern in future wing blade) and *CG43444* (I – enhancement in ommatidia of eye), have enhancements in the antenna and central leg disc. K-S are genes with relatively even expression in all discs (ubiquitous).

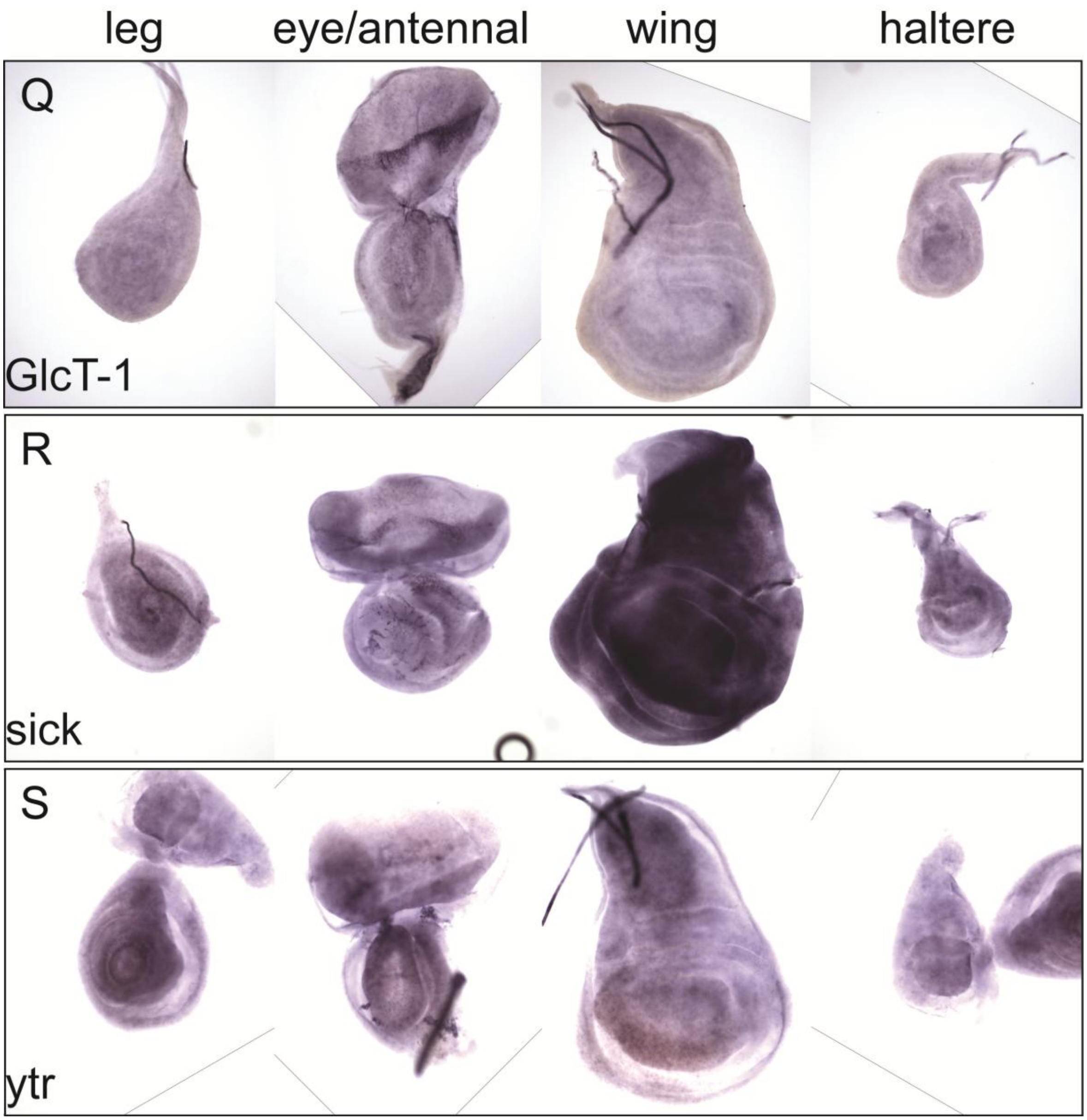

leg
eye/antennal
wing
haltere
A
bnl
B
CG9134
C
CG30371
D
CG43173


leg
eye/antennal
wing
haltere
E
PI31
F
boi
G
CG8486
H
CG32333


leg
eye/antennal
wing
haltere
I
CG43444
J
eIF6
K
CG3961
L
CG7695



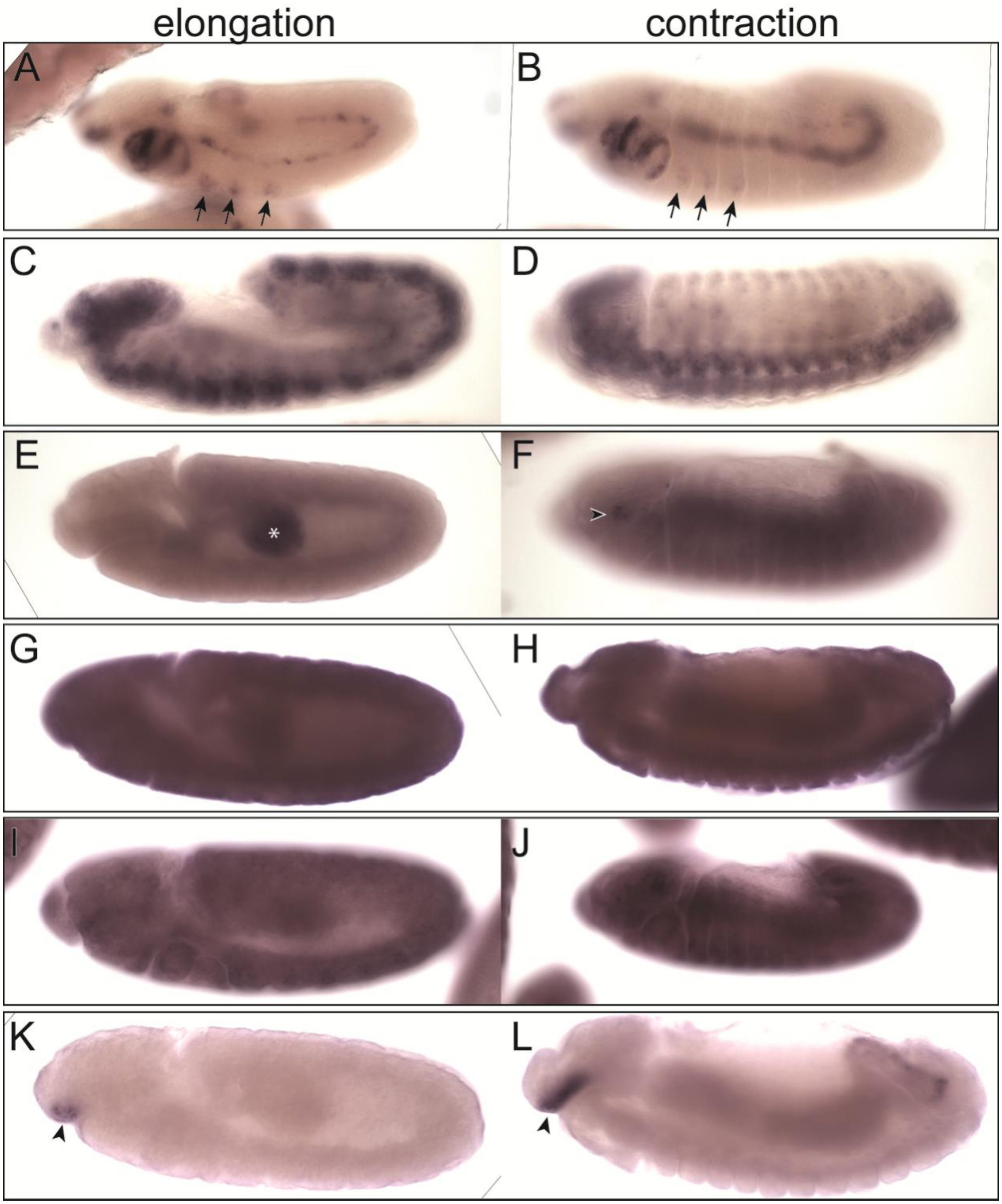
leg
eye/antennal
wing
haltere
M
CG9129
N
CG15012
O
CG18317
P
Cka

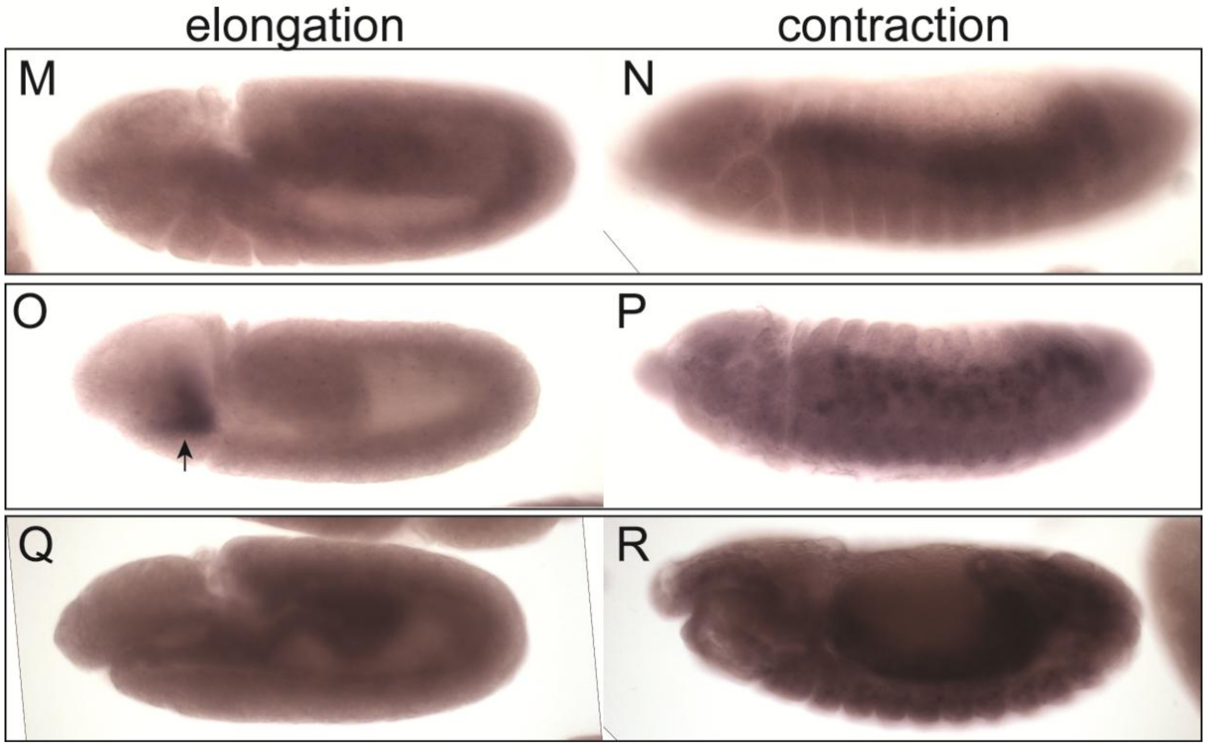


leg
eye/antennal
wing
haltere
Q
GlcT-1
R
sick
S
ytr

FIGURE 8: Expression of candidate genes in embryos

Wild-type embryos stained for expression of certain candidate genes. Two stages are shown for each, an early stage during or near the end of elongation, and a later stage during or near the end of contraction. Expression patterns varied, but, with the exception of *disco-r* (A-B; previously published in (Mahaffey, Griswold et al. 2001)) we did not identify expression of any candidate specifically in the leg disc primordia (arrows in A-B). C-D) *CG43444*, expression in the developing nervous system. E-F) *CG3961*, expression limited to germ cells early (asterisk) but broadened in later stages and enhanced in the presumptive optic lobe (arrowhead). G-H) *CG6841*, ubiquitous. I-J) *CG8486*, ubiquitous with enhancements in nervous system. K-L) *CG9134*, expression in the pharynx floor (arrowheads), later in the hindgut. M-N) *eIF6*, expression in mesoderm. O-P) *GlcT-1*, early expression in head segments (arrow), later expression in mesoderm and developing gut. Q-R) *mys*, ubiquitous expression with enhancements presumably in developing muscle tissue.


elongation
contraction
A
B
C
D
E
F
G
H
I
J
K
L


elongation
contraction
M
N
O
P
Q
R

mesodermal domain that resolved into spots of expression throughout the developing gut. Expression of *eIF6* was observed in the mesoderm of embryos. None of the patterned genes appeared to have expression in the developing disc primordia, as has been shown with *disco-r* (Fig 8A and (Mahaffey, Griswold et al. 2001)). The other nine genes (*CG15012*, *CG18317*, *CG12424*, *CG8486*, *CG6841*, *sick*, *RYBP*, *CSN1b*, and *Hexo*) were expressed ubiquitously.

RNAi Phenocopies in Leg Development

Candidate genes with mRNA detected in the developing fly leg discs were further analyzed for function in leg development by perturbing gene function using the Gal4 system to drive RNAi. The results of these experiments are described in Table 6. No RNAi lines were available for four genes (*ytr*, *GG43173*, *CG43174* and *Hexo1*).

RNAi with 11 candidates caused no discernible phenocopy with any driver. One of these genes was *brother of ihog* (*boi*), which shares some functions with a paralog, *interference hedgehog* (*ihog*) (Yao, Lum et al. 2006; Camp, Currie et al. 2010; Yan, Wu et al. 2010). Since redundant paralogs might prevent developmental abnormalities caused by RNAi, we examined the remaining candidates for potential paralogs (Table 7). However, only one other candidate that had no phenocopy with RNAi, *CG18317*, appeared to have a potential paralog in the genome. Another such gene, *CG3961*, did not have any clear paralogs, but did share common functional domains with other proteins. These proteins may provide redundant functions that account for the lack of phenocopy for this candidate.

Six genes yielded altered legs when expression was reduced using one or more of the Gal4 drivers. *Dll*-Gal4-driven RNAi caused more leg defects than were seen with the other

TABLE 6: Phenocopy from RNAi of listed gene driven by listed Gal4 driver

Gene	arm-GAL4	Dll-GAL4	Associated segment
CG3961	No effect	No effect	Femur
boi	No effect	No effect	Femur
CG43444	No effect	No effect	Femur
eIF6	No effect	Pupal semi-lethal; tarsi deleted and sex combs reduced	Femur
CG18317	No effect	No effect	Femur
GlcT-1	No effect	No effect	Femur
CG10947	No effect	No effect	Femur
sick	No effect	Lethal	Femur
Pl31	No effect	Pupal semi-lethal; legs missing one tarsal segment, arista deleted/reduced, wings hooded	Femur
CG9129	No effect	Antenna to leg	Tibia
CG7695	No effect	No effect	Tibia
CG32333	No effect	Antenna to leg	Tibia
CG30371	No effect	Antenna to leg	Tibia
CG8486	No effect	Antenna to leg	Tibia
CG9134	No effect	No effect	Tibia
bnl	Early larval lethal	No effect	Tibia
ckn	No effect	No effect	Tarsus
disco-r	No effect	antenna to leg	Tarsus
chinmo	Lethal	Lethal	Tarsus
Cka	Lethal	Tarsus deleted (only a nub of T1 is left), arista and antenna segment 3 deleted, wings hooded	Tarsus
CG13707	No effect	No effect	Tarsus
CG15012	No effect	No effect	Tarsus
RYBP	No effect	No effect	Tarsus
CSN1b	No effect	Legs and sex combs reduced at RT and tarsi variably deleted at 25C, arista reduced, wings hoo	Tarsus
mys	Embryo lethal	Lethal	Tarsus
CG6841	No effect	Pupal lethal; tarsus only have a few, fused segments; antennae reduced	Tarsus

TABLE 6: continued

Gene	dpp-GAL4	rn-GAL4	tsh-GAL4	Associated segment
CG3961	No effect	No effect	No effect	Femur
boi	No effect	No effect	No effect	Femur
CG43444	No effect	No effect	Wings held out	Femur
eIF6	No effect	Pupal semi-lethal; wings tiny and shriveled	Lethal	Femur
CG18317	No effect	No effect	No effect	Femur
GlcT-1	No effect	No effect	No effect	Femur
CG10947	No effect	No effect	No effect	Femur
sick	No effect	Wings variably serrate; sex combs affected	Pupal lethal	Femur
PI31	No effect	Shriveled wings	Pupal semi-lethal	Femur
CG9129	No effect	Wings shortened and hooded	No effect	Tibia
CG7695	No effect	No effect	No effect	Tibia
CG32333	No effect	Wings shortened and hooded	No effect	Tibia
CG30371	No effect	Wings shortened and hooded	No effect	Tibia
CG8486	No effect	Wings shortened and hooded	No effect	Tibia
CG9134	No effect	No effect	No effect	Tibia
bnl	No effect	No effect	Lethal	Tibia
ckn	No effect	No effect	No effect	Tarsus
disco-r	No effect	Wings shortened and hooded	No effect	Tarsus
chinmo	No effect	Very few progeny, but no obvious effect	Lethal	Tarsus
Cka	Legs shortened	Wings kinked	Lethal	Tarsus
CG13707	No effect	No effect	No effect	Tarsus
CG15012	No effect	No effect	No effect	Tarsus
RYBP	No effect	No effect	No effect	Tarsus
CSN1b	No effect	Wings shriveled	Lethal	Tarsus
mys	Trouble eclosing and walking	Semi-lethal; wings shriveled	Lethal	Tarsus
CG6841	No effect	Pupal semi-lethal; wings severely reduced/deleted, 3 tarsi missin	Pupal lethal	Tarsus

TABLE 7: Genes analyzed for RNAi effect, and the possibility of redundant paralogs

Observation of phenocopy is recorded simply as ‘Yes’ or ‘No’. The five candidates that gave antenna-to-leg transformation with the *Dll* driver are marked with an asterisk. Genes with potential paralogs are labeled as “Probable.” Genes with known redundant paralogs are labeled as “Yes” along with the name of the paralog. Genes are organized by segment of association and significance of candidate.

Gene	Phenocopy effect	Paralogs?	Associated segments
CG3961	No	May share functions with non-paralogs	Femur
boi	No	Yes - ihog	Femur
CG43444	Yes	No	Femur
eIF6	Yes	No	Femur
CG18317	No	Probable	Femur
GlcT-1	No	No	Femur
CG10947	No	No	Femur
sick	Yes	No	Femur
PI31	Yes	No	Femur
CG9129	Yes*	Probable	Tibia
CG7695	No	No	Tibia
CG32333	Yes*	No	Tibia
CG30371	Yes*	Probable	Tibia
CG8486	Yes*	Probable	Tibia
CG9134	No	No	Tibia
bnl	Yes	No	Tibia
ckn	No	No	Tarsus
disco-r	Yes*	Yes- disco	Tarsus
chinmo	Yes	No	Tarsus
Cka	Yes	No	Tarsus
CG13707	No	No	Tarsus
CG15012	No	No	Tarsus
RYBP	No	No	Tarsus
CSN1b	Yes	No	Tarsus
mys	Yes	Probable	Tarsus
CG6841	Yes	No	Tarsus

TABLE 8: Summary of RNAi results with *arm*-, *dpp*-, and *tsh*-Gal4 drivers

Candidates with no or lethal effects, as well as the phenocopies of exceptions can be found in Table 6. Survivors of *tsh*-driven *PI31* RNAi appeared normal.

Driver	# No Effect	# Lethal	Candidates With Exceptions
arm-GAL4	22	4	none
dpp-GAL4	24	0	Cka, mys
tsh-GAL4	16	8	CG43444, PI31

drivers (Table 6). RNAi of *CSN1b* with this driver resulted in legs of reduced size relative to overall body size, compared to wild-type (Fig 9B). These flies also had a reduction of the antennal arista segment (Fig 12C), and males had reduced sex-combs (Fig 11C). RNAi against *PI31* reduced the relative size of tarsal segments, occasionally deleting one of the middle tarsal segments (Fig 9C), and deleted the arista in the antennae (Fig 12B). Reduction of *eIF6*, *CG6841* or *Cka* expression resulted in fused and deleted tarsi, with the later two also deleting terminal structures, such as the claws and pulvilli (Fig 9D-F). *Dll* driven reduction of *eIF6* mRNA also caused reductions in the sex-comb tooth number (Fig 11B), while RNAi of *CG6841* and *Cka* altered or deleted distal antennal segments (Fig 12D-E).

Using the *rn* driver, RNAi against *CG6841* resulted in major leg defects; tarsal segments were deleted and fused, though terminal structures were still present (Fig 10). Reduction of *sick* expression in *rn*-Gal4 flies resulted in legs which appeared normal in every respect except sex-comb tooth numbers, which were, interestingly, either greatly decreased or increased (Fig 11D-E), suggesting *sick* may play role in controlling sex-comb tooth numbers. RNAi against *sick* driven by *Dll*-Gal4 was lethal prior to pupal formation, so we could not determine if there was an effect on leg development with that driver. RNAi of most of these six candidates, driven by either *rn*- or *Dll*-Gal4, also resulted in altered wing morphology (Fig 13), ranging from a slight reduction in size accompanied by a downward curvature (hooded) to near total loss of wing structures.

With few exceptions, the remaining drivers produced no effect of note or killed affected progeny before or shortly after pupal formation (Table 8). Amongst the exceptions was *CG43444*, which produced no effect when RNAi was driven by any driver except *tsh*.

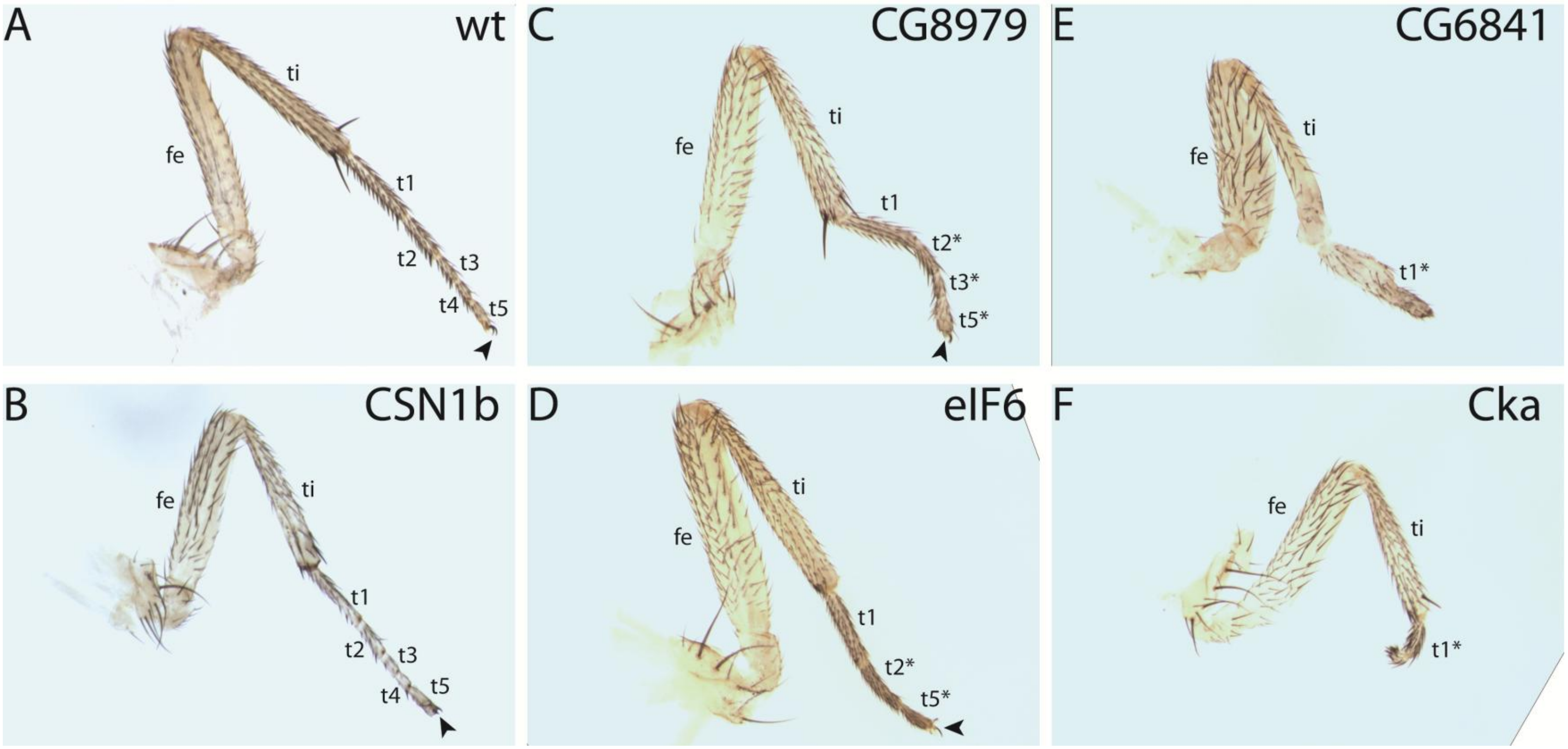
With this driver, RNAi caused adult flies to hold their wings out perpendicular to the body when at rest. Reduction of *PI31* with *tsh*-Gal4 reduced viability, but produced no visible abnormalities in surviving adults.

Reduced expression of three genes (*bnl*, *chinmo* and *mys*) resulted in early lethality with multiple drivers. In crosses with drivers that did not result in total lethality, progeny showed no overt disruptions of leg morphology.

Perhaps interestingly, RNAi of five of the genes we examined (*CG9129*, *CG32333*, *CG30371*, *CG8486* and *disco-r*) all caused a transformation of antennae towards leg identity when RNAi was driven by the *Dll* promoter. These also gave shortened, hooded wings with the *rn* promoter, but no effects were seen with the remaining drivers. None of these showed overt disruptions of leg morphology with any driver (however, see discussion of paralogous genes below).

FIGURE 9: Effect in legs of *Dll*-Gal4 driving RNAi

Second thoracic legs from flies with both *Dll*-Gal4 and UAS-RNAi insertions, labeled by targeted gene, in order of increasing severity. Arrowheads indicate presence of terminal structures, such as the claws and pulvilli. A) Wild-type. B) RNAi of *CSN1b* resulted in reduced leg size, but did not alter shape. C) RNAi of *PI31* resulted in reduction of size of tarsal segments, with occasional loss of a single segment. D) RNAi of *eIF6* resulted in loss and fusion of tarsal segments. E) RNAi of *CG6841* resulted in extreme loss and fusion of tarsal segments, such that one segment remained, missing terminal structures. F) RNAi of *Cka* was similar to that of *CG6841*, but more severe. Abbreviations: fe, femur; ti, tibia; t1-t5, tarsal segments 1-5; an asterisk indicates uncertain identity of segment due to loss and/or fusion.


A
wt
ti
fe
t1
t2
t3
t4
t5
B
CSN1b
fe
ti
t1
t2
t3
t4
t5
C
CG8979
ti
fe
t1
t2*
t3*
t5*
D
eIF6
ti
fe
t1
t2*
t5*
E
CG6841
ti
fe
t1*
F
Cka
fe
ti
t1*

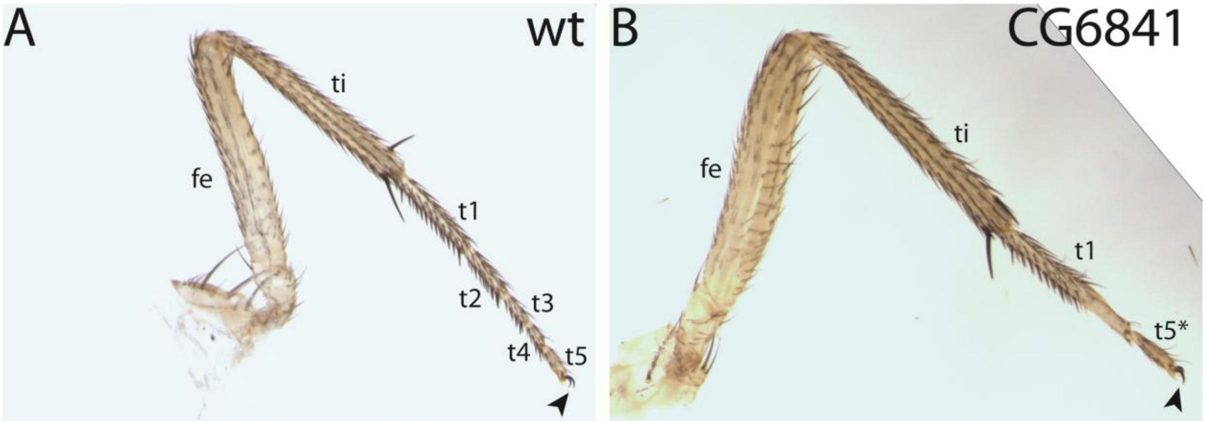


FIGURE 10: Effect in legs of *rn*-Gal4 driving RNAi

Second thoracic legs from flies with both *rn*-Gal4 and UAS-RNAi insertions, labeled by targeted gene. Arrowheads indicate presence of terminal structures, such as the claw and pulvilli. A) Wild-type. B) of *CG6841* resulted in loss and fusion of tarsal segments, such that two segments remained; leg bristles were also affected in the distal t1. Abbreviations: fe, femur; ti, tibia; t1-t5, tarsal segments 1-5; an asterisk indicates uncertain identity of segment due to loss and/or fusion.

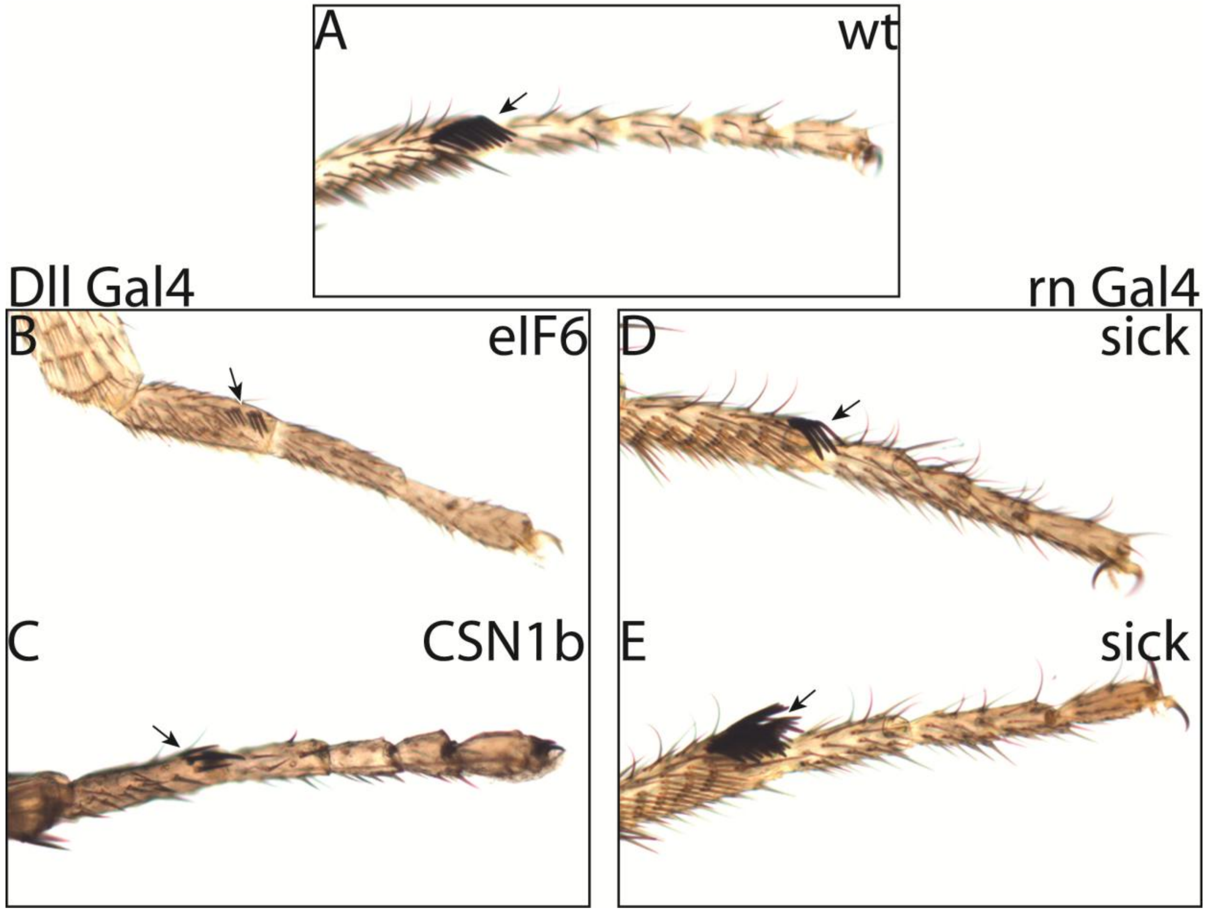


FIGURE 11: Effects of RNAi on development of sex-combs

First thoracic legs from flies with either *Dll*-Gal4 (B-C) or *rn*-Gal4 (D-E) driving RNAi, labeled by targeted gene. Arrows point to sex-combs. A) Wild-type. B-C) *Dll*-Gal4 driven RNAi of *eIF6* (B) and *CSN1b* (C) resulted in reduction of sex-combs. D-E) *rn*-Gal4 driven RNAi of *sick* resulted in either decrease (D) or increase (E) of sex-combs.

FIGURE 12: Effect in antennae of *Dll*-Gal4 driving RNAi

Heads from flies with both *Dll*-Gal4 and UAS-RNAi insertions, labeled by targeted gene, in order of increasing severity. Arrows point to arista, or where arista should have developed. A) Wild-type. B) RNAi of *PI31* resulted in loss of arista; remaining structures appear normal, but did not alter shape. C) RNAi of *CSN1b* resulted in loss of arista, and reduced size and some changes in shape of remaining structures. D) RNAi of *CG6841* resulted in loss of arista, and drastic changes to the appearance of antennal segment 3. E) RNAi of *Cka* resulted in loss of all but proximal structures. F) RNAi of *CG9129* (and four other candidates) resulted in transformation of antennae towards leg identity.

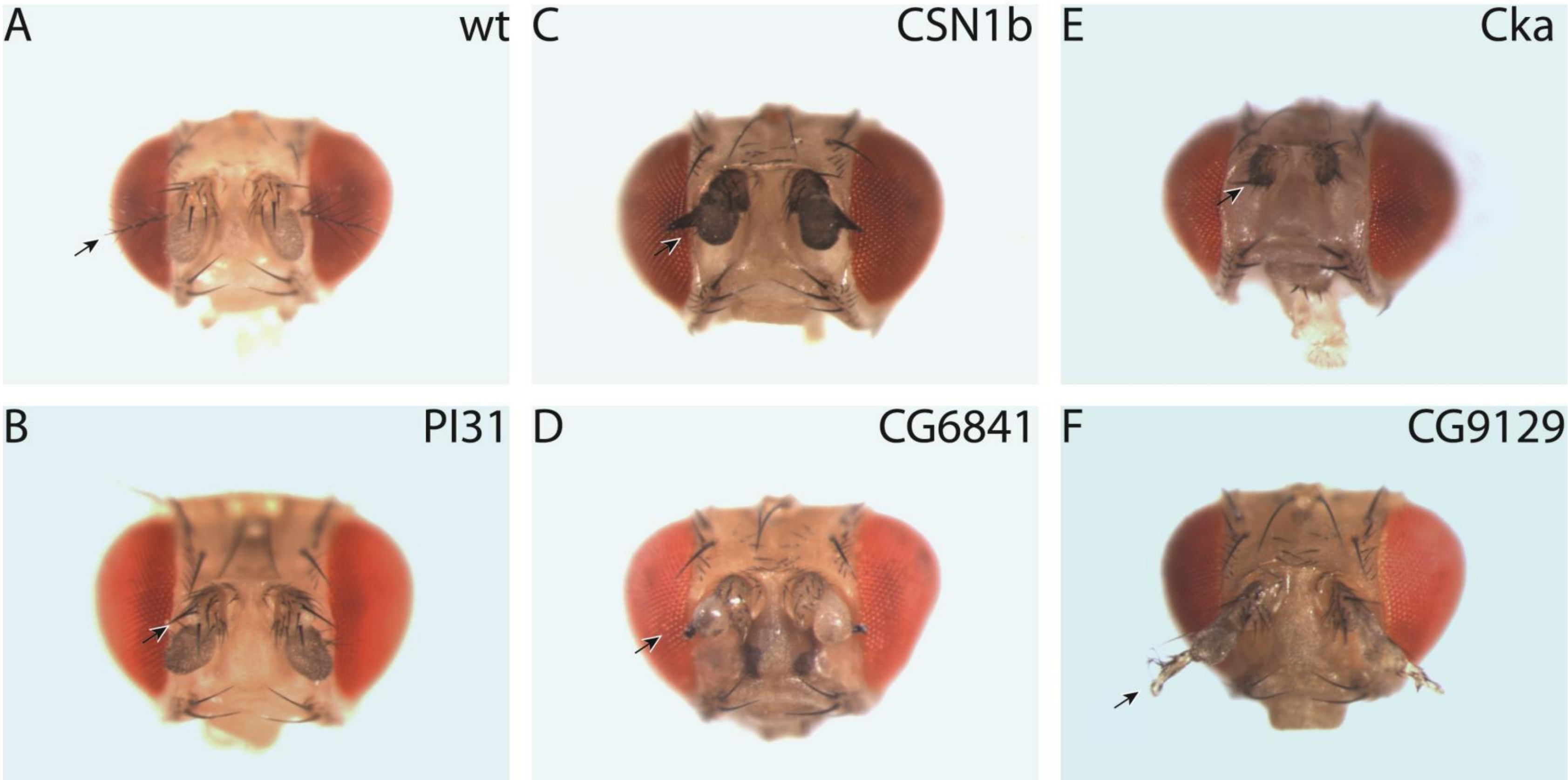
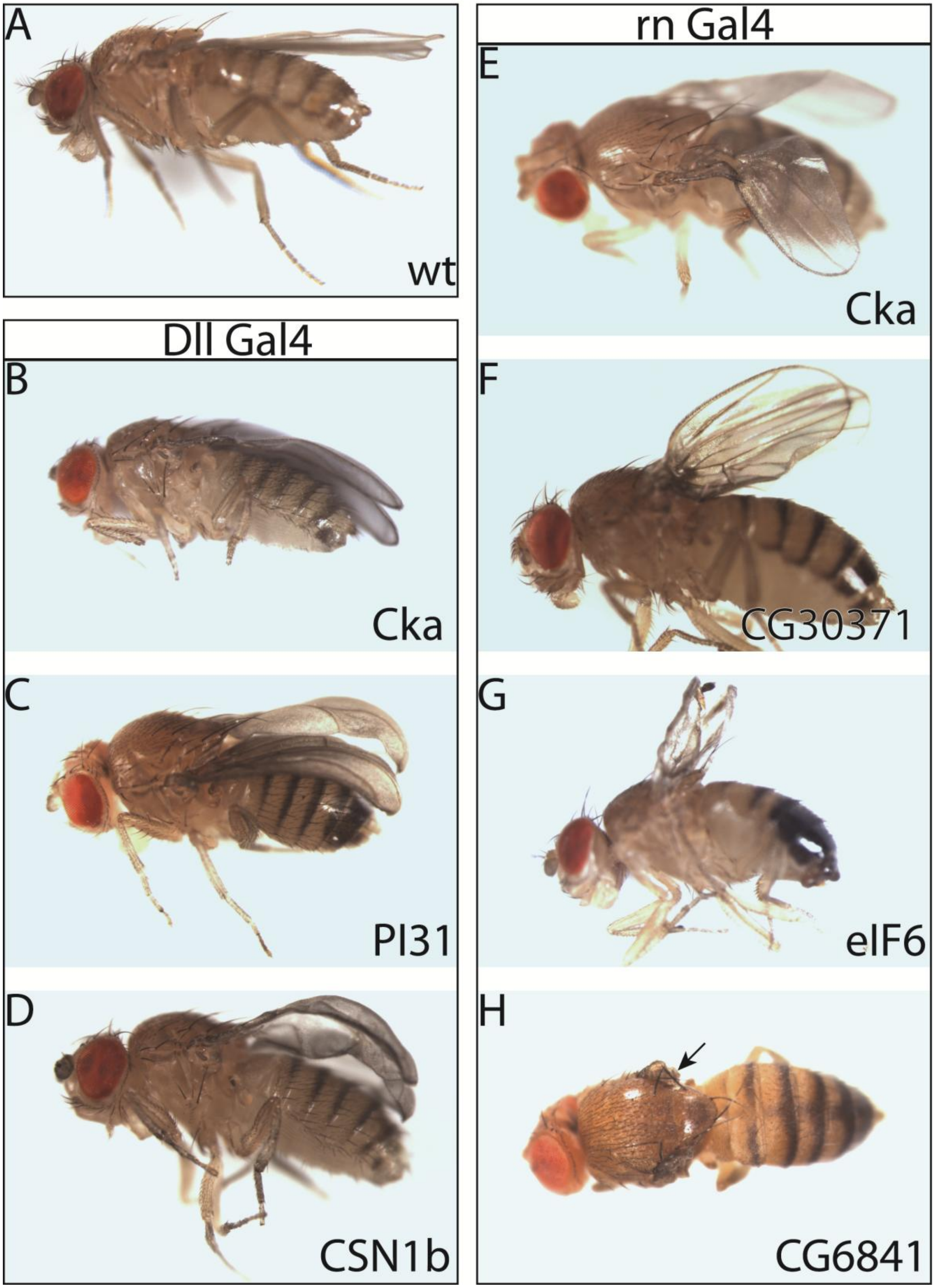

A
wt
C
CSN1b
E
Cka
B
PI31
D
CG6841
F
CG9129

FIGURE 13: Effects of RNAi on development of wings

Wings from flies with either *Dll*-Gal4 (B-D) or *rn*-Gal4 (E-H) driving RNAi, labeled by targeted gene, in order of increasing severity. A) Wild-type. B-D) *Dll*-Gal4 driving RNAi resulted in reduction of wing size, and an alteration of wing shape to be curved downwards (hooded). E) RNAi of *Cka*, driven by *rn*, resulted in a 'kinking' of wing shape. F) RNAi of *CG30371* (and four other candidates), driven by *rn*, resulted in a slight hooding and a reduction of wing size. G) RNAi of *eIF6*, driven by *rn*, resulted in small, shriveled wings. H) RNAi of CG6841, driven by *rn*, resulted in severe reduction of wings; remaining structure where a wing should be is indicated by an arrow.


A
wt
Dll Gal4
B
Cka
C
PI31
D
CSN1b
rn Gal4
E
Cka
F
CG30371
G
eIF6
H
CG6841

DISCUSSION

Using this GWA study, we were able to identify several new leg development genes. Although this technique has been used in the past to attempt to locate causes of heritable diseases or of variation in behavioral and quantitative traits (Mackay, Stone et al. 2009; Visscher, Brown et al. 2012), it has not, to our knowledge, been employed to identify development genes. Our study shows the usefulness of employing this technique to identify developmental genes that have thus far been missed with traditional genetics screens.

Interpreting positive results

RNAi against 15 out of 26 examined candidates produced some disruption of development. Of these, only five (*eIF6*, *PI31*, *Cka*, *CSN1b*, and *CG6841*) produced clear disruptions of leg development, while a sixth, *sick*, disrupted sex-comb development. The roles of these genes in leg development certainly bear further study.

In addition to these, three candidate genes (*bnl*, *chinmo*, and *mys*) were lethal with some drivers, though they yielded no leg abnormalities when expression was reduced. The weakest results came from *bnl*, RNAi of which was only lethal with the *arm* and *tsh* drivers. No other driver caused an effect. Since the other drivers have more restricted expression domains, it is possible that early, broad reduction of *bnl* is sufficient to kill, but the more restricted domains were insufficient to produce an effect. RNAi against *chinmo* was lethal with three drivers, *arm*, *tsh*, and *Dll*. The *rn* driver also seemed to reduce viability, but was not sufficient to produce leg defects in survivors. Like *chinmo*, RNAi of *mys* caused lethality in three drivers, and caused a reduction in viability with the *rn* driver. In this last case, survivors had shriveled wings, but no leg defects were noted. The *dpp* driver also had an effect on the development of *mys* RNAi flies, but the physical appearance of their legs was normal. So it seems that, if these genes do play a role in leg development, then the right balance has not been struck between reductions of mRNA that cause no effect and those that kill the affected flies.

We cannot, yet, dismiss the possibility that some of these positive results may be false positives. In particular, *bnl*, *chinmo*, and *mys* produced no leg abnormalities via RNAi. Although they seem to be important for survival, the lack of leg abnormalities may be because they have no role in leg development. Similarly, some of our other positive candidates may represent a role in tasks like cell survival rather than actually participating in leg pattering. Only further experiments, like those described below, will settle these questions.

Antenna-to-leg transformations

Perhaps one of the more interesting phenocopies we observed was the antenna-to-leg transformation caused by *Dll*-driven RNAi against five candidates (*CG9129*, *CG32333*, *CG30371*, *CG8486* and *disco-r*). It is important to consider that such a transformation is also caused by loss of *Dll* function (Sunkel and Whittle 1987; Cohen and Jurgens 1989), and that the Gal4 insertion yielding the *Dll*-Gal4 line generates a hypomorphic allele of *Dll* (Dong, Dicks et al. 2002; Emerald and Cohen 2004). While *Dll*-Gal4 does not cause such transformations alone at room temperature, using this driver does present a sensitized background where *Dll* expression might be reduced. Indeed, at 29 degrees, some reduction of the arista and final tarsal segment has been noted. Therefore, it is possible that the transformations of the antenna to leg we note are the result of a synergistic effect from the reduction of each candidate in the *Dll*-compromised background. If so, this would suggest that these candidates have some role in the *Dll* pathway, at least with respect to antennal identity specification.

These genes might also play a role in the *Dll* pathway for leg patterning, if the double-compromised system was insufficient to cause morphological abnormalities in the legs. One possible explanation for this insufficiency to cause leg defects would be the presence of redundant paralogs. Indeed, *disco-r*, one of the antenna-to-leg candidates, has a known paralog, *disco* (Mahaffey, Griswold et al. 2001). In support of the possibility that it is the compromised background that allows us to detect the transformation, RNAi against *disco* in the *Dll*-Gal4 background has been independently shown to cause similar transformations (Dey, Zhao et al. 2009), and interactions between *Dll* and *disco* have been previously postulated (Cohen, Wimmer et al. 1991). Loss of both *disco* and *disco-r* is necessary to affect leg development (Patel, Farzana et al. 2007), so compromising one gene, even in a *Dll*-compromised background might be insufficient to cause leg abnormalities while still causing defects in the antenna.

In keeping with this hypothesis, we found that three of the other four candidates that cause the antenna-to-leg transformation also have possible paralogs in the Drosophila genome (Table 7). However, this hypothesis does not address all questions. At least one *disco* gene appears to be necessary for proper antennal development, as well (Patel, Farzana et al. 2007). It might be that the homeotic functions of *Dll* are more sensitive to perturbation than the leg-patterning roles, or that the antennae themselves are more sensitive to RNAi perturbations. To answer these questions, we must first deterimine if these genes are part of the Dll pathway, and if so, where these genes fit into the pathway. Further, we cannot as yet rule out the possibility that there is something present in the background of some of the RNAi lines that is effecting the antenna-to-leg transformation, but the fact that an independent RNAi construct of *disco* (Dey, Zhao et al. 2009) has a similar phenotype is suggestive that this is not simply a background effect. We are continuing to examine this issue.

RNAi of these genes also produced a hooded wing phenotype under the *rn* driver. This is particularly interesting considering that *CG30371* and *disco-r* were reported to not have expression in the developing wing blade. However, closer examination of *disco-r* expression does suggest a domain in the wing discs that appears affected in *rn*-Gal4 RNAi (data not shown). It would also be interesting to determine if this wing defect has any connections to *Dll* or *rn* function in the wing, and what, if any, connections are shared by the wing and antennal roles of these candidates. As with the antenna-to-leg phenocopy, we cannot yet rule out the possibility of some background effect.

Interpreting negative results

For a small number of the candidates, we did not detect expression in our examinations of *in situ* hybridization. These data are corroborated by modEncode data (Roy, Ernst et al. 2010) from FlyBase (McQuilton, St Pierre et al. 2012). However, modEncode data do suggest that, while some of these genes are not expressed in the stages examined, there is expression at other stages, and we cannot rule out the possibility that this expression my play a role in leg development. It is also possible that these candidates were mis-identified because of the proximity of a SNP that was really affecting another, more distant gene, or that those SNPs represent false positives. However, identifying causative SNPs and their impact on development was beyond the scope of this study.

We also obtained negative results in some of our RNAi experiments. Amongst these candidates was *boi*. However, we know that *boi* shares some redundancy with its paralog, *ihog*, and it is likely that this masked any potential phenocopy. Our research identified two other ‘no effect’ candidates that possibly share redundant functions with other genes. These redundancies may account for the ineffectiveness of our attempts at RNAi against these genes. However, the presence of paralogs does not explain all of our candidates that yielded no phenocopy. Using *in situ* hybridization, we examined expression of some of these candidates in the RNAi background. Levels of mRNA appeared reduced, suggesting the RNAi is working. While it is possible that these genes have no effect on leg development or patterning, it is also possible that the RNAi lines selected were simply unable to reduce mRNA levels below that necessary for proper development.

Identifying *boi* as a candidate gene is of particular interest. Boi and Ihog have been identified as being critical for Hedgehog signaling(Lum, Yao et al. 2003; Yao, Lum et al. 2006; Camp, Currie et al. 2010; Yan, Wu et al. 2010; Zheng, Mann et al. 2010), a pathway known to be involved in Drosophila leg development. These proteins bind Hh, facilitating activation of the Hh signaling pathway. Several papers have examined Boi and Ihog in biochemical terms, and have suggested a role for them in Drosophila wing development, yet none have tied them directly to leg development (McLellan, Yao et al. 2006; Yao, Lum et al. 2006; Zhang, McLellan et al. 2007; McLellan, Zheng et al. 2008; Camp, Currie et al. 2010; Yan, Wu et al. 2010; Zheng, Mann et al. 2010). The role of *boi* in Hh signaling could easily provide an explanation for how changes in *boi* expression or function could alter leg segment proportions. In particular, it has been suggested that Boi and Ihog act to limit the range of Hh (Yan, Wu et al. 2010; Zheng, Mann et al. 2010). Any change that broadens or narrows the range of effect of Hh could, in turn, broaden or narrow the expression domains of *wg* or *dpp*, altering the domains of expression of the transcription factors responsible for P-D patterning (Fig 1). While this potential connection to the leg development model makes *boi* an intriguing candidate to pursue, reduction of *boi* expression caused no phenocopy in our RNAi experiments. It is likely that redundancy with active *ihog* was sufficient to prevent any gross abnormalities in leg development. However, this redundancy does not disqualify *boi* as a candidate for altering leg proportions. It is possible that subtle tweaks in *boi* expression or function could have effects on leg patterning without eliciting a compensatory response from *ihog*. Increased Dicer may help to determine if there is a role for *boi* in leg development. However, it is also likely that simultaneous RNAi of both *boi* and *ihog* may be necessary before clear defects can be observed. Mutant alleles are also available (Zheng, Mann et al. 2010), making clonal analysis a possibility as well.

Functions of candidates

The known or predicted functions of candidates that produced an RNAi phenocopy are summarized in Table 9. Although some candidates, such as *mys*, are members of pathways

TABLE 9: Protein domains and functions of candidates with RNAi phenocopy

Genes are organized by segment of association and significance of candidate. boi was included because of its known redundancy with ihog.

Gene Name	Protein Domains and Functions	Associated segment
boi	Hedgehog pathwhay member with Fibronectin and Immunoglobulin domains	Femur
CG43444	Zinc finger, CXXC-type.	Femur
eIF6	Translation initiation factor IF6	Femur
sick	involved in defense response to Gram-negative bacterium.	Femur
PI31	involved in response to DNA damage stimulus	Femur
CG9129	Unknown molecular function and processes.	Tibia
CG32333	Unknown molecular function and processes.	Tibia
CG30371	Possibly posseses serine-type endopeptidase activity	Tibia
CG8486	Unknown molecular function and processes.	Tibia
bnl	Invovled in tracheal system development	Tibia
disco-r	Zinc finger, C2H2-type	Tarsus
chinmo	C2H2 zinc finger domain and BTB/POZ domain.	Tarsus
Cka	Involved in dorsal closure and positive regulation of JNK cascade	Tarsus
CSN1b	proteasome, transcription repressor DNA-binding.	Tarsus
mys	Epidermal growth factor protein	Tarsus
CG6841	RNA-processing	Tarsus

with known connections to leg development, the rest are more unclear. The variety of functions, including DNA damage repair, transcription, RNA processing, translation, and immune response, are not easily connected directly to the model of leg development. Indeed, some have no predicted function. Each of these genes will require further examination before their place in the model is understood.

Future directions

We plan to continue examining most candidates in order to further confirm their role in leg development and to determine how they function in this process. The genes *eIF6*, *PI31*, *Cka*, *CSN1b*, *CG6841* and *sick* altered leg morphology, but we do not yet know if this altered morphology is accompanied by altered expression of known leg patterning genes. Using RNAi larvae, we hope to determine if reduction of candidate mRNA results in such changes. This would help to fit our candidates into this developmental pathway. Obtaining mutants of candidate genes would simplify future studies.

Three candidate genes (*bnl*, *chinmo*, and *mys*) were lethal with some, but not all drivers. Most of these lethalities were early, so we cannot say for certain if leg development would have been affected. Examination of embryonic leg markers (like *Dll* or *disco*) might yield some information. Delaying the expression of the Gal4 driver (particularly *Dll*, with *chinmo* and *mys* RNAi lines) by growing early stages at cooler temperatures, or by the use of the Gal4 inhibitor, Gal80, might prove sufficient to overcome any lethal effects, allowing for observation of the morphology of adult structures. Increasing Dicer with the non-lethal drivers may also offer some answers with these three candidates. In addition, driving expression of extra Dicer in our ‘no effect’ lines may be more effective in cases where mRNA was not sufficiently reduced by RNAi to produce a phenocopy.

We are also working to clarify the roles and functions of the antenna-to-leg candidates. Since most of these genes appeared to have a redundant paralog in the Drosophila genome, simultaneous RNAi against each candidate and its paralog may produce phenocopy effects that clarify if these genes actually have a role in leg development.

In addition to examining candidates in more depth, we are currently working to expand the scope of this study by adding measurements from the third thoracic leg (T3) to our data. With this information we hope to accomplish two goals. The first is to create a more robust study like the one presented above. By adding T3 data, it is our hope to increase the significance of SNPs associating with general leg development genes while reducing the significance of false positives. The second goal is to improve our ability to identify segment-specific leg development genes. Each of the three thoracic legs is slightly different, and these differences are associated with the expression of different Hox genes in each of these segments. Hox genes are transcription factors that provide segment identity along the anterior-posterior axis of the Drosophila body by activating targets in a segment-specific manner. This results in segment-specific changes in size, shape, and function of segmentally repeated structures, such as appendages (Mann and Carroll 2002; Pearson, Lemons et al. 2005). For example, it has been shown that the Hox gene *Ubx* is responsible for the differences between T3 and the other legs. Loss of Ubx in T3 alters the size and shape of the leg to resemble T2, while ectopic expression in T1 and T2 can alter the shapes of parts of these legs to resemble T3 (Stern 2003). Yet, each of these legs is formed by the same leg-patterning process. Questions remain about what Hox targets might act to alter the shape and size of each thoracic leg type, and how those targets interact with the leg-patterning system. By using GWAS, we hope to identify such genes by identifying candidates associated with SNPs that contribute to differences in proportion amongst the three thoracic legs.

Success of GWAS for identifying development genes

As we have shown, GWA studies are a useful method for identifying genes contributing to a measurable developmental trait. In addition to identifying new leg development genes, we believe we have also shown that this method can overcome some of the limitations, such as redundancy, presented by traditional developmental screens. Four of the five antenna-to-leg candidates possess potential paralogs in Drosophila. Yet, this phenocopy was only observed in the *Dll*-compromised background. If this phenocopy is the result of the RNAi, it is unlikely that these genes would have been identified in a mutagenesis screen, or even an RNAi screen (depending on the drivers used), due to the presence of the paralogs. However, these genes were highlighted in the GWAS, revealing a possible strength of this method as a tool in developmental genetics.

We believe this technique could be used for any measurable developmental trait. Even well studied developmental systems have open questions (ex: heart development (Medioni, Senatore et al. 2009; Munoz-Chapuli and Perez-Pomares 2010)). Using GWA studies could begin to provide answers and new areas of exploration in these systems.

ACKNOWLEDGEMENTS

We thank the TRiP at Harvard Medical School (NIH/NIGMS R01-GM084947) for providing the transgenic RNAi fly stocks used in this study.

References:

Altschul, S. F., W. Gish, et al. (1990). "Basic local alignment search tool." J Mol Biol **215**(3): 403-10.

Altschul, S. F., T. L. Madden, et al. (1997). "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs." Nucleic Acids Res **25**(17): 3389-402.

Ayroles, J. F., M. A. Carbone, et al. (2009). "Systems genetics of complex traits in Drosophila melanogaster." Nat Genet **41**(3): 299-307.

Briscoe, J. and P. Therond (2005). "Hedgehog signaling: from the Drosophila cuticle to anti-cancer drugs." Dev Cell **8**(2): 143-51.

Calleja, M., E. Moreno, et al. (1996). "Visualization of gene expression in living adult Drosophila." Science **274**(5285): 252-5.

Camp, D., K. Currie, et al. (2010). "Ihog and Boi are essential for Hedgehog signaling in Drosophila." Neural Dev **5**: 28.

Cohen, B., E. A. Wimmer, et al. (1991). "Early development of leg and wing primordia in the Drosophila embryo." Mech Dev **33**(3): 229-40.

Cohen, S. M. and G. Jurgens (1989). "Proximal-distal pattern formation in Drosophila: cell autonomous requirement for Distal-less gene activity in limb development." EMBO J **8**(7): 2045-55.

Dey, B. K., X. L. Zhao, et al. (2009). "Mutual regulation of the Drosophila disconnected (disco) and Distal-less (Dll) genes contributes to proximal-distal patterning of antenna and leg." Cell Tissue Res **338**(2): 227-40.

Dietzl, G., D. Chen, et al. (2007). "A genome-wide transgenic RNAi library for conditional gene inactivation in Drosophila." Nature **448**(7150): 151-6.

Dong, P. D., J. S. Dicks, et al. (2002). "Distal-less and homothorax regulate multiple targets to pattern the Drosophila antenna." Development **129**(8): 1967-74.

Emerald, B. S. and S. M. Cohen (2004). "Spatial and temporal regulation of the homeotic selector gene Antennapedia is required for the establishment of leg identity in Drosophila." Dev Biol **267**(2): 462-72.

Estella, C., R. Voutev, et al. (2012). "A dynamic network of morphogens and transcription factors patterns the fly leg." Curr Top Dev Biol **98**: 173-98.

Godt, D., J. L. Couderc, et al. (1993). "Pattern formation in the limbs of Drosophila: bric a brac is expressed in both a gradient and a wave-like pattern and is required for specification and proper segmentation of the tarsus." Development **119**(3): 799-812.

Kojima, T. (2004). "The mechanism of Drosophila leg development along the proximodistal axis." Dev Growth Differ **46**(2): 115-29.

Lecuit, T. and S. M. Cohen (1997). "Proximal-distal axis formation in the Drosophila leg." Nature **388**(6638): 139-45.

Lum, L., S. Yao, et al. (2003). "Identification of Hedgehog pathway components by RNAi in Drosophila cultured cells." Science **299**(5615): 2039-45.

Mackay, T. F., S. Richards, et al. (2012). "The Drosophila melanogaster Genetic Reference Panel." Nature **482**(7384): 173-8.

Mackay, T. F., E. A. Stone, et al. (2009). "The genetics of quantitative traits: challenges and prospects." Nat Rev Genet **10**(8): 565-77.

Mahaffey, J. W., C. M. Griswold, et al. (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**(1): 225-36.

Mann, R. S. and S. B. Carroll (2002). "Molecular mechanisms of selector gene function and evolution." Curr Opin Genet Dev **12**(5): 592-600.

McLellan, J. S., S. Yao, et al. (2006). "Structure of a heparin-dependent complex of Hedgehog and Ihog." Proc Natl Acad Sci U S A **103**(46): 17208-13.

McLellan, J. S., X. Zheng, et al. (2008). "The mode of Hedgehog binding to Ihog homologues is not conserved across different phyla." Nature **455**(7215): 979-83.

McQuilton, P., S. E. St Pierre, et al. (2012). "FlyBase 101--the basics of navigating FlyBase." Nucleic Acids Res **40**(Database issue): D706-14.

Medioni, C., S. Senatore, et al. (2009). "The fabulous destiny of the Drosophila heart." Curr Opin Genet Dev **19**(5): 518-25.

Morata, G. (2001). "How Drosophila appendages develop." Nat Rev Mol Cell Biol **2**(2): 89-97.

Morini, M., S. Astigiano, et al. (2010). "Mutually exclusive expression of DLX2 and DLX5/6 is associated with the metastatic potential of the human breast cancer cell line MDA-MB-231." BMC Cancer **10**: 649.

Munoz-Chapuli, R. and J. M. Perez-Pomares (2010). "Cardiogenesis: an embryological perspective." J Cardiovasc Transl Res **3**(1): 37-48.

Panganiban, G. (2000). "Distal-less function during Drosophila appendage and sense organ development." Dev Dyn **218**(4): 554-62.

Panganiban, G., S. M. Irvine, et al. (1997). "The origin and evolution of animal appendages." Proc Natl Acad Sci U S A **94**(10): 5162-6.

Patel, M., L. Farzana, et al. (2007). "The appendage role of insect disco genes and possible implications on the evolution of the maggot larval form." Dev Biol **309**(1): 56-69.

Pearson, J. C., D. Lemons, et al. (2005). "Modulating Hox gene functions during animal body patterning." Nat Rev Genet **6**(12): 893-904.

Phillips, M. D. and A. Shearn (1990). "Mutations in polycombeotic, a Drosophila polycomb-group gene, cause a wide range of maternal and zygotic phenotypes." Genetics **125**(1): 91-101.

Polakis, P. (2007). "The many ways of Wnt in cancer." Curr Opin Genet Dev **17**(1): 45-51.

Popov, V. M., K. Wu, et al. (2010). "The Dachshund gene in development and hormone-responsive tumorigenesis." Trends Endocrinol Metab **21**(1): 41-9.

Pueyo, J. I. and J. P. Couso (2005). "Parallels between the proximal-distal development of vertebrate and arthropod appendages: homology without an ancestor?" Curr Opin Genet Dev **15**(4): 439-46.

Roy, S., J. Ernst, et al. (2010). "Identification of functional elements and regulatory circuits by Drosophila modENCODE." Science **330**(6012): 1787-97.

Sanson, B., P. White, et al. (1996). "Uncoupling cadherin-based adhesion from wingless signalling in Drosophila." Nature **383**(6601): 627-30.

Shearn, A., G. Hersperger, et al. (1978). "Multiple Allele Approach to the Study of Genes in DROSOPHILA MELANOGASTER That Are Involved in Imaginal Disc Development." Genetics **89**(2): 355-70.

Speck, O., S. C. Hughes, et al. (2003). "Moesin functions antagonistically to the Rho pathway to maintain epithelial integrity." Nature **421**(6918): 83-7.

St Pierre, S. E., M. I. Galindo, et al. (2002). "Control of Drosophila imaginal disc development by rotund and roughened eye: differentially expressed transcripts of the same gene encoding functionally distinct zinc finger proteins." Development **129**(5): 1273-81.

Staehling-Hampton, K., P. D. Jackson, et al. (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**(6): 585-93.

Stern, D. L. (2003). "The Hox gene Ultrabithorax modulates the shape and size of the third leg of Drosophila by influencing diverse mechanisms." Dev Biol **256**(2): 355-66.

Sunkel, C. E. and J. R. S. Whittle (1987). "Brista: a gene involved in the specification and differentiation of distal cephalic and thoracic structures in Drosophila melanogaster." Development Genes and Evolution **196**(2): 124-132.

Sutherland, D., C. Samakovlis, et al. (1996). "branchless encodes a Drosophila FGF homolog that controls tracheal cell migration and the pattern of branching." Cell **87**(6): 1091-101.

Tautz, D. and C. Pfeifle (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**(2): 81-5.

Visscher, P. M., M. A. Brown, et al. (2012). "Five years of GWAS discovery." Am J Hum Genet **90**(1): 7-24.

Yan, D., Y. Wu, et al. (2010). "The cell-surface proteins Dally-like and Ihog differentially regulate Hedgehog signaling strength and range during development." Development **137**(12): 2033-44.

Yao, S., L. Lum, et al. (2006). "The ihog cell-surface proteins bind Hedgehog and mediate pathway activation." Cell **125**(2): 343-57.

Zhang, F., J. S. McLellan, et al. (2007). "Kinetic and structural studies on interactions between heparin or heparan sulfate and proteins of the hedgehog signaling pathway." Biochemistry **46**(13): 3933-41.

Zheng, X., R. K. Mann, et al. (2010). "Genetic and biochemical definition of the Hedgehog receptor." Genes Dev **24**(1): 57-71.

CHAPTER FIVE

Summary and Prospects

SUMMARY

The Drosophila gene *disconnected* (*disco*) was first identified for its role in establishing proper neural connections in Drosophila larvae (Steller, Fischbach et al. 1987). Since then, a redundant paralog, *disco-related* (*disco-r*), has been identified, and together these genes have been established as genetic Hox cofactors important for the development of ventral appendages and their larval derivatives (Mahaffey, Griswold et al. 2001; Robertson, Bowling et al. 2004; Patel, Farzana et al. 2007). The appendage development functions of *disco* genes appear to be complimentary to the body-wall/proximal appendage function of another Hox cofactor, *teashirt* (*tsh*) (Robertson, Bowling et al. 2004; Patel, Farzana et al. 2007).

One of the goals of my research was to expand our knowledge of *disco* genes beyond Drosophila and insects to identify conserved roles in development. This began with an extension of work begun previously in the lab, examining the function of the single Tribolium ortholog, *Tc-disco*. This research used a molecular marker (Tribolium *dachshund* ortholog, *Tc-dac*) to confirm that *Tc-disco* RNAi resulted in loss of the medial appendage domain in the developing Tribolium embryo. That work also examined the interactions of *Tc-disco* with the Tribolium *tsh* ortholog, *Tc-tiotsh*. Although the role for *Tc-tiotsh* in body-wall/proximal appendage was corroborated (Shippy, Tomoyasu et al. 2008), RNAi of *Tc-tiotsh* did not result in an expansion of *Tc-disco* beyond the developing legs, suggesting the Tribolium *tsh* ortholog does not repress *disco* expression, as is seen in Drosophila (Robertson, Bowling et al. 2004).

My research also sought to identify functions for *disco* orthologs in non-insect species. Prior research using *Caenorhabditis elegans* suggested that a predicted ortholog in this species is expressed in neurons. However RNAi and mutant analyses suggest that this predicted gene has no function. Attempts to observe expression of *disco* orthologs in the mouse (*Mus musculus*) were unsuccessful. Examination of expression in zebra fish (*Danio rerio*) was more fruitful, but zebra fish orthologs of *disco* were not identified in the developing pectoral fins, and use of zebra fish as a model for appendage development is questionable (Dahn, Davis et al. 2007; Davis, Dahn et al. 2007).

Finally, I presented the successful use of a genome wide association study (GWAS) to identify new genes involved in leg development. Several of these genes require further investigation before they can be integrated into a model for leg development. However, RNAi results of five candidates, including *disco-r*, indicate they may interact with the leg development gene *Distal-less* (*Dll*) in specifying antennal identity, if not leg development itself.

CONCLUSIONS AND LOOKING FORWARD

Leg patterning

The function of *Dll* in Drosophila leg development is critical (Reviewed in (Panganiban 2000; Morata 2001; Kojima 2004; Pueyo and Couso 2005; Estella, Voutev et al. 2012)). This role appears to be conserved in a number of other species (Akimenko, Ekker et al. 1994; Panganiban, Irvine et al. 1997; Beermann, Jay et al. 2001; Prpic, Janssen et al. 2003; Prpic and Tautz 2003; Angelini and Kaufman 2004; Kraus and Lufkin 2006) and has been used to argue support for *Dll* expression as a defining feature of the telopodite, also called the ‘true limb’ (Gonzalez-Crespo and Morata 1996; Kojima 2004; Estella, Voutev et al. 2012).

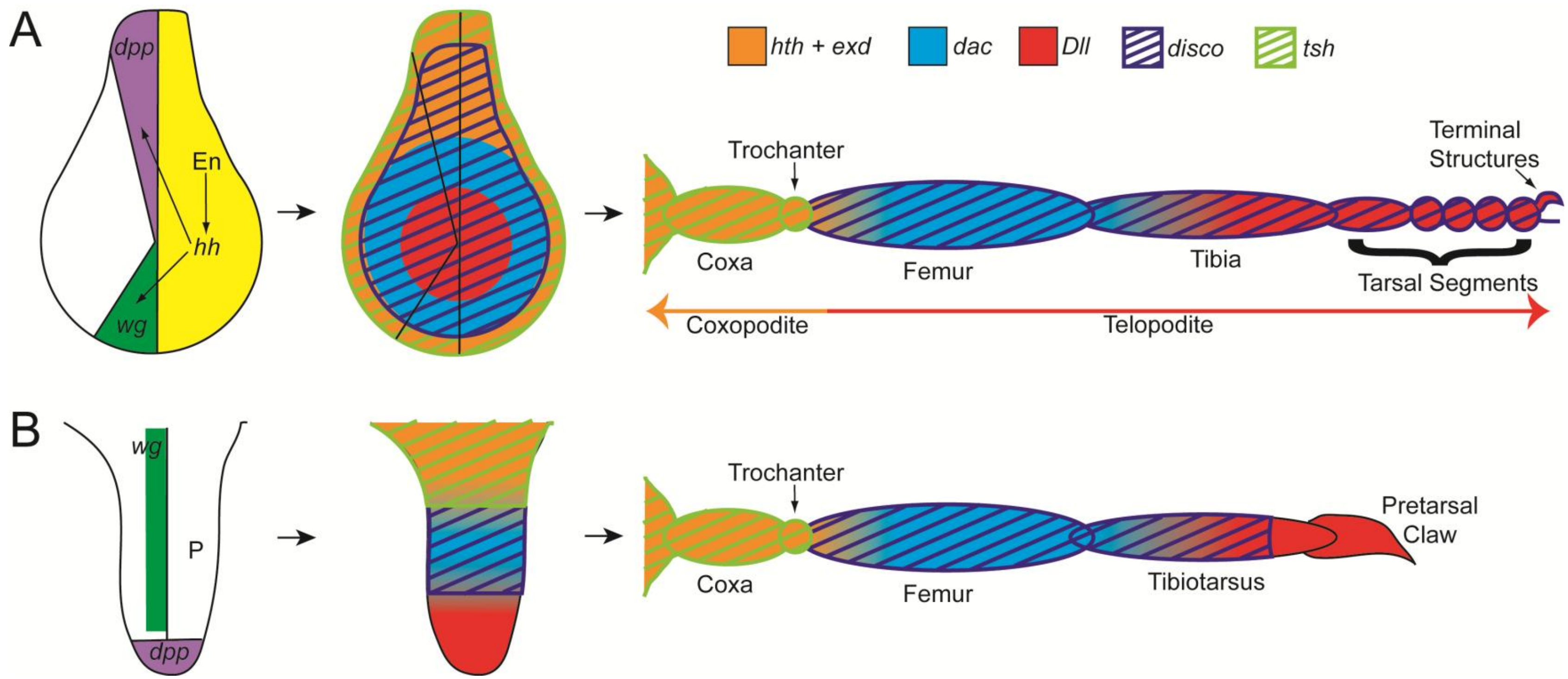
The role of *Dll* in the telopodite is juxtaposed to the combined function of *homothorax* (*hth*) and *extradentical* (*exd*) in the coxopodite, which is considered to be an extension of the body wall (Gonzalez-Crespo and Morata 1996; Kojima 2004; Estella, Voutev et al. 2012). Amongst the other genes associated with proper development of this proximal region is *tsh* (Gonzalez-Crespo and Morata 1996; Estella, Voutev et al. 2012). Both *Dll* and the *disco* genes are expressed in the early leg primordia (Cohen, Wimmer et al. 1991), and appear to be needed to pattern segments distal to the coxopodite (Cohen and Jurgens 1989; Patel, Farzana et al. 2007). This might suggest a role for the *disco* genes in helping *Dll* define the telopodite. In this model, I propose that *disco* and *tsh* establish mutually exclusive domains that serve to help define these broad leg features (Fig 1A), much as *Dll* and *hth*.

However, examination of *disco* expression and function in Tribolium might suggest otherwise. While function and expression of *Tc-Dll* appears to support its role in defining the telopodite of this species (Beermann, Jay et al. 2001), expression of *Tc-disco* in Tribolium embryos is limited to the medial appendage domain, and reduction of *Tc-disco* function results in loss or reduction of only medial portions of the legs (Fig 1B)(Patel, Farzana et al. 2007). Further, *Tc-tiotsh* does not appear to limit the *Tc-disco* domain, as it does in the Drosophila embryo. This seems to suggest that if *disco* genes help define the telopodite in Drosophila, then this feature is not conserved in all insects.

Some questions will need to be answered to confirm these conclusions. First, it has not yet been established that the expression domains of *disco* and *tsh* in the leg imaginal disc

FIGURE 1: Comparison of leg patterning between Drosophila and Tribolium

A) (Left) In Drosophila, leg discs are patterned when En expression in the posterior portion of the disc activates Hh signaling (yellow), which in turn activates *dpp* (purple) dorsally and *wg* (green) ventrally, in the anterior of the disc. (Middle) These morphogens then establish the pattern of concentric rings formed by *Dll* (red), *dac* (blue), and *hth/exd* (orange). Expression of *disco* (dark blue stripes) is also present in much of the disc, while *tsh* (lime stripes) is limited to the outer edge, which will form the proximal appendage. (Right) The leg segments to which these genes contribute can be seen. Further, *Dll*, *dac*, and *disco* appear to define the telopodite, while *hth/exd* and *tsh* contribute to the coxopodite. B) (Left) Ventral view of Tribolium limb bud. Expression of *dpp* (purple) is at the leg tip, while *wg* (green) is expressed in an anterior, ventral stripe. P = posterior. Proximal at top. (Middle) The expression patterns of *Dll* (red), *dac* (blue), and *hth/exd* (orange) resemble their Drosophila orthologs. *tsh* (lime stripes), as in Drosophila, is limited to the proximal portion, but *disco* (dark blue stripes) is only expressed in the middle leg domain. (Right) A representation of a Tribolium larval leg, which do not have a tarsal segment. The most distal segment is a claw-like segment called the pretarsal claw, while the next most distal segment has been assigned the designation tibiotarsus. Note that the areas defined by the combination of *Tc-Dll* and *Tc-dac* are believed to be part of the telopodite of the Tribolium limb.


A
dpp
En
hh
wg
hth + exd
dac
Dll
disco
tsh
Trochanter
Terminal
Structures
Coxa
Femur
Tibia
Tarsal Segments
Coxopodite
Telopodite
B
wg
P
dpp
Trochanter
Pretarsal
Claw
Coxa
Femur
Tibiotarsus

are mutually exclusive. While there is very little overlap between the two (Patel, Farzana et al. 2007), there is no evidence that either one represses the other in the leg imaginal disc. If this model were true, I would expect one gene to repress the other in leg development, much as *tsh* restricts *disco* expression in the embryo (Robertson, Bowling et al. 2004). It would make sense, given its domain of expression, that *disco* is one of many genes that help restrict the expansion of proximal genes into the telopodite domain.

Another question that should be addressed is the matter of interactions between *disco* and *dac*. In Drosophila, the domain of *dac* expression is part of the telopodite, and some models suggest *Dll* is required for *dac* activation (Giorgianni and Mann 2011). In Tribolium, distal expression of *Tc-dac* begins within the *Tc-Dll* expression domain (Prpic, Wigand et al. 2001), suggesting a similar relationship. If *disco* genes are necessary for telopodite formation and patterning, then they might play a role in controlling *dac* expression. Indeed, ectopic expression of *disco* in the wing imaginal disc of Drosophila is able to activate ectopic *dac* (Patel, Farzana et al. 2007), though it is not certain if this activation is direct, since *Dll* expression is also altered in these discs. To clarify this matter, it would be helpful to determine if the altered *Dll* domain, the ectopic *disco*, the combined function of both genes, or some other factor is responsible for activating *dac* in these altered wing discs, and what relationship this activity has to normal leg development. In Tribolium, it does not appear that the *disco* ortholog is required for *Tc-dac* activation, though *Tc-disco* may be required for its maintenance. To determine this, it should be ascertained whether the loss of the medial appendage due to *Tc-disco* RNAi is the result of reduced *Tc-dac* expression, or if the reverse is true instead. My own observations lead me to believe that *Tc-disco* is important not just for maintaining these cells, but also for maintaining the expression of *Tc-dac*. I am surprised that RNAi of *Tc-disco* did not result in total loss of *Tc-dac* expression, since *Tc-disco* appears to be expressed first, and I wonder if a null mutant of *Tc-disco* might never have this domain of *Tc-dac* expression.

While further work is still needed to fully understand all the interactions between *Dll* and *disco,* they have been studied to some extent, suggesting that *disco* requires *Dll* for activation in the leg primordia, and that each maintains the other in the developing leg disc (Cohen, Wimmer et al. 1991; Dey, Zhao et al. 2009). This seems to support a role for the *disco* genes in helping *Dll* define the telopodite. It remains to be determined to what extent these interactions hold true in the Tribolium appendage. Both genes are active before the limb buds take shape (Beermann, Jay et al. 2001; Patel, Farzana et al. 2007), but which is first, and whether either activates the other needs to be ascertained. It might be that early activation of *Tc-disco* is needed to help define the telopodite, even though it is not expressed in the entire structure all the way through leg development. But if this is true, this would not be unlike what has been seen with *Dll* in the beetle and fly (Prpic, Wigand et al. 2001; Estella, Voutev et al. 2012). I suspect *Tc-Dll* expression is first, and necessary for *Tc-disco* expression. Since both of these genes seem to be expressed prior to *Tc-dac,* it seems likely that they work together to activate and maintain *Tc-dac* expression. It might be that the reduced *Tc-disco* was sufficient to help activate *Tc-dac*, but not maintain it. This may be the result of the natural loss of *Tc-Dll* expression from the presumptive *Tc-dac* domain.

Finally, in order to establish true conservation, more species must be examined. The early expression of *Dll* orthologs in the as-yet unshaped appendages in several arthropod species, as well as the extent of their expression in the growing limb buds of these species (Gonzalez-Crespo and Morata 1996; Beermann, Jay et al. 2001; Prpic, Janssen et al. 2003; Prpic and Tautz 2003; Angelini and Kaufman 2004), seems to support a conserved role for *Dll* genes in allocating the appendage primordium and defining the telopodite. To my knowledge, Drosophila and Tribolium are the only arthropod species in which *disco* orthologs have been studied. Expanding the study of *disco* into other species will clarify whether this gene has a truly conserved role in appendage development. Further, it will help determine what this role is, and how it may relate to *Dll* and *dac* orthologs in these species, especially with respect to the formation and patterning of the telopodite. It would not be surprising to find that the *disco* orthologs in arthropods tend to be expressed in and to pattern the medial leg domain.

Evolutionary origins of legs

In addition to patterning legs in Drosophila, both *Dll* and *disco* are expressed in and required for proper development of the nervous system and various sensory organs (Steller, Fischbach et al. 1987; Sunkel and Whittle 1987; Cohen and Jurgens 1989; Lee, Freeman et al. 1991; Lee, Mukhopadhyay et al. 1999; Panganiban 2000). Indeed, *Dll* orthologs have been found to be expressed in the nervous systems of other arthropod species, including Tribolium (Beermann, Jay et al. 2001), the millipede, *Glomeris marginata* (Prpic and Tautz 2003), a cockroach (Blagburn 2007), and a crustacean (Williams, Nulsen et al. 2002). In 1997, Panganiban and colleagues examined the expression of *Dll* orthologs in several species from a variety of taxa, and found expression of these genes in limbs and other types of outgrowths.

Further, all of these species, including the limbless nematode, *Caenorhabditis elegans*, have expression of *Dll* in the developing nervous system (See also (Aspock and Burglin 2001)). This suggests that the role of *Dll* genes in these outgrowths may have originated in an ancestor through outgrowths of the nervous system. In other words, limbs may have originated evolutionarily as extensions of sensory organs, permitting early species to explore their surroundings without exposing their entire body to danger (See also (Pueyo and Couso 2005)). If true, I would expect to see other leg development genes associated with neuronal expression and function. So it is not surprising to find *dac* and *hth* orthologs (along with the *hth* cofactor, *exd*) are also expressed, not just in conserved regions of developing appendages, but also in the nervous systems of other arthropod species (Prpic, Wigand et al. 2001; Prpic, Janssen et al. 2003; Prpic and Tautz 2003; Angelini and Kaufman 2005; Blagburn 2007) This conservation does not seem to be limited to invertebrates. Expression and functional studies in vertebrates have also revealed appendage and nervous system roles for these genes in zebra fish, chicks, mice and frogs (Dirksen, Mathers et al. 1993; Akimenko, Ekker et al. 1994; Merlo, Zerega et al. 2000; Popperl, Rikhof et al. 2000; Waskiewicz, Rikhof et al. 2001; Hammond, Hill et al. 2002; Candiani, Kreslova et al. 2003; Kida, Maeda et al. 2004; Sanchez-Guardado, Irimia et al. 2011). As with *Dll*, orthologs of these genes are present in the nematode, where they contribute to development of the nervous system (Van Auken, Weaver et al. 2002; Colosimo, Brown et al. 2004; Yang, Sym et al. 2005).

Although most studies have focused on the genes traditionally associated with leg development, there is information that *tsh* genes function in the development of the nervous system and sensory organs of various species, including arthropods (Wu and Cohen 2000;

Caubit, Tiveron et al. 2005; Herke, Serio et al. 2005) and vertebrates (Caubit, Tiveron et al. 2005; Koebernick, Kashef et al. 2006; Wang, Lee et al. 2007). I have also found expression of potential *disco* orthologs in the nervous systems of C. elegans and the zebra fish, the latter of which is corroborated by another group (Lang, Patterson et al. 2009). Further, expression of mouse *disco* orthologs have been reported in the developing nervous system (Vanhoutteghem, Bouche et al. 2011). Altogether, these data reinforce the hypothesis that the ancestral origin of limbs was as outgrowths of the nervous system.

In this study are presented the results of a GWAS that successfully identified several new genes in Drosophila leg development. It would be of particular interest to determine if these genes have any developmental roles in the appendages and nervous systems of other species. Indeed most candidates have potential orthologs in other species (Table 1). Further, several of these genes may have expression in the nervous systems of embryos (Table 1). If these candidates have conserved roles in appendage and nervous system development, such function would provide further evidence for the evolution of appendages to have begun as outgrowths of the nervous system (Panganiban, Irvine et al. 1997; Pueyo and Couso 2005), and provide a bridge for further understanding of the interrelated functions of the nervous system and appendage development roles for these genes (Panganiban 2000).

To begin, it will be important to clarify the functions of these candidates in Drosophila leg development. Identifying which known leg development genes interact (either directly or indirectly) with these candidates will be helpful. From there, it can be determined

TABLE 1: GWAS candidates with the possibility of orthologs and neural expression

Potential orthologs of candidate genes were identified using BLAST (Altschul, Gish et al. 1990; Altschul, Madden et al. 1997) to search the encoded protein sequence against a database of translated DNA sequences on the NCBI website (www.ncbi.nlm.nih.gov). Orthologs are known for *disco-r* in both vertebrates and insects (Romano, Li et al. 2004; Vanhoutteghem and Djian 2004; Vanhoutteghem and Djian 2006; Patel, Farzana et al. 2007; Vanhoutteghem, Bouche et al. 2011) and for *boi* in vertebrates (Yao, Lum et al. 2006). Only CG43444 showed specific staining in the nervous system; the rest had either no such expression, or were ubiquitous.

Gene Name	Insect orthologs	Vertebrate orthologs	Embryonic neuronal expression
boi	Probable	Yes	No
CG43444	Probable	Probable	Yes
eIF6	Probable	Probable	No
sick	Probable	Probable	Possible
PI31	Probable	Probable	No
CG9129	No	No	No
CG32333	Probable	Probable	No
CG30371	Probable	No	No
CG8486	Probable	Probable	Possible
bnl	Probable	Probable	No
disco-r	Yes	Yes	Yes
chinmo	Probable	No	No
Cka	Probable	Probable	No
CSN1b	Probable	Probable	Possible
mys	Probable	Probable	Possible
CG6841	Probable	Probable	Possible

how those interactions contribute to leg patterning. Ascertaining if candidate orthologs act in leg development in other species could help in answering these questions.

In turn, helping to answer these questions may aid in determining if and how these genes contribute to the development of the nervous system. For example, as I have discussed, *disco* and *Dll* likely interact in leg development. However, it is not known if these interactions are also present in the developing nervous system. If such interactions do occur, then illuminating them will help broaden our understanding of *Dll* function in nervous system development. Further, this could help broaden our understanding of the evolutionary origin of appendages. If appendages originated as outgrowths of the nervous system, then I would expect to see the maintenance of networks of interaction between leg and nervous system development. I would also expect that these interactive networks would be relatively conserved between species.

***disco* genes in cancer**

In addition to providing evolutionary insights, understanding how these candidates function in leg development may help expand our knowledge of cancer. The leg imaginal disc is a dynamic tissue, which requires factors that both promote and limit cellular growth and proliferation to ensure appropriate proportions are reached with the rest of the body. Factors specifying identity are important not just for the disc as a whole, but for each of the future leg segments. To aid in appropriate identity and structure, factors are needed to provide appropriate cellular adhesion and interactions. In the Drosophila leg, the genes *hedgehog*, *wingless*, *Dll*, *dac*, and *hth*/*exd* are known to contribute to identity of and within the leg disc and to its growth (Lecuit and Cohen 1997; Panganiban 2000; Morata 2001; Kojima 2004; Pueyo and Couso 2005; Estella, Voutev et al. 2012). If these functions are conserved in vertebrates, then it would not be surprising to see that improper expression of these genes could result in diseases like cancer (Morata 2001). Indeed, misregulation of the Hh (Briscoe and Therond 2005) and Wg (Wnt) (Polakis 2007) pathways, as well as altered expression of *Dll* (*Dlx*) (Xu and Testa 2009; Morini, Astigiano et al. 2010; Tan, Cheung et al. 2010), *dac* (*dach*) (Popov, Wu et al. 2010) and *exd* (*pbx*) (Moens and Selleri 2006) have been identified in a variety of cancers.

Future studies in cancer research might benefit from examining not just the regulation of one gene or gene family, but also their connections to other known cancer genes. Such information could be gleaned from studies in model organisms, including insects. The well studied network of genes known to play a role in Drosophila leg development provides an excellent model. For example, *wg* and *Dll* work together to allocate and pattern the highly proliferative leg primordium in Drosophila and Tribolium (Grossmann, Scholten et al. 2009; Estella, Voutev et al. 2012). It seems possible that their orthologs might also work together to create aberrant proliferative tissues in vertebrates. Recognizing such connections might prove more useful in providing accurate prognoses and tailoring therapeutics.

As with other known leg factors, *disco* genes seem to play some role in specifying cell identity in Drosophila and Tribolium, since cells that are supposed to express these genes are lost when this expression is perturbed (Patel, Farzana et al. 2007). The fates of these cells are being explored in Drosophila, but it is still not clear what is happening to the *Tc-dac*-expressing cells in *Tc-disco* RNAi embryos. Primarily, it is not certain if these cells fail to proliferate or if they are killed. Understanding how *disco* might enact these roles could provide useful information in cancer studies. A 2004 study (Cui, Elsam et al. 2004) reported an elevated expression of *disco* ortholog *bnc1* in basal cell carcinomas. This is one of the first cell types in which *bnc1* expression was identified (Tseng and Green 1992), and it is needed for proper proliferation of these cells (Tseng and Green 1992; Tseng 1998; Zhang and Tseng 2007). So it is perhaps not so surprising that increased expression of *bnc1* in basal cells is associated with carcinomas. This might support a role for *disco* genes in proliferation, but does not rule out the possibility that *bnc1* also promotes survival. Since these functions also appear to be important in Drosophila and Tribolium leg development, these model organisms could be useful tools to piece together the pathway *bnc* genes might use to encourage proliferation and survival in similar tumors.

On the other hand, three studies have associated reduced expression of *bnc* genes with cancers. Two separate studies report hypermethylation of *bnc1* in lung, breast, prostate and colon cancers (Shames, Girard et al. 2006) and in leukemia (Dunwell, Hesson et al. 2009). The third study identified loss or reduction of *bnc2* expression in esophageal adenocarinomas (EAC), and showed a reduction in clonal proliferation when EAC cell lines were transfected with a *bnc2* construct (Akagi, Ito et al. 2009). In these cases, it would seem the *disco* orthologs prevent proliferation in some way. These effects may be tissue specific, which has been shown for *dac* ortholog function in cancer (Popov, Wu et al. 2010). If so, it would be interesting to understand how *bnc* genes promote a non-proliferative state in the normal tissues, and if there are comparable pathways active in Tribolium and Drosophila.

Conclusion

The goal of my research has been to expand our knowledge of factors contributing to appendage development. This has involved attempting to explore functions of disco genes in non-Drosophila species. This has also involved, with the help of my colleagues, the successful identifications of new genes involved in Drosophila leg development. Further work is needed to broaden our understanding of *disco* genes, in both vertebrates and invertebrates, before we can conclude how conserved *disco*'s role in leg development truly is. Similarly, the new genes identified in the GWAS presented here require further characterization to pinpoint their place in a model of Drosophila leg development, and what conserved functions they may possess. Broadening our understanding in these areas will likely provide insight into evolution, development, and cancer.

References:

Akagi, T., T. Ito, et al. (2009). "Chromosomal abnormalities and novel disease-related regions in progression from Barrett's esophagus to esophageal adenocarcinoma." Int J Cancer **125**(10): 2349-59.

Akimenko, M. A., M. Ekker, et al. (1994). "Combinatorial expression of three zebrafish genes related to distal-less: part of a homeobox gene code for the head." J Neurosci **14**(6): 3475-86.

Altschul, S. F., W. Gish, et al. (1990). "Basic local alignment search tool." J Mol Biol **215**(3): 403-10.

Altschul, S. F., T. L. Madden, et al. (1997). "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs." Nucleic Acids Res **25**(17): 3389-402.

Angelini, D. R. and T. C. Kaufman (2004). "Functional analyses in the hemipteran Oncopeltus fasciatus reveal conserved and derived aspects of appendage patterning in insects." Dev Biol **271**(2): 306-21.

Angelini, D. R. and T. C. Kaufman (2005). "Functional analyses in the milkweed bug Oncopeltus fasciatus (Hemiptera) support a role for Wnt signaling in body segmentation but not appendage development." Dev Biol **283**(2): 409-23.

Aspock, G. and T. R. Burglin (2001). "The Caenorhabditis elegans distal-less ortholog ceh-43 is required for development of the anterior hypodermis." Dev Dyn **222**(3): 403-9.

Beermann, A., D. G. Jay, et al. (2001). "The Short antennae gene of Tribolium is required for limb development and encodes the orthologue of the Drosophila Distal-less protein." Development **128**(2): 287-97.

Blagburn, J. M. (2007). "Co-factors and co-repressors of Engrailed: expression in the central nervous system and cerci of the cockroach, Periplaneta americana." Cell Tissue Res **327**(1): 177-87.

Briscoe, J. and P. Therond (2005). "Hedgehog signaling: from the Drosophila cuticle to anti-cancer drugs." Dev Cell **8**(2): 143-51.

Candiani, S., J. Kreslova, et al. (2003). "Cloning and developmental expression of amphioxus Dachschund." Gene Expr Patterns **3**(1): 65-9.

Caubit, X., M. C. Tiveron, et al. (2005). "Expression patterns of the three Teashirt-related genes define specific boundaries in the developing and postnatal mouse forebrain." J Comp Neurol **486**(1): 76-88.

Cohen, B., E. A. Wimmer, et al. (1991). "Early development of leg and wing primordia in the Drosophila embryo." Mech Dev **33**(3): 229-40.

Cohen, S. M. and G. Jurgens (1989). "Proximal-distal pattern formation in Drosophila: cell autonomous requirement for Distal-less gene activity in limb development." EMBO J **8**(7): 2045-55.

Colosimo, M. E., A. Brown, et al. (2004). "Identification of thermosensory and olfactory neuron-specific genes via expression profiling of single neuron types." Curr Biol **14**(24): 2245-51.

Cui, C., T. Elsam, et al. (2004). "Gli proteins up-regulate the expression of basonuclin in Basal cell carcinoma." Cancer Res **64**(16): 5651-8.

Dahn, R. D., M. C. Davis, et al. (2007). "Sonic hedgehog function in chondrichthyan fins and the evolution of appendage patterning." Nature **445**(7125): 311-4.

Davis, M. C., R. D. Dahn, et al. (2007). "An autopodial-like pattern of Hox expression in the fins of a basal actinopterygian fish." Nature **447**(7143): 473-6.

Dey, B. K., X. L. Zhao, et al. (2009). "Mutual regulation of the Drosophila disconnected (disco) and Distal-less (Dll) genes contributes to proximal-distal patterning of antenna and leg." Cell Tissue Res **338**(2): 227-40.

Dirksen, M. L., P. Mathers, et al. (1993). "Expression of a Xenopus Distal-less homeobox gene involved in forebrain and cranio-facial development." Mech Dev **41**(2-3): 121-8.

Dunwell, T. L., L. B. Hesson, et al. (2009). "Epigenetic analysis of childhood acute lymphoblastic leukemia." Epigenetics **4**(3): 185-93.

Estella, C., R. Voutev, et al. (2012). "A dynamic network of morphogens and transcription factors patterns the fly leg." Curr Top Dev Biol **98**: 173-98.

Giorgianni, M. W. and R. S. Mann (2011). "Establishment of medial fates along the proximodistal axis of the Drosophila leg through direct activation of dachshund by Distalless." Dev Cell **20**(4): 455-68.

Gonzalez-Crespo, S. and G. Morata (1996). "Genetic evidence for the subdivision of the arthropod limb into coxopodite and telopodite." Development **122**(12): 3921-8.

Grossmann, D., J. Scholten, et al. (2009). "Separable functions of wingless in distal and ventral patterning of the Tribolium leg." Dev Genes Evol **219**(9-10): 469-79.

Hammond, K. L., R. E. Hill, et al. (2002). "Isolation of three zebrafish dachshund homologues and their expression in sensory organs, the central nervous system and pectoral fin buds." Mech Dev **112**(1-2): 183-9.

Herke, S. W., N. V. Serio, et al. (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**(1): 27-34.

Kida, Y., Y. Maeda, et al. (2004). "Chick Dach1 interacts with the Smad complex and Sin3a to control AER formation and limb development along the proximodistal axis." Development **131**(17): 4179-87.

Koebernick, K., J. Kashef, et al. (2006). "Xenopus Teashirt1 regulates posterior identity in brain and cranial neural crest." Dev Biol **298**(1): 312-26.

Kojima, T. (2004). "The mechanism of Drosophila leg development along the proximodistal axis." Dev Growth Differ **46**(2): 115-29.

Kraus, P. and T. Lufkin (2006). "Dlx homeobox gene control of mammalian limb and craniofacial development." Am J Med Genet A **140**(13): 1366-74.

Lang, M. R., L. B. Patterson, et al. (2009). "Basonuclin-2 requirements for zebrafish adult pigment pattern development and female fertility." PLoS Genet **5**(11): e1000744.

Lecuit, T. and S. M. Cohen (1997). "Proximal-distal axis formation in the Drosophila leg." Nature **388**(6638): 139-45.

Lee, K. J., M. Freeman, et al. (1991). "Expression of the disconnected gene during development of Drosophila melanogaster." EMBO J **10**(4): 817-26.

Lee, K. J., M. Mukhopadhyay, et al. (1999). "Autoregulation of the Drosophila disconnected gene in the developing visual system." Dev Biol **214**(2): 385-98.

Mahaffey, J. W., C. M. Griswold, et al. (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**(1): 225-36.

Merlo, G. R., B. Zerega, et al. (2000). "Multiple functions of Dlx genes." Int J Dev Biol **44**(6): 619-26.

Moens, C. B. and L. Selleri (2006). "Hox cofactors in vertebrate development." Dev Biol **291**(2): 193-206.

Morata, G. (2001). "How Drosophila appendages develop." Nat Rev Mol Cell Biol **2**(2): 89-97.

Morini, M., S. Astigiano, et al. (2010). "Mutually exclusive expression of DLX2 and DLX5/6 is associated with the metastatic potential of the human breast cancer cell line MDA-MB-231." BMC Cancer **10**: 649.

Panganiban, G. (2000). "Distal-less function during Drosophila appendage and sense organ development." Dev Dyn **218**(4): 554-62.

Panganiban, G., S. M. Irvine, et al. (1997). "The origin and evolution of animal appendages." Proc Natl Acad Sci U S A **94**(10): 5162-6.

Patel, M., L. Farzana, et al. (2007). "The appendage role of insect disco genes and possible implications on the evolution of the maggot larval form." Dev Biol **309**(1): 56-69.

Polakis, P. (2007). "The many ways of Wnt in cancer." Curr Opin Genet Dev **17**(1): 45-51.

Popov, V. M., K. Wu, et al. (2010). "The Dachshund gene in development and hormone-responsive tumorigenesis." Trends Endocrinol Metab **21**(1): 41-9.

Popperl, H., H. Rikhof, et al. (2000). "lazarus is a novel pbx gene that globally mediates hox gene function in zebrafish." Mol Cell **6**(2): 255-67.

Prpic, N. M., R. Janssen, et al. (2003). "Gene expression in spider appendages reveals reversal of exd/hth spatial specificity, altered leg gap gene dynamics, and suggests divergent distal morphogen signaling." Dev Biol **264**(1): 119-40.

Prpic, N. M. and D. Tautz (2003). "The expression of the proximodistal axis patterning genes Distal-less and dachshund in the appendages of Glomeris marginata (Myriapoda: Diplopoda) suggests a special role of these genes in patterning the head appendages." Dev Biol **260**(1): 97-112.

Prpic, N. M., B. Wigand, et al. (2001). "Expression of dachshund in wild-type and Distal-less mutant Tribolium corroborates serial homologies in insect appendages." Dev Genes Evol **211**(10): 467-77.

Pueyo, J. I. and J. P. Couso (2005). "Parallels between the proximal-distal development of vertebrate and arthropod appendages: homology without an ancestor?" Curr Opin Genet Dev **15**(4): 439-46.

Robertson, L. K., D. B. Bowling, et al. (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**(12): 2781-9.

Romano, R. A., H. Li, et al. (2004). "Identification of Basonuclin2, a DNA-binding zinc-finger protein expressed in germ tissues and skin keratinocytes." Genomics **83**(5): 821-33.

Sanchez-Guardado, L. O., M. Irimia, et al. (2011). "Distinct and redundant expression and transcriptional diversity of MEIS gene paralogs during chicken development." Dev Dyn **240**(6): 1475-92.

Shames, D. S., L. Girard, et al. (2006). "A genome-wide screen for promoter methylation in lung cancer identifies novel methylation markers for multiple malignancies." PLoS Med **3**(12): e486.

Shippy, T. D., Y. Tomoyasu, et al. (2008). "Do teashirt family genes specify trunk identity? Insights from the single tiptop/teashirt homolog of Tribolium castaneum." Dev Genes Evol **218**(3-4): 141-52.

Steller, H., K. F. Fischbach, et al. (1987). "Disconnected: a locus required for neuronal pathway formation in the visual system of Drosophila." Cell **50**(7): 1139-53.

Sunkel, C. E. and J. R. S. Whittle (1987). "Brista: a gene involved in the specification and differentiation of distal cephalic and thoracic structures in Drosophila melanogaster." Development Genes and Evolution **196**(2): 124-132.

Tan, Y., M. Cheung, et al. (2010). "Upregulation of DLX5 promotes ovarian cancer cell proliferation by enhancing IRS-2-AKT signaling." Cancer Res **70**(22): 9197-206.

Tseng, H. (1998). "Basonuclin, a zinc finger protein associated with epithelial expansion and proliferation." Front Biosci **3**: D985-8.

Tseng, H. and H. Green (1992). "Basonuclin: a keratinocyte protein with multiple paired zinc fingers." Proc Natl Acad Sci U S A **89**(21): 10311-5.

Van Auken, K., D. Weaver, et al. (2002). "Roles of the Homothorax/Meis/Prep homolog UNC-62 and the Exd/Pbx homologs CEH-20 and CEH-40 in C. elegans embryogenesis." Development **129**(22): 5255-68.

Vanhoutteghem, A., C. Bouche, et al. (2011). "Basonuclins and disco: Orthologous zinc finger proteins essential for development in vertebrates and arthropods." Biochimie **93**(2): 127-33.

Vanhoutteghem, A. and P. Djian (2004). "Basonuclin 2: an extremely conserved homolog of the zinc finger protein basonuclin." Proc Natl Acad Sci U S A **101**(10): 3468-73.

Vanhoutteghem, A. and P. Djian (2006). "Basonuclins 1 and 2, whose genes share a common origin, are proteins with widely different properties and functions." Proc Natl Acad Sci U S A **103**(33): 12423-8.

Wang, H., E. M. Lee, et al. (2007). "Isolation and expression of zebrafish zinc-finger transcription factor gene tsh1." Gene Expr Patterns **7**(3): 318-22.

Waskiewicz, A. J., H. A. Rikhof, et al. (2001). "Zebrafish Meis functions to stabilize Pbx proteins and regulate hindbrain patterning." Development **128**(21): 4139-51.

Williams, T. A., C. Nulsen, et al. (2002). "A complex role for distal-less in crustacean appendage development." Dev Biol **241**(2): 302-12.

Wu, J. and S. M. Cohen (2000). "Proximal distal axis formation in the Drosophila leg: distinct functions of teashirt and homothorax in the proximal leg." Mech Dev **94**(1-2): 47-56.

Xu, J. and J. R. Testa (2009). "DLX5 (distal-less homeobox 5) promotes tumor cell proliferation by transcriptionally regulating MYC." J Biol Chem **284**(31): 20593-601.

Yang, L., M. Sym, et al. (2005). "The roles of two C. elegans HOX co-factor orthologs in cell migration and vulva development." Development **132**(6): 1413-28.

Yao, S., L. Lum, et al. (2006). "The ihog cell-surface proteins bind Hedgehog and mediate pathway activation." Cell **125**(2): 343-57.

Zhang, X. and H. Tseng (2007). "Basonuclin-null mutation impairs homeostasis and wound repair in mouse corneal epithelium." PLoS One **2**(10): e1087.